Theories of Memory

edited by

Alan F. Collins,
Susan E. Gathercole,
Martin A. Conway, and
Peter E. Morris

Memory Research Unit, Lancaster
University, UK

Copyright © 1993 by Lawrence Erlbaum Associates Ltd.
All rights reserved. No part of this book may be reproduced in any
form, by photostat, microform, retrieval system, or any other
means without the prior written permission of the publisher.
CHAPTER 3

Flexibility, Structure, and Linguistic Vagary in Concepts: Manifestations of a Compositional System of Perceptual Symbols

Lawrence W. Barsalou
Department of Psychology, University of Chicago, USA

INTRODUCTION

The concept of concept is notoriously slippery, taking diverse forms not only across the cognitive science disciplines, but also across perspectives within disciplines. In this chapter, I develop the view that a concept is a temporary construction in working memory, derived from a larger body of knowledge in long-term memory to represent a category, where a category, roughly speaking, is a related set of entities from any ontological type (e.g., robins, sweaters, weddings, mountains, plans, anxieties).\(^1\) Across contexts, a given person’s concept for the same category may change, utilising different knowledge from long-term memory, at least to some extent (Barsalou, 1982, 1987, 1989; Barsalou & Billman, 1989). For example, the concept that someone constructs for raccoons on one occasion might in part include the features furry, white stripes, and playful, whereas the concept constructed on another occasion might include the features furry, nocturnal, and scavenger. Later sections will develop this view of concept in greater detail.

By concept, I do not mean the objectively correct definition of a category that might exist independently of human observers (Frege, 1892/1952; Rey, 1983), nor do I mean the scientific definition of a...

cognitive representation of a category on a particular occasion,
regardless of its accuracy, although human concepts must be at least
somewhat accurate to be as useful as they are. By concept, I do not
necessarily imply conscious representations of categories. Certainly,
some information in a concept may be conscious on occasion, yet much
information may remain unconscious. Instead, I focus on the
information in concepts that allows people to classify exemplars during
perception, to process words semantically during language use, and to
reason about categories in induction, problem solving, decision making,
and other forms of thought. As is typical of information processing and
computational theory in cognitive science, I focus on accounts of concepts
that are sufficient to explain task performance (Barsalou, 1992a, pp.
8–12).

In particular, I focus on three properties of concepts that seem
particularly revealing of the cognitive mechanisms that produce them:
flexibility, structure, and linguistic vagary. Concepts are flexible. Rather
than existing as a stable set of features in different people, and in the
same person across contexts, a concept varies widely both between and
within individuals. Concepts are structured. Rather than being a list of
independent features, a concept is a hierarchical relational structure,
containing attribute–value sets, structural invariants, constraints, and
recursion. Concepts exhibit linguistic vagary. Rather than being coherent,
consistent, and complete, linguistic descriptions of conceptual content are unprincipled, haphazard, and incomplete. In the next three
sections, I address flexibility, structure, and linguistic vagary in turn.

In the process of examining these three properties, I consider their
implications for various accounts of concepts, especially connectionist
networks and frame theory. Although connectionist nets provide an excellent account of flexibility, they have difficulty representing structure, and they have little to say about linguistic vagary. Although
frame theory represents structure exquisitely, it fails to explicate the
cognitive mechanisms that produce structure, it provides no account of
flexibility, and it has little to say about linguistic vagary.

In the second half of this chapter, I propose a cognitive architecture
for explaining flexibility, structure, and linguistic vagary. In this
architecture, perceptual symbols—not linguistic symbols or amodal
propositions—constitute the cores of concepts. I describe how selective
attention extracts perceptual symbols from experience, and how
compositional mechanisms integrate them productively during
conceptual combination, imagery, and comprehension. I argue that the
amodal propositions found in most theories of concepts are a theoretical
extravagance, and that perceptual symbols provide a more conservative

and better motivated account of the phenomena that amodal
propositions explain. For example, I illustrate how perceptual symbols
can represent abstract concepts and memory for gist. I further suggest
that linguistic symbols become grounded in perceptual symbols, and
that pairs of linguistic and perceptual frames emerge from these
symbols to represent categories. Finally, I illustrate how interactions
between the linguistic and perceptual frames for a category produce
flexibility, structure, and linguistic vagary in the concepts constructed
for it.

**FLEXIBILITY**

Much evidence reveals the flexibility of human concepts. The
conceptualisation of an entity or set of entities can vary widely across
individuals and occasions. Numerous demonstrations of encoding
variability in the memory literature reveal flexibility, illustrating that
the encoding context of a word determines the conceptualisation of its
referent (e.g. Anderson & Ortony, 1975; Anderson et al., 1976; Barclay
et al., 1974; Greenspan, 1986; Thomson & Tulving, 1970; Tulving &
Thomson, 1973). For example, Barclay et al. (1974) presented subjects
with the word “piano” either in the context of producing music or in the
context of moving furniture. When subjects received “piano” in the
music context, cues relevant to the musical properties of pianos
functioned as optimal cues for retrieval, but when subjects received
“piano” in the moving context, cues relevant to the weight of pianos
functioned as optimal cues instead. One interpretation of this result is
that the concepts constructed for piano in different contexts contained
different features, such that different cues were later effective in
retrieving them.

Word on lexical access similarly demonstrates that people incorporate
different features into a concept depending on the encoding context.
Consider the word “newspaper” and note which of its features come to
mind. Now consider the word “newspaper” in the context of building a
fire. Whereas the feature flammable probably didn’t come to mind when
you considered “newspaper” in isolation, it probably did when you
considered it in the context of building a fire. Many researchers have
implemented such manipulations in experiments and observed large
effects on verification time (e.g. Barsalou, 1982; Conrad, 1978; Roth &
Shoben, 1983; Whitney, McKay, & Kellas, 1985). For example, Barsalou
(1982, Experiment 1) found that certain features were verified 145
milliseconds faster in a relevant context than in a neutral context. Such
a large difference in latency suggests that these features were inactive
in neutral contexts but became active in relevant contexts. Barsalou
Changes in context reveal still further flexibility. In another set of studies, we manipulated the context in which subjects performed typicality judgements (Barsalou & Sewell, 1984, in prep.; Sewell, 1985). To manipulate context, we had subjects judge the typicality of category exemplars from different points of view. For example, we asked some subjects to judge the typicality of various birds from the point of view of the average American citizen, and we asked other subjects to judge the typicality of birds from the point of view of the average Chinese citizen. This point of view manipulation had huge effects on prototypicality. For example, the ordering of birds by typicality completely inverted. Birds that were typical from the American point of view, such as robin and eagle, were atypical from the Chinese point of view. Birds that were typical from the Chinese point of view, such as swan and peacock, were atypical from the American point of view. Depending on the context, subjects completely inverted their conceptions of prototypicality. Many similar effects of context on conceptual content exist in the literature (e.g. Murphy, 1988; Roth & Shoben, 1983; Schwanenflugel & Rey, 1986).

The various forms of flexibility that we have been considering do not necessarily reflect differences in the underlying knowledge that people store in long-term memory for a category. Different individuals appear to have at least some highly similar knowledge for the same category, and the knowledge of a given individual appears to remain highly stable over time. We discovered this in the following experiment (Barsalou et al., in prep.). Recall the study in which we asked subjects to produce definitions for categories. Further recall that there were substantial differences both between and within subjects in the definitions they produced. To assess the stability of knowledge in long-term memory, we pooled all of the features that these subjects produced for the same category. For example, we pooled all of the features that subjects produced for bird. We then presented the full pool of features for each category to a new group of subjects, whose task was to specify whether each property was potentially true of its respective category. For example, subjects received each property for bird and assessed whether the feature was a feature of some bird. If different subjects have different knowledge, then they should differ in their judgments of feature validity. Features that are valid for one subject should be invalid for another. Similarly, if a given subject’s knowledge of a category changes over time, their assessments of feature validity should change as well. Features that are valid on one occasion should be invalid on another. Surprisingly, we found virtually perfect agreement both between and within subjects, with between-subject agreement being 97% overlap on the average, and within-subject agreement being 98%.
These results demonstrate that different people store very similar information for the same category in long-term memory, and that this information remains highly stable within individuals over time. The tremendous flexibility that we have seen in other experiments arises not from differences in knowledge, but from differences in the retrieval of this knowledge. On a given occasion, different people retrieve different subsets of features from their knowledge of a category. Across occasions, the same person retrieves different subsets. These varying subsets of features are what I am defining as concepts. Rather than being stable structures stored in long-term memory and retrieved as needed, concepts are temporary constructions in working memory.

One might question this definition of concept. Why is a concept the temporary representation of a category of working memory? Why isn’t it the stable knowledge of a category in long-term memory from which temporary representations are constructed? The critical issue seems to me to concern which of these two representations controls behaviour. Consider an experiment by McCloskey and Glucksberg (1978), in which subjects received the names of entities (e.g. “amoeba”) and were asked whether they belonged to a natural category (e.g. animals). Many entities that a subject admitted to the category on one occasion were excluded in a second session a month later, and vice versa. Because category membership can change significantly, there doesn’t appear to be a single representation, such as the stable knowledge of a category in long-term memory, that controls categorisation behaviour. Instead, long-term memory appears to contain a tremendous amount of loosely related and somewhat inconsistent information capable of producing contradictory behaviour on different occasions. If the knowledge of a category in long-term memory as a whole were controlling behaviour, we wouldn’t see the tremendous variability in performance that we do, not only in category membership but also in typicality, definitions, and probably most other categorisation tasks. Conversely, the temporary conceptualisations of categories constructed in working memory on particular occasions appear to be controlling categorisation behaviour. Because these temporary conceptualisations are doing the traditional work of concepts, I refer to them as concepts, and I use knowledge for referring to the body of information in long-term memory from which concepts are constructed.

Accessibility appears to be the critical factor that underlies which features are retrieved from knowledge of a category to construct a concept on a particular occasion. The particular features retrieved by a particular person in a particular context are those that are currently most accessible. Between individuals, different subsets of features for a concept are highly accessible on a given occasion, such that people retrieve different information. Within individuals, different feature subsets are highly accessible on different occasions, such that different information becomes active.

Three well-known factors appear to determine accessibility: frequency, recency, and context (Barsalou, 1987, 1989). Features processed frequently are likely to be highly accessible, as are features processed recently, and features associated with the current context. These three factors produce instability in the representations of a concept, when they produce differences in the accessibility of features. If two people have processed different features frequently, if they have processed different features recently, and if they are in different contexts, the features they retrieve for the same concept will differ. Although similar features for the category reside in each person’s long-term memory, the accessibility of these features differs, such that different subsets become active. However, frequency, recency, and context can also produce high levels of stability, when they cause people to converge on common features. If two people have processed similar features frequently, if they have processed similar features recently, and if they are in similar contexts, they will retrieve similar features.

Consider the following experiment, where a common context increased concept stability substantially (Barsalou et al., in prep.). In one condition, subjects judged typicality in a neutral context. For example, subjects simply judged the typicality of various vehicles. These subjects exhibited between-subject agreement of 0.45 and within-subject agreement of 0.81 in their judgements of typicality. Other subjects judged typicality in simple contexts. For example, these subjects judged the typicality of vehicles in the context of taking a vacation in the rugged mountains of Mexico. In the context condition, agreement climbed to 0.70 between subjects and 0.88 within subjects. This substantial effect illustrates that a simple and unfamiliar context can greatly constrain the retrieval process, increasing the accessibility of shared features relevant to the context and thereby causing subjects to represent the category more similarly. Context focuses retrieval, such that people establish common ground in linguistic and social interaction.

Accounting for Flexibility with Connectionist Models

Connectionist models provide a natural means of accounting for the dynamic accessibility I have just described (e.g. J.A. Anderson, 1983; Grossberg, 1987; Hinton, 1981; McClelland, Rumelhart, & the PDP Research Group, 1986; Rumelhart, McClelland, & the PDP Research Group, 1986). The basic learning mechanisms of connectionist models are acutely sensitive to the statistical properties of information, such
that they can account for the effects of frequency, recency, and context on accessibility. For example, the connection weight that links a feature to a concept generally increases as their co-occurrence frequency increases. Typically, but not always (Gluck & Bower, 1988), the more often a feature and concept co-occur, the stronger the connection between them becomes, such that each can activate the other to a greater extent. Recency has a similar effect. To the extent that a feature and concept have co-occurred recently, the strength of the connection between them increases. This property of connectionist learning can in fact be so dominant that it underlies the serious problem of catastrophic interference (McCloskey & Cohen, 1989; Ratcliff, 1990). Finally, connectionist nets are exquisitely sensitive to context. To the extent that the features of a concept occur in different contexts, they become associated with these contexts, such that a particular context biases activation towards its associated features. In general, because connectionist nets are so sensitive to the statistical properties of features, they are extremely well suited for explaining conceptual flexibility.

**CONCEPTUAL STRUCTURE**

In reviewing the flexibility of concepts and promoting connectionism as an account of it, I have assumed implicitly that feature list representations provide a satisfactory account of conceptual content. On this view, a concept contains a set of unrelated features, with each feature representing a possible property that exemplars from the respective category possess. For example, a feature list for *bird* might contain features for *wings, feathers, flies,* and *builds nests.* It is important to note that a feature list does not specify how its features are related conceptually. For example, this representation of *bird* does not specify that the relation between *wings* and *flies* is that *wings enable flying.*

A wide variety of models in cognitive psychology and cognitive science assume feature list representations (see Barsalou & Hale, 1992, for a review). Exemplar and prototype models of categorisation often assume that feature lists represent concepts (e.g. Barsalou, 1990; Brooks, 1978, 1987; Estes, 1986; Gluck & Bower, 1988; Hintzman, 1986; Homa, 1984; Jacoby & Brooks, 1984; Medin & Schaffer, 1978; Nosofsky, 1984; Posner & Keele, 1968; Rosch & Mervis, 1975). Similarly, theories of episodic memory generally assume that memory traces are lists of features (e.g. Eich, 1982; Gillund & Shiffrin, 1984, Figure 8; Hintzman, 1988; Murdock, 1982). Connectionist models also use feature list representations, where each processing unit represents a local feature or coarse codes several. For example, an input unit might represent the presence or absence of *red,* and a hidden unit might represent the co-occurrence of *red* and *large.*

There are compelling reasons for believing that feature lists provide grossly inadequate representations of concepts. As we shall see, feature lists provide only fragmentary bits and pieces of a concept's content and structure. In a sense, feature lists are a lot like a few fragments of a dinosaur skeleton, from which a palaeontologist attempts to infer its underlying, more complete structure. Although features represent valid pieces of a concept, they only represent independent bits and pieces, rather than a complete and coherent account (Barsalou & Hale, 1992; Murphy & Medin, 1985). For this reason, standard connectionist models, even though they provide an excellent account of conceptual flexibility, are inadequate models of concepts, along with all other models that assume feature list representations.

**Attribute-value Sets.** The general problem with all feature list models is their inability to represent conceptual relations. One conceptual relation that feature lists cannot represent adequately is the binding relation between attributes and values. Consider the representations of *banana, peach,* and *lime* in Fig. 3.1a. These three feature lists fail to capture the attribute-value relations that people know for these concepts. For example, the feature for *yellow* in *banana* fails to represent people's knowledge that the *colour of banana* is *yellow.* Instead, this feature simply represents the presence of *yellow,* omitting its binding to the attribute for *colour.* Furthermore, these feature lists fail to represent the fact that *yellow, orange,* and *green* are related, being contrastive values of the same attribute. Because feature lists do not represent conceptual relations between features, they cannot represent the binding relations between an attribute and its values, nor the contrastive relations between values of the same attribute.

