DOCUMENT RESUME

ED 105 514   CS 202 044

AUTHOR   Collins, Allan M.
TITLE   Studies of Human Memory and Language Processing.
INSTITUTION   Bolt, Beranek·and Newman, Inc., Cambridge, Mass.
SPONS AGENCY   National Inst. of Education (DHEW), Washington,
                D.C.
PUB DATE   Dec 73
NOTE   60p.

EDRS PRICE   MF-$0.76 HC-$3.32 PLUS POSTAGE
DESCRIPTORS   Computer Oriented Programs; *Computer Programs;
                *Computer Storage Devices; Decision Making;
                *Educational Research; Higher Education; Information
                Processing; *Memory; *Semantics

ABSTRACT
        The purposes of this study were to determine the
nature of human semantic memory and to obtain knowledge usable in the
future development of computer systems that can converse with people.
The work was based on a computer model which is designed to
comprehend English text, relating the text to information stored in a
semantic data base that is structured like the human brain. The study
compared the way information was organized within the computer model
with the way humans retrieve information from their own memory
structures. The results show that a network model is a viable
representation of numerical as well as semantic information. The
specific experimental results are summarized in the annotated
bibliography in the second section of the document. (Author/TS)

STUDIES OF HUMAN MEMORY AND LANGUAGE PROCESSING

National Institute of Education,
Project No. 1-0420

Project Director: Allan M. Collins

Institution: Bolt Beranek and Newman Inc.
                    50 Moulton Street
                    Cambridge, Massachusetts   02138

20 December 1973

2

TABLE OF CONTENTS

SUMMARY

The kinds of information education tries to convey for the
most part consist of facts, hypotheses, concepts, and procedures.
New material is always learned in the context of what the learner
already knows. In order for a child to learn that George Wash-
ington was the first president of the United States, he must
relate this information to his knowledge about the United States,
the office of the presidency, and what it means to be first.
Similarly, before one can learn how to multiply, one must know
about numbers and their order, how to add, etc. Simply put,
new concepts are built upon old concepts.

Most psychological studies of learning have concentrated
on how new associations are formed, strengthened, or forgotten,
as occurs in learning a list of arbitrarily paired words or
syllables. Although lear ig new associations obviously is
involved in a child's original learning of the names of things
and a student's learning of new words in a foreign language,
very little of the information that a student acquires in the
process of getting an education is acquired in this pairwise
fashion. More typically, the assimilation of new concepts, facts,
etc., involves the establishment of multiple interrelations
within a complex information structure.

Only recently has psychology taken up the question of how
conceptual information is represented in human memory, and the
processing used to retrieve it. Our work in this direction has
been based on a computer model developed by Quillian. Quillian's
program is designed to comprehend English text, relating the
text to information stored in a semantic data base that is
structured like human memory. We have been conducting

1

psychological experiments for the purpose of testing the impli-
cations of the model. Our goal has been both to understand the
nature of human semantic memory and to obtain knowledge usable
in the future development of computer systems that can converse
with people. The method we have been using is that of measuring
the time required for people to make such decisions as whether a
sentence is true or false, or whether an object named is in a
given category, or whether an equation is correct or incorrect.
Essentially, we test to see if the relations between decision
times that are implied by the way information is organized within
the computer model are in fact obtained when humans are given
tasks to perform that depend upon the retrieval of information
from their own memory structure.

The results of our studies have emphasized the feature-
matching aspects of human semantic processing. They have also
shown that a network model is a viable representation of numerical
as well as semantic information. The specific experimental results
are summarized in the Annotated Bibliography following this sec-
tion. The implications we have drawn from these studies, plus
other recent experiments, are presented in the theoretical paper
that makes up the bulk of this final report.

The knowledge derived from these experiments is being used
in the development of SCHOLAR, a new kind of interactive computer-
assisted instruction (CAI) system, whose knowledge about the world
is stored in a semantic network structured like human semantic
memory. Dialogue with SCHOLAR takes place in English which is
somewhat limited by SCHOLAR's syntactic capabilities. The system
uses its semantic network to generate the material it presents,
the questions it asks, and the corrections it makes. SCHOLAR

accepts questions and requests from the student, and generates
responses based on its semantic network, making inferences of
different types. The experiments have been particularly valuable
in developing SCHOLAR's inferential capabilities, and these
implications are discussed in the section of the theoretical
paper starting on p. 32. This work then is leading to computer
systems that have the kind of knowledge structure that enables
them to function like a human tutor.

AN ANNOTATED BIBLIOGRAPHY OF PAPERS PREPARED FOR THE PROJECT

Collins, A. M. & Quillian, M. R. Structure and Priming in
    Semantic Memory. Paper presented at the Psychonomic
    Society, November, 1971.

    This study measured reaction times for subjects to categor-
ize names of animals and plants with respect to three different
categories: "animal," "bird," and "mammal." There were four
kinds of lists, distinguished by the types of animals in each:
one kind contained only animals that were mammals, a second
kind both mammals and non-mammals, a third kind only birds, and
a fourth kind both birds and non-birds. The results indicated
that the category mammal is <u>not</u> intermediate between elephant
and animal in the way that bird is intermediate between robin
and animal. The results also showed that it takes longer to
decide about a robin or an elephant when there are non-birds or
non-mammals included in the lists. The findings are analyzed
with respect to a model of priming in semantic memory.

Collins, A. M. & Quillian, M. R. Comparing Concepts in Semantic
    Memory.

    This experiment measured reaction times (RTs) for $\underline{S}$s to
decide whether a sentence is true or false. The sentences were
constructed in pairs which appeared on succeeding trials. The
first sentence of the pair concerned a property of some object
(e.g., "An organ has keys") and was followed by a sentence
about the superordinate of the same object (e.g., "An organ is
a piano"). We found that a distinguishing property in the first
sentence acts to speed up the decision that two objects are
different in the second sentence. For example, $\underline{S}$s are faster

in rejecting as false a sentence like "An organ is a piano," if
the previous sentence was "An organ has pipes" rather than "An
organ has keys." Pipes are a distinguishing property of organs
and pianos, but both organs and pianos have keys. This result
supports the notion that comparisons of concepts involves com-
parison of the properties of those concepts.

Triggs, T. J. & Collins, A. M.   The Internal Representation of
    the Multiplication Table.   To be submitted to the
    Journal of Experimental Psychology.

    The experiment reported was performed to evaluate how
numbers and their magnitudes are internally represented in the
specific context of the multiplication table. The experiment
used a two response "correct-incorrect" paradigm where equa-
tions were displayed to subjects, half of which were correct
(3x4=12) and half of which were incorrect (6x4=18). The data
obtained from three well-trained subjects were compared to the
predictions from both a tabular organization and a network
model. Generally, it was found that the reaction time to
decide whether a product was correct or incorrect increased
for larger numbers, which suggests a tabular organization.
However, the fact that squares (e.g., 8x8=64) were responded
to faster than non-squares, but showed the same pattern of con-
fusions on the incorrect trials argues for a network model.

Collins, A. M. & Loftus, E. F.   A Spreading Activation Theory of
    Semantic Processing.   To be submitted to Psychological
    Review.

    Years ago, Quillian proposed a spreading activation theory
of human semantic processing. The theory viewed memory search

5

as activation spreading out from two or more concept nodes in
a semantic network until an intersection was found. The effects
of priming in semantic memory were also explained in terms of
spreading activation from the node of the primed concept. Since
Quillian's theory was proposed, there have been a number of ex-
periments investigating retrieval and priming in semantic memory.
In this paper, we attempt to show how an elaboration of Quillian's
basic theory can account for many of the results found. In the
first section, we briefly review the original theory and try to
correct a number of the common misunderstandings of this theory.
In the second section, we extend the theory in several respects,
and in the third section show how the extended theory deals with
the recent experimental findings.

A SPREADING ACTIVATION THEORY
OF SEMANTIC PROCESSING

Allan M. Collins
Bolt Beranek and Newman Inc.

Elizabeth F. Loftus
University of Washington

7

10

Table of Contents

INTRODUCTION

Years ago Quillian[1] (1962, 1965) proposed a spreading
activation theory of human semantic processing which he tried
to implement in computer simulations of memory search (1966)
and comprehension (1969). The theory viewed memory search as
activation spreading out from two or more concept nodes in a
semantic network until an intersection was found. The effects
of psychological set (or priming) in semantic memory were also
explained in terms of spreading activation from the node of the
primed concept. Rather than being a theory to explain data, it
was a theory designed to show how to build human semantic struc-
ture and processing into a computer.

Since the theory was proposed, there have been a number of
experiments investigating retrieval and priming in semantic
memory. In this paper, we will attempt to show how an elabora-
tion of Quillian's basic theory can account for many of the
results found. In the first section, we will briefly review the
original theory and try to correct a number of the common mis-
understandings of this theory. In the second section, we will
extend the theory in several respects, and in the third section
show how the extended theory deals with the recent experimental
findings.

QUILLIAN'S THEORY OF SEMANTIC MEMORY

The fact that Quillian's theory was developed as a program
for a digital computer imposed certain constraints on the theory,
which Quillian felt were unrealistic psychologically. We will
recount the theory as he proposed it, and then in our elabora-
tion we will remake the theory in psychological terms.

1

The theory made a number of assumptions about structure and processing in human semantic memory. Briefly they are as follows:

People's concepts contain indefinitely large amounts of information. Quillian used the example of a "machine." If you ask people to tell you what they know about machines, they will start telling you obvious facts; for example, that machines are man-made and have moving parts. But they can go on telling you less and less relevant facts, as for example that a typewriter is a machine and even that the keys on the IBM electric typewriters select the position of a ball that strikes the ribbon against the paper. The amount of information a person can generate about any concept in this way seems unlimited.

