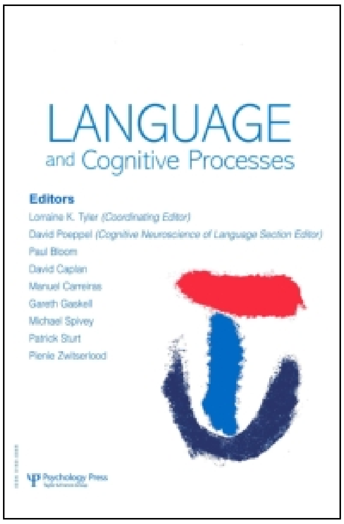

Language and Cognitive Processes
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713683153

Sentence comprehension: A parallel distributed processing approach

J. L. McClelland[a]; Mark St. John[a]; Roman Taraban[a]
[a] Department of Psychology, Carnegie Mellon University, Pittsburgh, USA

To cite this Article McClelland, J. L. , St. John, Mark and Taraban, Roman(1989) 'Sentence comprehension: A parallel distributed processing approach', Language and Cognitive Processes, 4: 3, SI287 — SI335
To link to this Article: DOI: 10.1080/01690968908406371
URL: http://dx.doi.org/10.1080/01690968908406371

LANGUAGE AND COGNITIVE PROCESSES, 1989, *4*, (3/4) SI 287–335

Sentence Comprehension: A Parallel Distributed Processing Approach

J. L. McClelland, Mark St. John, and Roman Taraban
Department of Psychology, Carnegie Mellon University, Pittsburgh, U.S.A.

In this paper, we review basic aspects of conventional approaches to sentence comprehension and point out some of the difficulties faced by models that take these approaches. We then describe an alternative approach, based on the principles of parallel distributed processing, and show how it offers different answers to basic questions about the nature of the language processing mechanism. We describe an illustrative simulation model that captures the key characteristics of the approach, and illustrate how it can cope with the difficulties faced by conventional models. We describe alternative ways of conceptualising basic aspects of language processing within the framework of this approach, consider how it can address several arguments that might be brought to bear against it, and suggest avenues for future development.

INTRODUCTION

What is constructed mentally when we comprehend a sentence? How does this constructive process occur? What role do words play in the construction process? How is the ability to construct such a representation acquired? These are some of the central questions that face any attempt to build a model of language processing.

In this paper, we present a view that differs from most existing notions about the general form of the answers to these questions. We briefly outline what we take to be a generic version of existing notions. Then, we point out some difficulties with these notions. After this, we present a sketch of an alternative that seems to us to have promise to address these problems. We illustrate the alternative by describing a preliminary model,

---

Requests for reprints should be addressed to J. L. McClelland, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213, U.S.A.

The work reported here was supported by NSF Grants BNS 86–09729 and BNS 88–12048, ONR Contracts N00014–86–G–0146 and N00014–86–K–0349, and an NIMH Career Development Award MH00385 to J. L. McClelland.

and consider how it gives different answers to some of the questions raised above, and note how it addresses the problems faced by models that give conventional answers to these questions. We examine some of the arguments, both theoretical and empirical, that have been taken as counting against this sort of alternative, and show that they can in fact be countered. Finally, we describe future directions for the further development of this approach.

CONVENTIONAL APPROACHES TO SENTENCE COMPREHENSION

The comprehension of sentences has been studied extensively, and there are many disparate views about the nature of this process. We do not mean to assert that all previous researchers have adhered to the views we describe in this section. However, quite a bit of work has been done which we believe either tacitly or explicitly adopts the views we describe here. We tend to cite the paper by Fodor and Pylyshyn (1988), because it articulates these views clearly. Where relevant, we will site works that apply these ideas and general texts where they are used or assumed.

What is Constructed when We Comprehend a Sentence?

It is typical to assume that what is constructed is an interconnected set of propositions (e.g. Clark & Clark, 1977), or *propositional representation*. The exact nature of these propositions varies from implementation to implementation, but in general they are taken to be symbolic expressions which have a combinatorial syntax and semantics (Fodor & Pylyshyn, 1988). According to Fodor and Pylyshyn, combinatorial representations are those which exhibit the following properties:

- They may be atomic or molecular expressions.
- If they are molecular, they have constituents which may be either atomic or molecular.
- The semantic content of a molecular expression is a function of the semantic content of each of the parts of the expression and of the organisation of the constituents.

What Role do Words Play in the Comprehension Process?

Implicit in many theories of comprehension is the notion that words have meanings, and that these meanings are the constituents of the meanings of the propositions that are constructed from sentences that contain these words. This view appears to underlie Fodor and Pylyshyn's (1988) principle of *compositionality:* According to this principle, "a word makes approximately the same semantic contribution to the meaning of every sentence in which it occurs." Let us use their example:

1. John loves the girl.
2. The girl loves John.

Fodor and Pylyshyn use these sentences to illustrate what they mean by compositionality. They ask us to consider the meaning of the word "loves" that appears in both of these sentences. They state that the relationship that John is said to bear to the girl in the first sentence is the same relationship that the girl is said to bear to John in the second sentence. This common relationship can be taken to be the meaning of the word "loves", and it occurs in the representation of the meaning of both of these sentences.

How does the Process of Constructing a Representation of the Propositions Underlying a Sentence Occur?

Often, this process is taken to be one of building a structural description using a system of *structure-sensitive* rules. Following Fodor and Pylyshyn, we take *structure-sensitive* to mean that the operations that apply to representations are sensitive to their form and not their content (Fodor & Pylyshyn, 1988). This means, for example, that they care only if some constituent is an item of the right very general type (N, V, NP, VP, etc.).

How is the Ability to Construct a Representation Acquired?

To the extent that we assume that the process of constructing representations of sentences proceeds by the use of structure-sensitive rules to structure the constituent expressions corresponding to words, it seems natural to assume that acquisition amounts to a process of determining what the rules are and what the constituent expressions are that words are used to designate. Researchers interested in acquisition of comprehension skill do not of course assume that the rules that are actually used in comprehension are the same rules that characterise the abstract linguistic competence of the speaker-hearer, but they are rules none the less.

Summary

In brief, the comprehension of sentences is generally taken to be the process whereby a listener uses a set of structure-sensitive rules to construct a propositional representation that constitutes the "meaning" of the sentence. The constituents of this representation include the meanings of the words in the sentence. Following Fodor and Pylyshyn's terminology, we call this view the *classical* view. These authors intend it to be taken as applying more broadly than to just the interpretation of sentences, but they make clear that language is a "paradigm of systematic cognition". We will not have anything to say about its broader applicability; instead, we will focus on the reasons why we feel that it may be worth seeking an alternative framework for addressing the problem of language comprehension.

PROBLEMS FOR THE CLASSICAL VIEW OF SENTENCE COMPREHENSION

Conceptual Guidance and Rule Conflicts

A central problem for the conventional view is the fact that sentence interpretations cannot in general be recovered correctly from structure-sensitive rules alone. Even those who try to go the farthest using structure-sensitive rules (Frazier, 1986; Marcus, 1980) are accurately aware of this problem. The problem is not just a curiosity; it comes up almost every time a prepositional phrase is encountered. Consider, for example:

3. The spy saw the policeman with binoculars.
4. The spy saw the policeman with a revolver.

In (3), most readers interpret the binoculars as the instrument used by the spy in seeing the policeman. In (4), most readers interpret the revolver as a possession of the policeman. This simple example illustrates clearly that it is necessary at a minimum to consider whether the object of the prepositional phrase is a plausible candidate for use as an instrument of the verb. In general, as the next example makes clear, it is also necessary to consider whether in fact the subject of the sentence might be the kind of actor that can use the instrument:

5. The bird saw the birdwatcher with binoculars.

Indeed, Oden (1978) has shown that every constituent of sentences like (3)–(5) can potentially influence the interpretation of the role of the prepositional phrase.

It is widely accepted that the *ultimate* interpretation that a sentence receives is affected by content. Many researchers accept this, but resist the idea that the initial processing of attachment ambiguities is influenced by non-syntactic content. Thus, for example, Frazier (1986) has proposed that initial parsing decisions are based on a purely syntactic mechanism that proposes its preferred alternative for consideration by semantic processes. Later in the paper, we review empirical evidence relevant to this claim. For the moment, we point out a more conceptual problem with it. The difficulty is that the decision as to which interpretation of an ambiguous sentence will win out in the end does not seem in general to be based on a simple yes–no decision about the acceptability of the supposedly syntactically preferred interpretation (Crain & Steedman, 1985). Thus in (5), it is not really plausible to argue that the interpretation in which the bird is using the binoculars as instrument is strictly blocked. For example, we have less difficulty accepting such an interpretation in "The bird saw its prey with binoculars", even if we find it somewhat odd for a bird to be using an instrument. Rather, it appears that the alternative interpretation is simply more plausible in the case of (5). It thus appears that more than one alternative interpretation must be evaluated for plausibility, thereby robbing the parser of any special role in providing a single alternative for consideration.

It is also important to note that it is not simply the case that decisions can either be made by syntactic rule or need to be left for semantic determination. As Marcus (1980) points out, language comprehenders have preferences for syntactic interpretation which must be seen as matters of degree, and therefore they sometimes win and sometimes loose when placed in conflict with other considerations. Very clear examples of this arise in sentences like (6) and (7):

6. We ate some food with some friends that we like.
7. We found a painting in the attic that was covered with cobwebs.

A structure-sensitive rule would allow us to parse (6) correctly, based on the idea that relative clauses should be taken to attach to the immediately preceding nounphrase rather than an earlier one, especially when, as in this case, attachment to the earlier nounphrase would violate the so-called "no-crossover" constraint. However, it is exactly this constraint that is violated in (7), where it is the painting, rather than the attic, which native speakers take to have been covered with cobwebs. Violating this constraint may make the sentence seem a bit awkward, but it does not prevent the cobwebs from attaching to the painting.

Contextual Shading as well as Selection of Word Meaning

The problem of word-meaning indeterminacy also poses a problem for conventional approaches. It is, of course, typical to assume that an individual word can have more than one meaning. The problem of sentence interpretation, then, is seen as one of selecting the right meaning from a set of possible meanings that are stored in a "mental lexicon". One problem with this is the potential combinatorial explosion that can result, as discussed below. Here we focus on a different problem: The problem is that it seems rather limiting to suppose that the range of meanings that a word can have is restricted in advance to the set of known usages of the word. Let us consider some examples:

8. The hostess threw the ball for charity.
9. The slugger hit the ball over the fence.
10. The baby rolled the ball to her daddy.

The distinctions between the meanings of ball as it appears in (8) and (9) seem well enough captured by the idea that the specification of a meaning for this word involves a selection of one of two alternatives, one that means something like "fancy dance" and one that means something like "spherical object". But in (10), it seems that the specification of the ball is somewhat different from the specification that we get from (9). It is possible to assert that here again we are selecting between two alternative meanings—one, let us say, in which the spherical object is smallish, hard and white, and the other in which it is larger, squishier, and probably multicoloured. But taken to its extreme, this view seems to lead to a vast explosion of lexical entries, one for each of the possible balls that we can envisage being implicitly described in a sentence. Is there to be a separate lexical entry for every shade of meaning that can be comprehended, for every word in the language?

A Similar Problem with Roles

A similar problem arises when we attempt to specify the set of thematic roles that are available to be filled by word meanings in the structural description that represents a sentence. In early work on roles (Fillmore, 1968), attempts were made to enumerate the set of roles that constituents could fill. However, this effort quickly ran into the problem that there are a large number of slight distinctions among roles, all of which have interpretive significance. The problem is so bad that many workers have taken the tack of assuming that for each verb there is an idiosyncratic set of roles. This is, of course, not terribly satisfactory either, because this simply obscures the broad commonality that does exist among, for example, the constituents which we would tend to call agents if we did not look too closely.

Implied Constituents

The notion that the representation of a sentence consists of an assemblage of representations of constituents of a sentence fails to provide any direct way of understanding why it is that many sentences convey implied constituents which native speakers do not need to hear mentioned. Thus in 11 and 12,

11. The boy spread the jelly on the bread.
12. The man stirred his coffee.

we can infer a knife and a spoon, respectively. That such inferred constituents are expected to be parts of the representations we form in listening to sentences, is indicated by the fact that we can refer to them as though they have been mentioned. Thus we can say, for example,

13. The boy spread the jelly on the bread.
    The knife was covered with poison.

and we can expect the reader to know that someone is in danger of being poisoned if they eat the sandwich.

Now, typically, it would be conventional to assume either that implied constituents are parts built into the representations of the lexical items (e.g. the knife is built into the representation of the verb *spread*) or that they are inferred by post-processes. However, it is by no means an easy task to decide when something should be built in; nor is it easy to decide when something should be inferred. We do not *always* stir coffee with a spoon, and we do not even *necessarily* spread jelly with a knife; therefore, drawing an inference in an all-or-nothing way can lead to over-commitment. We might draw inferences and assign them strengths, but there is no end to the inferences that we might draw. Should we draw all of them? Where should the line be drawn? These problems have plagued inference-based comprehension programs for years (Schank, 1981).