Many findings demonstrate the importance of attribute-value relations in cognition (see Barsalou, 1992b, for a brief review). Animals are sensitive to such relations, as the classic effects of intra-dimensional transfer, reversal shift, and other dimensional learning phenomena illustrate (e.g. Medin, 1976; Sutherland & Mackintosh, 1971). Animals learn not only which features predict reward, but also the attributes to which these features are bound. Garner's (1974, 1978) work on separable dimensions illustrates how readily people process features as values of attributes. For example, people can process visual stimuli with respect to their *shape* while ignoring their *colour.* Attribute-value relations play fundamental roles in theories of language. In phrase structure grammars, key attributes include the various syntactic...
in the input, as feature list theories assume. Instead, people bind features in the input to more general attributes. On encountering someone who is a white adult female, perceivers don’t simply note the presence of the features, *white, adult*, and *female*; instead they bind these features to attributes for *race, age, and gender*. People know these attributes and recognize specific features as values of them. Because feature lists do not represent the attribute–value relations pervasive in cognition, they are inadequate representations of concepts.

**Structural Invariants.** Feature lists suffer from another, equally damaging, limitation: They cannot represent structural invariants between features. Consider the feature list representation of *chair* in Fig. 3.1b. This feature list fails to represent the spatial relations between attributes, which everyone knows as part of their knowledge for *chair*. For example, this feature list does not represent the relations, ABOVE (*back, seat*), PERPENDICULAR (*seat, back*), ABOVE (*seat, legs*), and so forth, when a chair is in its conventional orientation. Such relations are structural invariants, because they represent relations between attributes that hold true across most exemplars of a category, although not necessarily across all. Across most chairs, for example, the seat and back maintain a relatively invariant structural relation to one another. Structural invariants are not only spatial but take numerous other forms as well, including relations of time, causality, and possession (see Barsalou, 1992b, for a brief review; see also Gentner, 1989). Because feature lists fail to represent these relations, they are inadequate as conceptual representations.

**Conceptual Constraints.** Still another critical relation that feature lists fail to represent are conceptual constraints. Whereas structural invariants hold true across most exemplars of a concept, conceptual constraints distinguish exemplars from one another (Barsalou, 1991, 1992b). Consider the concept for *ski vacation* in Fig. 3.1c. A structural invariant, not shown in the figure, exists between activity and location, stating that the activity occurs at the location. This is a structural invariant, because this relation remains relatively constant across most vacations, regardless of whether they are *ski vacations, surfing vacations*, and so forth. In contrast, consider the constraints that relate specific values of *activity* and *location*. If a planner selects *downhill skiing* as the value of *activity*, then this constrains the values for *location*. The planner can’t select any value for *location*, else they might end up trying to *downhill ski in Egypt*. Instead, *downhill skiing* as the value of *activity* constrains the values of *location* to be *mountainous*. As Fig. 3.1c illustrates, feature lists fail to capture such constraints.
Constraints are not structural invariants, because they do not hold true across most exemplars of a concept. For example, a *surfing vacation* certainly doesn't include the constraint that *downhill skiing requires a mountainous location*. Instead, this exemplar of *vacation* includes a very different constraint, *surfing requires an ocean*. Whereas structural invariants capture relatively constant relations between attributes, constraints capture relations between the values of these attributes in specific exemplars.

Note that we can't account for conceptual constraints simply with connection weights, as in a standard connectionist model. Whereas constraints in a connectionist model are statistical, representing a degree of co-occurrence or prediction, the constraints I have described are conceptual as well. To see this, consider the constraint between *downhill skiing* as a value of *activity*, and *mountains* as a value of *location*. Unlike a connection weight, this constraint is not simply statistical, nor is it symmetrical as connectionist weights often are. Instead, this constraint is a *requires* relation in one direction and an *enables* relation in the other. *Skiiing requires mountains*, whereas *mountains enable skiing*. For a model of concepts to be adequate, it must represent the conceptual nature of these constraints, as well as their statistical strength.

**Recursion.** Finally, feature list models fail to capture the recursive, decompositional nature of conceptual knowledge (Barsalou, 1992b; Minsky, 1977, 1985; Rumelhart & Ortony, 1978). To see this, consider the feature list representation of the concept for human in Fig. 3.1d. Although this feature list includes features for *arms, hands*, and *fingers*, it fails to represent the obvious fact that an arm decomposes into more specific parts, including hands, and that similarly a hand decomposes into more specific parts, including fingers. Instead, this feature list simply represents features at a single flat level that fails to acknowledge their nested partonomic structure. By recursion at this point, I simply mean the recursive *process* by which the content of a concept can be continually decomposed. As will become clearer in the next section, recursion more importantly refers to the recursive embedding of *frame structure* in a concept, analogous to the well-known recursive embedding of rules in phrase structure grammars (Chomsky, 1965, 1965).

In general, any component of a conceptual representation can be decomposed recursively into more specific features. In our planning protocols, we observed extensive decomposition of this sort (Barsalou, 1991). Consider the *location* attribute in the concept for vacation. Subjects often decomposed this attribute and its values into more specific attributes, such as *climate, accommodations*, and *distance from home*. Not only do people decompose attributes and values, they also decompose relations. Many theorists assume that relations such as *isa, part, cause*, and *requires* are conceptual primitives. However, Chaffin and his colleagues provide evidence that these relations decompose into more specific representational components (Chaffin, 1992; Winston, Chaffin, & Herrmann, 1987). Finally, it is easy to show that constraints decompose as well, such as the constraint *requires* having attributes for *source of the requirement, likelihood of being in effect, applicable conditions*, and so forth (Barsalou, 1992b). Because feature list representations cannot represent decompositional relations, they are once again inadequate.

**Frame Theory**

What kinds of representations are sufficiently expressive to account for the hierarchical relational structure of concepts? Traditionally, psychologists and computer scientists have adopted frame and schema representations to represent this structure. Frames and schemata, which I will consider equivalent, result from integrating the four types of relations we have just considered: If we integrate attribute—value bindings, structural invariants, constraints, and recursive decomposition, we obtain frame and schema representations of knowledge. On this view, a concept contains multiple levels of hierarchical structure, with components at one level decomposing recursively into more specific components at the next. At each level, attributes bind to values, structural invariants integrate attributes, and conceptual constraints integrate values. Frames potentially represent each of these components; in turn the components of these frames are represented by still more specific frames; and so forth.

Knowledge appears to be frames all the way down, with every component of a frame potentially decomposing recursively into a more specific frame. Similar to the self-similarity of fractals (Gleick, 1987), the same structure that occurs at a frame's most global level—attributes integrated by structural invariants, bound to values integrated by constraints—occurs at every more specific level of analysis as well (Barsalou, 1992b). Note that this analogy to fractals and self-similarity concerns the *form* of a frame and not its content. Clearly, a frame's content varies widely at different levels of analysis. In the frame for human body, the content for arm differs from the content for fingernail. Where self-similarity does exist is in the constant form that this content takes at each hierarchical level, namely, a configuration of attributes integrated by structural invariants, bound to values integrated by constraints. Self-similarity of form but not content also occurs in phrase
structure grammars, where the syntactic form of a rule remains constant at each recursive level, while its semantic content changes, for example the NP rule (Chomsky, 1957, 1965).

Elsewhere I have argued that these four relations—attribute–value bindings, structural invariants, conceptual constraints, and frame recursion—constitute sufficient conditions for frames and schemata (Barsalou, 1992b; Barsalou & Hale, 1992). Psychologists often criticize schemata as being excessively vague, yet it is possible to specify their contents precisely. Computer scientists rarely attempt to define frames, because they typically view them as an implementation detail, yet these four relations appear sufficient to describe these implementations. The hierarchical, relational representations that result from combining these four basic relations are dense, complex, and messy, but if one attempts to assess the content of human knowledge realistically, it takes this form. Feature lists only represent bits and pieces of this structure, analogous to the relation between a few dinosaur fossils and the dinosaur to which they once belonged.

Although frames and schemata have the expressive power to represent the structural properties of human knowledge, satisfactory accounts of their acquisition and use in human cognition do not exist. The processing environments that do exist have evolved primarily to serve formal and computational goals. For example, frames, which are sometimes viewed as an application of predicate calculus, can be processed with the rules of deductive logic (Hayes, 1979). However, psychologists tend not to believe that deductive theorem proving, at least in its pure form, constitutes a viable account of cognitive processing (Johnson-Laird, 1983; cf. Rips, 1985). A second popular processing environment for frames, and in fact the dominant environment, is the LISP programming language of artificial intelligence. Again, few psychologists view the list processing operations in LISP, such as CAR and CDR, as plausible cognitive processes.

Nevertheless, environments exist that offer viable processing accounts of frames. For example, production systems that operate on propositional networks, such as J.R. Anderson's (1983) ACT*, offer one plausible environment. Unlike standard formulations of deduction and LISP, the processing of frames with productions is naturally sensitive to similarity, frequency, recency, and context, thereby providing concepts with the statistical character so often observed for them. Although production systems have much potential as processing environments for frames, they have not been developed to produce the flexibility and structure of concepts reviewed here.

Recently, connectionists have attempted to develop networks that represent the conceptual relations in frames. As I document in the

Appendix for the interested reader, however, these attempts fall short in critical regards. The Appendix illustrates general problems that feature list theories face in trying to represent structure, and it suggests a set of relational criteria that any theory of representation should satisfy. Much remains to be done empirically in understanding how humans process frames dynamically, and in implementing these understandings theoretically in computational and formal systems. Just as perceptrons required a major innovation to evolve into more powerful systems (Rumelhart, Hinton, & Williams, 1986), so may a major innovation be necessary before connectionist nets evolve to represent the hierarchical structure of concepts satisfactorily.

Linguistic Vagary

As we have seen thus far, concepts are flexible and structured. As we shall see in this section, the knowledge from which concepts are constructed exhibits linguistic vagary. Frequently, psychologists attempt to construct representations of concepts, because doing so is instrumental to assessing issues of interest (e.g. Ashcraft, 1978; Barsalou, 1981; Glass & Holyoak, 1975; Hampton, 1979; Rosch & Mervis, 1975; Rosch et al., 1976). To represent concepts, psychologists have almost exclusively relied on linguistically oriented representations, borrowed from the other cognitive sciences. Most commonly, psychologists simply label the features of concepts with linguistic expressions, a practice adopted from linguistics (e.g. using “feathers” to represent a feature of bird). However, psychologists also frequently use propositional logic and predicate calculus from philosophy to represent concepts (e.g. RED (robin), ABOVE (head, neck)), as well as data structures and procedures from computer science (e.g. networks, productions).

In all of these approaches, language, in one form or other, lies at the heart of representing conceptual content: In linguistic representations, words and phrases label features. In logical representations, words and phrases label propositions, predicates, and arguments. In computational representations, words and phrases label the nodes and links of networks, as well as the conditions and actions of productions. In most representational schemes that psychologists have adopted, a strong linguistic element persists. From here on, I will use linguistic representation to mean any of these linguistically oriented representations.

The primary purpose of using linguistic representations has been to express the content of people's concepts. Note, however, that the concepts typically represented in traditional theory are assumed to be
stable structures in long-term memory, unlike the definition of concept adopted in this chapter. Consequently, my discussion of linguistic vagary here and elsewhere addresses the knowledge of categories in long-term memory, not temporary representations of them in working memory. In traditional theory, this knowledge is the same as concepts; in the theory developed here, this knowledge is the conceptual content from which concepts are constructed.

Linguistic representations of conceptual content in long-term memory face three serious problems, to which I refer collectively as linguistic vagary: First, we lack principled means for constructing linguistic representations of conceptual content; second, the representations we construct are haphazard; third, these representations are incomplete. These problems apply equally to feature lists, connectionist nets, frames, and most other models of concepts and memory in general use. In later sections, I shall argue that linguistic vagary is not a problem with our methodology for studying concepts, but instead arises naturally from the linguistic and perceptual symbols that underlie them.

Unprincipled Content. Imagine that we attempt to establish the conceptual content in long-term memory that someone has for a particular category, such as bird. How do we discover its representational elements? In general, psychologists, or most other cognitive scientists for that matter, do not have a priori, principled means for discovering this content, but rely instead on intuitive and blind empirical means for gathering linguistic descriptions of it. For example, theorists often develop representations of categories by describing exemplars to themselves intuitively. While introspecting on particular exemplars, they describe linguistically the conceptual elements that they observe. Frames in computational theories often appear to be generated in this manner, such as the restaurant script of Schank and Abelson (1977). Not surprisingly, there are often major differences between linguistic representations of categories that different theorists construct.

Psychologists frequently attempt to measure conceptual content in long-term memory more systematically. Essentially, this amounts to studying subjects' collective intuitions rather than studying one's own intuitions. For example, researchers often develop feature list representations by eliciting linguistic descriptions of features from subjects, as in the definition study described earlier (Barsalou et al., in prep.). Again, however, the discovery of conceptual content relies on introspection and spontaneous linguistic description, rather than on a principled theoretical rationale that guides its acquisition and description. It is not clear at this point what form a satisfactory approach would take, but the blind empirical methods generally in use are far from ideal.

A variety of well-known scaling programs can produce conceptual content in a somewhat less introspective manner. For example, multidimensional scaling identifies dimensions that structure stimulus sets, though with the interpretive help of an external viewer. Similarly, various clustering algorithms suggest the presence of features that exemplars of a category share. These approaches may offer a more systematic means of measuring conceptual content than simply describing our own introspections or collecting those of our subjects. Yet, typically, the basis of these scaling results resides in intuition as well, because subjective sorting data and subjective similarity ratings often provide the input to these programs. Additionally, the interpretation of dimensions and the number of dimensions interpreted also reflect subjective judgment. More seriously, these scaling techniques are compromised by the problems of haphazardness and incompleteness, which we shall consider shortly. Scaling results are haphazard, because they can vary considerably with context. Scaling results are incomplete, because they typically provide only a small amount of information about conceptual content, such as a few dimensions in a multidimensional scaling solution. For example, the dimensions of size and predacity, which structure the scaling solution for animal (Henley, 1969), hardly exhaust people's knowledge of this category.

Not only do we lack principled means for generating linguistic descriptions of conceptual content, we also lack principled means for recognising valid descriptions once they've been generated. Rather than having rigorous criteria for discriminating a true representational component from a fraud, we again rely on either our own intuitions or the intuitions of our subjects. Perhaps the most common method for verifying the validity of a representational component is consensus: if some critical percentage of subjects describe a component, say 25%, we include it in our representation. Note that if we asked subjects to recognise descriptions of components, subjects might verify nearly all of them as valid, as in our definitions experiment reviewed earlier. Regardless of this, consensus is hardly a principled means of verifying the validity of conceptual content. We have neither principled means for generating conceptual content, nor for recognising descriptions of it, and no solutions to these daunting problems are in sight.

Haphazard Content. A second aspect of linguistic vagary is the haphazard description of the conceptual content that resides in long-term memory. As we saw earlier, different people produce very
different linguistic descriptions when defining the same category, and the same person produces different descriptions when defining the same category on different occasions (Barsalou et al., in prep.). When context varies, linguistic descriptions vary still further, as Murphy (1988) found for descriptions of the same adjective in different noun contexts.