Information about a concept is differentially accessible. If one regards a concept as a node in a network, then there will be a number of links from the node to other concept nodes. These links are directed links, and they can have different criterialities. In turn, from each of the nodes linked to the first node, there will be links to other concept nodes. In the theory, the full meaning of any concept is the whole network as accessed from the node of the concept.

The links are not simply undifferentiated links, but must be complicated enough to represent any relation between two concepts. In the original theory, Quillian proposed five different kinds of links: a) superordinate and subordinate links, b) modifier links, c) disjunctive sets of links, d) conjuctive sets of links, e) a residual class of links, which allowed the specification of any relationship where the relationship (usually a verb relationship) itself was a concept. These different kinds of links could be embedded to any degree of depth, so that the format was designed to be flexible enough to express anything, however vague or specific, that can be expressed in natural language.

The search in memory between concepts involves tracing out in parallel (simulated in the computer by a breadth-first search) along the links from the node of each concept specified by the input words. The words might be part of a sentence or stimuli in an experimental task. The spread of activation constantly expands first to all the nodes linked to the first node, then to all the nodes linked to each of these nodes, etc. At each node reached in this process, an activation tag is left which specifies the starting node and the immediate predecessor. When a tag from another starting node is encountered, then an _inter-section_ between the two nodes has been found. By following the tags back to both starting nodes, the _path_ that led to the intersection can be reconstructed.

When an intersection has been found, it is necessary to _evaluate_ the path to decide if it satisfies the constraints imposed by syntax and context. The complicated kinds of decision rules that are invoked in this evaluation phase are described for comprehension of sentences in Quillian (1969) and for categorization tasks by Collins & Quillian (1972a). To give an example here, in a phrase like "the fall leaves" a path found between the verb concept "fall" and the noun concept "leaf" would be rejected because syntax requires a participial form of "fall" to fit that interpretation. If the path found is rejected, the search continues from where it left off.

Priming (or psychological set) involves the same tracing process that was described for memory search. When a concept is primed, activation tags are spread by tracing an expanding set of links in the network out to some unspecified depth. When another concept is present subsequently, it only has to make contact with one of the tags left earlier to find an intersection.

3

14

Common Misinterpretations of Quillian's Theory

There is a rich variety of misinterpretations of Quillian's theory around; many of them derive from reading Collins' (Collins & Quillian, 1969, 1970) simplifications of the theory. Fodor has been heard to say that the trouble with Quillian's theory is that it cannot be pinned down well enough to refute it. This is true in that the theory was designed to specify how to build a semantic processor, and not to specify predictions about experiments on semantic processing. So it turns out that many of the misinterpretations are created by experimenters when they try to pin down the theory in order to refute it.

One misinterpretation of the theory is that concepts consist of words or definitions. Originally properties were defined in terms of words, just as in a dictionary, but these were later replaced by pointers to other concepts. Most of the concepts have words attached as names, but they need not have. Quillian regarded words as pointers into the concept network, not as part of the network. The information that was initially coded into his network was dictionary information, but not because definitions or defining properties were thought to be the only content of concepts. They were regarded as the tip of the iceberg. In fact, the presumption was that anything one learns or infers is stored. Thus, a realistic network would have an enormous number of links into and out of most concept nodes.

Another assumption sometimes made about Quillian's theory is that all links are equal (Wilkins, 1971; Rips, Shoben & Smith, 1972). In Quillian's original theory, there were criteriality tags on links. In Collins & Quillian (1969, 1972a) links were assumed to have differential accessability or travel

4

times. The criteriality of a property derives from the frequency
of that property occurring with the concept. Hence, having lungs
is more criterial for being a human than having feet, because
there are more humans without feet than without lungs. Having
feet. in turn is more criterial than having warts. Accessability
or travel time is not the same thing, though it is highly cor-
related. The accessability of a property depends on how often a
person thinks about or uses a property of a concept, and hence
having feet may be a more accessible property of humans than
having lungs, even though it is less criterial. But it would be
absurb to argue that, even though lungs, feet, and warts are all
linked directly to the concept human, that these links are in any
sense equal. The same is true for the links between bird and
robin, chicken, or penguin.

Rips, Shoben, & Smith suggest that intermediate nodes are
necessary for a network model to explain the RT differences they
find in categorizing different birds. This makes the mistaken
assumption that all links are equal in any network model. In
fact, there is no reason to assume that all links are equally
criterial or accessible. It turns out that differences in ac-
cessibility are crucial to many different aspects of human
semantic processing, as Carbonell & Collins (1973) point out in
their discussion of importance (or accessibility) tags.

A related implication of the Rips, Shoben, & Smith paper
and also a more recent paper of Smith, Shoben, & Rips (in press)
is that feature models can account for data that network models
cannot. What is strange about this argument is that network
models were designed merely as a method of representing features
in a computer. Any process that can be represented in a feature

5

model should be representable in a network model; in particular,
the Smith, Shoben, & Rips model itself could be implemented in a
semantic network. If anything, network models are more powerful
than feature models, because it is not obvious how to handle
inferential processing or embedding in feature models.

Smith, Shoben, & Rips argued in favor of feature models,
because their data for comparison of concepts seemed to fit a
feature comparison process. What should be emphasized about
Quillian's theory is that the parallel search would inevitably
lead to just such a feature comparison process, though the pro-
cess would take place over a period of time as different connec-
tions are found. One way that Quillian's theory is different
from the Smith, Shoben, & Rips' model is that superordinate con-
nections, if they exist, would also be found and evaluated. That
is to say, that if a person learns that a chicken is a bird, or
that a bat is not a bird, in Quillian's theory this information
would be stored and used in the task of deciding whether either
a chicken or a bat is a bird. In Smith, Shoben, & Rips' theory
it would not be used. The distinction between these two theories
is so crucial that we will discuss it at length in conjunction
with the spreading activation theory's explanation of the Rips,
Shoben, & Smith results.

Perhaps the most prevalent misinterpretation of Quillian's
theory concerns the idea of cognitive economy. In this regard,
it is important to distinguish the strong theory of cognitive
economy, which Conrad (1972) takes issue with in her attack on
Collins & Quillian (1969) and the weak theory of cognitive
economy, which Quillian believed. As Conrad states it, she
rejects the "hypothesis that all properties are stored only once

6

in memory and must be retrieved through a series of inferences
for all words except those that they most directly define." This
is a statement of the strong theory of cognitive economy. But
Quillian never believed this was true, and Collins & Quillian
(1969) cautioned against making that interpretation of the theory.
Such a theory requires erasing information whenever it applies at
a more general level. If a person learns a robin can fly, and
then later that birds fly, the strong theory implies that it must
be erased from robin. The weak theory of cognitive economy
merely assumes that every time one learns that X is a bird, one
does not at that time store all the properties of birds with X
in memory. Thus, an inference will be necessary to decide that
X can fly, unless one encounters this fact directly. Hence,
Collins & Quillian, who were testing the weak theory of cognitive
economy, picked instances where people were not likely to have
encountered the general property with the specific instance (e.g.,
"A wren can fly.") The point of the experiment was to test
whether it was possible to measure inference time, when the weak
theory of cognitive economy implies that an inference is necessary.

Another common misconception of Quillian's theory shows up
in Juola & Atkinson's (1971) work on categorization. They as-
sume that in Quillian's theory, the search proceeds from the
instance to the category. To the contrary, Quillian's theory
assumes the search proceeds from both in parallel. However, if
one or the other is presented first, this gives the search from
that node a head start, which is the notion of priming. Their
experiment involves priming in a complicated way, which we will
discuss in the third section. But the predictions they ascribe
to Collins & Quillian are based on a misunderstanding of
Quillian's theory.

Anderson (in press) rejects a Quillian-like model of a
parallel search on the basis of his data, while acknowledging
that his data are compatible with "a parallel model whose search
rate is slower in proportion to the number of paths searched."
Anderson's argument implies wrongly that Quillian has made the
independence assumption for this parallel search. An independent
parallel search is like a race where the speed of each runner is
independent of the other runners. This is a common assumption in
psychology, because it makes it possible to assign an upper bound
to reaction time (see Sternberg, 1966). But everything we know
about physiology, as for example the prevelance of lateral inhibi-
tion in the CNS, makes the independence assumption unlikely for
any parallel process in humans. Furthermore, there is no dif-
ficulty for Quillian's theory if the parallel search rate depends
on the number of paths searched. Hence, Anderson's data are
perfectly compatible with Quillian's parallel search.

So that is what Quillian's theory is not, or at least some
of what it's not. Several other misconceptions were discussed
in Collins & Quillian (1972), in particular the notion that
Quillian's theory of memory is rigidly hierarchical, and
Schaeffer and Wallace's (1970) argument that Quillian's theory
predicts it will always take less time to compare concepts that
are close together in the semantic network. We will return to
many of these same papers in the third section in order to
describe how the extended version of Quillian's theory would
account for some of the results these experimenters have used to
reject Quillian's theory.

8

THE EXTENDED THEORY

In order to deal with the specific experimental results that have appeared in recent years, several more processing and structural assumptions must be added to the basic Quillian theory. These do not bend the theory, but merely elaborate it in a way that it can be applied to the kinds of experiments on semantic memory that have been performed recently. The elaboration may be wrong and the basic theory right, so that our mistakes should not be held against Quillian's theory.

Additional Processing Assumptions

There are four additional processing assumptions in the extended theory. These four assumptions transform the theory from computer terms to quasi-neurological terms, a la Pavlov. But all the assumptions of the original theory should be preserved despite the transformation, except that activation tags are to be considered as source-specific activation.