Combinatorial Explosion or Premature Commitment?

The multiplicity of alternative meanings of words and of possible roles, and the wide range of possible inferences which might follow from each possible combination of roles and meanings, becomes an extremely serious problem when we consider the implications for processing. Famous examples like

14. Time flies like an arrow.

remind us of the potential combinatorial explosion associated with the multiplicity of possible word-meaning and structural possibilities that arise in processing virtually every sentence. Models built in the classical tradition are forced to take one of two approaches to this problem: Either they can create a potentially exponential number of possible interpretations, or they can make an early commitment to pursue only a limited range of alternatives. In the extreme form, a single track is chosen, subject to backtracking if that track turns out to fail.

The Difficulty of Acquisition

As a final note, we remind the reader of the problem of acquisition. Several serious problems face anyone who attempts to build a model of acquisition of the rules and word meanings posited by the classical view:

- The rules are often overridden, as we saw above.
- The correct choice of rules is drastically underdetermined by the evidence available to the child.
- The feedback children receive on the correctness of their constructions is notoriously impoverished.
- A given sentence may have more than one perfectly acceptable interpretation. This makes it hard to know when to reject a rule as wrong or simply not always right.
- Correct performance requires not only the knowledge of the constraints but how much weight each one should be given.
- The child faces a very serious boot-strapping problem in learning to map sentences on to their meanings. This problem is reviewed by Gleitman and Wanner (1982).

These and other problems have led many psycholinguists committed to the view that acquisition involves learning rules to the view that acquisition is impossible. Instead, it has often been proposed that the rules of all languages are innate and that acquisition simply amounts to setting parameters where there are degrees of freedom. It has even been proposed (e.g. Chomsky, 1988) that it is not implausible to imagine that all concepts are innate.

Summary

We do not wish to make light of classical models. Such models do have considerable appeal, and they seem to us to capture approximately some of the general characteristics of natural languages. Indeed, there are regularities in the way we structure sentences which give clues to the ideas we wish these sentences to convey; and there are regularities in the ways in which we use words. These two facts seem consistent with the idea that words have meanings that are parts of the meanings of the sentences that they occur in and that the meanings of the wholes are constructed from these parts by structure-sensitive rules. Fodor and Pylyshyn (1988) are, of course, correct when they point to the productivity and systematicity of language, and it is no mean accomplishment of the classical view that it captures these essential characteristics of natural language.

But it is our view that the classical approach is destined to remain strapped with many, if not all, of the problems listed above. Of course, others have taken a different view, and many proposals have been made for augmenting or tuning classically based models of sentence processing. Thus, for example, we find that many current researchers working within the classical tradition allow lexical information associated with the heads of constituents to be referred to in parsing. Similarly, techniques such as beam search can be used to find a reasonable compromise between the combinatorial explosion that results from computing all possible parses and the premature commitment that arises from computing only one; essentially, one simply computes the best few and hopes that what turns out to be the correct interpretation happens to be among them.

It is, of course, possible that an accumulation of incremental fixes of this kind will ultimately provide an adequate framework for modelling the sentence comprehension process. But our bet is that it will not. Using lexical information associated with heads of constituents does not solve the whole problem of content sensitivity of parsing for reasons we have already tried to make clear, and beam search is just a way of eliminating some, but not all, cases of premature commitment without paying too high a cost in terms of maintaining multiple parses.

Our point is simply this: Models formulated in the classical framework face many serious problems—problems which authors like Fodor and Pylyshyn do not acknowledge when touting the virtues of the classical approach. While we acknowledge the achievements of the classical approach, we simply believe that it makes sense to explore the possibility of an alternative which deals directly with the difficulties that it faces, on the view that such an approach may turn out ultimately to lead to a superior overall account. The rest of this paper is an attempt to give the reader a sense of what this alternative may be like.

A PDP ALTERNATIVE

Denied Presuppositions

The PDP alternative which we will propose denies the point of departure, implicit in classical approaches, is necessary to require information to be displayed in structured form in the representation itself (van Gelder, in press). Rather, we ask only that the representations provide a sufficient basis for performing the task or tasks that are required of them. Thus, representations of sentences are not required to exhibit a specificially propositional format *so long as they can be used to perform the tasks we require*. Similarly, representations of knowledge about how to form representations are not required to take the form of rules *as long as this knowledge allows us to act in lawful ways as the environment demands*, and representations of word-specific knowledge are not required to have any visible internal structure representing the meaning of the word. Indeed, the knowledge of rules and of word-specific information may well be encoded in a densely compiled form, as long as this information can be used effectively to meet the imposed demands.

Nature of the Task

Our first step, then, must be to develop some conception of the nature of the imposed demands. At a general level, we think it is reasonable to think of the sentence comprehension task in the following terms. A sequence of words is presented, and the comprehender must form a representation which allows him to respond correctly when probed in various ways. In general, the probes can take a wide range of different forms, requiring actions, verbal responses, etc. Among the things we would expect is that we would be able to answer various questions using this representation. For example, on apprehending "The man stirred the coffee", we would expect a device that has understood this sentence to be able to give correct answers to many questions: Who did the stirring?, What did he stir?, What did he stir with?, and so on.

Of course, there are other aspects to language processing; for example, in processing language we have expectations for what the next word will be, and we can think of part of the task of language processing as the anticipation of the next word. Recently, Elman (1989) has applied an approach similar to the one we take here to this sequential anticipation task. In this task, listeners learn to construct representations that reflect the purely sequential structure of language. Here we focus on learning to construct representations that reflect the constraints sentences impose on the representations we form, in order to be able to answer simple questions about the events these sentences describe.

Given this conception of comprehension, we will need a model which can actually apprehend sentences and then respond correctly to a set of probes. Because we do not stipulate exactly what form the representations must take, we must rely on the adequacy of the performance of the model to determine if in fact its representations are adequate.

For the purposes of what follows, we will distinguish between the process of comprehension itself—the formation of a representation from a sentence—and the use of this representation to respond appropriately to probes. Our main interest is in the former, but for the reasons just given, the latter must be considered as well, or we have no measure of successful performance.

Constraint Satisfaction Processing

We think of the process of comprehension as a constraint satisfaction process (Rumelhart, Smolensky, McClelland, & Hinton, 1986b). In the comprehension of isolated sentences, there are two sorts of constraints: those imposed by the sequence of words, and those imposed by knowledge about how such sequences are to be interpreted. Both types of constraints are taken to be *graded*. They are assumed to act as forces shaping the formation of a representation, and to have magnitudes which determine their degree of influence. For our purposes, the sequence of words in the sentence can be instantiated as a sequence of patterns of activation over a set of processing units. As each new word comes in, we assume that it is used to update the sentence representation, which is also taken to be a pattern of activation over a set of processing units. In fact, if we consider the process at each time-step, it is useful to view it as a constraint satisfaction process in which there are two inputs: the sentence representation from the previous time-step and the new input. These two inputs are used to produce an updated sentence representation for the next time-step. The knowledge of how this updating is to be performed is stored in the connections that allow these inputs to update the sentence representation.

It may be worth noting that graded constraints can vary in magnitude from those that are so weak that they are very easily overridden to those that are so strong that they are nearly impossible to override even by a conspiracy of other quite strong constraints. Thus the existence of cases in which constraints are not overridden does not argue against the idea that they are graded; it just indicates that sometimes they can be very strong.

After each update of the sentence representation, it can be used to respond to one or more probes. Responding to these probes is also viewed as a constraint satisfaction process, where the goal is to produce externally

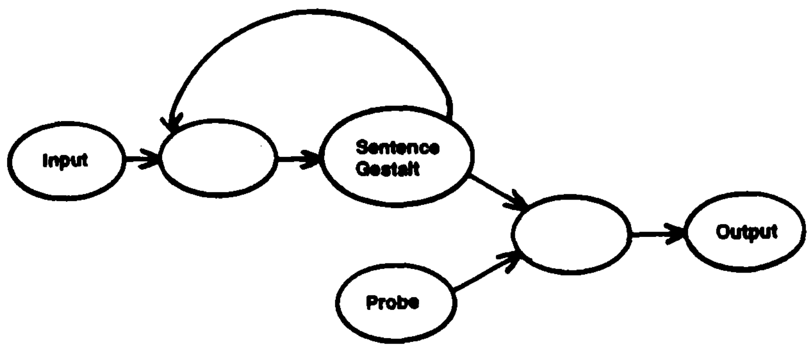

FIG. 1 A sketch of the present conception of the sentence comprehension mechanism. The ovals represent groups of units, and the arrows represent modifiable connections.

specified outputs in response to externally provided probes. There are now three sources of constraint: the sentence representation, the probe, and knowledge about what outputs should be produced for particular sentence/probe combinations. Both the sentence representation and the probe can be instantiated as patterns of activation over processing units, as can the desired outputs; and the knowledge of how to produce these outputs from the corresponding inputs can be encoded in the connections among the processing units.

So far we have outlined a general framework for sentence comprehension and for using the results of comprehension to respond to probes. A sketch of the network that instantiates this framework is shown in Fig. 1. In the figure, the ovals correspond to pools of units and the arrows correspond to connections. There is a pool of units for representing the successive words; a pool of units for representing the evolving sentence representation, or *Sentence Gestalt*; a pool for representing probes; and a pool for representing responses to the probes. The arrows represent connections, from each unit in the pool at the sending end of the arrow to each unit in the pool at the receiving end. The units in the unlabelled pools, which will simply be called "hidden units", serve to allow *combinations* of aspects of the patterns on the input side of these pools to constrain the patterns of activation on the output side.

Learning by Connection Adjustment

Three crucial questions remain. First, what determines the form of the sentence representation itself? Secondly, how is the form of this representation communicated to the inner part of the network? Thirdly, how is the knowledge acquired that governs the construction of the sentence representation from the sequence of words, and the production of appropriate outputs to sentence/probe combinations? The answer to all of these questions is the same: connection strength adjustment through error-correcting learning.

We assume that the output pattern actually generated by the network in response to each probe is compared to the correct output that is provided as part of the environment. The mismatch between the network's actual output and the correct answer is then used as the basis for connection strength adjustment, following the back-propagation learning procedure. This connection adjustment process occurs for connections in both the comprehension network and for connections in the readout network, gradually leading the network to learn both how to represent the information in each sentence and how to use it to respond to each probe.

Note that this connection strength adjustment process cannot actually result in perfect performance, because many of the sentences that the network sees are in fact ambiguous. Furthermore, early on during processing of a particular sentence, before the whole sentence has been presented, the network can only make its best guess as to the answers to certain questions, without any possibility that it can always be right. It turns out that it can be shown that the learning procedure is adjusting the strengths of the connections among the units in the network in the direction of minimising the discrepancy between the activation of each unit and the probability that it should be active, given the input that has been presented up to this point. And, indeed, the activations come gradually to reflect these probabilities reasonably well. In our simulations they tend to gradually approach an equilibrium, in which they jitter about the true probabilities based on the vicissitudes of the most recent set of training examples.

A MODEL ILLUSTRATING THE APPROACH

The model we describe here exemplifies the approach described above. It is in many ways highly simplified. It will not convince the reader that we have already succeeded in providing a complete alternative to conventional approaches. Rather, it provides a concretisation of the general approach as well as an illustration of some of the reasons for its appeal, which we hope will suggest that the further exploration of this new framework is worthwhile. The model is called the Sentence Gestalt or SG model. It is described briefly here (a fuller description is available in St. John & McClelland, in press).

The Environment

The model consists of a network placed in an environment consisting of sentence/event description pairs. The sentences are of but one clause, and they consist of a sequence of stripped-down constituents. Each constituent consists of a single contentive (noun, verb, or adverb) together with a single preposition or the verbal auxiliary element "was". For example, the English sentence "The school girl was kissed by the boy" is reduced to three constituents—"schoolgirl", "was kissed", "by boy". The most complex sentences involved dative passives like "The teacher was given a rose by the bus driver", with additional locative, manner, and/or instrumental prepositional phrases possible, depending on the verb.

The event descriptions are simple too; they consist only of a list of role-filler pairs. For "The schoolgirl was kissed by the boy", the list is: {agent: boy; action: kiss; patient: schoolgirl}. The roles are agent, action, object, recipient, location, manner, instrument, and what might best be called "accompanist" (as in "the bus driver ate the ice-cream with the schoolgirl").

While the sentences and the events they describe are both quite simple, the relationships which hold between them are not. For one thing, words used in a sentence may be ambiguous or vague, as in (15) and (16):

15. The pitcher hit the ball with the bat.
16. The adult ate something.

In both cases, the model is asked to do its best to recover the correct event description. In the latter case, the event description involves a specific adult and a specific something eaten, which may not be uniquely predictable (in the small world of the model, the adult might be a teacher or a bus driver; the something might be soup or a steak). The model must do its best based on the information given.