Furthermore, the linguistic descriptions of conceptual content that researchers infer using less direct methods are also haphazard. For example, the linguistic descriptions of dimensions in multidimensional scaling change with context (Sadler & Shoben, in press). Similarly, different scaling techniques produce different linguistic representations of the same categories. Gammock (1987) found that different scaling techniques produced somewhat different linguistic representations of the categories in a given conceptual domain. Finally, if one were to infer conceptual content from typicality judgements, the resulting linguistic descriptions of it would be haphazard. Recall the high typicality of swan and peacock in the category of birds from the Chinese point of view (Barsalou & Sewell, 1984; see also Roth & Shoben, 1983; Schwanenflugel & Rey, 1986). One might infer from this result that the feature graceful exists in people's concept of bird, because it is true of typical exemplars. However, because these exemplars are atypical from others points of view, graceful would not be inferred for bird in these contexts. Consequently, the linguistic descriptions inferred theoretically to represent bird are haphazard across contexts.10

In general, the linguistic descriptions of conceptual content vary with culture, individual, context, and task. This is true of indirect scaling methods as well as of direct linguistic report. In all cases, the conceptual content reflected in the resulting linguistic representation is influenced substantially by the circumstances surrounding its measurement.

The haphazard nature of conceptual content is a problem for traditional theories of concepts, because they assume that a concept resides as a stable structure in long-term memory. If we attempt to construct a linguistic representation for such a concept, how do we make sense of the haphazard linguistic descriptions that we acquire across contexts? How do we integrate them into a single representation? Are we justified in even attempting to do so?

One solution is to view concepts as temporary constructions in working memory that vary widely in the knowledge they incorporate from long-term memory. Adopting this view makes it unnecessary to integrate all possible conceptual content into a single coherent representation, because inconsistent and unrelated conceptual elements can reside together in long-term memory and rarely, if ever, be processed simultaneously. A second solution is to question our exclusive commitment to linguistic representations of concepts. As I shall argue in later sections, perceptual symbols constitute the cores of concepts, and the haphazardness of linguistic representations arises from the interaction of linguistic symbols with these perceptual cores.

Incomplete Content. A third aspect of linguistic vagary is that the linguistic descriptions we obtain for the conceptual content of a category in long-term memory are incomplete. As we saw earlier in our definitions experiment, a given subject only describes a very small subset of the information that they know for a category. If we trust the linguistic report of a single subject to define the conceptual content of category in long-term memory, we will surely be unable to develop a complete representation of it. Even if we combine the reports of all subjects, our representations are still likely to be incomplete. For example, the combined feature lists that we obtained by pooling our subjects' definitions lacked many features that could have been generated (e.g. eyes, hops, and eats worms for birds). Attempting to construct a complete linguistic representation of the conceptual content for a category is a sobering exercise, because one continues to discover new descriptions endlessly.

To a large extent, the problem of incomplete content reflects the distinction between stored versus inferred knowledge. Most likely, many of the linguistic descriptions that people produce of conceptual content are not stored in memory but are constructed spontaneously. Somehow, from existing knowledge, people produce new descriptions that they have never considered before. Assuming this to be true, how do we construct a complete linguistic representation of the conceptual content for a category? Can we construct a complete linguistic representation?

I suspect not for several reasons. First, the recursive decomposition of conceptual content appears endless, at least in principle. Where does decomposition stop? How deep is the decomposition of conceptual content? Often it seems that people can keep decomposing conceptual content forever, providing further structure for each component mentioned previously (Barsalou, 1992b). Rather than ending neatly in some set of terminal components, people seem endlessly able to produce further detail about the detail they just described. No one has yet observed people stopping at terminal elements, nor does anyone have any idea what these terminal elements might be. For every detail just described, further detail can be added. This recursive property of linguistic descriptions makes our theoretical representations of conceptual content appear arbitrary and whimsical. Regardless of the linguistic representation we construct for a category theoretically, we have only described part of its conceptual content. Moreover, we have relatively little perspective on what part of the total conceptual content
we have represented, on what other parts are missing, or on what the boundaries on this content are, assuming any exist. Once again, our linguistic representations of conceptual content appear flawed.

A second factor that produces incompleteness is people's profoundly creative ability for constructing linguistic descriptions that are relevant in the current context. In our protocol studies of planning, we often found people constructing amazingly ad hoc descriptions of attributes. Consider the category of companions in the plan for a vacation. One attribute of companions that subjects described frequently was the extent to which a possible companion will want to do the same vacation activities that I will want to do. After describing this attribute, subjects often evaluated possible companions with respect to it. Similarly, for the category of departures, subjects described the attribute the extent to which departing at this time will interfere with my work. Clearly, such attributes are context-dependent. In the context of a different event, the attributes described for companions and departures might well be different. Because the attributes constructed for categories are often context-dependent, and because people often seem to derive these attributes spontaneously, it appears impossible to develop a complete account of the conceptual content that describes a category. Because a category can always be considered in new contexts, it is likely to develop new attributes relevant to them. Descriptions of structural invariants and conceptual constraints are likely to exhibit this sort of context dependence as well. Not only does the recursive description of conceptual content present significant problems for completeness, so does its context dependence.

Finally, I shall argue later that the open-ended recursion and context dependence of linguistic representations reflect the perceptual symbols at the cores of concepts. Because perceptual symbols afford an indefinite number of linguistic descriptions, complete linguistic representations are impossible.

Summary. The unprincipled, haphazard, and incomplete character of linguistic representations for conceptual content could arise from the inadequacy of our methodological and theoretical tools. On this view, we lack principled means for generating and recognising valid descriptions of concepts, we don't know how to handle the haphazardness of the descriptions we acquire, and we don't know how to remedy their incompleteness. Perhaps we can solve these problems, if we develop better theories of linguistic representation and better methods of measuring it. Alternatively, linguistic vagary may not reflect problems with our methodology or theory. Perhaps our methodology, in manifesting linguistic vagary, is measuring linguistic descriptions of concepts accurately. Perhaps these descriptions are inherently unprincipled, haphazard, and incomplete.

Why might we resist such a conclusion? Why might we expect the linguistic representations of concepts to be otherwise? Perhaps we have these expectations, because the foundations of our theory rest heavily on traditions inherited from linguistics, philosophy, and computer science. We may expect linguistic representations of conceptual content to be principled, stable, and complete, because these are properties of similar representational systems in linguistics, philosophy, and computer science.

By focusing on linguistically oriented representations, we may be failing to see other mechanisms that are more central to the human conceptual ability, such as perceptual symbols and the compositional processes that operate them. In such a system, linguistic symbols may constitute the instruments that control the development and use of perceptual symbols, rather than constituting a closed representational system that exhibits coherency, consistency, and completeness. Viewing linguistic symbols as derivative of perceptual symbols raises new criteria for evaluating them. Rather than evaluating linguistic symbols with logical criteria, evaluating their functional roles in the storage, retrieval, integration, and conveyance of perceptual symbols may be more productive.

THE PERCEPTUAL LIFE OF CONCEPTS

In this section, I propose one possible account of the perceptual symbols that could underlie the human conceptual system. In the next section, I describe how linguistic vagary arises through the interaction of perceptual and linguistic symbols. Certainly, perceptual accounts of human knowledge have been suggested before (e.g. Miller & Johnson-Laird, 1976; Paivio, 1986), although the opinion that these attempts have failed is widespread. Often, however, the critiques of these positions assume overly simplistic views of perceptual representations, such as reference to physical objects or literal images, the assumption that perceptual representations must account for everything in cognition, and so forth. On considering more reasonable proposals, and on further inspection, there are many compelling reasons for believing that perceptual knowledge is central to concepts.

In the last 20 years, we have learned that the human imaginal ability is quite powerful (e.g. Finke, 1989; Kosslyn, 1980; Shepard & Cooper, 1982). Imagery research has demonstrated that people have impressive abilities to manipulate perceptual memories, and to process them in many of the same ways that perceptions are processed. Because
perception is arguably our most powerful and important ability, it is not surprising that imagery is powerful as well. Most importantly, it would not be surprising if the cognitive system took advantage of the resources associated with perception and imagery to represent concepts.

Furthermore, theorists are developing provocative new views of cognition that rest heavily on perceptual knowledge. Cognitive linguists are building theories of language around perceptual representations (Jackendoff, 1987; Lakoff, 1987; Lakoff & Johnson, 1980; Langacker, 1986, 1987; Talmy, 1988). Harnad (1987) argues that perceptual representations are central for solving the symbol grounding problem. Other authors of chapters in this volume view perceptual representations as central to cognition: Crowder and Schacter each report dissociations between perceptual processing and conceptual elaboration, which are consistent with my later argument that both perceptual and linguistic symbols underlie concepts. Johnson similarly includes a fundamental distinction in her view between perceptual and reflective processing. Baddeley reports that various forms of executive control do not interfere with image generation, which is consistent with my later argument that linguistic description should co-occur with perceptual processing, not interfere with it.

I will argue that adopting a perceptual view of concepts explains the daunting problems surrounding linguistic vagary. Rather than reflecting inadequacies in our theories or methods, linguistic vagary simply reflects the fact that perceptual symbols—not linguistic symbols—constitute the cores of concepts. I shall argue further that the structure and flexibility of concepts fall naturally out of a perceptual view as well.

The remainder of this section begins by describing how selective attention extracts schematic perceptual components from experience to produce perceptual symbols, and how compositional processes integrate these symbols during conceptual combination, comprehension, and imagery. The second half of this section addresses linguistic symbols and their relation to perceptual symbols. In the process, I provide accounts of symbol grounding, abstract concepts, and memory for gist and surface structure. Obviously, this theoretical proposal is speculative and requires much further articulation and empirical assessment, although evidence from various literatures can be marshalled to support it at this time. A very similar view of cognition was advocated nearly two decades ago by Huttenlocher and her colleagues (Huttenlocher, 1973, 1976; Huttenlocher & Higgins, 1978) and more recently by Mandler (1992).

Although I focus on visual perception in the examples that follow, I assume that my analysis applies to representations from all modalities. As I shall argue, perceptual symbols can develop for any aspect of human perceptual and introspective experience, including aspects of thoughts, emotions, proprioceptions, cognitive operations, and so forth. Examples along the way will illustrate these various types of perceptual symbols.

Perceptual Symbols and Their Composition

What form might the perceptual representations that underlie concepts take? An obvious possibility is that they are simply literal memories of conscious subjective states (i.e. analogue images). Each image represents the state of the perceptual system at a particular point in time, similar to the exemplars in some exemplar models. Although the simplicity of this view is appealing, empirical findings suggest that it is incorrect, and theoretical considerations suggest that a more productive form of schematic perceptual symbols is necessary.

Perceptual Structure. Rather than perceptual memories being analogue images, I will argue that they are hierarchical relational representations, i.e. structural descriptions (Palmer, 1975). Consider Kosslyn’s recent work on image generation, in which subjects generate images of letters and scenes (e.g. Kosslyn, Cave, Provost, & von Gierke, 1988; Roth & Kosslyn, 1988). If people store perceptual representations as analogue images, they should generate images of them in a single processing step, or perhaps in a “raster scan” from left to right, or top to bottom. Kosslyn and his colleagues found instead that people generate images component by component, in a systematic sequential manner. For example, people generate the image of a letter by constructing its line segments in their normal writing sequence. Similarly, people generate images of three-dimensional scenes, beginning with the closest component and constructing the remaining components from near to far. For people to generate images in this manner, their representations must be articulated in long-term memory by components hierarchically and relationally. Indeed, Kosslyn’s (1980) theory of image generation works exactly this way, beginning with a hierarchical relational representation in long-term memory, and using it to construct an image in working memory.

Further evidence for structured perceptual representations comes from Biederman’s (1987) work on geons. On this view, people represent components of a visual object with geons, namely, a small vocabulary of geometric solids, the categorisation of which remains largely invariant under rotation, depth, and occlusion (e.g. cones, rectangles, cylinders). For example, the representation of a chair might include rectangular geons for the seat and back, together with cylindrical geons for the legs,
integrated by spatial relations. Similarly, the representation of a suitcase might include a rectangular geon for the main container and a curved cylindrical geon for the handle. Biederman has accrued much evidence for this view, some of which suggests that abstract schematic components for geons are more central to perceptual representations than analogue images. For example, Biederman and Ju (1988) presented subjects with line drawings of objects that conveyed their geons, and found that subjects recognised these drawings as rapidly as colour photographs. Moreover, Biederman and Ju further found that this effect did not interact with the predictiveness of colour during categorisation. Specifically, colour photographs were not superior to geon diagrams for categories like banana, for which colour is diagnostic. Such evidence suggests that the underlying information for these categories used in object recognition is represented in schematic geons, rather than in analogue images.

Work by Mandler and her colleagues further suggests that people don’t store analogue images of perceptions in long-term memory (e.g. Mandler & Ritchey, 1977). After studying line drawings, subjects could not discriminate them from new drawings that contained various “surface” transformations of components in the old drawings. For example, subjects could not correctly reject a new drawing in which a component had been moved. If subjects had stored analogue images of the drawings, they should have rejected a wide variety of transformed drawings, but they did not. Instead, they stored the deeper “gist” of the pictures, given that they easily rejected new drawings whose gist differed from the old drawings.

Finally, work with the congenitally blind and the neurologically impaired offers further evidence against the analogue image view. Kerr (1983) found that congenitally blind subjects produce the same effects on rotation, scanning, and resolution tasks as sighted subjects. If analogue images were responsible for performance on these tasks, congenitally blind subjects should have been unable to perform them, or should have performed them differently. The identical patterns of data suggest instead that spatial representations, not perceptual images, are central to these tasks. Similarly, Farah, Hammond, Levine, and Calvanio (1988) found that spatial abilities were central to imagery tasks, although they found that sighted subjects used both imagery and spatial abilities, presumably because both were available. Certainly, one would expect that imagery could be produced on occasions when it is available and useful. However, the results from these studies indicate that imagery is not necessary for many perceptual tasks, further supporting the view that deeper, more structured, representations play central roles in perceptual processing.

Selective Attention and the Construction of Schematic Perceptual Components. How do the components of structured perceptual representations become established in long-term memory? Selective attention is one factor that appears central to this process. The basic idea is that the attentional system is capable of focusing strategic processing on various aspects of a perceptual experience and extracting them as individual components, while simultaneously tuning out other components to a large extent. Following much work on the processing of separable features (e.g. Garner, 1974, 1978), we know that this kind of focal processing occurs extensively, even though some contamination from irrelevant components often occurs (e.g. Malaru & Marks, 1980). For example, selective attention could focus on the geons of an object, the colours or textures of their surfaces, the overall shape of the object or its orientation, as well as numerous other compositional properties of visual experience. Furthermore, selective attention plays a critical role in constructing the relations between multiple components of an object (Treisman & Gelade, 1980). Detecting the conjunction of two components (i.e. their spatial and temporal correlation) often appears to require selective attention.

On this view, the components extracted from perceptions are diagrammatic or schematic, only representing the information relevant to the component (much like the abstract diagrammatic components in cognitive grammars; e.g. Jackendoff, 1987; Langacker, 1986, 1987; Talmy, 1988). The representation of a particular geon on a particular occasion contains little, if any, information about colour, and conversely, the representation of colour contains little information about shape. Instead, selective attention primarily extracts information about shape for geons, and primarily extracts information about colour for colour representations. In this way, components are not literal images of their counterparts in perception, but are schematic representations of them.12

Selective attention may extract a wide variety of complex perceptual components, not just simple ones having to do with shape and colour. For example, selective attention may construct schematic representations of an organism’s behaviour, such as a dog biting something or wagging its tail. In these schematic representations, multiple states are represented over time, as in cognitive grammars, establishing the sequence of states that specifies the behaviour. For example, biting could be represented as a series of three schematic states, showing a mouth closed next to an object, followed by the mouth open, and then the mouth around the object. Similarly, wagging could be represented by a short cyclic series of states showing a tail in different orientations. Again, much irrelevant information may not be stored in these schematic representations, including colour, texture, and so forth.
Schematic information from other modalities besides vision can also be extracted in this manner. For example, barking could be represented not only by a visual schematisation of a dog's mouth movements, but also by a schematisation of the relevant sounds, extracted from events that contain them. Similarly, schematisations of pricky and salty can be extracted from the tactile and gustatory modalities. Moreover, schematisations of purely introspective events are possible as well. By selecting the aspects of introspective experience that are associated with anxiety, tranquility, anger, and so forth, schematic representations of affective states develop. Similarly, schematic representations can develop for various computational states (e.g., idea, goal), as well as for operations on such states (e.g., attend, search, remember, forget, rehearse, compare; cf. Johnson, this volume). Anything in conscious experience that can be selected by attention is a potential schematic component.

