THEORIES OF THE BRAIN; BEHAVIOR, THE MIND, ROBOTS AND COGNITIVE PROCESSES, RESPECTIVELY

John T. Lamendella
San Jose State College

Many transformational grammarians have become convinced that the field of linguistics is, in reality, a branch of cognitive psychology and that a transformational grammar should be viewed as a cognitive theory of language. This position has led at least some linguists to wonder how one might validate claims of cognitive relevance; how one could choose between theories which equally well accounted for the emperical data but with different psychological implications. Having asked these questions, linguists join the ranks of psychologists, physiologists, psycho-physiologists, neuro-psychologists, philosophers, and others who also do not know what they mean when they talk about cognitive processes.

In this paper I will attempt to outline a framework for discussing questions like "What is a cognitive process?" "What would a formal theory of cognitive processes be like?" "What is the relationship between theories of the brain, theories of behavior, theories of mental processes, theories of robots, and theories of cognitive processes?" The distinctions I make apply only to an idealized world of scientific inquiry inhabited by strawmen who construct theories purely of one type rather than the messy types of theories which have a way to turning up in the real world. What I give here should be taken as a plea for the establishment of a conceptual framework in terms of which we can understand what a given real world theory is a theory of.

The first class of theories I would like to distinguish involves an anatomical and physiological description of the structures and processes of the nervous system. The theoretical component of such descriptions would for the most part be hypotheses filling in the gaps of empirical observation. Physical description theories would include definitions of morphological units such as the neuron, spinal cord, and the VI cranial nerve. In addition, they would describe functional units such as the auditory pathways, the extra-pyramidal motor system, and the limbic system. The term "functional" as used in this context reflects a concern with the spatio-temporal organization of activity in the anatomical units defined. The auditory pathways are defined as a "functional" entity on the basis of sequenced patterns of neural activity traversing specified structures in the nervous system over time. For the neurophysiologist, structure and function are inseparable and it would make no sense for the physiologist to talk of auditory pathways apart from such structures as the cochlear nuclei, medial geniculate body, etc. Hypothetical functional entities such as the sodium pump involved in the conduction of the nerve impulse are properly posited by the physiologist only

when they are construed as having some reasonably direct physical realization rather than merely being a way of looking at what happens.

The notion of functional component found in many information processing models of cognitive processes is quite different. Thus, we might discuss a functional entity mousetrap apart from any spatio-temporal physical manifestation. Something is a mousetrap provided it catches mice. What physical form it has is irrelevant. Nevertheless, it should be pointed out that while a physiologist's definition of function involves a physical manifestation, physiological descriptions don't reduce to anatomical structures in any simple manner. Thus, for example, the auditory pathways have no direct physical existence as a distinct entity.

The question is, could a physiological description of neural processes count as an explanation of cognitive processes. Certainly some scholars believe it is only because we don't know enough about the brain that we cannot give adequate physiological accounts of cognitive processes. However, just as a physiological description reduces to an anatomical description only in an oblique, as yet formally undefined way, so a description of cognitive processes reduces to physiological description in some even more oblique, even more undefined way.

The concerns of the idealized psychologist are quite different from those of the idealized physiologist. For example, while a physiological theory would be concerned with the biochemical and/or structural modifications which underlie memory storage in the brain, it would not be concerned with the fact that people store information about the world and use this information in particular ways for particular reasons. If reverberating circuits, facilitated synaptic transmission, network structures, protein synthesis, and glial processes were all involved in the storage of information about the world, the physiologist, as physiologist, would not notice. There is no physiological basis for identifying these diverse physical structures as functionally equivalent in the sense in which they all function to store information. This is the psychologist's notion of function. Even if a totally adequate description of physiological processes, existed, a description of cognitive processes would still have to be given by the field of psychology.

A very popular approach to the study of cognitive processes within the real world field of psychology has been to ignore them entirely and describe instead the external behavior which is the result of cognitive processes. Theories within this approach can take several forms. Early behaviorism and much so-called neo-behaviorism exemplify an approach I will call behaviorial taxonomy. Within this framework, a description of a specified domain of human behavior involves a system of predictions of the associations among external events and states; i.e., an inventory of stimulus-response pairs. The two major sources of evidence for the construction of such taxonomies are observation of overt behavior in relation

to external events and experimentally derived conclusions about the correlations between observed events.