Constituents may also be left out of sentences, as in (17):

17. The bus driver stirred the coffee.

Here the network is expected to understand that the event being described involved an instrument, which in the case of stirring is always a spoon.

Role assignment is made difficult in two ways. First, both active and passive constructions are used. Though there are semantic constraints that often make a correct interpretation of passives possible, this is not always the case, as in sentences like (18):

18. The teacher was given the rose by the bus driver.

In this and other kinds of cases, the corpus was structured so that the two human participants were equally likely to serve as agent or as recipient, thereby forcing the model to rely on the syntactic cues in the sentence. (Note that the difficulty here is further increased by the fact that the model does not distinguish "gave" from "given". We simply use a single form for each verb throughout, because for most verbs the past and past participle are the same in English.)

The other source of role assignment difficulty arises from the ambiguity of surface role cues. Prepositions and word-order information provide some cues, but these cues are often quite ambiguous as to the roles that they signify. Thus in (19) and (20),

19. The bus driver ate the steak with the teacher.
20. The bus driver ate the steak with the knife.

the semantics of the role-filler must be considered in determining whether the object of the with-phrase is an instrument or an accompanist.

The actual set of sentence–event pairs that the model sees is generated as follows. First, an action is selected at random from a set of possible actions. Then, an agent is selected from a set of possible agents who might perform the action. Following this, an object, an instrument, or indirect object if applicable, and other roles are filled. An illustration is given for the action "eat" in Fig. 2. Note that the selection process is inherently probabilistic, and that there are complex dependencies. Given, for exam-

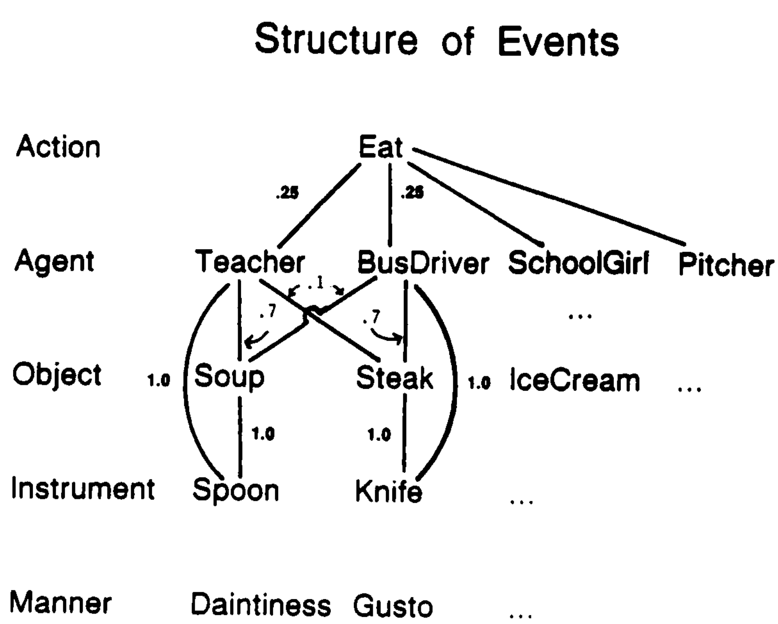

FIG. 2 Structure of the event generator for the action *eat* used in training the SG model.

ple, that the action is eat and the agent is bus driver, the object is probably steak ($P = 0.7$) but may be soup ($P = 0.1$); the instrument depends on the object eaten, the manner on the agent of eating.

This procedure produces an ensemble of event descriptions which are strongly constrained. These constraints can be absolute or *hard*, so that, for example, animal bats do not show up at all as the instruments of hitting; or they may be *soft*, so that, for example, steak is the preferred but not the unique object of eating for the bus driver. Note that the constraints are fairly complex, in that they depend on particular conjunctions of verbs and role fillers. Steak is the preferred food only of the bus driver, the knife is the instrument of eating when the food is steak but not soup, etc.

The assignment of words to events is also probabilistic. Thus, the bus driver in the eating example might be described with the words *bus driver* or with the word *adult*; the steak might be described as *steak* or as *food*; the instrument as *knife* or as *utensil*. Thus, the actual specific participants in the events can only be inferred by using information from context. Sometimes, the sentence contains sufficient information to remove all uncertainty with respect to a particular participant (as in "the bus driver ate the food with the knife"; the food can only be steak), but other times not (as in "the bus driver ate the food"). Even here some answers may be more likely than others, though in some cases there may be at least two equally likely alternatives (in "the adult ate the food", soup and steak are equally likely).

Sometimes, whole constituents are simply left out of sentences describing events in which their referents appear. Thus, the knife can be left out of the sentence on the bus driver eating steak. The model adheres to the conventions that subject and verb are always mentioned (however vaguely), but other constituents may go unmentioned, depending on the specific actions.

The Task and the Interface to the Environment

The model's task is to process the sequence of constituents that represents a particular sentence and, as each constituent comes in, to update a representation which is intended to allow it to respond to probes querying its comprehension of the event described by the sentence. To assess the model's performance, we can actually probe it after each constituent has been processed.

Each input constituent consists of a content word and possibly a preposition or "was". Each such word is represented by a single unit. Thus there is a unit for "bat" (regardless of meaning), a unit for "gave", a unit for "adult", a unit for "was", "with", "by", etc. Altogether, there were units for 58 words.

A similar localist representation scheme was also used for probes and responses. Responding to a probe can be thought of as completion: filling in a member of a role–filler pair, when probed with either the role or the filler. Note that the fillers are now concepts rather than words, and that fillers in particular events are always specific concepts, rather than superordinate categories. There were a total of 45 concept units, covering actions, manners, and noun-concepts, including persons, places, and things.

Given this scheme, we often find sentence–event pairs where a word is used (say "bat") that corresponds to two distinct concepts (baseball bat/flying bat). As in real languages, it is not the word itself which tells us which of the concepts is intended; the correct answer must be derived by making use of cues provided by the context in which the word occurs.

In some simulations using the model, St. John and McClelland included a few units representing superordinate concepts in addition to the units for specific concepts. In this case, a concept is represented not by a single unit, but by a set of units representing the specific concept and its superordinate features. Thus, for example, there are units for person, for male and female, for adult and child. The bus driver is an adult male and the teacher is an adult female, etc.

In considering the task of the network, it is worth noting that there is not always a single right answer. Indeed, early on in a sentence, just after the presentation of the first constituent, there is a great deal of indeterminacy; the initial nounphrase need not even describe the agent of the sentence. Nevertheless, it is possible to view each constituent, as it is presented, as imposing constraints on the possible event-descriptions that might be correct. In this context, we can characterise the task of the network as being one of indicating, in response to each probe, what the range of possibilities might be, and of giving an indication, by the activations that it assigns to the completions of the various probes, of its estimate of the probability associated with each.

Network Architecture and Processing

The architecture of the network, as shown in Fig. 1, can be treated as consisting of two basic parts. One part is the actual comprehension mechanism itself, the part that reads in the constituents sequentially and updates the sentence representation; the other part is the output mechanism, that performs the probe completion task. The sentence gestalt units are in both parts, and form the interface between the two.

Processing occurs as follows. At the beginning of a sentence, the pattern of activation on the sentence gestalt units is set to all 0's, and the unit or units representing the first input constituent in the input pool are turned on. Activation feeds from the SG units (via the feedback loop) and the input units to the hidden units in the comprehension part of the system, and from these it feeds on again to the SG units, where the initial SG representation of all 0's is replaced by a new pattern of activation reflecting the influence of the first constituent of the sentence. This representation is now part of the input at the next time-step, when the next constituent is input in place of the first. This process continues to the end of the sentence.

Each of the units inside the network is a simple logistic processing unit; that is, the activation that a unit takes on is equal to the logistic function of its net input, where the net input is simply the sum over all connections coming to the unit of the input on each connection. The input on each connection is just the product of the activation of the sending unit at the end of the connection × the weight on the connection. Activations range from 0 to 1; weights are floating-point numbers initialised in a range between +/− 0.3, and adjusted according to the learning procedure described below.

Processing in the output network is also quite simple, and can occur at any point during or after the presentation of a sentence. The two inputs to the output network are the pattern on the SG units and the pattern on the probe units. This pattern consists of a single unit on, representing either a queried role or the queried filler. Activation feeds forward from the SG units and the probe input units to a set of hidden units and then from these to the probe output units, where the pattern is taken to represent the network's response to the probe.

Learning

Learning in the network occurs via the back-propagation learning procedure. When a probe is presented, the response to the probe can be compared to the response that would be correct for the current sentence–event pair, and a measure of error called cross-entropy can be computed. Back-propagation is used to adjust the connection strengths so as to minimise this measure (see St. John & McClelland, in press, for details).

It is important to note that the minima in this measure occur at those points where the activations of units in particular situations represent the probabilities that the units should be on in these situations. We think of the activations of the output units as representing the probability that the unit should be on. The training procedure can be seen as trying to find an ensemble of connection weight values that allow the network to get these probabilities correct.

In training the network, we followed the procedure of presenting a complete set of probes after the presentation of each constituent of each of a large number of training sentences. The complete set of probes consisted of a role probe and a filler probe for each role–filler pair in the event description for the sentence–event pair currently being processed.

This training procedure was intended to approximate the situation in which a language learner has just witnessed an event, so that he already has a description of it; and hears a sentence spoken about that event. We imagine that as the learner processes the sentence, he is continually (implicitly) asking himself, "how well does the machinery that I have for language comprehension allow me to describe correctly the event I have just witnessed". The question is posed in the form of the set of probes, and the answer is the set of responses to the probes. The mismatch between the responses to the probes and the correct responses dictated by the description then serves as the basis for learning.

This procedure has two interesting characteristics. First, it does not provide the learner with any specific alignment between the constituents of the sentence and the corresponding constituents of the event description. Thus it forces the network to discover the solution to the bootstrapping problem mentioned earlier for itself. Secondly, the procedure requires the network to do its best at each time-step to predict all of the constituents of the event from what is has seen so far. If learning reaches the global minimum in the error measure described above, then the activations will always reflect the best achievable estimates of the probabilities that the units should be on at each point in the processing of every sentence.

Several different runs of the model have been undertaken. The one from which we report results here involved 630,000 training trials, each involving the presentation of an independently generated sentence–event pair. Some sentence–event pairs occur often, whereas others occur extremely rarely. The low-frequency items also involve relatively atypical role–fillers, and it is the process of learning from rare errors to overcome the tendency to give the most typical answer that makes learning take such a very long time. A fuller discussion of the time-course of acquisition is provided by St. John and McClelland (in press).

Results

After training, the model was first tested on a set of 55 randomly generated sentences that are unambiguous given the hard constraints built into the corpus. That is, although each of these sentences actually contained at least one ambiguous word or unspecified filler, the hard constraints built into the corpus were enough to allow it to respond correctly to all probes. For example, "The teacher ate the soup with the utensil" is unambiguous because the only utensil that could be used for eating soup is a spoon. After the presentation of each sentence, we tested the full set of probes for the role–filler pairs in the event described by the sentence. The network activated all of the correct output units more strongly than any output units it should not have activated on more than 99% of the probes.

The network was also tested specifically on several sets of sentences designed to assess its ability to handle different aspects of the comprehension task. The tasks are broken down into two broad categories, having to do with role assignment on the one hand and specification of the identity of role fillers on the other. With regard to role assignment, St. John and McClelland probed with fillers from the events described by test sentences and examined the roles assigned to these fillers. The use of both syntactic and semantic constraints was examined. Thus, for a sentence like "The schoolgirl stirred the kool-aid with the spoon", semantic constraints must be used to determine that the spoon is an instrument, and not, for example, an accompanist of the schoolgirl (cf. "The schoolgirl stirred the kool-aid with the teacher"). In other sentences, syntactic constraints were examined. Thus, for the sentence "The bus driver was given the rose by the teacher", the order of the constituents, together with the presence of the passive marker and the preposition "by", are necessary to determine the correct role assignments of "bus driver" and "teacher", because either could play the role of agent or recipient. In tests involving five sentences of each of four types (active, passive, crossed with a need to rely on semantic or syntactic constraints), all of the fillers were assigned to the correct roles. The top part of Fig. 3 illustrates a passive syntactic role assignment case. Examples illustrating the other kinds of cases may be found in St. John and McClelland (in press).

For the specification fillers, three distinct variants were considered: The first is the straightforward resolution of word ambiguity, in which the network is asked simply to choose between two alternative and quite distinct interpretations of the fillers of one or more roles. For example, in "The pitcher hit the bat with the bat", the subject, object, and prepositional phrase object are all ambiguous words in the corpus, but each is sufficiently constrained by the context, given the training experience of the model, to yield a unique interpretation. The middle portion of Fig. 3 illustrates what happens in this case. When we probe for the agent, the model activates the concept unit corresponding to the baseball-playing pitcher; when we probe for the patient, it activates the concept unit corresponding to the flying bat; and when we probe for the instrument, it activates the baseball bat. This occurs because, in training, people but not pitchers occur as agents of hitting, flying bats but not baseball bats occur as patients of hitting, and baseball bats but not flying bats occur as instruments of hitting.