Finally, it is important to note that selective attention also plays the important role of profiling information central to meaning, as illustrated in the following examples from Langacker (1986). To represent hypotenuse, the line that corresponds to the hypotenuse in a schematic right-angled triangle is profiled by attention, relative to the other two lines that form the right angle in the background. The hypotenuse in the foreground and the right angle in the background are both essential to the meaning of hypotenuse. Similarly, consider the sentence,

The hill falls gently to the bank of the river.

To represent falls in this sentence, attention travels downward along a schematic perceptual representation of hill in the background, whereas an upward attentional trajectory represents rises in the sentence.

The hill rises gently from the bank of the river.

In sum, selective attention is important, first, for extracting schematic information from perception, and second, for highlighting central information within the schematic representation.

Component Storage. Once a component has been selected, a memory of it becomes established in long-term memory. From extensive work on the control of memory encoding, especially work on depth of processing (Craik & Lockhart, 1972), it is clear that selective attention determines the information encoded into memory, at least to a significant extent. Because each stored component has been selectively attended, other aspects may not be stored with it, or only stored minimally. For example, the storage of a particular geon may minimally include information about its colour and texture. As a result of processing a perceptual image in this analytic manner, it becomes stored as a collection of separate components, integrated by whatever relations happen to have been processed selectively between them. The final long-term memory of the image is a compositional structure, integrated hierarchically and relationally.

Contrary to what I have suggested, people often seem to store irrelevant information from perceptual experiences implicitly (Hasher & Zacks, 1979, 1984). For example, Jacoby and Hayman (1987) found that readers stored irrelevant information about fonts that biases subsequent word recognition. Consequently, why shouldn't people store information about colour and texture when they attend selectively to information about the shape of a component? One answer is that irrelevant information is stored automatically to some extent across the processing episodes of a component (Jacoby, 1991), but that it typically cancels itself out through interference, leaving only the common relevant information that is strengthened through repetition (Thorndyke & Hayes-Roth, 1979; Watkins & Kerkar, 1985). In processing different cylinders on different occasions, for example, their common shape becomes strengthened through repetition, but their different colours cancel each other out through interference. In contrast, when selective attention focuses on a particular colour, the common colour becomes strengthened across repetitions, and the different shapes cancel each other out.

Irrelevant information about a schematic perceptual component may be retrieved on occasion by a cue that contains it, but typically the selected information dominates processing, allowing transfer to new contexts. To see this, imagine someone learning to classify cylinders and spheres for the first time, where the surface pattern on the cylinders and spheres varies widely (e.g., green and beige stripes, blue and grey checks). Across experiences of cylinders, the common shape is strengthened, whereas the irrelevant patterns cancel each other out. However, if a new cylinder has the same pattern as a previous cylinder, the irrelevant pattern from the previous exemplar may be retrieved and speed classification (assuming that this pattern never occurred for spheres). In a sense, the irrelevant pattern constitutes diagnostic evidence, although definitionally irrelevant, for cylinder. In contrast, imagine receiving a new cylinder having a pattern never seen before. Because shape information has been attended selectively for previous cylinders, and because it has been strengthened by its commonality across exemplars, it produces transfer to the new cylinder, even though its surface pattern is new. What may strike the reader as an obvious point is important, because it is the converse of the point that implicit...
memory researchers typically make: Whereas implicit memory researchers often stress that the storage of irrelevant context information allows implicit transfer to similar contexts, at least as important is the point that the selection of relevant information allows transfer to completely new contexts.

Once a type of entity becomes familiar (e.g. *chairs*), slow strategic processing may not be necessary to process its components sequentially (e.g. *seat, back, legs*). Instead, the hierarchical relational representation of the entity in long-term memory becomes automatized to a large extent, such that all of its components can be processed in parallel (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). Note that efficient parallel processing of components does not imply that the perception is processed as an analogue image; instead, the individual components established initially through selective attention are now capable of each controlling their own processing from long-term memory, bypassing the need for sequential strategic processing in working memory. Indeed, Biederman (1987) has found that the time taken to recognize a familiar object is relatively independent of the number of geons it contains, suggesting that they are processed in parallel.

**Perceptual Symbols.** Once a perceptual component becomes established in long-term memory, it can function as a symbol. I adopt the traditional view that a symbol is something that designates something else, with appearance of the former bringing to mind the latter (Goodman, 1968; Huttenlocher & Higgins, 1978). For example, words are symbols, because they designate various referents in the environment and thought. On hearing “chair” and “anxiety,” speakers of English can usually establish reference correctly.

In the modern world of cognitive science, a second layer of symbols, beyond linguistic and perceptual “surface structure,” is often assumed to provide a deeper “language of thought” (Fodor, 1975; Pylyshyn, 1973, 1984). Theorists tend to represent this symbolic level with propositional logic, predicate calculus, networks, procedures, and other schemes that have strong origins in linguistic representation, as we saw earlier. Typically, these representational languages express the *propositions* that form the bedrock of knowledge and memory. Rather than containing information extracted from perceptual experience, propositions are generally assumed to constitute an amodal, more abstract representation of conceptual structure. Nearly every psychological theory that has attempted to represent human knowledge has adopted this approach, excluding connectionism (e.g. Anderson & Bower, 1973; J.R. Anderson, 1983; Barsalou, 1991, 1992a,b; Gentner, 1989; Kintsch, 1974; Murphy, 1988; Newell & Simon, 1972; Norman, Rumelhart, & the LNR Research Group, 1975; Smith, Osherson, Rips, & Keane, 1988; Thibadeau, Just, & Carpenter, 1982; van Dijk & Kintsch, 1983; but see Larkin & Simon, 1987). Moreover, artificial intelligence is heavily committed to representing knowledge with amodal propositions (e.g. Charniak & McDermott, 1985; Collins & Michalski, 1989; Lenat & Guha, 1989, Schank, 1975). Theories of meaning in linguistics and philosophy generally adopt this approach as well, excluding cognitive linguistics.

Most importantly, the symbols in these languages of thought bear an arbitrary relation to their referents. Because these symbols are amodal and abstract, they bear no similarity to the entities that they designate. Unless one knows the convention that links a symbol to its referent, the link cannot be established. In predicate calculus, for example, unless one knows the convention that links a predicate and its arguments to the world, reference cannot be established by examining the predicate in isolation. Similarly, in a natural language, establishing the link from a word to its referent is usually impossible unless one knows the convention between them (e.g. knowing the convention that, in French, “chien” refers to *dog*). Indeed, the arbitrary character of linguistic symbols is widely believed to constitute a linguistic universal.

Notwithstanding the bias of modern cognitive science, not all symbols function in this arbitrary manner (Goodman, 1968; Huttenlocher & Higgins, 1978). For example, drawings and icons have transparent links to their referents. By examining the characteristics of these symbols, people can establish reference on many occasions without knowing specific conventions. Symbols such as drawings function analogically, because mapping their characteristics to the characteristics of potential referents enables successful reference. For example, successfully mapping characteristics in the drawing of a dog to analogous characteristics in the perception of a dog enables viewing the picture as a symbol of the dog.\textsuperscript{14}

Symbols further include the schematic perceptual components extracted by selective attention from perceptual and introspective experience. *These representations are symbols, because they can designate referents.* For example, schematic perceptual components for *cylinder, bite, anger,* and *search* can each designate referents in the environment and introspection. Moreover, the links between these symbols and their referents should be relatively transparent. Because these symbols were extracted analytically from the same type of perceptual experiences that constitute their referents, links between them can be established analogically. From this point on, I will use the expression *perceptual symbol* to mean schematic perceptual components that designate referents in the environment and thought.
A critical problem is specifying how people establish the relations between perceptual symbols and their referents (Goodman, 1968). I do not address this problem here, but simply note the following points: First, contextual information may greatly constrain this mapping process. For example, a schematic sphere in a schematically imagined grocery store is likely to be an orange, but on a schematically imagined tennis court is likely to be a ball. Second, multiple perceptual symbols for the same referent may act in concert to eliminate the ambiguity in any particular one of them. For example, perceptual symbols for sound and bounce will disambiguate whether a schematically imagined sphere is an orange or a tennis ball as it hits the ground (not to mention perceptual symbols for texture, malleability, and so forth). Third, because people construct perceptual symbols themselves, rather than receiving them from a teacher, they designate the relations between perceptual symbols and their referents and do not have to figure them out. A variety of such factors is likely to facilitate the mapping of perceptual symbols to their referents.

Perceptual Compositionality. As the result of establishing perceptual symbols in long-term memory, the cognitive system acquires a vocabulary of compositional elements that can be assembled to represent objects, events, introspective experiences, and so forth. To see this, consider a cylinder; now assume that it is red; now assume that it has a pitted surface. One account of how you constructed this series of representations is that you first retrieved a perceptual symbol for cylinder that only had a default colour and texture (or possibly none at all), and that you subsequently revised your representation by retrieving perceptual symbols for red and pitted that modified your representation perceptually. Indeed, much conceptual combination may reflect the constructive evolution of perceptual representations (e.g. a red pitted cylinder, a rusted blue Cadillac, an upside-down orange waterfall, a gentle dog bite, brief intense anger, rapid forgetting, a vivid memory).

A wide variety of operations appear to underlie the composition of perceptual symbols. As we just saw, perceptual symbols for adjectives can transform perceptual symbols for nouns. Typically, the perceptual symbols for adjectives may be accompanied by procedures that specify the nature of the transformation. For count nouns such as cylinder and car, the procedure associated with red (and possibly associated with colour in general) only transforms the external surface to the specified colour. In contrast, another procedure associated with colour may specify that the entire volume of a mass noun is typically transformed to the specified colour, as in red milk, red wax, and red sand. If a colour is known to cover a particular area, the transformation may only occur in that area, as in:

The robin isn’t red, it’s green.

As these examples illustrate, transformations provide one form of attribute-value relation in perceptual composition, where the type of transformation constitutes the attribute (e.g. colour change), and the specific transformations constitute possible values (e.g. red, blue).

Insertion appears to be another important operation that underlies perceptual composition. Consider two examples adapted from Langacker (1986), with the first being a static spatial relation:

The bird is above the kite.

The perceptual symbol for above is a region of three-dimensional space that contains two schematic regions, one with a higher vertical position than the other. To represent the bird being above the kite, perceptual symbols for bird and kite are inserted into the upper and lower subregions in the perceptual symbol for above. Insertion is also central to the representation of processes, as in:

The hiker crossed the river.

Here, the perceptual symbol for cross contains a series of more specific perceptual symbols for each successive state in the process: The first state contains a schematic entity on one side of a schematic region, several states where the entity is in a series of positions in a path across the region, and a final state where the entity completes the path on the other side of the region. To represent the hiker crossing the river, a perceptual symbol for hiker is inserted into the moving entity, and a perceptual symbol for river is inserted into the region. As these examples illustrate, insertion provides a second form of attribute-value relation in perceptual composition, where entities in a perceptual symbol constitute attributes, and more specific perceptual symbols inserted into them constitute values. Later discussions of frames address attribute-value relations in perceptual composition further.15

Little, if any, empirical evidence exists that people compose perceptual symbols in this manner, yet it is easy to imagine how they could. Indeed, much work in cognitive linguistics attempts to represent composition in this way, and already it is apparent that this approach has considerable expressive power. We shall encounter further examples of perceptual composition in later sections.
To the extent that perceptual composition does in fact occur, it appears to have the desirable property of proceeding rapidly, as the following three observations suggest: First, people easily comprehend rapidly unfolding perceptual events that are novel, such as the events in a movie or series of pictures (Gernsbacher, Varner, & Faust, 1990). Visual comprehension appears to occur through the matching of components in immediate perception to corresponding perceptual symbols. Because people have little difficulty comprehending novel visual events, this process cannot depend on matching perceived events to literal replicas of them in memory, because such replicas do not exist for novel events as a whole. Instead, this process is more likely to rely on matching components of the perceived event to perceptual symbols in memory that are composed in novel ways through productive comprehension mechanisms. If this account is correct, it follows that the perceptual composition process must proceed rapidly, because people can comprehend rapidly unfolding novel events.

A second reason for believing that perceptual composition proceeds rapidly comes from intuitions about imagery. Clearly, one can imagine novel events that have never occurred, and that therefore cannot be retrievals of non-compositional exemplars experienced previously. For example, one might imagine a counterfactual or a future event. In either case, the ease with which such scenarios are often envisioned suggests that the perceptual composition process can combine perceptual components rapidly.

Finally, provocative work by Potter and her colleagues is consistent with the view that perceptual composition underlies linguistic comprehension. Potter and Faulconer (1975) found that subjects match pictures of exemplars to linguistic categories at least as fast as they match the linguistic names for them (see also Rosch, 1975). Similarly, Potter, Valian, and Faulconer (1977) and Von Eckardt and Potter (1985) found that subjects match visual probes to sentences as fast as word probes. Moreover, Potter et al. (1986) found that subjects integrate pictures that replace nouns in a sentence as fast as they integrate the words themselves. All of these results are consistent with the view that subjects are constructing compositional perceptual representations during language comprehension in a rapid manner. On this view, subjects integrate the pictures rapidly into linguistic comprehension, because the primary representation being constructed for the text is perceptual (see also Morrow, Bower, & Greenspan, 1989; Morrow, Greenspan, & Bower, 1987).

In summary, the perceptual symbols that I assume underlie concepts are not literal analogue images, but are schematic perceptual components extracted through selective attention. Rather than being a library of “pointillist” images, these symbols constitute compositional elements that can be combined productively to form novel representations. Moreover, this compositional system appears relatively fast, able to keep pace with the rapid unfolding of perceived events, to produce imaginary events, and to construct the perceptual representations that underlie linguistic comprehension. As I shall suggest in later sections, perceptual frames organise perceptual symbols, and linguistic symbols often control the composition of perceptual symbols.

Linguistic Symbols and Their Relation to Perceptual Symbols

Linguistic Symbols as Perceptual Memories. What role do linguistic symbols play in a compositional system of perceptual symbols, and what form might they take? I assume that linguistic symbols in the cognitive system are simply perceptual memories of linguistic symbols encountered in the environment. For example, perceptual memory of “bird” in its spoken form constitutes an internal linguistic symbol within the cognitive system. Once this memory becomes active, it acts as a conventional symbol by designating perceptual symbols and entities in the environment. Similarly, perceptual memory of “bird” in its written form constitutes another internal linguistic symbol that designates these referents. From this point on, linguistic symbol will mean memories of linguistic symbols. I will not use linguistic symbol when referring to linguistic symbols in the environment, but will refer to these as external linguistic symbols. In other words, linguistic symbol will be shorthand for memory of a linguistic symbol.