1) When a concept is processed (or stimulated), activation spreads out along the paths of the network in a decreasing gradient. Thus, activation is like a signal from a source that is attenuated as it travels outward.

2) The longer a concept is processed (either by reading or hearing it, or by thinking about it), the longer activation is released from the node of the concept at a fixed rate.

9

3) Activation decays over time and/or intervening
activity. This is a non-committal assumption that
activation goes away gradually by some mechanism.
Assumptions 2 and 3 impose a limitation on the amount
of activation that can be allocated in priming more
than one concept, because the more concepts that are
primed, the less each will be primed.

4) With the assumption that activation is a variable
quanitity, the notion of intersection requires a
threshold for firing. The assumption is that activa-
tion from different sources summates, and that when
the .ummation at the point of intersection reaches
threshold, the path in the network producing the
intersection will be evaluated.

Additional Structural Assumptions

There are two additional structural assumptions in the ex-
tended theory. These are generalizations of Loftus's (1973b)
arguments that semantic memory is organized primarily into noun
categories and that there is a dictionary (or lexicon) separate
from the conceptual network.

(5) The conceptual (semantic) network is organized
along the lines of semantic similarity. The more
properties two concepts have in common, the more
links there are between the two nodes via these pro-
perties, and the more closely related are the concepts.
This means that different vehicles, and different
colors, and different verbs of motion will all be

highly interlinked with each other through their com-
mon properties. This also implies that white things
(e.g., chalk, cauliflower, doves, and clouds) and
things that move (e.g., clouds, fire engines, robots,
and gazelles) are not closely interlinked, despite
the one property they have in common. In these
terms semantic relatedness is based on an aggregate
of the interconnections between two concepts.[2]

It follows from this assumption, together with
earlier assumptions, that if you prime vehicles,
the activation at any type of vehicle will accumulate
from many neighboring nodes, because of the many
interconnections between different vehicles. In
other words, to the degree fire engine is primed
it will prime ambulance and police car, etc., and
each of these in turn will prime each other. But
if you prime white things, clouds will not prime
chalk, or cauliflower, or doves, to any great ex-
tent, because of the paucity of interconnections
between these concepts. Instead, clouds will tend
to prime sky and rain, etc. Hence the same amount
of activation will be diffused among a greater
number of concepts.

(6) The names of concepts are stored in a lexical
network (or "dictionary") which is organized along
lines of phonemic (and to some degree orthographic)
similarity. The links from each node in the lexical
network are the phonemic properties of the name,
specified with respect to their position in the word.
The properties stored about names are assumed to be

11

22

the properties Brown & McNeill (1966) found people
could identify about words on the "tip of their
tongue." Each "name" node in the lexical network
must be connected to one or more concept nodes in
the semantic network.

Loftus's (1973b) data lead to the further assumption
that a person can control whether he primes the lexical
network or the semantic network, or both. For example,
a person can control whether to prime (a) words in the
lexical network that sound like bird, or (b) concepts
in the semantic network related to bird, or (c) words
in the lexical network corresponding to the concepts
in (b). This control over priming can be thought of
in terms of summation of diffuse activation for an
entire network (perhaps in a particular part of the
brain) and source-specific activation released from a
particular node. Thus, (a) would derive from activa-
tion of the lexical network together with the word
bird, (b) would derive from activation of the semantic
network together with the concept bird, and (c) would
derive from activation of both networks together with
the concept bird. .

RECENT EXPERIMENTS

In this section, we will discuss how the theory deals with
the different kinds of experiments that have been done recently.
The five types of studies we want to apply the theory to are:
(1) several experiments by Loftus (Freedman & Loftus. 1971;
Loftus, 1973a, 1973b); (2) Collins & Quillian's (1971) and

Juola & Atkinson's (1971) categorization tasks, (3) Collins &
Quillian (1970a) sentence-priming experiments, (4) the Conrad
(1972) sentence-verification experiment, and (5) the Rips,
Shoben, & Smith (1973), the Smith, Shoben, & Rips (in press)
and the Rosch (1973) categorization experiments, where the
effects of typicality were investigated. This is not meant to
exhaustively cover the literature, but to deal with the major
kinds of findings available that the Quillian theory has not
been applied to.

The Loftus Experiments

There are three sets of Loftus experiments we want to dis-
cuss in terms of the spreading activation theory. The first of
these is an experiment by Freedman & Loftus (1971) where sub-
jects had to name an instance of a category which began with a
given letter or which were characterized by a given adjective.
For example, they might be asked to name a fruit that begins
with "A" or a fruit that is red. On some trials the category
was shown first and on some trials second. Reaction time (RT)
was measured from the onset of the second stimulus.

The result from this first experiment that concerns us
here is that subjects were faster when the category (e.g., fruit)
was given first, then when either the letter or the adjective
were given first. This basic result was later replicated with
adjectives even when the adjective was a closer associate to
the instance named than the category noun (e.g., sour is a
closer associate of lemon than is fruit). This experiment is
a priming experiment in that one concept is activated before the
other.

13

The explanation in terms of the theory is this. When a
noun, such as fruit is presented first, the activation spreads to
nodes connected to fruit, among which are instances such as apple
pear, peach, orange, lemon, etc. But these concepts are all
highly interlinked with each other (though some, such as orange
and lemon, are more closely interlinked than others). Thus, the
total amount of activation is spread among a relatively small
number of closely interlinked concepts. However, when an adjec-
tive or letter is presented first, say red or A, the activation
spreads to a much wider set of concepts, which are not particularly
interlinked with each other. (See Assumption 5 in the Extended
Theory.) To the degree an adjective such as red, leads to priming
a set of interlinked concepts, it is the various color concepts,
such as orange and pink. But the task requires the subject to
respond with an instance of the category, and these will receive
relatively little priming from an adjective or letter. Because
priming the noun leads to a greater accumulation of activation
on the instances, these are closer to their threshold for firing,
so that it takes less stimulation, and hence less time, to trigger
an intersection when the second stimulus is presented.

Freedman and Loftus explain their finding in terms of en-
tering the category when a noun is presented, and entering a
cluster within the category when the adjective or letter is
presented. On this theory, if the noun is presented first, the
subject can enter the category immediately, and need only choose
the correct cluster when the adjective or letter is presented.
But if the adjective or letter is presented first, the subject
must wait until the category is presented, because the cluster
is specific to the category. (However, Loftus (1973b) has re-
vised this explanation for the letter stimulus in her dictionary-
network model.)

The Freedman and Loftus explanation is not altogether
different from the explanation offered here. Entering a category,
translated into spreading activation theory terminology, is
stimulating the node of the category. Also, the structural as-
sumptions in their explanation are akin to the structural
assumptions made here, though our theory is less rigidly hier-
archical. The rigid hierarchy gets into trouble with errors
such as one we encountered where a subject named Ben Franklin,
given the pair "president" and "F." He recognized his mistake
afterwards, so the problem was not that Ben Franklin was thought
to be an instance of the category "president." But in an activa-
tion theory, Ben Franklin is a very likely intersection starting
at "president" and "F," because he is so closely linked with the
concept "president" and some of its foremost instances, such as
Washington. It is a general problem of category search models
that they cannot deal with such errors.

Perhaps the major advantage of the spreading activation
theory over the Freedman and Loftus explanation is that it ties
this result to a parallel result in a quite different experiment
by Loftus (1973a). In a categorization experiment, Loftus found
that $\underline{S}$s are faster in deciding that a shrimp is a seafood, for
example, if the category is presented first and the instance
second rather than in the opposite order. But the opposite order
is better when the category is insect and the instance is butter-
fly, for example. The difference between the two cases is
represented schematically in Figure 1. Shrimp is a relatively
high-frequency response in naming instances of seafoods, but
seafood is a low-frequency response in naming superordinates of
shrimp. Contrariwise, butterfly is a low-frequency response in
naming instances of insects, but insect is a high-frequency

15

response in naming superordinates of butterfly. (In Figure 1, the length of the arrow indicates the difference in accessability, so that more activation will be carried along a shorter arrow. The box indicates the stimulus which, if presented first, produces the shorter RT.) The spreading activation theory explains this asymmetry in RT, in the following way: when seafood or butterfly are presented first, more activation spreads to the concept presented second, and thus it takes _less_ time to reach the threshold for an intersection. In the case where shrimp or insect are presented first, the second concept is primed less, and it takes longer to reach threshold.

By comparing the diagrams for this experiment with the diagram for the Freedman and Loftus study, it can be seen that the two results are exactly parallel. Based on our structural assumptions, fruit primes apple more than does red. Hence, the shorter RT occurs when fruit is presented first. Similarly, because seafood primes shrimp more than vice versa, and butterfly primes insect more than vice versa, the shorter RT occurs when seafood and butterfly are presented first. Some kind of spreading activation notion is quite compelling to account for the Loftus (1973a) results, and the theory offered here encompasses the order effect in both the Loftus study and the earlier Freedman and Loftus experiment within a single framework.