The second variant is concept instantiation. The sentence "The teacher kissed someone" illustrates a particularly interesting case, because the someone cannot be resolved uniquely given the context but can be resolved

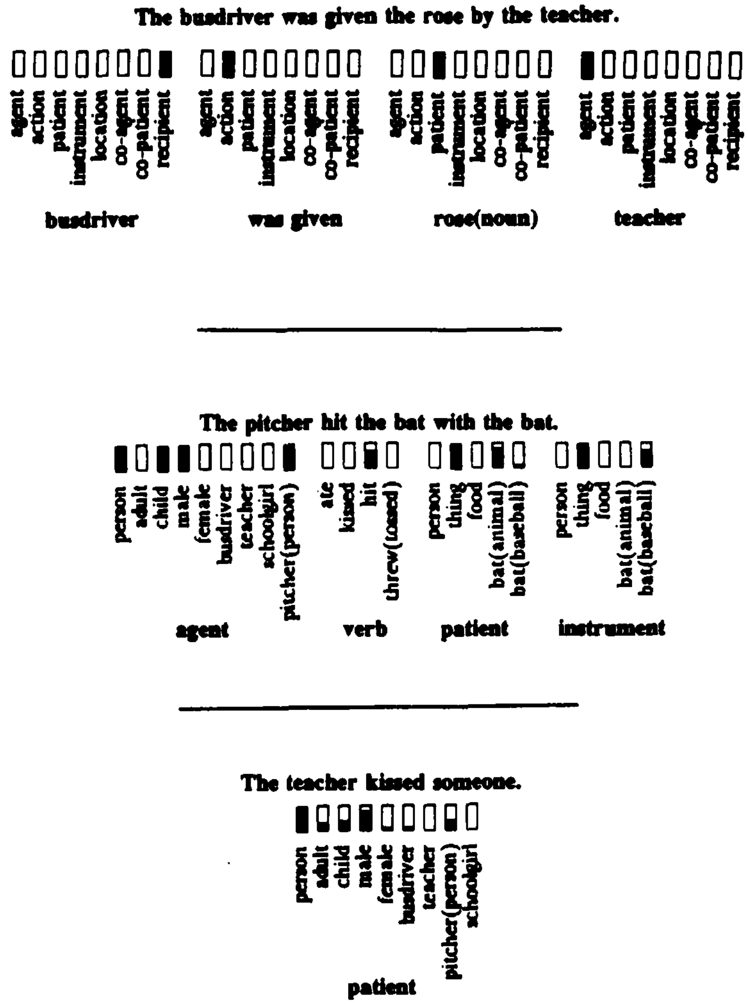

FIG. 3 Activations of relevant output units in response to the indicated probes after presentation of the sentences shown.

partially. In the experience of the network, the teacher is a female, and the event generator is constrained so that kissing is always a heterosexual activity; but the teacher is just as likely to kiss the pitcher (a child) or the bus driver (an adult). Thus we would expect the model to be able to identify the someone as a male but not to determine his age or whether specifically it was the pitcher or the bus driver. In the bottom panel of Fig. 3, the output produced in response to a probe for the patient in "the teacher kissed someone" is shown, where the context partially specifies the filler. Here we can see that the sex, but not the age, is clearly specified. (There appears to be a slight preference for the pitcher over the bus driver. These preferences often reflect the effects of specific training trials that occurred just prior to testing.)

The third type involves what might be called "inference of implicit arguments", because in this case the sentences contained no overt indication even that there was a filler of a particular role. For example, in "The teacher ate the soup", there is no instrument mentioned; but during training, the eating of soup always occurred with the use of a spoon, and so the spoon is inferrable in this context. In this case (not shown) the model learns to fill in spoon when probed for the instrument.

The model was tested with five different example sentences for each of these three types of filler specification cases. In all cases it performed correctly (see St. John and McClelland, in press, for further details and examples).

HOW DOES THE MODEL WORK?

In this section, we begin by following the time-course of processing one example sentence, to give the reader a feeling for the step-by-step processing activity that occurs in the model. We then return to the questions raised at the beginning of this paper, to see how the model gives very different answers to each of these questions.

The sentence we shall study is "The adult ate the steak with daintiness". The sentence is interesting, in that there are three different sources of information as to the identity of the subject. One of these is the word *adult* itself. The second is the fact that the adult is eating *steak*. And the third is the adverb (with *daintiness*); in the model's experience it is only the teacher (a female) who ever eats with daintiness. As we shall see, the example illustrates the model's ability to make use of a variety of cues of varying strength, spread throughout the sentence, to identify a particular constituent.

After the presentation of each constituent (adult, ate, steak, with daintiness), we can examine the response of the network to probes assessing the fillers of the agent, action, instrument, and patient roles (see Fig. 4). Later, we will return to consider the pattern of actication over the SG units, which provides the representation of the whole sentence.

We consider first the response to "agent", because it is here that we see the effects of several constituents operating most clearly. After the presentation of "adult", the model takes the agent to be an adult person; there is some activation of both male and female, and of both bus driver and teacher, the only two adults in the set. There is a slight bias favouring male. Child is included to illustrate that it is not active at any point. There is little

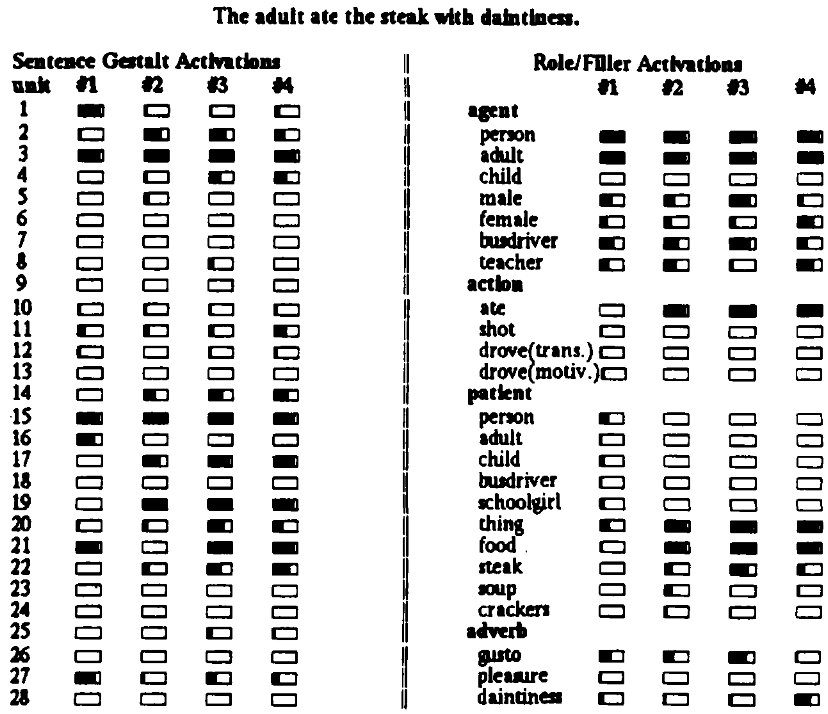

FIG. 4 Activation of a subset of the sentence gestalt units (on the left) and of relevant output units in response to the indicated probes (on the right) after presentation of each constituent of the sentence "The adult ate the steak with daintiness". The columns labelled #1, #2, etc., refer to the state after each of the successive constituents (#1 = adult, etc.).

change after the presentation of the verb, because this does not really provide any constraints on the identity of the adult (the teacher and the bus driver appear equally often in sentences involving eating). The presentation of "steak", however, produces a shift in the direction of male and bus driver. This shift is reversed (though not completely) when the final constitutent, "with daintiness", is presented.

For the other roles, the reader will note that the model performs in a generally sensible way. The one slight problem appears in the case of the patient. We see the activation of "steak", which was quite strong just after the presentation of the steak constituent, weaken considerably when "daintiness" is presented. We will return to a consideration of this specific aspect of the model's performance later.

Given the overall success of the model, let us now ask what kinds of answers do we get to the questions raised at the begining of this paper when we use a model of this sort?

What is Constructed when We Comprehend a Sentence?

In this case, the answer is not "a structural description". What is constructed is a pattern of activation which permits the performance of a specific task or tasks. In this case, the task is to provide a basis for completing role–filler pairs; but one can imagine a wide variety of other uses as well. Whatever the tasks were that we were called upon to use the results of comprehension to perform, a model with the general structure of the one used here could be used to learn to perform that task.

Given this, it becomes a matter of empiricial research to ascertain just how a network will choose to use its units in learning to perform the tasks that it is given to perform. We know from other connectionist research that the answers to these questions are dependent both on the specific tasks the network is asked to perform, and on the details of network architecture (Hinton, 1986; McClelland, in press). In this instance, just perusing the pattern of activation in the sentence gestalt at each successive presentation of a new input constituent, we can see two things. First, many of the units take on graded activations, and several of these seem only partially correlated with particular role–filler activations. This suggests that the activations of particular output units in response to particular probes are generally determined by the joint influence of a number of hidden units; thus they provide a distributed, coarse-coded representation of the role–filler information conveyed by the sentence (cf. Hinton, McClelland, & Rumelhart, 1986).

What Role do Words Play in the Comprehension Process?

In the present model, as each word is presented, it changes the pattern of activation in the sentence gestalt. In this case, we see each word as exerting constraints on the representation. It will be noted that these constraints can in general influence the responses to all of the probes we might present after presentation of a word. Thus the presentation of "ate" affects not only responses to probes for the action but also probes for the patient; and the presentation of steak and daintiness each influence responses to probes for the agent, the patient, and the manner. Thus a word is a clue that constrains the interpretation of the event as a whole.

The influence that a particular word will have on the comprehension process depends on what has already been presented. But, there is a systematic contribution that each word makes. This systematic contribution is represented by the set of connection strengths from the input unit that represents a particular word to the set of hidden units inside the comprehension part of the network.

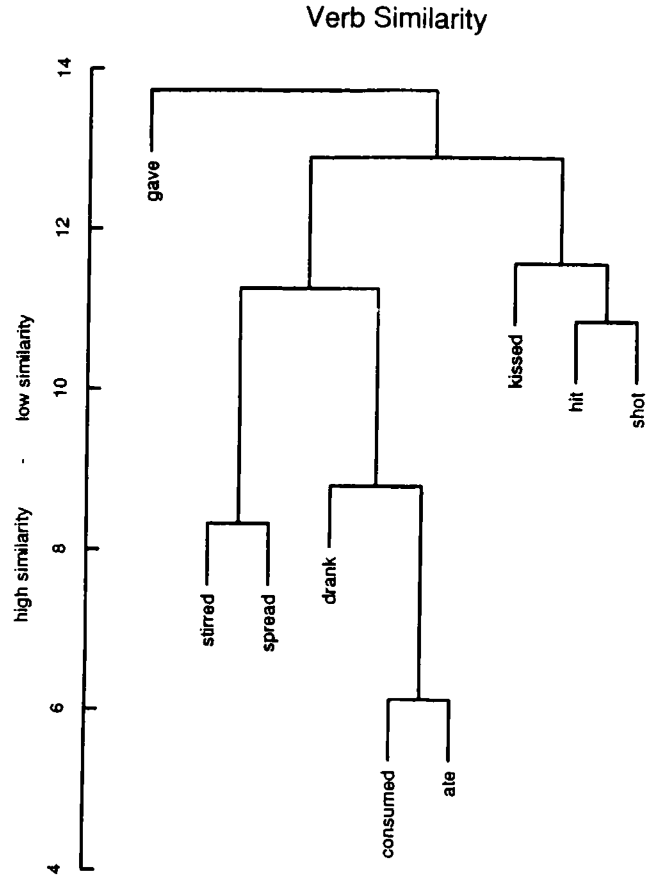

FIG. 5 Cluster analysis of the weight vectors emanating from each word input unit to the hidden units in the comprehension part of the SG model, for the units representing the 11 unambiguous verbs shown. The vertical position of the horizontal bar joining two branches indicates the similarity of the leaves or branches joined.