I do not assume that deeper, more conceptual forms of linguistic symbols develop that correspond to something along the lines of predicate calculus, propositions, or any other type of amodal arbitrary symbols. Instead, I assume that linguistic symbols only exist in the cognitive system as memories of external linguistic symbols.

Analogous to the perceptual representations of objects and events, I assume that linguistic symbols are hierarchical relational structures, not literal analogue images. Indeed, as reviewed earlier, Kosslyn et al. (1988) found that people produce images of letters in a compositional manner. More significantly, the elements of the linguistic symbols are known to be highly compositional. Clearly, the phonemes that underlie spoken words are compositional, with a small vocabulary of roughly 50 phonemes serving to produce somewhere around 50,000 words in a given language. Similarly, in orthographies, written letters combine productively to form larger written units. For these reasons, we have
propositions, because these symbols carry no information in their structure that specifies their referents (e.g., Soarle, 1980). A common assumption is that the meanings of such symbols are the other amodal symbols to which they are logically related, but this simply pushes the problem back a level, because at some point some symbols must be grounded in the environment for the states of such a system to have meaning. Another frequent solution is to rely on an external agent, such as a programmer, to specify the mappings, but again, this is hardly a satisfactory, principled solution.

Linking linguistic symbols to perceptual symbols provides a natural solution to the symbol grounding problem (Harnad, 1987, 1990). As we have seen, linguistic symbols designate their perceptual symbols by convention. In the conventions of English, for example, the linguistic symbol “cylinder” designates a perceptual symbol for cylinder. In turn, this perceptual symbol designates referents in the environment and introspection by analogy: Entities that satisfy the schematic conditions of the perceptual symbol are instances of it. Perceptual symbols ground linguistic symbols, because they link them with referents in the environment and introspection.

This approach to symbol grounding also provides a simple solution to the representation of concepts that are difficult to describe with linguistically oriented representations. Again consider the different meanings of “red” in red hair, red clay, red wine, and red roses, and imagine trying to represent each of these redes propositionally. Putting these differences into amodal, predicate-like propositions is obviously difficult. Although one could imagine a propositional representation that captures these differences, it seems more parsimonious and transparent to assume that perceptual symbols represent them instead, grounding each sense of “red” in a different perceptual symbol, as in:

My accountant Jimmy, who has red hair, dyed it the colour of red wine.

Grounding ambiguous, difficult-to-describe terms in perceptual symbols offers a natural account of their semantics. 16

Perceptual Symbols for Abstract Concepts. No doubt the reader is wondering just how far such an account can go. Although it may work for concrete objects and events, how can it account for abstract concepts? This question has been raised many times before, and typically, the conclusion is that perceptual representations don’t do a very good job of accounting for abstractness. However, two important considerations must be borne in mind. First, even if perceptual symbols do not represent all concepts, they may nevertheless be responsible for
representing many of them. In addition, it may be these concepts that
the human brain evolved to process in non-technological environments.
Consequently, accounting for the perceptual nature of these concepts
could provide significant insight into the basic mechanisms underlying
human cognition. The second factor to consider is that our under-
standing of computational architectures has developed substantially
in recent years, and employing the new tools that this approach affords
may produce insights into the problems surrounding abstract concepts.

Already, we have considered one computational mechanism—
selective attention—that greatly increases the ability of an intelligent
system to construct perceptual symbols. Through the process of
extracting components selectively from perception, an intelligent
system acquires the perceptual symbols that enable a compositional
system of perceptual knowledge. Concepts could be extracted from
perceptual experience in this manner, yet seem abstract because so
much perceptual experience has been stripped away. Consider the
sentence:

The kind bouncer removed the patron from the bar.

In this sentence, “kind” might be construed as an abstract term that is
difficult to represent perceptually. Imagine, however, how it might be
extracted from perceptual experience. Perhaps people selectively
extract the facial expressions of people who are referred to as “kind”, as
well as their actions during social interaction. In addition, people may
selectively extract aspects from their introspective states when they are
referred to as “kind”. In comprehending the sentence about the “kind
bouncer”, the reader retrieves a perceptual symbol of a bouncer (e.g. a
large muscular male) and transforms it with the perceptual symbols just
described for “kind”, producing a complex perceptual symbol that
includes facial expressions, body movements, and mental states.
Moreover, if people experience different forms of kind in different types
of people (e.g. kind mother, kind doctor), they may store different
perceptual symbols for kind that are contextualised, as we saw earlier
for red. Depending on the person that kind modifies, different perceptual
symbols apply. The point is that it generally seems possible to identify
schematic perceptual components in experience that could become
perceptual symbols for abstract concepts.

A second computational mechanism, operations on perceptual
symbols, provides an additional solution to the representation of
abstract concepts. To see this, consider the representation of function
words, such as “a” and “the”. How is it possible to represent the meanings
of such words perceptually? Rather than being grounded in perceptual
symbols, the meanings of these words may instead be grounded in
operations on perceptual symbols. Consider “the” in the sentence:

I took the car to a junkyard.

On hearing “the car”, the listener assumes that the speaker has a
particular car in mind, because “the” is definite. Computationally, the
listener could treat “the” as an instruction to retrieve the perceptual
symbol for a particular car identified previously in communication. In
other words, the meaning of “the” is grounded in an operation to be
performed on a perceptual symbol. Conversely, the determiner “a” might
specify that the listener retrieve any perceptual symbol for an instance
that is plausible in the current context, or possibly for a typical instance,
as in:

Gwen put a plant in her office.

Here, the perceptual symbol retrieved for “a plant” could be the
perceptual symbol for any plant, or for a typical plant. As these examples
illustrate, some abstract concepts may be grounded in operations on
perceptual symbols, rather than in these symbols themselves.17

Operations on perceptual symbols are also important for representing
abstract nouns and verbs. Consider “truth”. When people define “truth”,
they typically roll their eyes and take considerable time to produce a
definition. Often, they imagine a scenario that exemplifies “truth” and
describe it linguistically, attempting to refine their description into a
generic definition. For example, someone might recall the event of her
son telling the truth about where he had been one afternoon, and then
attempt to produce a definition from it, such as:

Truth is when someone makes a claim that corresponds to what really
happened.

Two aspects of this process could reflect perceptual processing. First,
the scenario constructed initially could be perceptual. Second, the
definition ultimately established could be grounded in perceptual
symbols and operations on them. Specifically, “someone” could be a
perceptual symbol for a person, “claim” could be a perceptual symbol for
a state of the world (e.g. one’s son at school), and “what really happened”
could be another perceptual symbol for a state of the world (e.g. one’s
son at school). Finally, “corresponds” could be an operation on perceptual
symbols, similar to “a” and “the”, which specifies a match between the
perceptual symbols for the claimed and actual states of the world. Much
work in cognitive grammar similarly attempts to develop perceptual symbols of abstract concepts (e.g. Talmy's 1988, account of cause and related concepts).

As these examples illustrate, a compositional account of perceptual symbols embodied in a computational architecture exhibits considerable potential for explaining abstract concepts. The selective extraction of components from perceptual and introspective states, together with the extraction of cognitive operations, provide many opportunities for grounding abstract concepts. To account for an abstract concept, one examines the perceptual and introspective situations to which it applies, and attempts to extract the critical aspects into perceptual symbols.

A critical problem for this view is explaining how people identify these critical aspects in the first place. A wide variety of strategies may be relevant. Again, Markman's (1989) strategies for discovering the mappings between linguistic symbols and their referents may be central. Perhaps more importantly, the sophisticated coordination of selective attention between conversationalists may facilitate the acquisition of perceptual symbols for abstract concepts (Tomasello, Kruger, & Ratner, in press). As Tomasello et al. show, adults are able to direct children’s attention to the critical aspects of a situation that are relevant to meaning. If such coordination allows the joint selection of perceptual symbols for abstract concepts, then once an individual knows an abstract concept, they can teach it to someone else.

**Accounting for Gist with Perceptual Symbols.** The primary reason that cognitive scientists believe amodal propositions form the bedrock of knowledge and memory is that they explain memory for gist. As has been known for some time, people retain the conceptual gist of sentences and pictures for long durations, after quickly forgetting their surface structure. For example, Sachs (1974) found that people quickly forget whether a sentence was in the active or passive voice, but continue to remember its semantic content. Similarly, Mandler and Ritchey (1977) found that people quickly forget surface information in pictures, but continue to remember their gist. Amodal propositions have always seemed central to cognition, because they provide representations of the conceptual gist that remains after perceptual surface forms are forgotten (Anderson & Bower, 1973; Kintsch, 1974; for a brief review, see Barsalou, 1992a, pp. 250–254). In addition, amodal propositions have long been viewed as providing a means for integrating perceptual information from different modalities about the same entity (e.g. the spoken and written forms of “dog”).

As we saw earlier, amodal propositions are arbitrary symbols, because their structure provides no clues to their referents, which must instead be identified through cultural convention. But then just what is the structure of these amodal propositions? What form do they take? Although this form is typically assumed to be something along the lines of predicate calculus, we have no direct evidence that anything in human knowledge takes this form. Furthermore, we do not have a compelling account of how amodal propositions enter the cognitive system. Essentially, we have no idea how they might arise through perception, how they might originate in thought, or how they might evolve through evolution.

Because the form and origin of amodal propositions remain mysteries, ascribing these representations to cognition is a theoretical extravagance. We have no direct evidence for amodal propositions, having instead only the indirect evidence that they provide one possible account of memory for gist. If some other account can explain the phenomena that amodal propositions are supposed to explain, and if this account explains these phenomena more parsimoniously and plausibly, then amodal propositions become unnecessary.

In this spirit, gist can be easily accounted for by a compositional system of perceptual symbols. Imagine hearing the sentence,

A cat chased a dog.

Further imagine that this sentence is converted immediately into perceptual symbols for cat, dog, and chase. Specifically, the perceptual symbols for cat and dog might only include information about their shape, size, and movement, excluding other information about their colour, speed, and so forth. Similarly, the perceptual symbol for chase (in the spirit of cognitive linguistics) represents two entities moving along a common path, with the goal of the entity at the back being to contact the entity in front, and with the goal of the entity in front being to avoid contact. Finally, the perceptual symbols for cat and dog are inserted into the back entity and the front entity, respectively.

Once this schematic perceptual representation has been established, it represents gist. On a subsequent memory test, subjects can correctly say that they didn’t hear the sentence,

The dog chased the cat.

because its schematic representation fails to match the schematic representation of the original sentence. In contrast, subjects cannot decisively reject the passive version of the original sentence,

The dog was chased by the cat.
Because its schematic representation matches the schematic representation of the original sentence.

As this example illustrates, we can account for the memory of gist and the forgetting of surface structure without using amodal propositions. Instead, we can account for them readily with a compositional system of perceptual symbols. Most importantly, we can imagine how such representations might develop (i.e. through the operation of selective attention on perceptions), and we can imagine the schematic form they might take (i.e. from examining the content of perceptions). In contrast, we have no idea how amodal propositions originate, and the best account of their content is predicate calculus. Both parsimony and plausibility favour the perceptual account of gist.

The other reason for adopting amodal propositions—linking related perceptual representations from different modalities—is clearly less parsimonious than a purely perceptual account. For example, to link the spoken and written forms of a word (e.g. "dog"), a simple link between them will suffice. There is no reason to add a third amodal representation between them, if its only purpose is to establish a linkage. Clearly, semantic content of the word must be represented, but this content could be represented by perceptual symbols rather than amodal propositions. Again, the perceptual solution is more parsimonious and plausible than the propositional solution.

**EXPLAINING LINGUISTIC VAGARY, STRUCTURE, AND FLEXIBILITY**

Now that we have attempted to do away with amodal propositions, we are left with two systems: A system of linguistic symbols grounded in a compositional system of perceptual symbols. How might these two systems interact to produce the properties of concepts we considered earlier?

**Linguistic Vagary**

Linguistic vagary is the problem that linguistic representations of conceptual content in long-term memory are unprincipled, haphazard, and incomplete. We have no principled means for generating or verifying linguistic representations of this content; these representations vary with context; and they are impossible to specify exhaustively.

In a system of linguistic symbols grounded in perceptual symbols, all three aspects of linguistic vagary emerge naturally, as captured by the maxim:

>A picture is worth a thousand words.

More technically, this maxim can be restated as:

A perceptual representation affords an infinite number of linguistic representations.

In the following subsections, I show how this maxim explains linguistic vagary.

**Unprincipled Content.** To see how this maxim explains the unprincipled linguistic descriptions that people produce for conceptual content, imagine describing the perception of red wine. As we all know, wine connoisseurs use a ridiculous number of linguistic descriptions for taste, bouquet, and colour (e.g. "fruity", "big", "complex", "subdued", "woody", etc., for taste). The potential number of linguistic descriptions seems endless, with new ones continually entering the vocabulary. Similarly consider the possible linguistic descriptions for red hair and red roses. For centuries, authors have waxed poetic about such perceptions, producing myriad linguistic descriptions that stimulate our imaginations and stir our emotions. For nearly any aspect of perception, extensive and often limitless ways of describing it exist linguistically. Sometimes, these descriptions refer strictly to perceptual character, as in "bright red", "dull red", and "greyish red". Other times these descriptions are more figurative, as in "soft red", "torrid red", and "royal red". Regardless, people apply linguistic descriptions liberally and divergently to perceptual experience.

For this reason, people's linguistic descriptions of concepts may strike us as unprincipled. Because so many linguistic descriptions can be applied to the perceptual symbols that underlie a concept, we should have little reason to believe that these descriptions will be highly constrained or easily predictable. Essentially, any linguistic symbol, or combination of linguistic symbols, whose schematic perceptual conditions are satisfied in the perception being described, can constitute a linguistic description of it. Rather than there being a principled set of linguistic representations for a concept, anything expressible linguistically that applies to its perceptual symbol is a potential descriptor. Finding a principled, consistent, and complete set of linguistic representations for a concept may therefore be difficult, if not impossible, as much failed effort on their behalf suggests.

There must be principled factors that determine the mapping of linguistic descriptions to perceptual symbols, yet these factors are probably very different from the logical principles or meaning postulates.
that theorists often seek to establish within the closed world of linguistic symbols and propositions. Instead, the principled factors that link linguistic symbols to perceptual symbols may concern people's strategies for coordinating attention easily and reliably (Markman, 1989; Tomasello et al., in press). Moreover, these factors may be far from absolute, allowing many possible mappings, rather than producing a highly predictable and narrow set.

Although there may be no privileged linguistic primitives, there may well be perceptual primitives (e.g., geos, colours, textures, transformations). However, we should be wary of seeking perceptual primitives through linguistic analysis, because different languages may map linguistic symbols onto these primitives differently, depending on historical and cultural factors. Certainly, some perceptual distinctions may be so salient as to create strong statistical patterns in the mappings across languages, yet the mapping of linguistic symbols into the components of perception appears to be relatively unconstrained, at least in principle. Bypassing language, therefore, and attempting instead to measure the primitive components of perception more directly seems potentially more informative. Although language is perhaps our easiest means of tapping perceptual knowledge, and even though language may often be correlated with the structure of perceptual knowledge, we should not let its convenience prevent us from seeking less convenient but more direct measures that enable stronger conclusions.

**Haphazard Content.** Following the maxim that a perceptual representation affords an infinite number of linguistic representations, we should expect that linguistic descriptions of conceptual content will be haphazard. Depending on the context, people may describe the same perceptual symbol differently. Whereas a particular description might be salient in one context, it might be less salient in another. For example, “wings” might be a more salient description than “beak” when the perceptual symbol for *robin* is examined in the context of flying, but “beak” might be more salient in the context of eating. Different contexts may often cause people to focus selective attention on different aspects of the same perceptual symbol, such that the aspects they describe vary. Because so many aspects of a perceptual representation can be selected, and because selection is likely to vary extensively with context, linguistic descriptions are likely to be haphazard. Even though the perceptual representation remains constant, the linguistic descriptions of it change.