Recently Loftus (1973b) has found two different ways in which presenting a letter acts differently than an adjective in variations of the Freedman and Loftus paradigm. This has led her to develop a dictionary-network model, which we will translate into spreading activation terms. The first difference between presenting a letter and an adjective appeared when she

16

FREEDMAN AND LOFTUS (1971) EXPERIMENT
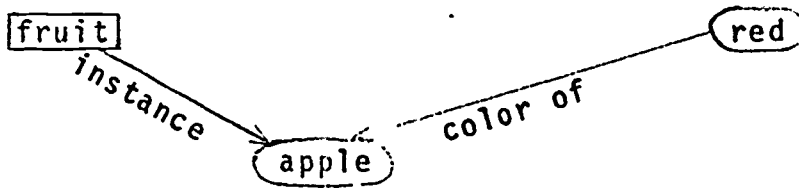
LOFTUS (1973a) EXPERIMENT

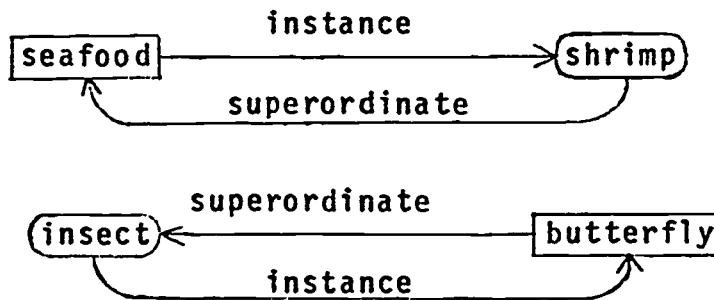

FIGURE 1.   SCHEMATIC REPRESENTATION OF SIMILARITY
             BETWEEN THE FREEDMAN & LOFTUS (1971)
             EXPERIMENT AND THE LOFTUS (1973a)
             EXPERIMENT (WHERE A SHORTER ARROW
             REPRESENTS A GREATER PRIMING EFFECT)

compared RT in two conditions: one where noun-adjective (e.g., fruit-red) and noun-letter (e.g., fruit-A) trials were randomly intermixed, and one where noun-adjective and noun-letter trials were separated into blocks. In all cases the noun preceded the adjective or letter. The results of this experiment are shown in Figure 2. It is clear that when the subject knows a letter is coming, he can prepare for it. But in the mixed condition, the subject apparently prepares for either kind of trial the same way he prepares for an adjective trial, since adjective trials take the same amount of time in either case. The theory's description of semantic processing on the adjective trials is the same as in the Freedman & Loftus experiment, with the amendment that only the semantic network and not the lexical network would be diffusely primed in this condition.

Loftus describes what must happen on noun-letter trials in the blocked condition as follows:

> "The first step of the process is entering the category.
> The next step is a quasi-parallel simultaneous search
> towards the Dictionary. That is to say, the subject
> traces some number of pathways leading from category
> instances to the Dictionary representations of those
> instances. This step can be started during the in-
> terval between the presentation of the category name
> and the restricting letter if the subject knows a
> letter is coming."

This is essentially the spreading activation explanation, if the dictionary is taken to be a lexical network. Rather than saying that "the subject traces some number of pathways," which suggests a conscious tracing process, the present theory would say that activation spreads along some number of pathways, because the subject has activated the lexical network in addition to the semantic network. (See assumption 6 in the Extended Theory.)
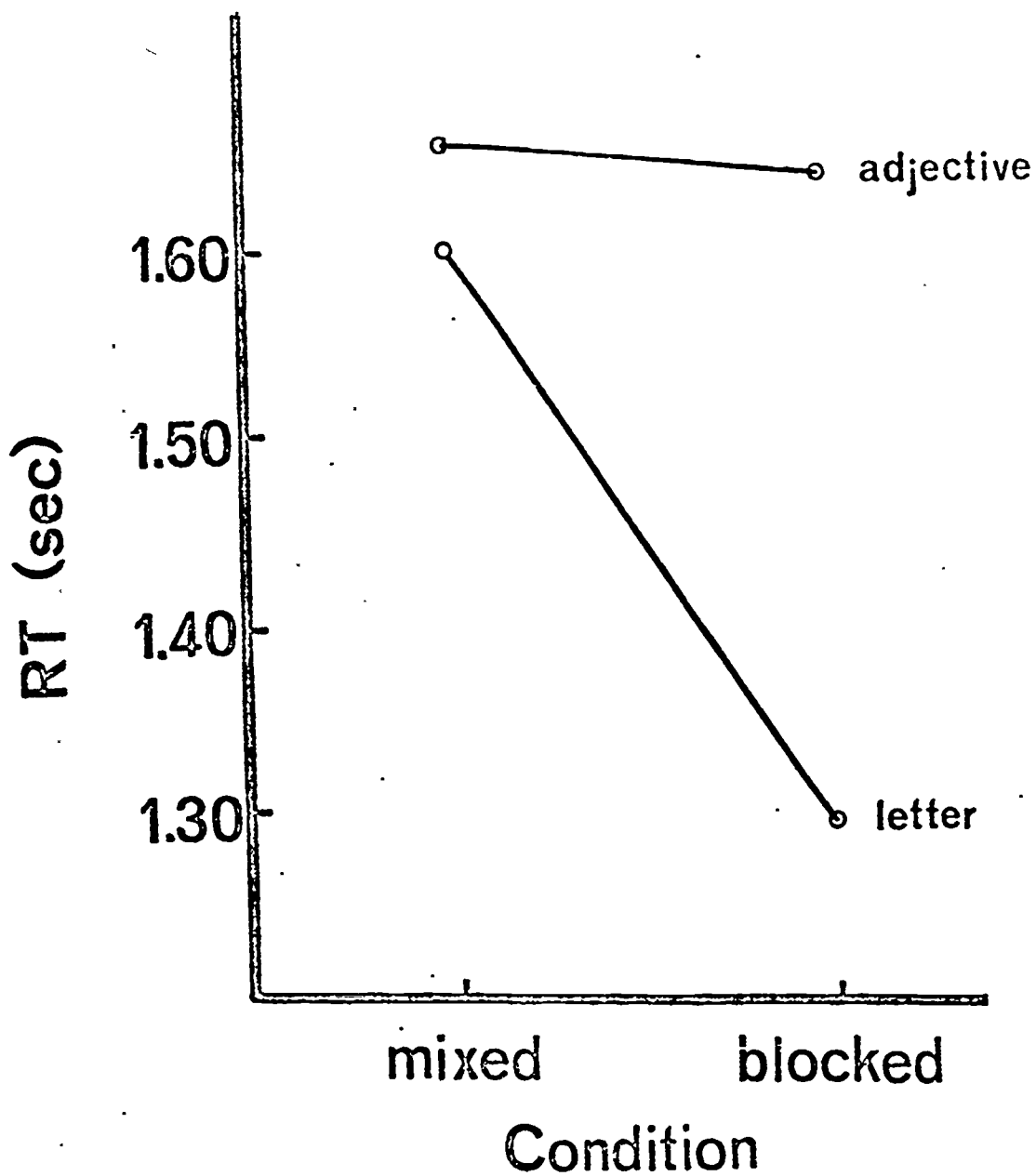
FIGURE 2.   REACTION TIME FOR NOUN-ADJECTIVE AND
            NOUN-LETTER STIMULI IN MIXED AND
            BLOCKED CONDITIONS

Hence, in the present explanation, the subject's control is re-
duced to diffusely activating whole networks rather than specific
pathways (in addition to the specific nodes activated by the
stimuli in the experiment). The difference in the results for
the noun-letter trials in the two conditions then depends on
whether the subject primes both networks (as in the blocked
condition) or only the semantic network (as in the mixed con-
dition). The reason he only activates the semantic network in
the mixed condition may either be because of a principle of
least effort (hence he could speed up his RT if he tried), or
becuase there is less activation available to the semantic network
if both are primed (hence he will be slower on noun-adjective
trials if he primes both).

As can be seen in Figure 2, the subject is much slower on
noun-adjective trials than on noun-letter trials in the blocked
condition. This is probably accounted for by the fact that an
intersection on a noun-adjective trial occurs in the semantic
network, and requires the further step of retrieving the corres-
ponding name in the lexical network. On the other hand, the
intersection on a noun-letter trial occurs at the name in the
lexical network. Therefore, the name does not then need to be
retrieved.

The second result that shows up the difference between ad-
jectives and letters, was predicted by Loftus from the dictionary-
network model. In this experiment she presented three stimuli,
either in the order noun, adjective, and letter, or in the order
noun, letter, and adjective. For example, the three stimuli
might be animal, small, and M, for which an appropriate response
is mouse. The prediction was that the subject should be faster

20

when the adjective is presented before the letter, and this was
the result found. The reasoning is as follows: when the adjective
appears first, activation will spread from a small set of in-
stances (maybe one) in the semantic network to the lexical
network, since a letter can be expected just as in the blocked
condition. When a letter is presented first, activation will
spread from a small set of instances in the lexical network
back to the semantic network where an intersection with the
adjective presented will occur. Then the subject must return
again to the lexical network to retrieve the name, so there is
an extra transit necessary in this condition.

There are other expeiments of Loftus (Loftus, 1972; Loftus
& Loftus, in press) involving repetition of the same category on
several trials in the Freedman & Loftus paradigm. These have
also required an explanation in terms of spreading activation
that decays gradually over trials (or time). We will not repeat
Loftus's arguments here, though we do want to emphasize the
widely different effects that a spreading activation theory can
encompass.

Two Categorization Experiments

Collins & Quillian (1971) looked at how RT changes over
trials in a categorization task, when the category remained con-
stant for 12 trials in a row.[3] They constructed four different
kinds of lists in each of two domains, "birds" and "mammals."
For our purposes here, we ignore the difference in results between
birds and mammals (which is discussed in Collins & Quillian,
1972a), and discuss the commonalities in terms of the bird
domain. The four kinds of lists for the bird domain were as

21

follows: (1) For the category "bird" there was a narrow list
consisting of six birds and six non-animals. (2) For the
category "bird" there was a wide list consisting of six birds,
three non-bird animals, and three non-animals. (3) For the
category "animal" there was a narrow list, consisting of six
birds and six non-animals (exactly the same as for 1). (4) For
the category "animal" there was a wide list consisting of three
birds, three non-bird animals, and six non-animals. In Figure 3
are shown the RTs for the same bird names (or narrow instances)
occurring in the same positions in the four different kinds of
lists. The same description applies to the mammal lists, with
the substitution of mammal for bird and non-mammal for non-bird
in the description. The results clearly show that subjects were
faster on the bird and mammal names in narrow lists than they
were in wide lists, after some exposure to the list.