To examine these contributions, St. John and McClelland extracted the vector of connection weights emanating from each word input unit to this first layer of hidden units. These feature vectors were then entered into a hierarchical cluster analysis (separate analyses were performed for the nouns and verbs). The analysis for the verbs (Fig. 5) displays clearly that the model has captured the similarity structure among the "frames" represented by these verbs as used in our training corpus. The verb "give" is the only dative verb in the corpus, and is clustered separately from all the others. The verbs "ate", "drank", and "consumed" all take animate things as subjects and inanimate things (food) as their objects; the verbs "stirred" and "spread" each take a human subject, food as an object, and a spoon or

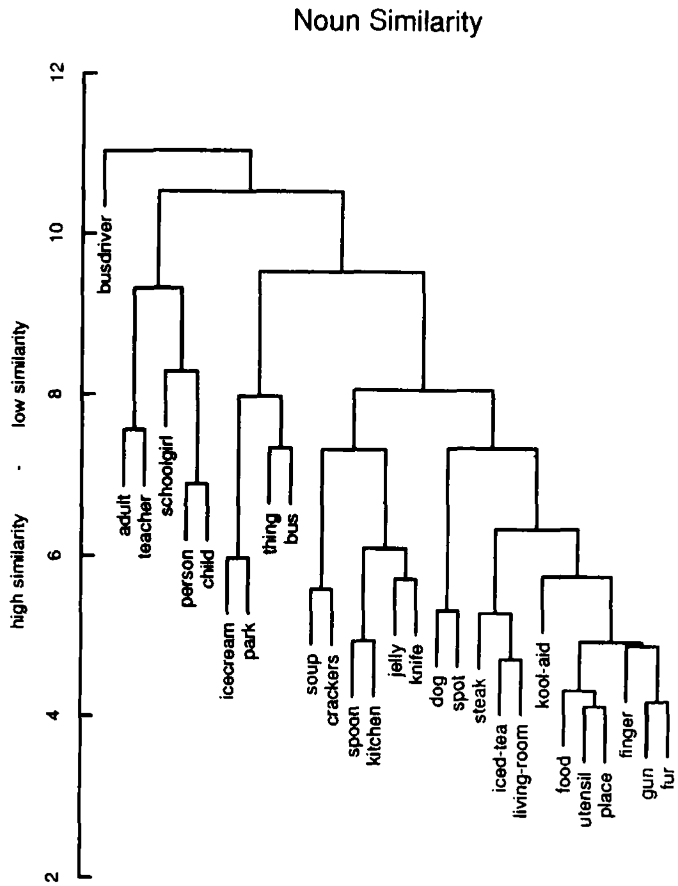

FIG. 6 Cluster analysis of the weight vectors emanating from each word input unit to the hidden units in the comprehension part of the SG model, for the units representing the unambiguous nouns shown. The vertical position of the horizontal bar joining two branches indicates the similarity of the leaves or branches joined.

a knife as the instrument; and "hit", "kicked", and "kissed" are all passivisible in the corpus (unlike the food-related verbs), and all involve a patient that may be animate.

The analysis for the nouns (Fig. 6) is less clear; it appears that there are two organisational principles that are both at work. Sometimes, nouns cluster by meaning. Thus all the human nouns cluster separately from the rest of the nouns. However, at a finer grain, the nouns sometimes appear to cluster by co-occurrence in the same events. Thus ice-cream clusters with park because in our corpus ice-cream is eaten in the park, and that is the only thing that ever happens in the park. Once again, the model appears to be picking up what might be called the frames that the nouns enter into, rather than their individual meanings *per se*. Of course, the details of this depend on the particular training corpus; in ordinary life, much happens in parks besides the eating of ice-cream. In general, it seems likely that noun-frames are much weaker than verb-frames; but to the extent that such frames do exist, they can be captured by models such as this.

How does the Process of Constructing a Representation of a Sentence Occur?

In the connectionist model, there is no separation of the structure-sensitive rules and the lexical content of words. The process is inherently susceptible to guidance by content as well as structural information.

In some sense, the model represents the strongest possible alternative to a modular approach. Not only are all different sources of constraint taken into account simultaneously, the knowledge underlying each source of constraint is inextricably interwoven in the connections.

There are limits on the kinds of processing ambiguities which our model actually deals with, and on the kinds of information that it brings to bear in dealing with them. The first of these limits is due primarily to the simplicity of the task we have asked the network to perform; we can study role ambiguities, because we can probe for the fillers of particular roles, but we cannot study attachment ambiguities, because the role–filler completion task is too limited. If the language of the probe completion task were enriched, however, it should be possible for the network to learn to deal with more interesting structural ambiguities as well. To make this concrete, suppose we had trained the network to complete head–role–filler triples, rather than simply role–filler pairs. For example, consider again sentences like (3) and (4):

3. The spy saw the policeman with the revolver.
4. The spy saw the policeman with the binoculars.

Suppose that the probes specified a head noun or verb, and a role, and the task was to fill in the filler. In the case of Y, we would want the network to fill in the concept corresponding to revolver when probed for the modifier of *policeman*, but not when probed for the instrument of *saw*. In the case of (4), we would want the network to fill in *binoculars* when probed for the instrument of saw, but not when probed for the modifier of *policeman*. In principle, we see no reason why some version of this approach should not be able to extend to multiclause sentences, though clearly this is a matter for further research.

A complete model would, of course, also provide some representation of context prior to the beginning of the current sentence. As the work of Crain and Steedman (1985) and Altmann and Steedman (1988) makes clear, constraints arising from the tendency for multisentence texts to maintain referential coherence can influence attachment ambiguity resolution. We would expect that a model that somehow kept a record of prior context could learn to exploit it for the resolution of attachment ambiguities.

Of course, what we have just described are hoped-for extensions, not real results at this point. Some relevant research, indicating an ability to use context to resolve ambiguities of reference in multisentence texts, has been carried out (see, e.g. Allen, 1987; 1988; Miikkulainen & Dyer, 1989), but to our knowledge nothing relevant has yet been done with regard to the resolution of structural ambiguity.

How does Acquisition Work?

Acquisition works by a process of gradual connection strength adjustment. This is quite different from the formulation of a system of explicit rules. Certain problems are avoided right from the start, such as the question of when to form a rule, and when to simply list exceptions. However, it would certainly not be accurate to suggest that the model we have presented here is a *tabula rasa*, acquiring knowledge of language without any prior structure. Indeed, the input is parsed for the model into constituents and words; and the role–filler representation of the event descriptions and the set of concepts used in the output network are predetermined as well. Finally, the structure of the network is preordained, and tailored to the task. These features of the model were not adopted out of any belief that their adoption was necessary but simply out of a desire to establish a simple illustrative model. Just how much prior structure has to be built in, and in what way it is built in, remain basic and central issues for connectionist models in this and a number of other domains.

CAN THE PDP APPROACH SOLVE THE PROBLEMS WITH CONVENTIONAL MODELS?

Earlier we enumerated a set of problems with conventional models. Here we consider how they are or could be solved in models of the kind we have considered here.

Conceptual Guidance and Rule Conflicts

The problem of conceptual guidance is naturally solved by the integrated handling of both syntactic and content-based constraints on processing. The problem of rule conflicts is dealt with by the connection adjustment process. That process assigns strengths to the features so that the correct interpretations are achieved across the entire corpus.

Contextual Shading as well as Selection of Word Meaning

This characteristic of PDP models is not illustrated so clearly by the present model because of its use of local representations for concepts. We can see this kind of thing to a limited degree in such examples as "The adult ate the steak with daintiness". Though "teacher" and "female" are ultimately more active than "bus driver" and "male", the fact that it is a steak that is eaten definitely shades the activations in the network with maleness; the model seems only too natural in its ability to capture stereotypes like the one immortalised in the phrase, "real men don't eat quiche", and to use innuendo in shading its representations.

The use of local representations for concepts makes it possible to see contextual shading only in the relative degree of activation of the few superordinate feature units that were included in the model. However, this use of local representations is not inherent in the connectionist approach and we adopted this usage here only for ease of testing and to avoid building undue amounts of knowledge into the concept representations. However, an earlier model that did use distributed representations does illustrate shading effects on a grander scale (McClelland & Kawamoto, 1986). In that model, concepts were represented by fully distributed patterns. The model was trained to interpret a variety of sentences involving breaking one object with another, and all but one of the objects that could occur as the instrument shared a feature indicating that the object was hard. The one exception, the ball, was encoded as soft, and the model correctly treated it as such when it occurred in most contexts. However, when it was used to break other objects, the model shaded the representation, giving it the feature hard instead of soft; this happened just because things that break other things were typically hard, and the model became sensitive to this fact. It is worth noting that the resulting pattern was not one of the existing patterns on which the model had been trained but an extension by the model of the ensemble of possible concepts.

The Similar Problem with Roles

The shading of concept representations that is captured in McClelland and Kawamoto's model has been applied to roles by Touretzky and Geva (1987). The idea is simply that the set of possible roles is not some fixed set of N alternatives but an extensible set with a rich similarity structure, such as is naturally captured by distributed representations.

Implied Constituents

The handling of implied constituents is not a problem in the model. It is quite natural for the model to learn that events involving eating steak always involve a knife as the instrument. There is no special "inference step" required to fill in the knife. This is in part a direct result of the fact that there is no prior stipulation that a particular part of the representation of the sentence corresponds to the internal reflex of each particular constituent of the sentence. It's just that events described by sentences with "ate" as the verb and "steak" as the object always involve knives as instruments. The probabilistic nature of many implied constituents is not a problem either, because of the inherently graded nature of the activation process, coupled with the fact that intermediate activation values directly reflect probabilities intermediate between 0 and 1.

Combinatorial Explosion or Premature Commitment

The model avoids combinatorial explosion by keeping multiple alternatives implicit in the single pattern of activity over the sentence gestalt. It avoids the catastrophic side-effects of premature commitment because its graded activations can be adjusted as each new constraint is introduced. In a sense, it does make commitments as each new constituent is encountered, but these are not all-or-none choices, but simply continuous shifts in the pattern of activation. Thus commitments made can be reversed without any backtracking. It is true that some constituents cause a more marked adjustment of the SG representation than others. These marked adjustments can be related to experimental data on reading times if we make the simple assumption that larger adjustments take longer to make. This assumption holds in systems that adjust their activations continuously (McClelland, 1979) rather than in a single time-step. We view these continuous systems as more realistic that the discrete time-step system used here; as with the use of localist representation, the use of discrete time in the illustrative example model is simply a matter of greater tractability.

The notion that larger changes in the SG are associated with longer reading times provides a natural way of accounting for a lot of reading time data in which a slowdown in processing is observed in one condition relative to another. In these cases, it seems unnecessary to invoke the notion of reprocessing, which is often associated with theoretical discussions of these effects in the experimental literature coloured by the classical framework. But, it does sometimes happen that language readers and listeners experience a true garden path. By a true garden path, we mean the strong feeling that something has gone awry—that the sentence no longer makes any sense at all. This occurs when one reaches the word "fell" in a sentence like this one of Bever's (1970):

21. The horse raced past the barn fell.

The question is, why is it, on the story that we are telling, that subjects are not at this point able to recover, based on the constraints on the interpretation imposed by the final word?

The answer is that the word "fell" by itself may not exert a strong enough influence to reorganise the SG representation all on its own. This argument implies that if the sentence continues in a way that imposes additional input favouring the correct interpretation, the garden-path effect might actually be overcome. Thus we expect that the garden-path effect can be ameliorated if the reader just continues to read and the sentence goes on and provides additional relevant constraints:

22. The horse raced past the barn fell into the ditch.

Obviously, this is a topic for careful research, rather than the jaded intuitions of theoretically biased psycholinguists; but we feel the example supports our sense that even in such dramatic cases as (21), we may not really be forced to reprocess, as long as there is subsequent information that allows us to overcome the effects of what has come earlier in the sentence.

The Problem of Acquisition

The use of gradual connection adjustments in the model helps it overcome some of the problems conventional approaches face in learning to interpret sentences. First, the strengths of constraints imposed by various words on the interpretation process are naturally graded and are brought gradually into balance by the connection adjustment process. Secondly, the solution to the bootstrapping problem emerges naturally through the exposure of the model to the statistical properties of an ensemble of sentence–event pairs. It is true that the sentence "the boy kissed the girl" could map on to the event of a boy kissing a girl in two different ways; but these alternatives are further constrained by other sentences. Thus in every sentence where the subject of the verb "kiss" is girl, there is a girl in the event and she is the agent.

We do not wish to suggest at all that the problems of acquisition are fully solved by the present model; the sentences and events are highly simplified, and the pre-parsing of sentences into words and constituents, together with the pre-structuring of events into role–filler pairs certainly makes things easier for the model.

Our only claim is that the connectionist learning procedure we have used does have some signficant advantages over rule-learning approaches. As noted above, it remains for further research to establish how much support these procedures require from pre-existing structure and how much they can induce from the environment.

Comparison with Other Connectionist Models

As many readers will doubtlessly be aware, a number of researchers have noted the appeal of the PDP or connectionist modelling framework for capturing aspects of language processing, and have chosen to build models within this framework. Our approach certainly falls well within this tradition, but differs from most of the previous work in one important way. Earlier models have tended to incorporate, in one way or another, specific characteristics of conventional models of language directly into the internal representations used. Our earlier work is no exception to this rule. For example, the interactive activation model of visual word recognition (McClelland & Rumelhart, 1981; Rumelhart & McClelland, 1982) incorporated units that represent letters and words; the TRACE model of speech perception (Elman & McClelland, 1986; McClelland & Elman, 1986) incorporates phonemes and words. McClelland and Kawamoto (1986) use semantic features such as "animate", "inanimate", etc., and gave over slots within the model to each of several possible representational roles. By taking these steps—reifying some aspects of conventional representations inside connectionist networks—these and other models have in fact fallen prey to some of the difficulties that are faced by conventional models. The most serious problem is the problem of combinatorial explosion. Unlike conventional models, in which the problem of combinatorial explosion is only faced during processing, these connectionist models have to face it in advance, and construct networks of units and connections that will allow all possible inputs to be processed.