**Incomplete Content.** Finally, we should expect that linguistic descriptions of conceptual content will be incomplete. Because there are virtually limitless ways to describe a perceptual representation, we should never expect that any linguistic description of conceptual content, either intuitive or theoretical, will capture all of them. This view readily explains the incompleteness problems associated with recursion and goal-relevance. For recursion, we simply need to assume that the level of decomposition expressed linguistically depends on how selective attention is applied to a perceptual representation. If attention focuses only on large global chunks of the perceptual representation, then the hierarchical depth of linguistic description will remain shallow. As attention focuses increasingly on the fine details of the perceptual representation, the recursive decomposition that follows increases the hierarchical depth of linguistic description. Without analysis of perceptual representations, it is hard to imagine how recursion could proceed, because it is not clear what source of information would provide the detail for further decomposition. In a perceptual representation, however, it is often possible to extract additional perceptual detail not described previously. Because it is difficult, if not impossible, to exhaust the potential detail of perception, recursive description can continue indefinitely.  

In addition, there are an indefinite number of potential relations between the parts of a perceptual representation that can be described linguistically. For example, an indefinite number of spatial relations exist between all possible pairs of points in the perceptual symbol for a horse's body. Although some of these may be established in memory, many may not be represented explicitly but be easily computable from examination of the perceptual symbol (e.g., a horse's mane is above its stomach, a horse's stomach is above its hocks). Depending on which pairs, triples, etc., of points are selected, and depending on the nature of the relation(s) specified between them, an indefinitely large set of linguistic descriptions becomes possible.

For goal relevance, we simply need to assume that a perceptual symbol can be described differently depending on a person's current goal. Consider the perceptual symbol for the seat of a chair. This symbol can be described linguistically as “supporting sitting”, as “supporting standing”, as “providing cover for a child in an earthquake”, as “a place to store newspapers”, and so forth. As a perceptual symbol is perceived in new contexts, new functions arise for it, depending on what the current goal suggests.

Furthermore, a perceiver can construct endless numbers of linguistic attributes for a seat, depending on the current function. For example, if someone wants to stand on a chair, its “height”, “sturdiness”, and “mobility” may become relevant, but if someone wants to rest a glass of water on a chair, its “flatness” may become relevant instead. Because a
perceptual symbol can be viewed as serving many goals, the linguistic
descriptions of its functions and attributes remain an open, not closed,
set.

In summary, linguistic vagary falls naturally out of an architecture
in which linguistic symbols are grounded in a compositional system of
perceptual symbols: Because the linguistic description of perceptual
symbols is relatively unconstrained, linguistic vagary ensues. As I shall
describe next, the structure of concepts also emerges naturally from this
architecture.

Structure
As we saw earlier, people's knowledge of a category in long-term memory
is not a "flat" set of features. Instead, this knowledge contains
hierarchical, relational frames, built from attribute-value sets,
structural invariants, constraints, and recursion. In the architecture I
am proposing, the knowledge of a category in long-term memory
contains two levels of frame structure: a perceptual frame and a
linguistic frame. A perceptual frame represents the perceptual symbols
generally shared by the exemplars of a category, as well as the spatial
and temporal relations between them. For example, the perceptual
frame for a chair represents the spatial configuration of perceptual
symbols for its seat, back, and legs, whereas the perceptual frame for
chase represents not only a spatial configuration of perceptual symbols
for chaser and chasee, but also the changes in this configuration over
time. In addition, these perceptual frames may be recursive, with the
selection of a particular perceptual symbol within a perceptual frame
revealing a more specific configuration of perceptual symbols that
constitutes a more specific frame, and so forth.

In contrast, a linguistic frame is the integrated network of linguistic
symbols that is grounded in a perceptual frame. Consider the linguistic
and perceptual frames for dog. The perceptual frame might include
perceptual symbols fur, ears, head, fur, tail, and many other schematic
components integrated spatially and temporally. In turn, the linguistic
frame contains linguistic symbols for some of these perceptual symbols,
as well as linguistic expressions for some of the relations between them.
For example, the linguistic symbols "ears" and "fur" might be grounded
in their respective perceptual symbols, and the linguistic expression
"the ears are attached to the head" might be grounded in the spatial
relation between ears and head. As we shall see shortly, the linguistic
frame serves to organise constancy and variability in the perceptual
frame, and to provide compositional access to it during retrieval.

Frame Creation. When is a new pair of linguistic and perceptual
frames created? This is a difficult problem, although the following sorts
of heuristics appear responsible.21 First, whenever a newly discovered
set of entities appears to share a common configurational structure, a
new pair of frames is established for that type of thing (cf. Biederman,
1987; Tversky & Hemenway, 1988). For example, if a new type of animal
is discovered, a pair of frames is established that captures its common
configuration of parts. Similarly, if one of these components can be
described recursively as containing its own configuration of parts, a new
pair of frames is established for this component. For example, if all
instances of this new animal share a common leg structure, a pair of
frames becomes established for the configuration of leg components.
Note that the linguistic frame is likely to lag behind the perceptual
frame, with only some perceptual symbols in the perceptual frame
developing linguistic counterparts.

Second, whenever a set of entities predicts a criterion of interest, a
pair of frames is established for that set. For example, if people are
interested in the set of foods that best predicts weight loss, they may
establish a pair of frames for the category foods to eat on a diet. Note
that the perceptual frame for such categories may often not contain
spatially integrated components. For foods to eat on a diet, the
perceptual attributes for taste and fillingness might simply be
represented as independent perceptual symbols extracted from the
gustatory and proprioceptive experiences of eating. Furthermore, some
of the critical attributes primarily appear to be linguistic, having no
perceptual counterparts (e.g. number of calories, nutritional value).
For attributes like these, people may simply associate linguistic (and
numeric) values with particular exemplars (e.g. 130 calories for one cup
of non-fat yoghurt), and then use these values to predict the criterion,
which might be perceptual symbols extracted from the perception of
body weight. In some cases, a perceptual representation may not be
necessary for linguistic processing.22

Third, whenever an individual of any set becomes familiar and
established in the perceiver's world model (Barsalou, 1991, pp.
53–57), a pair of frames becomes established for the individual. For
example, if an instance of cat becomes a pet, a pair of frames mayecome established for it. Not everything in the generic frames for
cat is necessarily duplicated in the frames for the individual; instead,
the frames for the individual may primarily contain information that
has been selectively extracted from experiences of that individual.
Should information that has not been selectively extracted become
relevant, the frames for cat (or another related category) could be
consulted.
Frame Development. A pair of linguistic and perceptual frames develops in long-term memory to represent the constancy and variability of a category's instances. Consider the pair of frames that might develop for dog. In the perceptual frame, the relative constancies of dogs are represented by perceptual symbols for its characteristic body parts (e.g. ears), behaviours (e.g. barking), and so forth. In turn, linguistic symbols become grounded in these perceptual symbols, such as the words "ears" and "bark". Because the specific ears and barking that dogs exhibit vary considerably, however, this pair of frames must evolve to represent variability as well as constancy. To do this, the linguistic symbols associated with perceptual constancies, such as "ears" and "barking", become attributes, whereas the linguistic symbols that represent variability become their values. For example, the linguistic symbols "floppy" and "upright" might specify values of "ears", whereas the linguistic symbols "deep" and "shrill" might specify values of "barking". These linguistic symbols for values become grounded in perceptual symbols that specialize their respective attributes in the perceptual frame. For example, the perceptual frame for dog can be specialized to represent the attribute values of upright ears and shrill barking by simply inserting the corresponding perceptual symbols into the frame (if not already present as defaults). Note that perceptual grounding solves the problem of representing relative size, as is needed to represent short versus long tail, simply by scaling the length of the tail relative to the body size in the perceptual frame. Similar to the context-dependent representations of red and kind described earlier, perceptual grounding provides a natural solution for the context-dependent forms that attributes and their values often take.

Analogous to attributes and values, structural invariants and constraints represent constancy and variability, respectively: Whereas structural invariants represent relations that remain relatively constant across the exemplars of a category, constraints capture systematic variability. In the perceptual frame for dog, the spatial relation between a dog's ears and its head constitutes a spatial invariant. Analogously, in the perceptual frame for bite, the temporal relation between perceptual symbols for a mouth closed, then open, and then closed on an object constitutes a temporal invariant.

In general, it appears that relations like these are often not represented in the corresponding linguistic frames. Instead, a more general vocabulary of linguistic expressions exists to describe these relations as they become relevant. In English, these expressions often incorporate spatial prepositions such as "above", "in", and "attached", temporal prepositions such as "before", "after", and "simultaneously", as well as many other linguistic devices. To the extent that structural invariants are described frequently for a particular perceptual frame, linguistic counterparts may become grounded in them. To the extent that such relations aren't described linguistically, however, they can always be constructed ad hoc, simply by examining the perceptual frame and applying linguistic expressions for relations whose schematic perceptual conditions are satisfied.

Although structural invariants in a perceptual frame may not have counterparts in the corresponding linguistic frame, counterparts may often exist in another frame at a higher taxonomic level, or in the frame for a more typical exemplar. Although people may not linguistically represent the fact that a dog's ears are attached to its head, such a relation may exist in the linguistic frame for animal (e.g. "an animal's ears are attached to its head"). Similarly, because humans are typical animals from our egocentric perspective, we may represent many structural invariants in the linguistic frame for human (e.g. "the thigh bone is connected to the knee bone"). In either case, these linguistically expressed relations can be generalized to other concepts either through inheritance or analogy.

Constraints, too, can exist in both the perceptual and linguistic frames for a category. Rather than capturing variability across exemplars, however, constraints capture systematic covariances. In the perceptual frame, they manifest themselves as co-occurring specialisations of the perceptual symbols that represent attributes. Consider the constraint, "a dog that growls will bite". In the perceptual frame for dog, this relation may typically be absent, because it is not a structural invariant across dogs. Instead, the default sound represented in the perceptual frame for dog might be silence or barking, and the default mouth action might be closed or panting. To represent the constraint between growling and biting, these defaults are replaced with a perceptual symbol for growl, preceding a perceptual symbol for bite. In other words, the constraint is represented by specialising the default perceptual frame with the temporal sequence of perceptual symbols that represents the constraint. In the linguistic frame, a corresponding linguistic description might also exist, "a dog that growls will bite", which is grounded in the specialised form of the perceptual frame.

Because many constraints are probabilistic rather than absolute, it is necessary to have some means of representing these weaker relations, as in "a dog that growls may bite". This weaker form of the constraint may be represented by following barking in the perceptual frame with two separate values for mouth action: In one instance of the specialised frame, the mouth does not bite, and in the other it does, representing the possibility of either occurring. To represent the probability of a constraint, linguistic descriptors such as "often" and "rarely" could
modify its linguistic form, based on the relative numbers of memories for each event stored with the perceptual frame.

Finally, recursion in the pair of frames for a concept may often develop in synchrony: As an existing component of a perceptual frame is decomposed into more specific perceptual symbols, linguistic descriptions may often become grounded in them, although probably lagging behind. In the perceptual frame for horse, if the perceptual symbol for a part (e.g. rear leg) is broken down into further perceptual symbols for its parts (e.g. hock), new linguistic symbols may become grounded in them (e.g. “hock”).

Linguistic Control of Perceptual Composition. Once established, a linguistic frame enables compositional control over a perceptual frame. Consider the ability to vary the values of colour in the perceptual frame for horse. Because linguistic symbols for these values can be accessed as a set from “colour” in the linguistic frame, they can generate a rapid sequence of perceptual symbols that specialise colour in the perceptual frame. By accessing “sorrel”, “bay”, “pinto”, and “palomino”, one can imagine a sorrel horse, then a bay horse, a pinto, a palomino, and so forth. Furthermore, these sets of associates can easily be extended metaphorically to new objects, increasing productivity further (cf. Lakoff, 1987; Lakoff & Johnson, 1980). For example, one can imagine pinto-patterned wall paper or a palomino-coloured car.

Not only do linguistic frames allow a person to control perceptual composition as they imagine possible states of the world, linguistic frames also allow cooperating individuals to control each other’s composition process during linguistic interaction. As we saw earlier, the conceptual combination that is ubiquitous in language allows conversationalists to convey states of the world that are not immediately present. Imagine someone saying:

My child’s room has pinto-patterned wall paper.

Using these linguistic symbols to control perceptual composition enables imagining with some veridicality the state described.

Finally, linguistic control of perceptual composition may endow humans with much more control over the environment and their minds than other animal species. In the absence of a linguistic system, the perceptual composition process may be much more dependent on perceptual experience. The primary way in which novel compositions of perceptual symbols may occur is to be driven into these states through perception. In contrast, once perceptual symbols have linguistic counterparts, perceptual composition can be controlled extensively through the manipulation of linguistic symbols during thought or linguistic interaction, even though the corresponding states of the physical world are absent. Linguistic frames may provide much greater compositional control over perceptual frames than is possible in their absence.

Furthermore, the ability to control perceptual composition through linguistic manipulation may greatly increase the accumulation of knowledge in a culture, and therefore the amount of knowledge available to an individual. As is well known, humans, unlike other animals, accumulate knowledge across generations. This is known as the ratchet effect, analogous to a ratchet wrench that rotates backwards to prepare for each new turn without undoing the turns accumulated thus far. As each new generation comes along, it builds on the knowledge of previous generations. For example, most people do not learn to cook from scratch, but borrow heavily from the previous experience of others. Similarly, in building a new electronic device, engineers don’t start from scratch but build on accumulated knowledge of electronics. In contrast, when animals learn to hunt or scavenge for food, their learning almost always starts from scratch. Animals exhibit little, if any, cumulative cultural learning.

Again, the presence of a linguistic system for manipulating perceptual composition may make the difference. Because others can describe objects, procedures, and introspective states to us linguistically, we can benefit from their experience, rather than having to experience everything for ourselves. If we have the requisite frames for constructing the perceptual representations that other people’s linguistic descriptions convey, our knowledge grows beyond our experience. Furthermore, if these bodies of knowledge exist in a written form, an oral tradition is not necessary for conveying them, and the ratchet effect increases, such that the knowledge we can acquire without experience grows dramatically.

Flexibility

In an architecture of linguistic symbols grounded in perceptual symbols, the flexibility of concepts in working memory can arise from many sources. Imagine having to define dog. Because the retrieval cue, “dog”, is a linguistic symbol, it first accesses other linguistic symbols in the same linguistic frame rapidly, including lexical associates for “fur”, “tail”, “ears”, “companion”, “bites”, “barks”, and so forth. Much flexibility can occur at this point, to the extent that the strength of particular lexical associates differs both between and within individuals as a function of frequency, recency, and context. In situations where only a
quick sampling of information is retrieved for a category from a linguistic cue, lexical associates such as these may constitute the primary content of the concept constructed (Moss, 1991).