The experiment was designed to test the implication from
spreading activation theory that the instances in a narrow list
would all tend to prime the same concepts, whereas in wide lists
the same amount of activation would be spread over a wider range
of concepts. Thus, because the instances in a wide list are less
interconnected with each other, there would be less accumulation
of activation at any one concept. But this explanation doesn't
hold when the category was "bird" (or "mammal"), because the non-
bird animals added to the wide list did not replace any bird
names, but merely replaced some non-animals.

The fact that subjects were faster on narrow lists for the
category "bird" (or "mammal") can be explained, however, in
terms of pattern matching on features. Put in pattern-matching
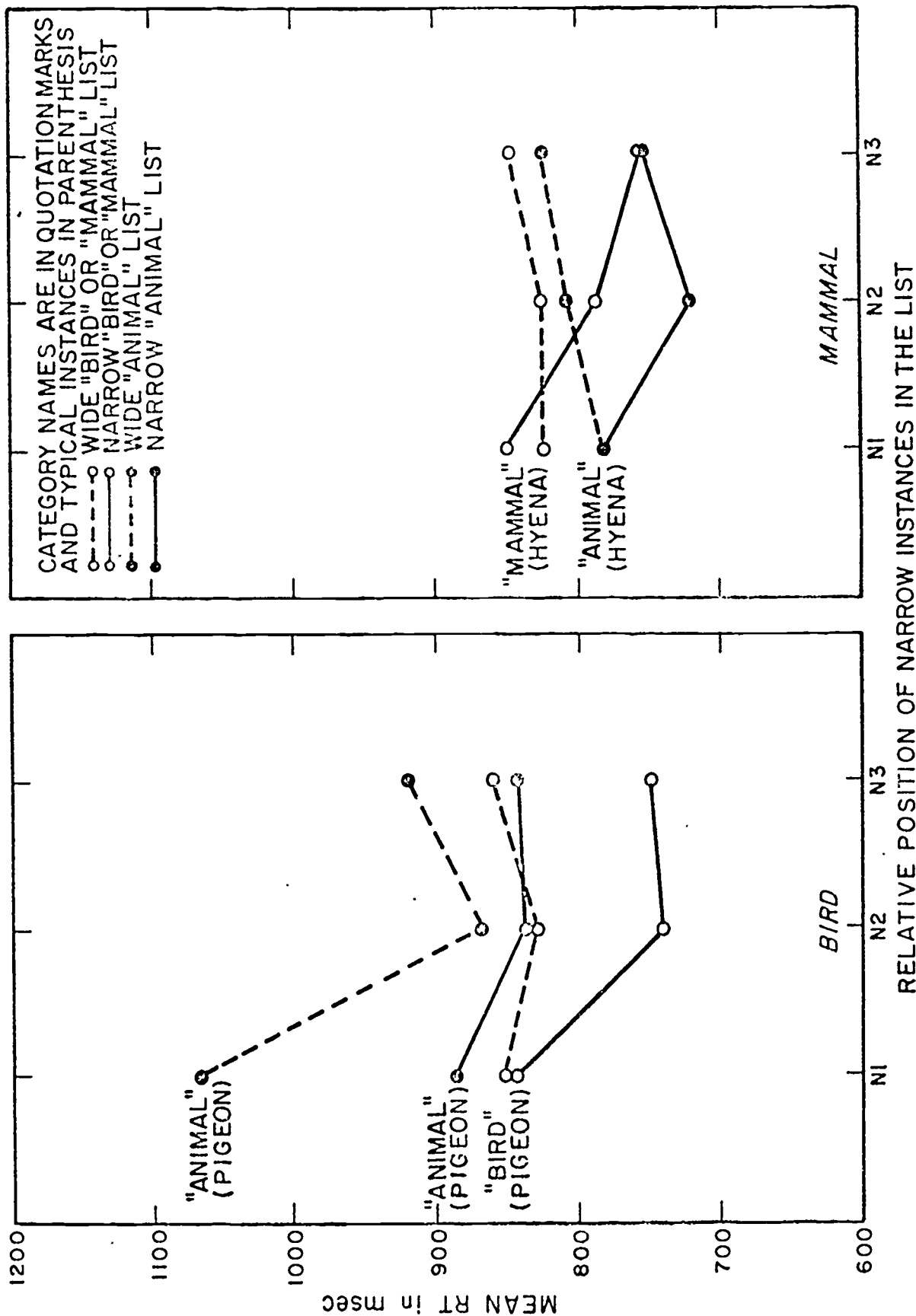terms, the set of "Yes" instances and the set of "No" instances

22

FIGURE 3.   AVERAGE REACTION TIMES FOR NARROW INSTANCES AT DIFFERENT
             RELATIVE POSITIONS IN NARROW AND WIDE LISTS

23

in narrow lists are easily discriminable; e.g., all birds on one
hand and all non-animals on the other. But in wide lists a
class of not so discriminable instances was added to the "No"
instances, in particular other animal names were added which are
hard to discriminate from birds. Thus, the RT for responding
"Yes" to a bird name would increase, because more features must
be compared to make the discrimination (or in our terms the sub-
ject must set a higher threshold for responding "Yes"). This is
the kind of result Smith, Shoben, & Rips (in press) model was
designed to deal with, though they have not developed the kind
of variable threshold necessary to handle this particular result.

The same explanation will not work very well though for the
category "animal," because the birds and non-bird animals in the
wide lists were all "Yes" responses. Birds are no more discrim-
able from non-animals than are non-bird animals, such as mammals.
But in this case, the priming explanation given above works
quite well. (For "Yes" responses there were six birds in the
narrow list and three birds and three non-bird animals in the
wide list. Thus, activation would be more thinly spread in a
wide list.) We will return to the role of pattern matching in
the theory when discussing the Smith, Shoben & Rips model. At
this point, we only want to argue that these results require both
a pattern-matching explanation and priming explanation.

Spreading activation theory also implies that there should
be an accumulation of activation from stimulating the same cate-
gory on repeated trials. Other evidence from the experiment
suggested there was such an effect, but a question arises in that
case as to why there was no decrease for narrow instances in a
wide list (see the Figure 3).

24

First, notice in Figure 3 that subjects did become faster
on bird names in the wide list when the category was "animal."
(In the narrow list the RT was faster on trial N1 because the
subject already had seen one bird name.) The probable reason
is that subjects primed two categories in the wide list, bird
and animal, rather than making the inference each time that a
bird is an animal. As a general strategy then, subjects may have
been activating multiple categories to deal with wide lists.
The priming of categories other than the one given can account
for the sharp decreases in RT of about 150 msec that Collins &
Quillian found for wide instances in a wide list. But what
happens to narrow instances in a wide list if the subject primes
multiple categories?

If other categories are primed, the increased number of
categories should act to slow down RT to any one category ac-
cording to the theory (in particular the given category which
applies to the narrow instances). This is because less activa-
tion will spread out from each category (see assumption 3 in
the Extended Theory) and with less activation it will take
longer to reach threshold. Hence, this slowdown will act to
offset the speedup derived from repeating the category twelve
times.

An increase in RT with multiple categories has been found
by Juola and Atkinson (1971) in a task where subjects had to
decide whether a stimulus word belonged to one of a variable
number (1 to 4) of prespecified (target) categories. They com-
pared this task with a task where subjects decided if the
stimulus word was the same as one of a variable number (1 to
4) of target words. Their experiment was designed to distin-
guish between two kinds of models, one they attribute to

25

Landauer & Freedman (1968) and one they attribute to Collins &
Quillian (1970b). In most respects, their results fit the model
they derived from Landauer & Freedman, but since the spreading
activation theory provides an alternative explanation for their
results we want to compare their two models with the theory.

The model Juola & Atkinson derived from Landauer & Freedman
is very similar to what Landauer & Meyer (1972) call the "category
search model." It assumes the subject searches through instances
of the categories in memory seeking a match for the stimulus word.
Such a model predicts that as the number of categories or words
in the memory set increases, RT for the category-matching task
should increase at a greater slope than RT for the word-matching
task. This is because each additional target category adds more
instances that must be searched, whereas each additional target
word only adds one, the word itself. This result was essentially
what Juola & Atkinson found.

The other model assumed that subjects perform the category-
matching task by retrieving their stored category for the stimulus
word, and comparing this to the given categories to see if it
matches one of them. This model would predict that the slope for
the two tasks should be about the same, and the intercept for the
category-matching task should be greater than for the word-matching
task. Their results clearly reject this model. Though attributed
to Collins & Quillian, this model is quite different from Quillian's
theory, because the semantic search in Quillian's (1966) theory is
assumed to spread in parallel from both categories and instances.
When the categories are given first as Atkinson & Juola did, then
activation would spread out from the categories before the instance
even appeared.