Two approaches have been taken to avoid this problem. The first is to exploit the idea of coarse coding, in which units do not correspond to specific representational entities, but to conjunctions of pre-specified features of possible representational entities. It is well established that this can limit the problem in some domains, and some models have been successful using this approach (c.f. Touretzky & Hinton, 1988), but we have found it unworkable for sentence processing (St. John & McClelland, 1987). The problem is that the number of features that need to be pre-specified is unbounded, as is the number of these that may need to be conjoined to represent any particular specific proposition. Again the combinatorics can be devastating.

The other approach—the one that is taken here—is to let connectionist learning procedures do the work of determining what the internal representation should be like. Of course, the approach has only become possible since the advent of learning rules that allow internal representations to be formed in response to task demands (Rumelhart, Hinton, & Williams, 1986a). This approach has one serious disadvantage—learning can be very slow. But it does have some very important advantages. It allows the network to learn to form representations that make efficient use of the available representational resources, given the task as specified by the inputs that are presented and the outputs that the network is expected to generate in response to these inputs. A number of researchers have been exploring this approach. Elman (1989), for example, has developed a simple network that learns to anticipate the next element of a sequence of words. In so doing, it learns to form an internal representation of each word that reflects the sequential constraints in the sequences. If the words form sentences, and the sentences reflect co-occurrence constraints—such as the fact that a certain verb requires an animate subject and an inanimate object—the representations come to reflect the relevant semantic distinctions.

Just how far this kind of approach can be taken remains to be seen. Some (e.g. Touretzky, 1988) have suggested that more structure needs to be built into the network, and the facts of learning speed do suggest that something must be done before the approach can be successfully extended to large corpuses of complex sentences (see also Miikkulainen & Dyer, 1989). At present, it is too early to tell just what will be the most effective way to provide for an extension of the approach.

ARGUMENTS AGAINST THE PDP APPROACH

Several different types of arguments might be given in favour of conventional approaches and against the approach that we have taken. Here we consider three that seem particularly central. In all three cases, we believe that the arguments are less compelling than the proponents of alternatives have alleged.

Systematicity and Productivity

In their critique of connectionist models, Fodor and Pylyshyn (1988) point out that an inherent feature of the conventional approach is the fact that it accounts for the systematicity and productivity of language. These characteristics follow directly, they point out, if we assume that our cognitive apparatus makes use of a combinatorial syntax and semantics; they also claim that these characteristics do not follow from the PDP approach. Let us examine these characteristics.

Systematicity refers to the fact that if a speaker can understand a sentence like "John loves the girl" and (let us say) "Bill dislikes the teacher", then he can also understand other sentences, such as "John loves the teacher", "Bill dislikes John", etc. In other words, sentences are not just isolated, unanalysed wholes but are composed of parts which can be recombined to produce other sentences that the speaker will understand.

To test the capability of a model such as ours to exhibit systematicity, we generated a new corpus, containing 10 persons and 10 actions. Each of the actions could be done by any person to any person so that there were a total of 1000 possible events. Each could be expressed in an active or passive sentence for a total of 2000 possible sentences.

We trained the same network described above with all but a randomly chosen 250 of the possible sentences; then, after training, we tested it on the remaining 250 sentences. A stringent accuracy criterion was adopted: A sentence was scored correct only if the unit representing the correct person or action was more active than any other unit in response to probes for the actor, action, and patient. The model got 97% of these novel sentences correct.

Now obviously this is but the first step in demonstrating that connectionist networks can exhibit systematicity. The corpus is finite, and 87.5% of it was used during training. Nevertheless, there is considerable systematicity in the model's performance.

Now, it might be noted that systematicity is not in fact an inherent attribute of the model that we have proposed. In fact, it is probably true that the model could learn to treat sentences as unanalysed wholes, if in fact the constituent structure of sentences had no relation to their actual meanings. Thus, we could train the model to produce arbitrary answers to probes given in conjunction with each member of a list of sentences. Learning would be slower than if the sentences had a systematic structure, and it would require more units and connections, but in principle this could be done.

The question arises, then, as to why it is that languages turn out to be systematic. One possibility is that it is an inherent characteristic of human cognition to be systematic in just this way, and this is the crux of Fodor and Pylyshyn's argument. On this view, human languages are systematic because our minds force them to be that way, due to their use of a combinatorial syntax and semantics. Indeed, Fodor and Pylyshyn argue that connectionist models cannot explain the systematicity of language, because connectionist models do not inherently impose a combinatorial syntax and semantics.

Our response to this argument has two parts. First, we question the implicit assumption that the explanation for the systematicity of language must lie with inherent characteristics of the mechanisms of thought.

Indeed, we would look to the tasks that humans must perform with linguistic stimuli, and the experiences from which they learn, for at least some part of the explanation. If the task is to form representations of events which themselves have a combinatorial structure, then a computational mechanism that becomes systematic may be simply discovering the systematicity in the environment.

Systematicity in the environment is probabily only part of the story. It has been the force of much work in linguistics and elsewhere to argue that the systematicity in the environment is much less than the systematicity that is imposed by the human observer. Indeed, we suspect that there is some truth to this claim.

This leads us to our second point, which is that in fact connectionist models *do* tend to impose systematicity, even though they do not have a combinatorial syntax and semantics. Networks do not, as a matter of fact, simply memorise individual input–output pairs and treat each one as an isolated individual case; the generalisation experiment just described is just one of a very large number of relevant demonstrations that in fact they do exhibit a tendency to behave systematically. A tendency towards behaving systematically is in fact a characteristic that our model shares with mechanisms which simply stipulate a combinatorial syntax and semantics.

From here on, the argument simply gets tendentious. One side can claim that connectionist models do not in their present form exhibit *enough* systematicity; and while this may be true, it places any in-principle argument against systematicity in connectionist models in considerable doubt. Or it might be claimed that a tendency to behave systematically has been snuck into the model by some slight of hand. It is in fact true that particular choices of details of network architecture do influence the degree of systematicity; and indeed it is quite important to get a clearer picture of what aspects of network architecture are conducive to good generalisation. We do not doubt that evolution may have shaped our cognitive structures so as to make them more likely to be able to act systematically; but we see no reason to suppose that it has done so by endowing them in advance with an explicitly combinatorial syntax and semantics.

What about productivity? Productivity is of course intimately linked to systematicity; it refers to the fact that we can understand many sentences that we have not actually heard before. The experiment just described addressed this point, though again, in a fairly limited way. It is generally assumed that humans can comprehend an infinite number of sentences, while in our experiment the corpus was indeed only a bit larger than the training set.

Other research on the productivity of connectionist networks is currently underway. Servan-Schreiber, Cleeremans, and McClelland (1988) have shown that a simple network architecture first introduced by Elman (1988) can learn to accept all and only the grammatical tokens of a simple finite state language. Because in the case of this finite state language the corpus is infinite, we have the first clear indication that a network can learn from finite experience to process an infinite corpus. The handling of long-distance dependencies is currently under rather active exploration (Servan-Schreiber et al., 1988; Elman, 1989). Extension to comprehension, rather than mere acceptance of grammatical tokens, awaits as the next challenge.

There is an aspect of the productivity of language that appears to be better explained by our connectionist approach than by conventional approaches. This is the use of context to shade meanings of concepts as they are instantiated in particular events which may be contextually appropriate. The example of the ball from McClelland and Kawamoto illustrates this. In another case, they presented their model with the sentence "The doll moved". This sentence was novel to the model. Among the features that the model had learned were associated with "doll" was inanimacy. However, in interpreting this sentence the model "animated" the doll. This is because, in all of the sentences that the model had been trained on, the subject of the sentences of the form "X moved" were always animate. It seems to us that this interpretive liberty on the part of the model is entirely correct and appropriate, and illustrates a productivity that extends far beyond the capabilities of conventional models.

Beyond Compositionality

We have discussed two out of the three characteristics Fodor and Pylyshyn (1988) claim language has that are captured by conventional approaches. The third characteristic is compositionality: The idea that a word contributes the same thing to the meaning of all of the sentences in which it occurs. In the introduction, we criticised the notion of compositionality, indicating that in fact it represents an impoverished view of the comprehension process. In our illustrative model, a word does always exert the same influence on the net input to the first set of hidden units in the comprehension part of the model. But, due to the non-linearities in the hidden units at that layer in the network, and due to the concurrent influence of inputs from context, the actual impact of the word can differ greatly from context to context. The word exerts the same *force* on the representation at each occurrence, but this force is combined with those applied by context, thereby allowing for context sensitivity.

It might be argued that the model is too sensitive to context, in that in fact it allows context sometimes to override the correct interpretation of a word. This happens, for example, with the steak in the example presented above: After the presentation of *with daintiness*, the activation of *steak* on probing for the patient is weakened. In fact, at earlier points in learning, the model actually activates soup more strongly than steak after *with daintiness* is presented.

This behaviour must be taken as an error, but it is an error of the kind that people often make. For example, Erickson and Mattson (1981) asked subjects to answer the question:

How many animals of each kind did Moses take on the ark?

Most subjects said two, and noticed nothing amiss, even though they had been warned to look out for trick sentences. Indeed, many subjects could not pinpoint the problem with this sentence even after being told that there was something wrong. Apparently, the constraints imposed by the word *Moses* itself are not sufficiently strong to override those imposed by the context.

Some may view errors of the Moses type as abberations, though the effect is easy to produce with other examples. To us it is a reflection of the fact that the doctrine of compositionality misrepresents the contributions of words to an understanding of the meanings of sentences.

In sum, we do not see any reason to suppose that the observed degree of systematicity, productivity, and compositionality of human language need be attributed to inherent structural characteristics of the kind that Fodor and Pylyshyn have advocated. In part, the systematicity of the world may be to blame; beyond this it is clear that networks do tend to impose systematicity; productivity does not appear to be beyond the power of PDP models like the one we have considered here; and there are aspects of the expressive capability of human language that go beyond what can be captured in a combinatorial syntax and semantics which seem naturally to follow from a PDP account.

Lexical and Syntactic Autonomy

We turn now to a set of considerations that arise from psychological experiments, where it is claimed that at least during some initial stage of processing, both lexical access (i.e. activation of the possible meanings associated with words) and syntactic processing (i.e. assigning attachment relations between sentence constituents) are autonomus processes. These claims run directly counter to the basic tenets of the approach that we have taken here, because the approach assumes that these processes are inextricably intertwined with each other and with the exploitation of contextual influences. Clear evidence for autonomy would therefore undercut our approach completely. So let us see, what is the evidence?

Lexical Access. In well-known experiments (Swinney, 1979; Tanenhaus, Leiman, & Seidenberg, 1979), subjects had to listen to a spoken text containing an ambiguous word (such as BUGS) and were probed for a lexical decision immediately after the offset of the word with another word related to either meaning of the ambiguity. The oft-cited result of such experiments is the finding that decisions to words related to either meaning of the ambiguity are faster than decisions to unrelated words, indicating that both meanings are initially accessed; only later is the ambiguity resolved to fit the context, so that the contextually appropriate reading is the only one that remains active.

There are two points to be made. The first is that a recent meta-analysis (St. John, 1988) of a total of 19 studies, using both lexical decision and word naming methods, reveals that in fact there is a reliable advantage for the contextually appropriate reading, even at an immediate test. The pattern exhibited in Fig. 7 from the seminal experiment of Swinney (1979) is exemplary of the general pattern of the results.

The second point is that this pattern is very close to what is found in a simulation of the processes of settling on an interpretation of an ambiguous word in a PDP model of the disambiguation process (Kawamoto, 1985; 1988; see Fig. 7). Kawamoto's model differs from the illustrative model described here in three crucial ways. First, it uses a continuous, gradual activation process, so that units gradually settle into their final state, rather than being thrust into a state in a single step. Secondly, it makes use of full recurrence in the connections among the units, so that units within the same part of the system feed back on each other. Thirdly, it does not actually simulate the full process of sentence interpretation, but only considers the process of settling on an interpretation of an individual word as a joint function of contextual and phonological input. We view Kawamoto's model as an attempt to characterise the fine-grain temporal processes involved in lexical access that are more coarsely approximated in the SG model.

Now, Kawamoto's model most clearly does not assume that the process of accessing meaning is autonomous; as in the SG model, both contextual and input-based constraints influence the process from the start. However, what happens in the model is that at first both of the possible meanings consistent with the input word are activated. It is only as the activation process continues, that one interpretation is gradually pushed out and the other comes to dominate completely. Thus it appears that the empirical evidence is quite similar to what should be expected on the PDP account.