To the extent that processing is deeper, the lexical associates accessed initially may begin to activate the perceptual symbols in which they are grounded. As a result, a perceptual representation of the concept begins to develop, with several factors producing flexibility: First, if different people access different linguistic symbols initially, then the different perceptual symbols that ground them produce flexibility. For example, if one person accesses the linguistic symbol for “barks”, whereas another accesses the linguistic symbol for “bites”, different perceptual symbols will become active. Second, different individuals may have the same linguistic symbol grounded differently, depending again on the factors of frequency, recency, and context. For example, if one person has typically experienced dogs that bark deeply, whereas another person has typically experienced dogs that bark shrilly, the same linguistic symbol “barks” may be grounded in perceptual symbols that differ. Third, constraints may produce flexibility. For example, the perceptual symbol for deep barking may activate the constraint that deep barking dogs tend to be large, whereas the perceptual symbol for shrill barking may activate the constraint that shrill barking dogs tend to be small. As the result of activating a constraint, the values for size in the perceptual frame vary, thereby producing flexibility.

During the process of constructing a perceptual representation, the linguistic description of it may become increasingly coherent. Whereas the initial burst of lexical associates may be relatively splintered, the structural coherence of the developing perceptual representation may cause subsequent linguistic description to be more text-like. For example, perceptual symbols for the body parts of a dog may assemble coherently, such that subsequent linguistic description follows the body’s spatial organisation. This increasing coherence does not mean that the perceptual representation becomes a familiar exemplar (e.g. a known dog). Instead, the perceptual representation may be unlike an exemplar, because many details are not represented (e.g. ‘colour’), and because multiple values for some attributes are considered (e.g. perceptual symbols for both upright and floppy associated with the perceptual symbol for ears).

Once a perceptual representation exists, further flexibility may result from linguistic description of it, as people begin to describe aspects not specified by the initial burst of lexical associates. As we saw earlier for linguistic vagary, a tremendous amount of flexibility is possible at this point in the construction of a concept, with the linguistic descriptions produced being relatively unprincipled, haphazard, and incomplete.

Finally, as these new linguistic descriptions are produced, they may begin this entire cycle all over again. New lexical associates become active, whose perceptual symbols modify the developing perceptual representation further, bringing new constraints to bear, along with additional linguistic description. Because this procedure can continue indefinitely, and because it contains so many degrees of freedom, tremendous potential exists for flexibility.

As we saw earlier, however, this flexibility depends critically on the vagueness of the retrieval cue that begins the process. Because the cue, “Define dog”, is extremely vague, and because a tremendous amount of information in memory is potentially relevant, extensive flexibility is not surprising. As we also saw earlier, however, when the construction process is constrained, concepts become more stable. For example, if we present new subjects with potential properties of dog that other subjects generated previously, the new subjects now agree unanimously that nearly all of these properties are potentially true. Although only a fraction of subjects mentioned barks in their definition of dog, they all agree that it is potentially true of dogs. In many such cases, subjects may agree because they all share the same lexical associates for dog, even though their initial bursts of associates differ when constructing definitions. In other cases, a unanimous property may not reside in the system of lexical associates, but may instead be verified by checking it against a perceptual representation. For example, imagine that a subject produces the relation “the ears are attached to the head” in a definition of dog. Just as this subject may have produced the property while describing a perceptual representation rather than having a linguistic expression for it, so may other subjects verify it by consulting their perceptual representations. Subjects may exhibit high stability when their systems of linguistic and perceptual symbols are driven into a common state by a highly specific cue.

Similarly, we saw that a common context increased stability. For example, concepts of vehicle were more stable when subjects considered this category in a specific context (e.g. vacationing in the rugged mountains of Mexico) than when they considered it in a neutral context. Patterns of lexical associates may provide one source of increased stability. For example, the linguistic symbol for “mountains” may be associated with the linguistic symbols for “truck”, “jeep”, and so forth, such that any subsequent construction of a perceptual representation is biased toward these exemplars. Constraints in perceptual frames may also provide increased stability. For example, subjects may construct a schematic perceptual representation for a rocky, washed-out, dirt road in the mountains, which activates perceptual symbols for vehicles encountered previously in this situation. As a result of lexical associates
and perceptual constraints, the concept used to judge typicality is more
constrained than it is in a neutral context, where any information
retrieved about vehicles could apply.

In summary, the explanation of flexibility in the architecture I have
proposed depends on a wide variety of factors. The structure of linguistic
and perceptual frames must be considered, as must the relations
between them. Similarly, the nature of the retrieval cue must be
considered, which could be vague or specific, linguistic or perceptual.
Not only does this architecture have the potential to be flexible, it also
has the potential to be stable.

CONCLUSION

I have proposed that concepts arise from an architecture in which
linguistic symbols are grounded in a compositional system of perceptual
symbols. Methodologically, we may primarily see the linguistic side of
this process, because we rely so heavily on the language of our subjects
to assess concepts. It does not follow, however, that the linguistic
symbols subjects report provide the foundation of their conceptual
systems. Instead, perceptual symbols may be fundamental, because
they ground linguistic symbols, and because they contain a limitless
wealth of information essential for flexibility. We should not let the
convenience of a methodology dictate our preferred theoretical view.
Just because linguistic symbols provide the easiest means of measuring
concepts, as well as of representing them theoretically, it does not
necessarily follow that the underlying representations of concepts are
linguistic. Instead, the core representations could take some other form,
such as perceptual symbols, which is more difficult for both subjects and
theorists to articulate.

As we have seen, previous theorists have argued that perceptual
representations provide the foundation of human knowledge. In recent
work on concepts, however, researchers, including myself, have largely
adopted linguistically oriented representations, failing to question this
orientation and entertain other possibilities (e.g. Barsalou, 1992a,
Chapters 7, 8, 9). Even if perceptual symbols turn out not to be
important, they should be considered seriously, because doing so is likely
to produce significant progress in understanding human concepts. At
the least, we may reach a clearer understanding of perception’s role
in the human conceptual ability, and we may establish a clearer and better
justified account of amodal propositions, should they turn out to play a
more central role than I have suggested.

Connectionism is not the only solution to the ills of amodal symbol
systems. Perceptual symbols offer an additional orthogonal solution. By
discovering the relations between linguistic symbols and the perceptual
symbols that ground them, by developing psychologically plausible
mechanisms for representing and processing the hierarchical relational
structure of frames, and by finding ways to do all of this in a statistical
processing environment, we may move closer to a satisfactory account
of human concepts.

ACKNOWLEDGEMENTS

This chapter was supported by funding from the Army Research
Institute (MDA 903-90-K-0112) and the National Science Foundation
(IRI-8609187). I am grateful to Richard Billington for discussion on
chaos theory and frames, from which the ideas in this chapter developed.
I am also grateful to Janellen Huttenlocher for helpful discussion, and
to Martin Conway, Gary Dell, Raymond Gibbs, Robert Goldstone, Koen
Lamberts, Gordon Logan, Arthur Markman, Gregory Murphy, Vinzenz
Morger, Nora Newcombe, Edward Shoben, Edward Wisniewski, and
Brian Ross and his graduate seminar on concepts for helpful comments
on earlier drafts of this chapter.

REFERENCES

University Press.
DC: Winston.
Anderson, R.C., Pichert, J.W., Goetz, E.T., Schallert, D.L., Stevens, K.V., & Trollip,
Ashcraft, M.H. (1976). Property norms for typical and atypical items form 17
Comprehension and semantic flexibility. Journal of Verbal Learning and Verbal
Behavior, 13, 471–481.
Neisser (Ed.), Concepts and conceptual development. Ecological and
intellectual factors in categorization (pp. 101–140). New York: Cambridge
University Press.


Barsalou, L.W. & Sewell, D.R. (in prep.). Constructing categories from different points of view.

Barsalou, L.W., Sewell, D.R., & Spindler, J.L. (in prep.). *Flexibility and stability in concepts*.


3. A COMPOSITIONAL SYSTEM OF PERCEPTUAL SYMBOLS


3. A COMPOSITIONAL SYSTEM OF PERCEPTUAL SYMBOLS


APPENDIX
Representing Conceptual Relations in Feature Lists and Connectionist Models

Theorists sometimes attempt to represent conceptual relations in feature lists. It is important to review these attempts, because doing so illustrates the fundamental limitations of feature lists, and the necessity of using frame-like representations to capture the structure of human concepts. In this brief review, I will focus primarily on the use of features in connectionist models to represent conceptual relations, although all of my conclusions apply equally to the use of feature lists in other models as well, including exemplar models, prototype models, memory models, and so forth. In addition, I will focus mostly on localist connectionist nets, although distributed nets offer no solutions to the problems I raise, as far as I can tell. In the following six subsections, I review six connectionist attempts to represent conceptual relations with features, and the problems for each. In the final subsection, I summarise the lessons we have learned from these attempts, and suggest some basic relational criteria that a representational system must satisfy.

Mutually Inhibitory Features

One way to represent attribute-value relations is to integrate mutually exclusive values of the same attribute with inhibitory connections (Rumelhart, Smolensky, McClelland, & Hinton, 1986). On this view, for example, each feature for a colour has an inhibitory relation to every other colour feature (e.g. the feature for **green** has inhibitory relations to **blue**, **brown**, **orange**, **yellow**, etc.). The basic idea is that when one of these features is active, it inhibits all others. Because only one of these features can be the current value of the attribute, only one of them is active at a time. For example, if **green** is active, then the features for all other colours should be inactive, because **green** is the current value of the attribute for **colour**. In this way, fields of mutually inhibitory features come to function as attribute-value sets, representing these sets implicitly rather than explicitly.

One problem with this account is that inhibition between features is neither necessary nor sufficient for attribute-value relations. Inhibition is not necessary, because two values are frequently active simultaneously for the same attribute. For example, the **colour** of a dalmation is simultaneously **white** and **black**. When **white** is active, it does not necessarily inhibit all other colours. Nor is inhibition between features sufficient for attribute-value relations, because inhibitory

relations can exist between features that are not values of the same attribute. For example, a feature for **snow** might inhibit a feature for **beach**, even though they are values of different attributes, **activity** and **location**, in the frame for **vacation**.

A second problem with feature inhibition is that it does not account for the representation of attributes. Only the values of an attribute are represented, with there being no representation of the attribute to which they bind. Because of these two problems, feature inhibition does not provide a satisfactory means of implementing attribute-value relations.24

Jointly Active Features

A second approach to representing conceptual relations is to represent them as configurations of jointly active features. Consider representing the relation, **ABOVE** (**circle**, **triangle**). In this approach, separate features exist for **ABOVE**, **circle**, and **triangle** in the network, along with features for other relations (e.g. **FRONT**, **LEFT**) and other objects (e.g. **square**, **octagon**). When a particular relation is present in the environment, the features corresponding to its relation and arguments become active simultaneously to represent the relation. For example, the features for **ABOVE**, **circle**, and **triangle** become active to represent **ABOVE** (**circle**, **triangle**). For other relations, different configurations of relation and argument features become active.

One serious problem with this approach is that it fails to specify the argument bindings within a relation. When **ABOVE**, **circle**, and **triangle** are active simultaneously, they do not distinguish between the relations **ABOVE** (**circle**, **triangle**) and **ABOVE** (**triangle**, **circle**), because this representation has no way of specifying the bindings between the relation and its arguments.

A second serious problem, described in compelling detail by Fodor and Pylyshyn (1988, pp. 22–23), is the failure of this approach to specify argument bindings between relations. To see this, imagine that two relations are present simultaneously in the input, **ABOVE** (**circle**, **triangle**) and **FRONT** (**square**, **octagon**). If the features for **ABOVE**, **FRONT**, **circle**, **triangle**, **square**, and **octagon** are active simultaneously, the system has no way of distinguishing among the 12 possible relations that are consistent with the activation of these 6 features.25

Relational Features

A third approach to representing conceptual relations in feature lists is to construct relational features. A relational feature is simply a single
feature that represents an entire conceptual relation (e.g., Hinton, 1989; McClelland, Rumelhart, & Hinton, 1986). To see this, consider how we might use relational features to handle the feature binding problem (Fodor & Pylyshyn, 1988, pp. 22–28). For the relation, ABOVE (circle, triangle), features could exist for circle-subject and triangle-object. When the system encounters this relation, features for ABOVE, circle-subject, and triangle-object become jointly active and represent the binding relations unambiguously. Handling the between-relation binding problem, however, requires more complex features. Imagine representing the pair of relations ABOVE (circle, triangle) and FRONT (square, octagon). To represent these relations unambiguously, we need relational features such as circle-subject-of-ABOVE and square-subject-of-FRONT. The complexity of relational features must increase still further to handle multiple instances of the same relation, as in ABOVE (circle, triangle) together with ABOVE (square, octagon). Here, relational features, such as circle-subject-of-ABOVE-circle-triangle, become necessary, even though they redundantly repeat the relation whose component they are representing.

It is important to see that these relational features are non-compositional. Rather than being compositional, a feature such as circle-subject-of-ABOVE-circle-triangle is “pointillist”, representing a single state of affairs in a way that is totally independent of the features for circle, subject, ABOVE, and triangle. Whereas most other attempts to represent relations with features are compositional at least in some sense, this one is not. Instead, for any possible relation, whatever the type, an independent feature can be defined, because this approach allows us to put whatever information we want in a feature, no matter how complex, and no matter what other features do or do not exist.

There are four devastating problems for this approach. First, the number of relational features explodes exponentially with the number of non-relational features being related (Fodor & Pylyshyn, 1988, pp. 24, 34). Consider all of the possible relations for above. A book can be above a table, a pencil can be above a table, an apple can be above a table, a mouse can be above a table, an apple can be above a mongoose, and so forth. To represent every possible above relation would require a relational feature for every possible pair of objects and events in the world.

Second, the representational system has to anticipate all possible relations, no matter how nonsensical, in order to represent any one that could possibly occur. For example, it would have to represent ABOVE (apple, mongoose) a priori, to recognize this relation should it ever occur.

The third problem is related to the second: Relational features are not productive. Since Chomsky (1957), we have known how critical it is that theories of cognition be productive, at least in certain regards. Relations constitute one important form of productivity in human cognition. People can recognize all sorts of relations that they have never seen before and most likely do not have represented a priori in memory. Again, consider above, which has two attributes that become bound to two entities, one of which must have a higher vertical position than the other. Using a very simple frame representation that can be bound to any pair of entities in space, it is possible to account for all above relations with an extremely simple mechanism. Rather than requiring an indefinite number of relational features, we only need a single relational frame. Not only is this frame efficient, it is able to represent above relations that have never been encountered before and that aren’t already stored explicitly in the cognitive system.

A fourth problem for relational features concerns the absence of relations between related features (Fodor & Pylyshyn, 1988, p. 40, footnote 26). Imagine that a system contains relational features for ABOVE-book-table, ABOVE-bird-house, and ABOVE-shovel-sidewalk. Clearly, these three relational features are related, because they all express an above relation. However, this representation treats them independently and has no way of knowing they’re related, even though people can recognize their relatedness immediately. Similarly, consider the features for book, table, and ABOVE-book-table. Although these three features are related when a book happens to be above a table, the representation again treats them independently and has no way of knowing they’re related. Because relational features themselves are unrelated in a feature list, they, too, fail to capture essential relations between features. Although we can put whatever information we want in a feature, we can’t represent the conceptual relations between related features unless we adopt a more expressive representation.

Ordered Feature Modules

A fourth approach to representing conceptual relations is to dedicate modules of features to attributes and their values. To see this, consider how connectionist models often attempt to represent verbs (e.g., Mikkullainen & Dyer, 1991; McClelland & Kawamoto, 1986). One module is dedicated to representing the verb, with different patterns of activation across the module’s features (i.e., units) representing different verbs. A second module is dedicated to representing the agents of the verbs, with different patterns of activation across the module’s features representing different agents. Similarly, further modules might represent other verb attributes, such as theme, patient, instrument, source, and so forth. In each case, the module represents an attribute,
and the patterns of activation across the module represent its values. In addition, connections exist between modules, which allow the system to represent co-occurring patterns of verbs and their arguments (e.g. EAT (lion, meat), BUY (child, gum, penny)).