These theories account for internal processes only in the sense that the association between input stimuli and output responses of an organism are taken to ultimately involve processes in the nervous system. The term cognitive process, often used to show merely that one is not a naive S-R behaviorist, has come to be applied by naive S-R behaviorists to things such as creativity, personality, emotion, motivation, and other previously unmentionable topics. The attempt has been made to "theoretically" account for these phenomena solely in terms of overt behavior without involving internal processes and structures in any theoretically relevant way. Most psychologists have recognized the failure of these attempts and many neo-behaviorists have shifted their theoretical interests beyond behavioral taxonomy. Many neo-behavioral approaches now attempt to incorporate covert intervening variables which mediate between external stimuli and responses.

Before discussing the status of these formulations as descriptions of cognitive processes, it would perhaps be helpful to consider a historical development in the field of linguistics which began with another brand of behavioral taxonomy and also ended with a concern for describing covert structures.

Although it had no aspirations of psychological relevance, structuralism within the field of linguistics adopted the methodological strictures of behaviorism and shunned all "mentalism" in describing language behavior. The structural linguist's main concern was the development of analytical techniques which could be used as discovery procedures for objectively getting at the structure of language data.

A phrase structure grammar describes a language by assigning to each sentence in that language a labeled bracketing which gives its syntactic structure. The language is defined to be a set of sentences. For the structuralists, this meant the sentences found in their corpus of data. During the 1950's in the field of linguistics, there developed the realization that judgments of native speakers about the grammaticality of sentences could provide evidence which was just as "empirical" and "scientific" as that obtained by copying down whatever native speakers happened to say, taking their utterances to be ipso facto grammatical. Perhaps Chomsky's most significant contribution to linguistics was the reintroduction of the notion that native speakers make mistakes and utter ungrammatical sentences. This revelation was significant in at least two respects. First, it led to a redefinition of what was acceptable data for the linguist and a concern for language structures which are never manifested in overt speech production. Once the linguist couldn't rely on the native speaker's sentence production, but instead on what the native speaker intuited to be grammatical, it was only a short step to a concern for the native speaker's intuitions about which sentences seemed related and which

didn't. For sentences such as open the door, empirical evidence could now be found for the understood you which your fourth grade English teacher told you was there. This eventually led to the distinction between syntactic surface structures and deep structures. Secondly, the elimination of mistakes from the data and the addition of sentences which had never been uttered but whose grammaticality the native speaker would affirm, led eventually to the competence/performance distinction. In its current form the distinction states that a transformational grammar characterizes the knowledge native speakers have of their language (that is, linguistic competence) rather than the use to which this knowledge is put (that is, linguistic performance). This is the principal basis for Chomsky's claim that linguistics is a brancy of cognitive psychology.

Should a transformational grammar be construed as a theory of cognitive processes because it posits abstract structures which are never manifested in overt language behavior? Should neo-behaviorist formulations be considered to describe internal processes because they posit covert mediating responses?

It is not by purpose here to argue questions about particular theories, although my biases should be clear. What I would like to do is claim that it is fruitless to discuss such questions without some general conceptual framework which tells one what kind of thing a cognitive process is and, in a general fashion, what would constitute a theory of such a thing.

Consider a type of theory, call it an abstract characterization theory, which is still within the realm of empirical description but goes beyond the behavioral taxonomy approach. Such a theory, in addition to accounting for the structure of observed behavior, defines the implicit relationships in the structure of the data. It contains an emperical taxonomy theory as a component but utilizes other types of evidence as well. In particular, it employs experimental evidence based on intuitions and introspections about the relationships holding between the overt structures. By describing these data, it may be said that this type of theory characterizes the facts which account for the overt form of the data. An abstract syntactic structure rule such as "Every sentence consists of a noun phrase followed by a verb phrase" may correctly assert that people possess this bit of information as part ot their linguistic knowledge but in no way could this statement be taken to constitute a description of this knowledge nor of the way it is integrated into memory structures in either a physiological or psychological sense.

We might compare the types of theories distinguished so far in terms of their treatment of the cognitive processes underlying human arithmetic behavior.

First, the physiologist would have nothing special to say about the cognitive processes involved when people add and subtract unless, for example, there were specialized cortical neurons involved in the storage and utilization of arithmetic information.

While we might agree that a behavioral taxonomy of arithmetic behavior catalogs interesting data, we would not be satisfied that these data help explain the underlying cognitive processes which were responsible for the behavior in question. An abstract characterization theory of arithmetic knowledge already exists as a brancy of number theory and posits rules such as $1 + 1 = 2$. This theory not only describes overt human arithmetic behavior (after eliminating irrelevant performance variables such as lapses of attention, stupidity, etc.) but characterizes peoples' knowledge of arithmetic relationships. For example, that $10 + 2 = 7 + 5$. It generates an infinite set of arithmetic expressions including novel expressions which have never been produced. While it is true, in a sense, that number theory accounts for what people know about arithmetic, no one is tempted to say that it describes arithmetic knowledge in the mind and is therefore a branch of cognitive psychology.