Autonomous Syntax? A number of studies have been reported indicating that syntactic preferences initially determine the outcome of on-line parsing processes, so that sentences in which the context eventually requires an alternative interpretation are processed more slowly than those in which the content is consistent with the syntactic bias. A variety of

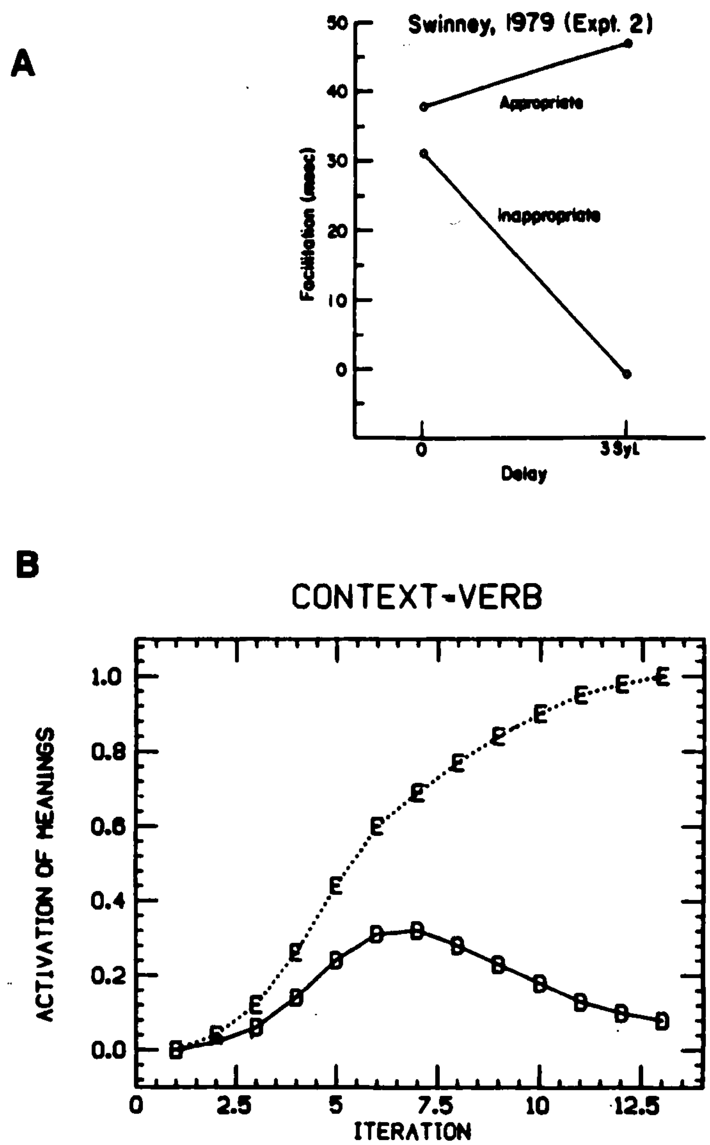

FIG. 7 (a) Data from Swinney (1979); (b) activations of contextually appropriate and inappropriate meanings of the ambiguous word *wind* (from Kawamoto's distributed model of ambiguity resolution). The context specifies that a *verb* is expected, and the two curves represent activations of patterns corresponding to the contextually appropriate meaning (E for encircle) and the contextually inappropriate meaning (D for draft). (a) is reprinted from McClelland (1987) and (b) from Kawamoto (1985).

constructions have been examined in studies of this type. One of these is the reduced relative construction, in sentences like (23) and (24):

23. The actress sent the flowers was very pleased.
24. The florist sent the flowers was very pleased.

Another is the N-V-N-PP construction, as in (25) and (26):

25. The spy saw the policeman with the binoculars, but . . .
26. The spy saw the policeman with the revolver, but . . .

Using the first kind of construction, Rayner, Carlson, and Frazier (1983) found that subjects have difficulty processing the reduced relative clause in both cases, even though in one of the examples (the actress sent the flowers) semantic constraints are said to favour the idea that the actress would be the recipient rather than the sender of the flowers as is required in the reduced relative interpretation.

Such a finding is, in our view, not particularly telling in indicating whether there is some *initial* syntactic process that favours one interpretation over the other, or whether, alternatively, there is simply a strong *weight* associated with the syntactic preference to treat a N-V-N sequence as actor–action–object. It certainly is the case that the initial part of the sentence

The actress sent the flowers . . .

is unambiguously interpreted by native speakers as indicating that the actress is the sender not the recipient of the flowers; plausible continuations might involve a recipient (herself, perhaps?) or another clause. Thus it appears that the syntactic cues are simply overriding in this case. Similar arguments apply to many of the materials used in the subsequent study by Ferreira and Clifton (1986).

In the second kind of construction, it was found (Rayner et al., 1983) that there was an advantage for sentences of the form of (25), in which the prepositional phrase is ultimately attached to the verbphrase, compared to sentences of the form of (26), in which the prepositional phrase is ultimately attached to the nounphrase. However, a series of experiments (Taraban, 1988; Taraban & McClelland, 1988; 1990) has now established several important findings regarding this particular construction. Taraban and McClelland (1988, experiment 1) established three basic points. First, the materials used by Rayner et al. generally had a bias such that the part of the sentence preceding the disambiguating word (revolver or binoculars, in this case) tended to favour the VP attachment of the prepositional phrase. Secondly, other materials are easily constructed in which this attachment preference is reversed. Thirdly, studies of on-line processing using the word-by-word reading task developed by Just, Car-

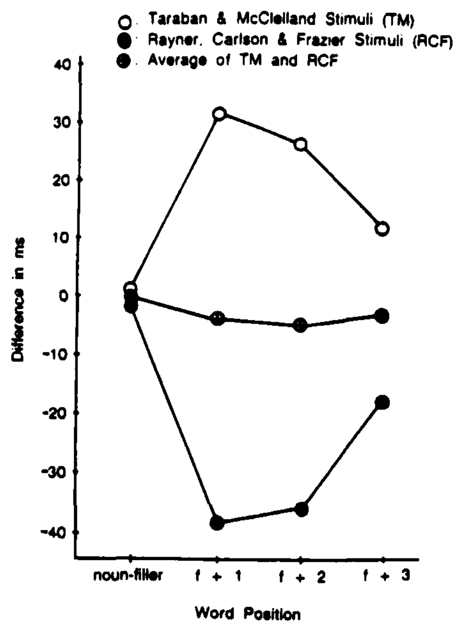

FIG. 8 Reading time advantage (negative numbers) or disadvantage (positive numbers) for sentences requiring a verbphrase attachment of a prepositional phrase compared to matched sentences requiring a nounphrase attachment. Rayner et al.'s (1983) stimuli are biased so that subjects expect the PP to attach to the VP. Taraban and McClelland's stimuli lead subjects to expect the PP to attach to the preceding NP.

penter, and Woolley (1982) revealed that the finding reported by Rayner et al. (1983) only holds with the VP-attachment-biased materials, and is reversed with the NP-biased-materials (Fig. 8). With VP-attachment-biased materials (Rayner et al.'s materials), there is a reading time advantage for noun-fillers that accords with the VP attachment bias, which totals about 100 msec and is distributed over the three words following the noun-filler. However, with NP-attachment-biased materials (Taraban and McClelland's sentences), there is an approximately equal and opposite pattern; averaging the two types of materials, there is virtually no overall advantage for either type of attachment. Thus, the study indicates that content, rather than any general syntactic preference, appears to determine initial attachment preferences in this kind of construction.

Another experiment (Taraban & McClelland, 1990) addressed the source of the content-based influences on processing of the prepositional phrase. One possibility that is often considered is the idea that the verb may provide a basis for expectations about possible arguments that might influence the course of processing; these expectations could still be attributed to the workings of an autonomous syntactic process which nevertheless consulted syntactic information in the lexicon. In this experiment, Taraban and McClelland demonstrated, however, that the content of the object NP also influenced performance. For example, in sentence (27),

27. The dictator viewed the masses from the . . .
28. The dictator viewed the petition from the . . .

subjects expected a locative PP, attaching to the verb, indicating the place from which the viewing was to occur; whereas in (28), they expected a source of the petition, attaching to the object NP. When these expectations were violated, there was a slowdown in processing.

The experiments by Taraban and McClelland demonstrate that the content of the main verb, as well as that of the post-verbal object NP, can influence on-line processing decisions about PP attachment. They do not address whether or not the content of the *subject* nounphrase can also influence on-line parsing decisions, although it is known from Oden (1978) that it can influence the choice of the ultimate attachment. However, work reported by Tanenhaus, Carlson, and Trueswell (this issue) suggests that the semantic characteristics of the subject *can* influence on-line processing decisions in structures similar to (23) and (24) (see Tanenhaus et al. for further details). Thus, it would appear that evidence is accumulating in favour of the view that all parts of a sentence can influence on-line processing decisions about every other part.

Another of Taraban and McClelland's (1988) experiments considered the possibility that the disruption in processing that is occurring in these sentences is due to specific expectations for particular fillers rather than expectations concerning the role and/or attachment of the prepositional phrases. Though a small effect for particular fillers was found, the largest effect appeared to be due to violations of expectations for the role of the prepositional phrase. Violations of expected attachment had no further disruptive effect over and above that attributable to the inevitable concomitant violation of the subject's role expectations (see Taraban & McClelland, 1988, for details). These findings are certainly consistent with the SG model, in that there is no separate representation of the syntactic form of a sentence; there is, instead, direct processing into a representation which can be used to answer questions about the roles of the participants in the event that is described by the sentence.

In summary, the evidence from Taraban and McClelland's PP attachment studies seems consistent with the view that content can indeed play a role in setting up expectations for the roles played by the objects of prepositional phrases, and that these expectations can govern the initial processing of these phrases as they are encountered on-line in sentence processing. Tanenhaus et al.'s findings indicate that the effects of content are not restricted to lexical information that might be stored directly with heads of phrases. Though it is very clear that syntax often exerts an overriding influence, there is no reason to suppose on the basis of the studies reviewed here that it occupies a privileged or autonomous position in the initial processing of sentences. Instead, it appears that content as well as syntax can influence initial attachment and role assignment preferences.

Further arguments against the autonomy of syntax come from the research of Crain and Steedman (1985), Altmann and Steedman (1988), and Altmann (1988). These papers argue that attachment decisions can be governed by referential processes triggered by context presented prior to the sentence containing the ambiguity. Taken together with Taraban and McClelland's results, these results help paint a general picture in which syntax is far from autonomous.

Altmann and Steedman (1988) point out that the findings on attachment ambiguity resolution are consistent with a view they call "weak interactivity", in which a syntactic module constructs all possible parses and the candidate that best satisfies all of the constraints is selected by subsequent processes sensitive to content, referential coherence, etc. They point out that such a weak interactivity account is probably not distinguishable empirically from plausible versions of strongly interactive accounts, in which conceptual/referential modules in the language processing system instruct modules specialised for syntactic processing.

The view taken here goes beyond even strongly interactive accounts, in proposing that the syntactic and conceptual aspects of processing are in fact inextricably intertwined. Perhaps this view might best be called an integrative as opposed to an interactive account. Interactivity suggests separate systems that exert simultaneous mutual influence (cf. McClelland, 1987; Rumelhart & McClelland, 1981), even though they construct separate representations of different kinds of information. In the present approach, there is but a single integrated system in which syntactic and other constraints are combined in the connection weights, to influence the construction of a single representation reflecting the influences of syntactic, semantic, and lexical constraints.

Neuropsychological Dissociations

This integrative approach is actually quite different from the position one of us has taken in previous publications (McClelland, 1987). We have adopted it here, not out of any strong *a priori* commitment, but because it has turned out to work well in capturing the phenomena considered in this paper. Indeed, the notion that there is a separate module for syntax is so ingrained in theoretical treatments of language processing, that it is difficult even for us to be fully comfortable with abandoning it. But, the successes of the SG model in dealing with some of the central difficulties facing conventional approaches, coupled with the fact that the empirical evidence is beginning to favour at least some form of an interactive account, makes us feel that it is worthwhile to see if indeed there is any real basis for this very general acceptance of some form of modularity.

In this connection, it is worth considering evidence from neuropsychology, because some of the most often-cited evidence for the view that there are separate processing systems for syntactic and conceptual information come from neuropsychological dissociations. It is generally claimed, for example, that Wernicke's aphasics have a general deficit in the comprehension of word and sentence meaning, which interferes with their understanding of all sentences regardless of their syntactic complexity, whereas Broca's aphasics have a specific deficit in the ability to make use of syntactic information for comprehension. Such dissociations invite a modularist approach, in which one part of the system is specialised for the use of content information and the other for the use of syntactic cues in comprehension. Could such findings possibly be consistent with the framework considered here, in which syntactic and content-based influences on processing are inextricably intertwined in the structure of the language processing mechanism?