26

The spreading activation theory explains the increase in RT
as the number of categories increases in terms of the limited amount
of activation that can be allocated (see assumption 3 in the
Extended Theory). As the number of categories increases, the
amount of activation that spreads out from any one category decreases.
Thus, the longer on the average it will take to trigger an inter-
section with the correct category. In the word-matching task, each
of the words given as targets will be directly activated; and with
so few words, they will all be quite close to threshold. So vary-
ing the number of target words in the word-matching task should not
have as much effect as varying the number of target categories in
the category-matching task. Furthermore, as Juola & Atkinson point
out, there are two aspects of their data (namely, the nonlinearites
for positive responses and the recency effects in the serial
position curves) that fit much better with a parallel model, such
as spreading activation theory, than they do with a serial model,
like the one they derive from Landauer & Freedman.

Loftus (1973a) points out that the category search model has
real difficulty in handling the asymmetry in RT she found between
categorizing a shrimp as a seafood, for example, and a butterfly
as an insect. (See discussion of Figure 1.) If a person decides
whether an instance is a member of a category by scanning through
the names of the instances stored with that category, it should not
matter which order the category and instance are presented in. The
category-search model nicely predicts her results when the category
is presented first; but the other model that Atkinson & Juola
discuss handles Loftus's results when the instance is presented
first. Thus, a combined model that postulates one strategy when
the category is presented first and the other strategy when the
instance is presented is a possible alternative to explain both
the Juola & Atkinson and the Loftus results. But it is not clear
with such a model why categorization time should depend on the con-
text of the list as shown in Figure 3.

27

Sentence Priming Studies

Collins & Quillian (1970a) first studied the implications of
spreading activation theory by measuring reaction time (RT) for
subjects to decide whether a stimulus sentence is true or false.
In this particular experiment they presented two related sentences
on succeeding trials, and looked at the effect of the first sentence
on RT to the second sentence. For instance, a sentence like "A
canary is a bird" might be preceded by either "A canary can fly"
or "A canary is yellow" for different subjects.   Because flying
is a general property of birds and b.ing yellow is specific to
canaries, the theory predicts that the decision about whether "A
canary is a bird" would be primed more by seeing "A canary can fly"
than by seeing "A canary is yellow." This is because activation
spreading out from both canary and flying would prime bird if the
first sentence is "A canary can fly," whereas only activation spreading
out from "canary" would prime bird if the first sentence is "A canary
is yellow." In fact, there were 12 such predictions for the various
kinds of sentence pairs presented, and all 12 were confirmed when a
blobal correction was made for differences between sentence types
(without the correction, 11 of 12 were confirmed).

One later, more exploratory study using the same method,
proved less successful, because most of the RT differences were too
small to be significant. But the one result that came out bears
on the various models we have been discussing.

One condition of the experiment compared how long it took
subjects to decide that two similar things were different, if the
preceding sentence refers u to a distinguishing (D) property, a
common (C) property, or an irrelevant (I) property. For example,
consider the case where the second sentence was "Sugar is salt."
The sentence "Sugar is sweet" refers to a distinguishing property,
since salt is not sweet; the sentence "Sugar is white" refers to

a common property, since both sugar and salt are white; and the
sentence "Sugar is refined" refers to a property irrelevant for
deciding whether sugar is salt, since it was assumed most people
do not know whether salt is refined. Because the theory, like the
Smith, Shoben & Rips (in press) model assumes that comparing two
concepts like sugar and salt, involves comparison of properties
(see Collins & Quillian, 1972a), the prediction was that priming
a distinguishing property would speed up the distinction and
priming a common property would slow down the distinction, in
both cases with respect to the irrelevant property. This predicted
order for three types of sentence pairs was significant using a
Page (1963) test for linear ranks (L(48) = 599, p <.01). This
result confirms the notion in the theories of Shaeffer & Wallace
(1970), Collins & Quillian (1972a), and Smith, Shoben & Rips (in
press) that concept comparison involves some kind of feature com-
parison, and argues against the category-search model or the hybrid
model suggested in the last section.

The Conrad Study

Using a similar true/false RT technique for sentences, Conrad
(1972) found results which she interpreted against Collins &
Quillian's (1969) theory of semantic processing. Aside from the
misinterpretation of Quillian's theory mentioned earlier, the
results of her study are quite close to what Quillian's theory would
predict given Conrad's methodology.

In her first experiment, which was like the Collins & Quillian
(1969) study, she selected 2-level and 3-level hierarchies from the
common culture (e.g., shark —> fish —> animal) and properties
associated with the objects at different levels. Then she con-
structed sentences with instances, such as shark, from the lowest

level and properties from all three levels. The results Collins
& Quillian found were that RT increased as the property was farther
removed from the instance in the hierarchy. The reason for the
increases in RT according to spreading activation theory is that
as the instance and property are farther apart in the hierarchy,
it takes activation longer to spread between them and trigger an
intersection (and perhaps to evaluate the path found as well).

Unlike Collins & Quillian, Conrad broke down the properties
in her sentences into three groups on the basis of the frequency
(high, medium, low) with which people generated each property,
given the different objects in the hierarchies. Another difference
from Collins & Quillian's study is that she collected data over
5 days by repeating all the sentences each day.

Her results in the first expriment generally replicated the
increases in RT that Collins & Quillian found as the property was
farther removed in the hierarchy from the instance. However, the
increases she found were much smaller on the average than those
Collins & Quillian found. (There was one reversal in her data for
the high-frequency properties, but as she noted that reversal did
not occur when the 2-level and 3-level hierarchies were looked at
separately.) Because people store a property with whatever
instance it is linked to in a sentence (given the weak theory of
cognitive economy), Conrad's repetition of sentences over 5 days
should lead to the smaller RT increases she found. This is because
an inference necessary on the first day would be less likely on
the second day, etc. Conrad, in fact, reports a large level-by-day
interaction.

In general, she found that the higher the frequency of the
property, the smaller the increases between levels. Though not a

30

specific assumption of the theory, it is quite natural to assume that it is high-frequency properties that are more likely to be stored at several levels in the hierarchy. For example, leaves are more likely to be stored as a property with particular types of trees than is bark, because leaves are a higher-frequency property. So her results in the first experiment fit the kind of theory presented here quite nicely.

It was her second experiment that looks damaging for Quillian's theory, but here she made a crucial methodological change that she probably did not realize the implications of. The change was that she presented the subject of the sentence a second before the predicate, and this turned the experiment into a priming study. In the study she presented only properties true of the highest level nodes, together with objects at different levels in the hier- archy. Therefore, she predicted from Quillian's theory that the lower-level objects would take longer to confirm, since it would take activation longer to spread between lower-level objects and higher-level properties. But by presenting the object a second before the property, activation would spread out from the object to its superordinates. Hence, with this method, no clear increase is predicted; and by using only high-level properties she insured that Ss would prepare as best they could by priming the superordinates. In any case, this particular experiment has real methodological problems as a test of Quillian's theory, and it is less strong evidence <u>against</u> spreading activation theory than her first experiment is evidence <u>for</u> the theory.

Typicality and the Smith, Shoben & Rips Model

In recent experiments, Rips, Shoben & Smith (1973); Shoben
& Rips (in press); and Rosch (1973) have shown that reaction
time in a categorization task corresponds very closely to sub-
ject's ratings of how typical the instance is of the category.
For example, robins and sparrows are considered typical birds
whereas chickens and geese are not. The effect of typicality on
RT is quite large even when frequency of the particular instances
in the language is controlled for. Like Smith, Shoben, & Rips,
we would argue that the typicality effect is one more manifesta-
tion of the fact that semantic similarity speeds up positive
decisions and slows down negative decisions. Such an effect has
been found repeatedly (Collins & Quillian, 1969, 1970b, 1972b;
Shaeffer & Wallace, 1969, 1970,; Wilkins, 1971). While Landauer
& Meyer (1972) argue that similarity effects in negative deci-
sions are either questionable or artifactual, the evidence now
seems so overwhelming that any viable theory must account for
them. They are very damaging for the category search model or
the hybrid model suggested earlier.

Semantic similarity effects suggest that pattern matching
is involved in making semantic judgments, not simply in cate-
gorization tasks, but in all of semantic processing (Collins &
Quillian, 1972a). Schaeffer & Wallace (1970) have suggested
one kind of pattern-matching model, but it is not clear how it
would work in other than the specific task they used. Collins
& Quillian (1972a) have argued for a hybrid pattern-matching and
path-evaluation theory based on Quillian's theory. Their theory
is quite complicated and not all the decision strategies are
well specified. Recently Smith, Shoben & Rips (in press) have

32

suggested a two-stage pattern-matching model, based on their
categorization experiments. This model has the virtue that it is
quite clear and explicit. It is not complicated like the Collins &
Quillian theory, and it agrees quite well with the reaction time
data for categorization judgments. Because it is such a viable
model, we want to emphasize how it differs from Quillian's theory
and point out what we think are its inherent difficulties.

We will briefly describe the Smith, Shoben, & Rips model
and then enumerate its differences from the Collins and Quillian
theory, which we will try to specify more precisely. In the
Smith, Shoben & Rips model, the meaning of a concept is assumed
to be represented by semantic features of two kinds: defining
features and characteristic features. Defining features are
those that an instance must have to be a member of the concept;
and the model assumes that features can be more or less defining.
Characteristic features are those that are commonly associated
with the concept, but are not necessary for concept membership.
For example, wings might be a defining feature of birds and
flying a characteristic feature, since all birds have wings but
not all fly. In a categorization task, the model assumes that
the two concepts are first compared in Stage 1 with respect to
all their features, both characteristic and defining. If the
match is above a positive threshold, the subject answers "Yes";
if it is below a negative threshold, the subject answers "No";
and if it is in between, the subject makes a second comparison
in Stage 2 based on just the defining features. If the instance
has all the defining features of the category, the subject says
"Yes" and otherwise says "No." If the subject can decide on
the basis of Stage 1, his RT will be faster than if he decides
on the basis of Stage 2.