One problem for this approach is that all of the necessary modules must be specified a priori. The origins of these modules remain unspecified, and the system has no ability to develop new modules as needed. For this approach to be viable, clear accounts are necessary of how a priori modules originate. Clearly, some modules could be innate for attributes that have acquired significance through evolution. However, a system must also have the ability to create new modules, given people's proclivity for constructing new attributes. For example, Barsalou (1992b) found that people frequently construct new attributes as they become necessary, such as amount of work disruption for the category of departure times when planning a vacation. Similarly, any value of an attribute can become an attribute itself. For example, means of locomotion is an attribute for animal, taking values such as legs, wings, and fins; however, each of these values can in turn be an attribute, taking still more specific values. Legs, for example, could be an attribute of mammal, taking different values across dog, monkey, and elephant. Given the ubiquity of attribute construction, a viable system of representation requires the ability to construct new attributes as they become necessary.

A second problem for the ordered module approach is that it fails to solve the between-relation binding problem noted earlier. Although the system can correctly bind the components of a single relation by assigning each to its appropriate module, it cannot distinguish between ambiguous interpretations of two or more simultaneous relations. Imagine that someone hears the sentence:

"Ann bought a sawhorse to build a workbench for Bill."

and extracts the relations, BUY (Ann, sawhorse) and BUILD (Ann, workbench, Bill). The system has no way of knowing whether Ann or Bill is the agent of BUY, whether sawhorse or workbench is its theme, and so forth. Instead, both verbs and their arguments are superimposed in a single ambiguous pattern. This will always be a problem when multiple relations must be represented simultaneously, as when representing the spatial relations between the parts of an artifact or animal (e.g. simultaneously representing the spatial relations between the seat, back, and legs of a chair).

A third problem for the ordered module approach is that it requires external interpretation, the problem addressed in the next section. The form it takes here is that external interpretation of the modules is necessary for setting up and using the system. Nothing in the structure of the system per se specifies that one module represents the agent, another the theme, and so forth; nor does anything intrinsic to the structure of the modules determine their function in the system. Clearly, patterns of features may come to specify that agents are animate, that instruments are inanimate, and so forth. Nevertheless, these inductions do not determine the initial assignments of interpretations to modules (e.g. agent to the first module), nor do they affect the subsequent assignment of modules to segments of the input (e.g. agent to the first word, theme to the third, etc.). Instead, these interpretations and assignments are handled externally to the system.

External Interpretation

A fifth approach to representing conceptual relations in feature lists involves an external interpreter that projects relations onto feature sets. Imagine a connectionist net that represents mammal with one pattern of activation over a set of features (i.e. processing units), and that represents horse with a second pattern of activation over the same features. Because the features active for mammal constitute a subset of the features active for horse, an external interpreter can infer that an isa relation links these two concepts, namely, ISA (horse, mammal).

There are two serious problems for this approach. First, the relation between feature lists is ambiguous. Imagine instead that we observed one pattern of active features for mammal and a second pattern of active features for head, with the active features for head being a subset of the active features for mammal. We would not infer the relation ISA (head, mammal), because we should infer PART (head, mammal) instead. In general, considerable ambiguity can arise between sets of active features, with a variety of possible relations existing between them.

The second serious problem lies in how the ambiguity between feature sets is resolved. If ambiguity is resolved by an external interpreter, important issues are finessed by not handling them directly within the system. For example, an external interpreter might assign isa relations, as in feature comparison models of semantic memory, or an interpreter might assign part relations, as in certain connectionist models that attempt to represent paronymic structure (e.g. Smolensky's, 1988, cup of coffee example, pp. 16–17). The key problem is that these conceptual relations lie outside the representation in an external interpreter. Because the external interpreter typically is not part of the cognitive theory proper, and because the relations it computes do not lie explicitly in the representation, relations cannot be processed.
as information in their own right, in the same manner as information directly in the representation.

The problem of external interpretation sometimes takes another form in connectionist models. In some connectionist research, hidden weights are submitted to scaling programs, such as multidimensional scaling and hierarchical clustering (e.g., Elman, 1990; Sejnowski & Rosenberg, 1987). These scaling programs identify subsets of features that are then interpreted as being related to one another in various ways, such as through isa or part relations. Because these scaling programs and their solutions are not part of the connectionist architecture, however, the isa and part relations they discover do not exist in the connectionist representation. Although patterns of activation across features may be similar to one another in various ways, the relations between them remain ambiguous, until interpreted externally.

Vector Superimposition

A sixth approach to representing conceptual relations in connectionist nets is to superimpose the components of these relations as vectors over a common set of processing units (Hinton, 1988; Pollack, 1990; Smolensky, 1990; van Gelder, 1990). To store the relation, ABOVE (circle, triangle), the pattern for relation ABOVE might be stored in the network, followed by the patterns for subject-circle and object-triangle. All three components of the relation would be stored as distributed representations, superimposed as vectors over the same processing units. Later, the network can be cued for each component by presenting it with the partial information that specifies the relevant vector. For example, the network might be presented with subject-, which matches the subject-circle vector and reinstates circle as the entire pattern becomes active. Entire relations can be retrieved, simply by cueing the relation and each of its arguments. For example, ABOVE (circle, triangle) could be retrieved by the series of cues: relation-, subject-, object-

Problems that we have seen earlier are problems for vector superimposition as well. To represent multiple relations, the system needs to store relational features that can become unwieldy and explode in number exponentially. To store ABOVE (circle, triangle) and FRONT (square, octagon) unambiguously, vectors such as subject-circle will not work, because circle could be the subject of either relation. Consequently, vectors such as circle-subject-of-ABOVE are necessary. However, these more complex vectors will not work when multiple tokens of the same relation must be superimposed, as in ABOVE (circle, triangle) and ABOVE (square, octagon), because circle-subject-of-ABOVE could be the subject of either relation. Again, we need vectors like circle-subject-of-ABOVE-circle-triangle, which contains the relation itself.

There is also much potential for the abuse of external interpretation in such systems. For example, to cue the system for relations and the vectors that constitute them, an external agent needs to know the content of the network. If only one relation has been stored, the external cueing system needs to know the number and names of arguments to cue. If multiple relations have been stored, then the system needs to know how to cue the components in each relation, which may require cues such as subject-of-ABOVE or subject-of-ABOVE-circle-triangle, as we just saw. If the system does not know which relations are in the network, it could apply all cues exhaustively to see which provide strong, coherent responses, suggesting that they had been stored. However, the system still requires external knowledge of what the system could contain. Another possibility is that relational components can simply be stored according to sequence, with the first component being associated with position 1, the second with position 2, and so forth. The problem here is that the network must then be cued in sequence to retrieve its contents in an orderly manner that specifies which components go together. If content addressable search is attempted, the problems of ambiguous bindings again come into play.

SUMMARY

As we have seen, attempts to represent conceptual relations with feature lists raise at least as many problems as they solve. It may well be possible to overcome these problems and to develop a satisfactory connectionist account of conceptual relations. If so, it should satisfy the following five criteria:

1. Attribute learning. A representational system must have the capacity to develop new attributes, rather than relying on an a priori module for each (e.g. not having to build an a priori module for the attribute transmission type into the network that represents car).

2. Productive relations. For a given type of relation, a representational system must have the ability to construct new instantiations of it, rather than having to pre-store all of them (e.g. not having to store all possible instantiations of ABOVE a priori).

3. Unambiguous and efficient binding. A representational system must have means of binding arguments both within and between
relations that is unambiguous and efficient (e.g. without having to employ features like circle-object-of-ABOVE).

4. Relational similarity. A representational system must be sensitive to the similarity between tokens of the same relation (e.g. the similarity between ABOVE (circle, triangle) and ABOVE (square, octagon)), as well as to the similarity between a relation and its arguments (e.g. the similarity between ABOVE (circle, triangle), circle, and triangle).

5. Internal completeness. A representational system should not have to rely on external interpretation for the representation of its conceptual relations.

One might believe that the only way to satisfy these criteria is to adopt a language along the lines of predicate calculus, as advocated by Fodor and Pylyshyn (1988), yet this approach is fraught with problems, as we saw in the main body of this chapter. Moreover, as Rumelhart (personal communication, July 1991) has observed, the fact that people's behaviour exhibits knowledge of conceptual structure does not mean that this structure exists literally as predicate calculus expressions in the cognitive system, as Fodor and Pylyshyn propose. Instead, such structure may arise as emergent properties of simpler mechanisms (e.g. a compositional system of perceptual symbols). What remains is to discover a satisfactory implementation of such structure in a statistical processing environment.

NOTES
1. For the definition of represent that I am using, see Barsalou (1992a, Ch.3). For a more extensive definition of category, see Barsalou (1992a, pp. 170-171).
2. Quotes around a word or expression represent either its auditory or visual surface form (e.g. "piano" represents the spoken or written form of this word). A word or expression in italics represents either its intensional or extensional meaning (e.g. piano represents a concept of pianos, or a set of physical pianos in the world).
3. Note that some of the features in a concept may not be stored in long-term memory but may be inferred (e.g. inferring that birds have hearts). However, casual inspection of the features in this experiment suggests that generally they were well-known features of categories with a high likelihood of being stored (e.g. birds have wings). Clearly, though, stored information can always be used to infer new information, and the stability of such inferences is an important issue.
4. For additional factors that determine accessibility, see Barsalou and Billman (1989, pp. 195-199).
5. Technically, the models of Estes, Medin and Schafer, and Nosofsky use frames to represent categories, rather than feature lists, because each feature is a value on an attribute. However, because there are only two binary values per attribute, and because the attributes have no theoretical status other than to enable weighting, these representations are essentially feature lists (Barsalou, 1992b, pp. 22-25).
6. As connectionist models have amply demonstrated, feature list models readily account for statistical relations between features. The problem is that they have difficulty representing different types of relation between features, and in particular the various types of conceptual relations described in this section. See Barsalou and Hale (1992) for further discussion.
7. The feature lists in Fig. 3.1 are not intended as accurate or complete accounts of the features in a concept, but simply serve to illustrate the absence of particular conceptual relations. Indeed, the extreme difficulty of identifying accurate and complete accounts of concepts, should one attempt to do so, is the theme of the section on linguistic vagary.
8. The reader may wonder why these relations can't be represented by relational features such as colour-yellow. As described shortly, the Appendix raises serious problems for this approach, as well as for other feature list approaches to representing conceptual relations.
9. The reader may wonder why such relations can't be represented by relational features such as back-above-seat. As mentioned in Note 8, and as described shortly, the Appendix raises serious problems for this approach, as well as for other feature list approaches to representing conceptual relations.
10. By haphazard I don't mean that the content of concepts is completely random. Certainly, some content remains stable across contexts (Barsalou, 1982; Gennari, 1987). By haphazard I simply mean that context can produce partial changes in content, although often to a sizeable extent.
11. Relational in this discussion means that the perceptual components at a given hierarchical level are integrated by spatial relations (i.e. structural invariants). For example, the lips, teeth, and tongue of the mouth are related by spatial invariants at a particular hierarchical level within representation of the face. Similarly, the parts of a tooth (e.g. crown, root) are related spatially at the next lower level of hierarchical analysis.
12. Because these schematic representations are perceptual, they may of necessity contain information about irrelevant components. For example, the representation of red may be stored as a colour patch that has a roughly circular shape, simply because a perceptual representation must have some shape. Note, however, that this shape may play no role in the application of the schematic colour. In imagining a red banana or a red cloud, the default circular shape is immediately abandoned as the red is applied by a procedure that "paints" the modified object (e.g. banana). Similarly, a schematic shape might have a default colour, such as white or grey, simply because a perceptual representation may require some colour. Again, however, this default may be abandoned rapidly when modified, as in red cylinder.
13. This example makes no assumptions about whether the underlying account of category learning is an exemplar or abstraction model (Barsalou, 1990).
14. Although specific conventions for establishing the referents of particular
drawings may typically not be necessary, general cultural conventions about
the representational roles of drawings may be, given that some cultures do
not view drawings symbolically.
15. Clearly, more complete analyses are necessary to handle the ambiguous
nature of above and cross. Rather than attempting to provide complete
accounts, these examples simply serve to illustrate the insertion process in
perceptual composition.
16. Another example of this problem is mental rotation, which theorists can
represent in a propositional system, but which is represented much more
easily and naturally in a compositional system of perceptual symbols that
includes rotation as a possible transformation.
17. Clearly, other uses of “a” and “the” work differently. For example, in “Today
we discussed the cell in my biology class”, “the” specifies the operation of
retrieving the generic perceptual symbol for cell, not any particular instance
of it.
18. Further definitions of this perceptual symbol for chase are in order. For
example, goal could be represented as a perceptual symbol for the state of
the world desired by an agent (e.g. the cat touching the dog). In turn, desire
can be grounded in a perceptual symbol extracted from the introspective
state of having a positive affective attitude towards achieving a state, and
agent can be grounded in a perceptual symbol for entities that control their
own movement. Clearly, more careful accounts must be given, but the point
is that such accounts appear possible with perceptual symbols, as much
work in cognitive linguistics suggests.
19. Note that other symbolic systems besides language can also describe the
perceptual symbols that underlie a concept, including gestures, drawings,
and numbers.
20. It is important to distinguish between linguistic description of perceptual
symbols and linguistic description of immediate perception. Because
perceptual symbols are schematic, thereby omitting much irrelevant detail
by definition, they may not afford much opportunity for the recursive
description of detail (although they do permit infinite description of
relations, as described in the next paragraph). Where the recursive
extraction of detail seems most likely is either from immediate perception,
or from well-established perceptual memories having high resolution. For
example, when perceiving an object or event directly, a viewer can always
focus attention on increasingly fine detail and produce linguistic
descriptions of it recursively. Similarly, a detailed perceptual memory of an
object or event might afford recursive linguistic description.
21. The following ideas were developed in collaboration with Koen Lamberts.
22. Note, however, that a perceptual account could be given of calories and
nutritional value, using perceptual symbols for quantity (see Lakoff &
Johnson’s, 1980, discussion of more). Perceptual symbols for an increasing
pile of substance, associated with corresponding perceptual symbols for
increasing body weight, could represent calories. Similarly, perceptual
symbols for an increasing pile of substance, associated with perceptual
symbols for a healthy individual (not a sickly one) could represent
nutritional value.
23. A sorrel is a light brown horse, a bay is dark brown, a pinto has large
irregular patches of two colours (e.g. white and brown), and a palomino is
golden.
24. As far as I know, mutually inhibitory features have not been extended to
the representation of structural invariants and hierarchical recursion.
Failure to represent these additional relations successfully would constitute
a third argument against this approach.
25. This value of 12 possible relations assumes that no object occurs more than
once in the input (e.g. relations such as ABOVE (circle, circle)), and similarly
that pairs of the same relation do not occur (e.g. ABOVE (circle, triangle)
together with ABOVE (square, octagon)). A third problem, then, for this
approach is that it has no means of representing multiple tokens of the same
feature, unless the system is extended in some way, perhaps as in
McClelland’s (1986) programmable blackboard model.
26. It is important to note that traditional propositional systems suffer from
problems of external interpretation as well. Because propositions are
arbitrary amodal symbols, they typically have no meaning unless
interpreted externally (i.e. the symbol grounding problem). The problem
here for connectionist nets is somewhat different, because relations that
don’t exist anywhere in the net, such as isa and part, are attributed to
relations between states of activation. In a propositional system, at least
symbols for these relations exist.