In fact, the notion that the difference between Wernicke's and Broca's aphasics can be characterised in terms of syntax and semantics is being called into question from several different vantage points. First, Milberg, Blumstein, and Dworetzky (1988) have recently reported that both Wernicke's and Broca's aphasics differ from normals in lexical access, though the differences are complementary. Normals show a graded decrement in priming as primes are increasingly distorted, but Broca's aphasics show priming only when the prime is undistorted, and Wernicke's aphasics show priming over a wider range of distortion than normals. This suggests that Wernicke's aphasics may be suffering from something akin to undamped activation, whereas Broca's aphasics are suffering from overdamping.

Other studies suggest that Broca's and Wernicke's aphasics both show comprehension deficits, and that the deficts differ more between aphasics who speak different languages than they differ between different types of aphasics who speak the same language. For example, Bates, Friederici, Miceli, and Wulfeck (1985) studied groups of Broca's aphasics, Wernicke's aphasics, and normal controls who were native speakers of each of three different languages—English, German and Italian. They found that within each language, Broca's and Wernicke's aphasics both showed deficits in the use of morphological cues, and that the degree of preservation of the use of these cues correlated with the extent of reliance on these cues in the speaker's language. Italians rely much more on agreement and much less on word order than English speakers, and the Italian aphasics showed the least impairment in the use of subject–verb agreement to mark agency, whereas English aphasics showed the greatest impairment. German is intermediate between the two languages in the extent of normal reliance on word order *vs* agreement cues, and showed an intermediate degree of disruption of the use of agreement with damage. The findings of this study are consistent with the idea that both aphasic groups show the greatest deficits in the use of cues that are relatively weaker in their native language (Bates & Wulfeck, in press; McDonald and MacWhinney, in press), and tend to run counter to the notion that Broca's and Wernicke's aphasia differentially impact mechanisms specific to syntactic and semantic aspects of comprehension, respectively.

We do not mean to suggest that there is no basis at all for the idea that there may be specific dissociations of aspects of linguistic knowledge that call into question the idea that content and syntactic constraints are as fully integrated as they are in the approach that we have taken. There are several studies which support the idea that there are particular deficits in the use of closed-class words that are restricted to Broca's and not to Wernicke's aphasics, which have yet to be reconciled with the type of account suggested by Milberg et al.'s findings, as well as other evidence which has been taken as favouring the existence of autonomous syntactic structures (Caplan & Hildebrandt, 1988). Our only claim here is that the neuropsychological evidence is not completely clear-cut, and there is room for a consideration of the idea that there may indeed be a single processing system that is simply thrown out of regulation in slightly different ways in Broca's and Wernicke's aphasics. The model we have proposed does not of course offer any insight into this differential disruption, but the model is compatible with the idea that there is a single system which uses syntax and content together to guide the language comprehension process.

FUTURE DIRECTIONS

In this paper, we have described an alternative to traditional models of language processing. We have tried to indicate how this alternative may allow us to solve many of the problems facing traditional approaches, and how it may provide different ways of conceptualising basic aspects of the problem of comprehension. We have also indicated that many of the arguments against the type of approach we have taken can be countered. Of course, the facts are not all in, but given what is known at this time the approach seems to us to be at least as viable as any other that we know of. The model we have offered is far from the final word, and many problems need to be addressed. Our only goal has been to suggest that there may be some basis for optimism that further development of the approach might be successful.

There are several further steps that need to be taken. First, we need to find ways of improving the rate of learning; as things stand, learning is unduly slow, especially given the small size of the corpora that we have used in our training experiments. Secondly, we need to extend the framework to force the construction of representations that can answer more sophisticated questions than merely the completion of role–filler pairs. As previously noted, the role–filler completion task that we have used here has several limitations; the role–filler pair language is insufficiently structured, and the localist representation of concepts lacks the reliance on distributed representations which is one of the strengths of the PDP framework. Thirdly, our long-term goal is to move in the direction of capturing the influence of broader, extra-sentential context on sentence processing. Ultimately, the approach will stand or fall on its ability to capture the pervasive influences of these extra-sentential factors.

REFERENCES

Allen, R. B. (1987). Several studies on natural language and back-propagation. *Proceedings of the IEEE First Annual International Conference on Neural Networks*, Vol. II, 287–298. Boston.

Allen, R. B. (1988). Sequential connectionist networks for answering simple questions about a microworld. *Proceedings of the Tenth Annual Conference of the Cognitive Science Society*, pp. 489–495. Montreal.

Altmann, G. (1988). Ambiguity, parsing strategies, and computational models. *Language and Cognitive Processes*, *3*, 73–97.

Altmann, G. & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition*, *30*, 191–238.

Bates, E. & Wulfeck, B. (in press). Crosslinguistic studies of aphasia. In B. MacWhinney & E. Bates (Eds), *The crosslinguistic study of sentence processing*. New York: Cambridge University Press.

Bates, E., Friederici, A., Miceli, G., & Wulfeck, B. (1985). Sentence comprehension in aphasia: A cross-linguistic study. Manuscript, University of California, San Diego.

Bever, T. G. (1970). The cognitive basis of linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language*. New York: John Wiley.

Caplan, D. & Hildebrandt, N. (1988). *Disorders of syntactic comprehension*. Cambridge, Mass.: MIT Press.

Chomsky, N. (1988). Lecture presented at the University of Pittsburgh, Fall, 1988.

Clark, H. H. & Clark, E. V. (1977). *Psychology and language: An introduction to psycholinguistics*. New York: Harcourt Brace Jovanovich.

Crain, S. & Steedman, M. J. (1985). On not being led up the garden path: The use of context by the psychological parser. In D. Dowty, L. Karttunen, & A. Zwicky (Eds), *Natural* *language parsing: Psychological, computational, and theoretical perspectives*. Cambridge: Cambridge University Press.

Elman, J. L. (1988). Finding structure in time. CRL Technical Report 8801. Center for Research in Language, University of California, San Diego.

Elman, J. L. (1989). Representation and structure in connectionist models. CRL Technical Report 8903. Center for Research in Language, University of California, San Diego.

Elman, J. L. & McClelland, J. L. (1986). Exploiting the lawful variability in the speech wave. In J. S. Perkell & D. H. Klatt (Eds), *Invariance and variability of speech processes*. Hillsdale, N.J.: Lawrence Erlbaum Associates Inc.

Erickson, T. D. & Mattson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior*, *20*, 540–551.

Ferreira, F. & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25*, 348–368.

Fillmore, C. (1968). The case for case. In E. Bach & R. T. Harms (Eds), *Universals in linguistic theory*. New York: Holt, Rinehart and Winston.

Fodor, J. A. & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*, 3–71.

Frazier, L. (1986). Theories of sentence processing. In J. Garfield (Ed.), *Modularity in knowledge representation and natural language processing*. Cambridge, Mass.: MIT Press.

Gleitman, L. R. & Wanner, E. (1982). Language acquisition: The state of the state of the art. In E. Wanner & L. R. Gleitman (Eds), *Language acquisition: The state of the art*. Cambridge: Cambridge University Press.

Hinton, G. E. (1986). Learning distributed representations of concepts. *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, Amherst, Mass.

Hinton, G. E., McClelland, J. L., & Rumelhart, D. E. (1986). Distributed representations. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds), *Parallel distributed processing: Explorations in the microstructure of cognition*, Vol. I. Cambridge, Mass.: Bradford Books.

Just, M. A., Carpenter, P. A., & Wooley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General, 111*, 228–238.

Kawamoto, A. H. (1985). Dynamic processes in the (re)solution of lexical ambiguity. Doctoral dissertation, Department of Psychology, Brown University.

Kawamoto, A. H. (1988). Distributed representations of ambiguous words and their resolution in a connectionist network. In S. L. Small, G. W. Cottrell, & M. K. Tanenhaus (Eds), *Lexical ambiguity resolution: Perspectives from psycholinguistics, neuropsychology, and artificial intelligence*. San Mateo, Calif.: Morgan Kaufmann.

Marcus, M. P. (1980). *A theory of syntactic recognition for natural language*. Cambridge, Mass.: MIT Press.

McClelland, J. L. (1979). On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review*, *86*, 287–330.

McClelland, J. L. (1987). The case for interactionism in language processing. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading*, pp. 1–36. London: Lawrence Erlbaum Associates Ltd.

McClelland, J. L. (in press). Parallel distributed processing: Implications for cognition and development. In R. Morris (Ed.), *Parallel distributed processing: implications for psychology and neurobiology*. Oxford: Oxford University Press.

McClelland, J. L. & Elman, J. L. (1986). An architecture for parallel processing in speech recognition: The TRACE model. In M. R. Schroeder (Ed.), *Speech recognition*. Basel: S. Krager AG.

McClelland, J. L. & Kawamoto, A. H. (1986). Mechanisms of sentence processing: Assigning roles to constituents. In J. L. McClelland, D. E. Rumelhart, & the PDP Research

Group (Eds), *Parallel distributed processing: Explorations in the microstructure of cognition*, Vol. II. Cambridge, Mass.: Bradford Books.
McClelland, J. L. & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception, Part 1: An account of basic findings. *Psychological Review*, *88*, 375–407.
McDonald, J. & MacWhinney, B. (in press). Maximum likelihood models for sentence processing. In B. MacWhinney & E. Bates (Eds), *The crosslinguistic study of sentence processing*. New York: Cambridge University Press.
Miikkulainen, R. & Dyer, M. G. (1989). A modular neural network architecture for sequential paraphrasing of script-based stories. Technical Report UCLA-AI-89-02. University of California, Artificial Intelligence Laboratory, Computer Science Department, Los Angeles.
Milberg, W., Blumstein, S., & Dworetzky, B. (1988). Phonological processing and lexical access in aphasia. *Brain and Language*, *34*, 279–293.
Oden, G. (1978). Semantic constraints and judged preference for interpretations of ambiguous sentences. *Memory and Cognition*, *6*, 26–37.
Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, *22*, 358–374.
Rumelhart, D. E. & McClelland, J. L. (1981). Interactive processing through spreading activation. In C. Perfetti & A. Lesgold (Eds), *Interactive processes in reading*. Hillsdale, N.J.: Lawrence Erlbaum Associates Inc.
Rumelhart, D. E. & McClelland, J. L. (1982). An interactive activation model of context effects in letter perception, Part II: The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, *89*, 60–94.
Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986a). Learning internal representations by error propagation. In D. E. Rumelhart, J. L. McClelland, & the PDP Research Group (Eds), *Parallel distributed processing: Explorations in the microstructure of cognition*, Vol. I. Cambridge, Mass.: MIT Press.
Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986b). Parallel distributed processing models of schemata and sequential thought processes. In J. L. McClelland, D. E. Rumelhart, & the PDP Research Group (Eds), *Parallel distributed processing: Explorations in the microstructure of cognition*, Vol. II. Cambridge, Mass.: Bradford Books.
St. John, M. F. (1988). Hitting the right pitch: A meta-analysis of the processing of ambiguous words in context. Manuscript.
St. John, M. F. & McClelland, J. L. (1987). Reconstructive memory for sentences: A PDP approach. Paper presented to the Ohio University Inference Conference, *Proceedings Inference: OUIC 86*. University of Ohio, Athens, Ohio.
St. John, M. F. & McClelland, J. L. (in press). Learning and applying contextual constraints in sentence comprehension. *Artificial Intelligence*.
Schank, R. C. (1981). Language and memory. In D. A. Norman (Ed.), *Perspectives on cognitive science*. Norwood, N. J.: Ablex.
Servan-Schreiber, D., Cleeremans, A., & McClelland, J. L. (1988). Encoding sequential structure in simple recurrent networks. CMU-CS-88-183. Carnegie Mellon University, Computer Science Department, Pittsburgh.
Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, *18*, 645–659.
Tanenhaus, M. K., Leiman, J. M., & Seidenberg, M. S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior*, *18*, 427–440.

Taraban, R. (1988). Content-based expectations: One source of guidance for syntactic attachment and thematic role assignment in sentence processing. Ph.D. dissertation, Carnegie Mellon University, Pittsburgh.

Taraban, R. & McClelland, J. L. (1988). Constituent attachment and thematic role assignment in sentence processing: Influences of content-based expectations. *Journal of Memory and Language*, *27*, 597–632.

Taraban, R. & McClelland, J. L. (1990). Parsing and comprehension: A multiple-constraint view. In D. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds), *Comprehension processes in reading*. Hillsdale, N.J.: Lawrence Erlbaum Associates Inc.

Touretzky, D. S. (1988). Connectionism and PP attachment. In D. Touretzky, G. Hinton, & T. Sejnowski (Eds), *Proceedings of the 1988 Connectionist Mosels*. San Mateo, Calif.: Morgan Kaufmann.

Touretzky, D. S. & Geva, S. (1987). A distributed connectionist representation for concept structures. Paper presented to the 9th Annual Conference of the Cognitive Science Society, Seattle.

Touretzky, D. S. & Hinton, G. E. (1988). A distributed connectionist production system. *Cognitive Science*, *12*, 423–466.

van Gelder, T. (in press). Compositionality: Variation on a classical theme. *Cognitive Science*.