33

There are several minor differences between the model and
the spreading of activation theory that could be patched over by
changing their model slightly. The difference in wording between
comparing features in their model and finding links between pro-
perties in our theory is really a non-difference. But the
distinction between defining and characteristic features has
the inherent difficulty, pointed out through the ages, that
there is no feature that is absolutely necessary for any
category.[4] For example, if you remove the wings from a bird, it
does not stop being a bird. Furthermore, people have no ability
to make such a distinction consistently, either from time to
time or from one person to another. Smith, Shoben, & Rips
recognized that features are more or less defining (or criterial),
but they were forced into making the artificial distinction
between defining and characteristic in order to have a two-stage
model. But the model could be revised to work without the two
stages and make essentially the same reaction time predictions.

The revision is this: if features are compared over time,
as in Quillian's theory, then the longer the process goes on,
the more features that will be compared (assuming features have
different accessibilities). The comparison process can have a
positive threshold and a negative threshold just as before, and
features can be weighted by their criteriality. If the match
at any point in time is above threshold, then the subject says
"yes"; if the match falls below the negative threshold the
subject says "no"; and otherwise he goes on comparing features.
Finally, if he is running out of relevant information, he says
"I don't know" (see Carbonell & Collins, 1973). This is simply
the Bayesian decision model common in the reaction time literature
(see, for example, Stone, 1960, or Fitts, 1966).

34

This revised model is one of the decision strategies that
Collins & Quillian (1972a) suggested people use, if they do not
have a superordinate connection stored. They suggested people
might use this strategy for deciding whether sheep, minks, or
cats are farm animals. We will call this the Bayesian strategy.
Collins & Quillian argued that there is an asymmetry between the
positive and negative thresholds in the Bayesian strategy, be-
cause often a mismatch on just one fairly criterial feature forms
a basis for saying "No," whereas most of the highly criterial
features must match in order to say "Yes" (though there may be
criterial features which the person does not know about).

Thus there is agreement that a decision process similar to
the one that Smith, Shoben & Rips postulate does occur for some
categorization decisions. But there is a fundamental disagree-
ment because they argue that all categorizations judgments are
made by comparing features of the instance and category. In
Table 1, we have listed some other decision strategies we think
are applied depending upon what connections are found in the
memory search. These are the same as those suggested earlier
by Collins & Quillian (1972a) with the addition of one strategy
(the exclusive-or strategy) that we found necessary in program-
ming a categorization subroutine in a computer (Carbonell &
Collins, 1973). We will describe each of these decision
strategies and give examples of when they are used.

The Wittgenstein strategy is a variant of the Bayesian
strategy described above, so we will describe it first. It
was based on Wittgenstein's (1953) observation that to decide
whether something is a game, for example, a person compares it
to similar things, that are known to be games. We would argue

35

TABLE 1

Decision strategies people use in categorization judgments

"Yes" Decision	"No" Decision
1) Superordinate connection	1) Negative superordinate
2) Bayesian strategy	2) Distinguishing property from Bayesian strategy
3) Wittgenstein strategy	3) Distinguishing properties from Wittgenstein strategy
	4) Exclusive-or strategy
	5) No intersection

36

that such a strategy might be used in deciding whether a beaver
is a mammal or a sled a vehicle, because people know so little
about the properties of mammals and vehicles. But they do know
some instances, so they apply the Bayesian strategy to the most
similar instance, that is known to be a mammal or a vehicle.
The thresholds change in applying the Bayesian strategy to
another instance rather than the category itself. Th negative
threshold is moved so as to be symmetrical with the positive
threshold, and therefore, matching properties count just as much
toward a positive decision as mismatched properties count toward
a negative decision. As Collins & Quillian point out, in de-
ciding whether a stagecoach is a vehicle, it might be compared
to a car. But even though a stagecoach does not have a motor,
which is highly criterial for being a car, a stagecoach may
still be a vehicle. So in the Wittgenstein strategy, the nega-
tive threshold must not be so stringent as in the Bayesian
strategy.

The other decision strategies all involve superordinate
links and this is where we are in most fundamental disagreement
with Smith, Shoben & Rips. While people may not have learned
many superordinate relations (for instance, that a beaver is a
mammal or a sled is a vehicle), there are many they have learned
(for instance, that a wren is a bird and a beaver is an animal).
The implication of the Smith, Shoben & Rips model is that even
when people have such information stored, they do not use it in
categorization judgments. Spreading activation theory assumes
the person evaluates whatever connections he finds, including
superordinate links.

37

Before explaining the particular decision strategies that
use superordinate links, it might be well to repeat several re-
levant assumptions about the memory structure and memory search.
First, we want to emphasize that superordinate links differ in
accessibility. We also assume that accessbility depends on use.
If a person frequently uses the link that a robin is a bird, and
less frequently uses the link that a chicken is a bird (assuming
approximately equal frequency for chickens and robins), then the
accessibility of bird from robin will be greater than from
chicken. Because of this, accessibility will be highly correlated
with typicality ratings. All the factors acting to make a chicken
or a goose atypical birds, will also act to make the use of the
superordinate link from chicken or goose to bird infrequent. It
is because they are atypical that the superordinate link is weak,
and this will act to slow down RT in making categorization judg-
ments about atypical instances.

Because the spreading activation search does not distinguish
between different kinds of links, connections through properties
may or may not be found before a superordinate connection, if
one exists. In fact, the theory treats superordinate links es-
sentially as very highly criterial properties. To understand
the theory's decision strategy in simplest terms, it can be
viewed as the Bayesian strategy (of which the Wittgenstein
strategy is a variant), where superordinate links are treated
like other properties except that they have higher criteriality.
Because they are so highly criterial, particular superordinate
connections can put the decision over either the positive or
negative threshold. That is why we have listed particular
decision strategies based on superordinate connections separately
in Table 1.

38

The most obvious decision strategy based on superordinate
information is one where a positive or negative connection is
found. We would argue that a person could never confirm that a
whale is a mammal, or that a sponge is an animal, or that a bat
is not a bird, without finding a superordinate link. In the
Smith, Shoben, & Rips model, these difficult (and slow) decisions
would be made in Stage 2 on the basis of defining properties.
But people generally have no idea what the defining properties
of a mammal, an animal, or a bird are. Even if they know that
one of the most criterial properties for being a mammal is that
it bears its young alive, it seems highly unlikely that they
know whether whales (or beavers for that matter) bear their
young alive. Neither of us has any idea what properties of a
sponge make it an animal, but if asked in an experiment whether
a sponge was an animal, we would answer "yes," and we would be
comparatively slow about it. The reason we would answer "yes"
is simply that we have been told at one time that a sponge is an
animal. Similarly, we have been told that robins, sparrows,
chickens, and geese are birds at one time or another, and it is
hard to believe we would not use this information if asked to
make a categorization judgment. We have also been told that a
bat is not a bird, and if we had not been told, we fear we might
have responded "Yes" if asked whether a bat is a bird in a
categorization experiment. The fact that there are cases where
people must use superordinate information to make correct cate-
gorization judgments makes it unlikely that they do not use such
information in other cases where they could make the decision
simply by pattern matching on features. This is one of the
strongest arguments for a hybrid theory.

There are two other ways we have used superordinate information in a computer program that makes categorization judgments. Both lead to negative judgments. The first occurs when the two concepts have a common superordinate, with mutually exclusive-or links into the common superordinate. We have labeled this the exclusive-or strategy in Table 1. As an example, consider the question of whether a mallard is an eagle. Since a mallard is a duck, and ducks and eagles are mutually exclusive birds, the answer is "No." We would argue that people use just such a strategy in deciding whether a sparrow is a wren, or a robin is a finch. In both cases, the concepts are semantically similar and the literature cited earlier shows that RT is slow in such judgments. According to the Smith, Shoben, & Rips model, which is designed to handle the semantic similarity effects, the decision would be made in Stage 2 on the basis of defining properties. But what are the defining properties of a wren that a sparrow does not have, and what are the defining properties of a finch that a robin does not have. In fact, the latter example is a trap that most people would answer wrong, because in fact the American robin is not a robin at all, but a finch. But most people think of these as different kinds of birds and will say "No" on that basis (though they may say "Don't know" because of lack of information). This is the exclusive-or strategy.

The failure to find an intersection was a hypothesis Collins & Quillian (1972a) suggested might account for people's introspections in making a judgment such as whether a cafeteria is a dog. Subjects are very fast to respond "No" in such cases, and they give as a reason afterwards that a cafeteria is one sort of thing and a dog is a very different sort of thing. The

40

Smith, Shoben, & Rips model, and our revised version of it, would
account for such an introspection and the fast RTs. However,
it turned out in writing a categorization subroutine which traces
along superordinate links, that if there is no common super-
ordinate the answer should be "No." For example, if the question
were "Is Surinam a person," the answer would be "No, it is a
place," because there is no intersection between place and person.
It seems probable therefore that people use a failure to find a
common superordinate as negative evidence in this decision. It
should be pointed out that the subroutine is fast to say "No"
when there is no intersection, but slow when it finds a common
superordinate (as in the exclusive-or condition_, because it must
then back up to see if it can find a basis for saying "No" rather
than "I don't know."   This is the same pattern as found in the
RT data, but the Smith, Shoben, & Rips model, and our revised
version of it, predict the same result.

We would like to close this section by raising the question
of why one should believe such a complicated theory when the
Smith, Shoben, & Rips model is simpler and predicts the RT data
as well or better. We have tried to stress the inherent difficul-
ties that their model has in ignoring superordinate information
and in relying on defining properties. Probably experimental
tests can be devised that will show up those difficulties. We
will suggest a couple of such tests, but first we might point
out that the results of the Loftus (1973a) categorization experi-
ment described earlier do not fit the Smith, Shoben & Rips model
very well. If a person is merely comparing features between the
instance and the category, then it should not matter whether the
instance or category is presented first (i.e., shrimp or seafood;
butterfly or insect). It is the asymmetry in the superordinate

41

connections that predicts the asymmetry Loftus found in RT, and
it is hard to imagine how one could have an asymmetry of that
kind in comparing features of two concepts.

One experiment that might show that superordinate links
correlate with typicality would be a generation experiment.
Subjects could be timed to generate the superordinates of
typical and atypical instances (controlling for word frequency).
It is our prediction that subjects would be slower to generate
bird from chicken, for example, than from robin. If links are
all the same, as Rips, Shoben, & Smith (1973) imply, then it
is hard to see how they would explain such a result. They could
postulate a checking process, based on their model, where the
given instance is compared to the superordinate generated. But
the predicted result falls out of our theory more naturally.
By confining themselves to modelling one particular paradigm,
they gain simplicity at the cost of generality.

Another experiment that might show difficulties with their
model is a categorization task. The categories and instances
used are based on their multidimensional scaling of birds and
animals on the one hand, and mammals and animals on the other.
As both Collins & Quillian (1971) (see Figure 3) and Rips,
Shoben, & Smith (1973) report, subjects are faster at deciding
bird names are in the category bird than in the category animal,
whereas they are slower at deciding if mammal names are in the
category mammal than in the category animal. Collins & Quillian
argue that this is the way people learn the superordinates;
that pigeons are birds and lions are animals. Smith, Shoben &
Rips argue that it is based on shared featurs, and they show
by their scaling solution that most birds are closer to bird

42

than animal, and most mammals are closer to animal than mammal.
But there are several bird names that are closer to animal than
bird (in particular goose, chicken, and duck; pigeon is equi-
distant), and there are several mammal names that are closer to
mammal than animal (in particular, deer, bear, and lion; horse
is equidistant). We would predict that even for those instances
the above pattern would hold, whereas a pure feature-matching
theory, such as the Smith, Shoben, & Rips model makes the op-
posite prediction. So this is a second possible test of the two
theories. There are undoubtedly many other tests.

Finally, we want to explain why we have been led to such a
complicated theory. In trying to write computer programs that
answer different types of questions (for example, "Is Surinam a
person?" "Whit is a chicken?", "Is Atlanta a port?"), it becomes
apparent that any decision procedure which gives correct answers
must be flexible enough to deal with many different configurations
of knowledge. This is because people have incomplete knowledge
about the world (see Carbonell & Collins', 1973, discussion of
open vs closed worlds), and they often do not have stored
particular superordinate links or criterial properties. Any
realistic data base for a computer system will have this same
kind of incomplete knowledge. Therefore, perhaps our strongest
criticism of the Smith, Shoben, & Rips model is that it breaks
down when people lack knowledge about defining features.

CONCLUSION

We have extended Quillian's spreading activation theory of
semantic processing in order to deal with a number of experiments
that have been performed on semantic memory in recent years.

43

The result is a fairly complicated theory with enough generality
to apply to results from many different experimental paradigms.
The theory can also be considered as a prescription for building
human semantic processing in a computer, though at that level
many details are omitted about decision strategies for different
judgments that arise in language processing (see Quillian, 1969;
Collins & Quillian, 1972a; Carbonell & Collins, 1973).   A
modern-day semantic theory should not only predict experimental
data, it should also discuss things like a reasonable person.

FOOTNOTES

1. Quillian's published version of the 1965 paper, appearing
    in Behavioral Science in 1966, did not include his theory
    of preparatory set or priming.

2. Semantic relatedness is a slightly different notion than
    semantic distance, though the two terms are sometimes used
    interchangeably. Semantic distance is the distance along
    the shortest path, and semantic relatedness (or similarity)
    is an aggregate of all the paths. Two concepts may be
    close in distance, say by a path through white, and still
    not be closely related, because that is the only path.
    Our use of "close" to refer to both relationships is ad-
    mittedly confusing. In this paper we shall use close to
    refer to relatedness or similarity though in some tasks
    (Quillian, 1966) it is only distance that matters.

3. There were actually 14 trials with the same category, but
    the last two trials were constructed by different rules
    that would complicate the description.

4. There is for living things a biologist's taxonomy which
    categorizes objects using properties that are not always
    those most apparent to the layman. Thus, there are
    technical definitions that are different from popular
    definitions, but this is not true in most domains.
    There is no technical definition of a game or a vehicle
    or a country that is generally accepted.

45

Anderson, J. R. Retrieval of propositional information from
    long-term memory. _Cognitive Psychology,_ in press.

Brown, R. W. & McNeill, D. The "tip of the tongue" phenomenon,
    Journal of Verbal Learning and Verbal Behavior, 1966,
    5, 325-337.

Carbonell, J. R., & Collins, A. M. Natural semantics in artifi-
    cial intelligence. _Proceedings of the Third International_
    Joint Conference on Artificial Intelligence, Stanford,
    1973, 344-351.

Collins, A. M. & Quillian, M. R. Retrieval time from semantic
    memory. _Journal of Verbal Learning and Verbal Behavior,_
    1969, 8, 240-248.

Collins, A. M. & Quillian, M. R. Facilitating retrieval from
    semantic memory: The effect of repeating part of an
    inference. _Acta Psychologica,_ 1970a, 33, 304-314.

Collins, A. M. & Quillian, M. R. Does category size affect
    categorization time? _Journal of Verbal Learning and_
    Verbal Behavior, 1970b, 9, 432-438.

Collins, A. M. & Quillian, M. R. Structure and priming in
    semantic memory. Paper presented at the Psychonomic
    Society, St. Louis, 1971.

Collins, A. M. & Quillian, M. R. How to make a language user.
    In E. Tulving & W. Donaldson (Eds.), _Organization of_
    Memory. New York: Academic Press, 1972a.

46

Collins, A. M. & Quillian, M. R.   Experiments on semantic memory
     and language comprehension.   In L. W. Gregg (Ed.),
     Cognition in Learning and Memory, New York: Wiley,
     1972b.

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

Fitts, P. M. Cognitive aspects of information processing: III
     Set for speed versus accuracy.   Journal of Experimental
     Psychology, 1966, 71, 849-857.

Freedman, J. L. & Loftus, E. F.   Retrieval of words from long-
     term memory.   Journal of Verbal Learning and Verbal
     Behavior, 1971, 10, 107-115.

Juola, J. F. & Atkinson, R. C.   Memory scanning for words versus
     categories.   Journal of Verbal Learning and Verbal
     Behavior, 1971, 10, 522-527.

Landauer, T. K. & Freedman, J. L.   Information retrieval from
     long-term memory: Category size and recognition time.
     Journal of Verbal Learning and Verbal Behavior, 1968,
     7, 291-295.

Landauer, T. K. & Meyer, D. E.   Category size and semantic
     memory retrieval.   Journal of Verbal Learning and
     Verbal Behavior, 1972, 11, 539-549.

Loftus, E. F.   Category dominance, instance dominance and cate-
     gorization time.   Journal of Experimental Psychology,
     1973a, 97, 70-74.

47

Loftus, E. F. How to catch a zebra in semantic memory. Paper
    presented at the Minnesota Conference on Cognition,
    Knowledge, and Adaptation, Minneapolis, 1973b.

Loftus, E. F. Activation of semantic memory, American Journal
    of Psychology, in press.

Loftus, G. R. & Loftus, E. F. The influence of one memory re-
    trieval on a subsequent memory retrieval. Memory and
    Cognition, in press.

Quillian, M. R. A revised design for an understanding machine.
    Mechanical Translation, 1962, 7, 17-29.

Quillian, M. R. Word concepts: A theory and simulation of
    some basic semantic capabilities, CIP Paper No. 79,
    Carnegie Institute of Technology, 1965.

Quillian, M. R. Semantic Memory. Ph.D. Thesis, Carnegie In-
    stitute of Technology, 1966. Reprinted in M. Minsky (Ed.)
    Semantic Information Processing, Cambridge, Massachusetts,
    M.I.T. Press, 1968.

Quillian, M. R. The teachable language comprehender: A simula-
    tion program and theory of language. Communications of
    the ACM, 1969, 12, 459-476.

Rips, L. J., Shoben, E. J., & Smith, E. E. Semantic distance
    and the verification of semantic relations. Journal of
    Verbal Learning and Verbal Behavior, 1973, 12, 1-20.

48

Rosch, E. On the internal structure of perceptual and semantic
    categories. In T. E. Moore (Ed.), Cognitive Development
    and Acquisition of Language. New York: Academic Press,
    1973.

Schaeffer, B. & Wallace, R. Semantic similarity and the compari-
    son of word meanings. Journal of Experimental Psychology,
    1969, 82-343-346.

Schaeffer, B. & Wallace R. The comparison of word meanings.
    Journal of Experimental Psychology, 1970, 86, 144-152.

Smith, E. E., Shoben, E. J., & Rips, L. J. Comparison processes
    in semantic memory. Submitted to Psychological Review.

Sternberg, S. High-speed scanning in human memory. Science,
    1966, 153, 652-654.

Stone, M. Models of choice-reaction time. Psychometrika, 1960,
    25, 251-260.

Wilkins, A. J. Conjoint frequency, category size, and categori-
    zation time. Journal Verbal Learning and Verbal Behavior,
    1971, 10, 382-385.

Wittgenstein, L. Philosophical Investigations (Translated by
    G. E. M. Anscombe). Oxford: Blackwell, 1953.