An abstract characterization theory posits maps which relate formal objects, but there are no processes defined by the theory. A transformational grammar describes a sentence in part by assigning it a set of syntactic phrase markers, related by a set of maps called transformations. There is no logical or chronological ordering between deep structure trees and surface structure trees in a transformational grammar; these are simply two formal objects which are part of the structural description of the sentence. An abstract characterization theory also does not posit any functional entities. It is simply a system for cataloging empirical data of both the overt behavior type and the abstract relationship type. No psychological claims are made by such theories.

The next class of theories to be distinguished I will call functional description theories, there being at least three different levels of such theories which could be construed as "modeling" some aspect of the organism. Each of these three subtypes of theory have different goals, are constructed on the basis of different evidence, and model the organism in a different sense. The first class of theories, call it a behavioral isomorphism theory, contains the other two; the second, call it a psychological isomorphism theory, contains the third, what I will call a psycho-physical isomorphism theory. Figure 1 summarizes the types of theories posited and their interrelationships.
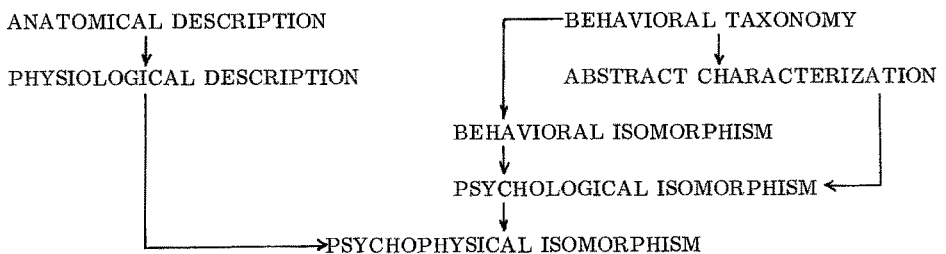

Figure 1

It has been suggested that if one wanted to understand how a bird flies, a good way to approach finding out would be to build one. However, not everything which can fly qualifies as a model of a bird. At a minimum, the model would have to fly like a bird rather than a jet plane or a helicopter. In order to qualify as a model of some organism, the minimal constraint on a theory must be <u>behavioral isomorphism</u> within a specified domain. A model whose motive power was a rubber band could qualify as a behavior isomorphism theory of a bird, for the domain "flying behavior", if the model flew the way a bird does. In order for a chess playing program to qualify as a behavioral model of human chess playing, it would have to show the same pattern of developing skills, make an occasional blunder and once in a while allow an opponent to retract really dumb moves. A behavioral isomorphism theory of a bird would not have to have feathers and be a biological organism, but however the behavior was generated, it would have to be isomorphis with what a bird does. Notice, of course, that it is the generation or definition of the behavior which is at issue rather than the production of the behavior. A running computer program which produces behavior may be flashier than the same program written on a piece of paper, but it is the same theory in either case.

A behavioral isomorphism theory bears a special relationship to a behavioral taxonomy theory for the domain of definition in that the former would incorporate the latter by including in its repetoire of behavior all of the empirical data about the behavior of the organism and its responses to stimuli. This is really the only constraint on a behavioral isomorphism theory. The internal processes of the model could be developed by pragmatic trial and error and it wouldn't matter what form they took. It makes no sense to ask what kind of evidence would verify the aptness of its internal structure. The only relevant question to be asked is: Does it define the behavior of the organism in question?

A behavioral taxonomy of arithmetic behavior would simply consist of a list of associations of the sort: When Ss are presented with the stimulus "How much is one and one?", the response "two" is elicited with a .83 probability. Given this understanding of the overt behavior, the behavioral isomorphism theorist would have the task of constructing an automaton which would "behave" in exactly this way. Such models already exist. An adding machine is a fairly good behavioral isomorphism theory of human arithmetic behavior, except insofar as it fails to make the right sorts of mistakes.

What distinguishes a behavioral isomorphism theory from a behavioral taxonomy is that the former is an automaton containing a set of devices each of which embodies one or more of the functions which account for either the generation or production of the desired behavior. Its functional entities are things like input device, output device, matching process, memory storage device, etc. In a behavioral isomorphism theory there are processes defined which are chronologically ordered. No one could build a robot using only the infor-

mation in a behavioral taxonomy, but a behavioral isomorphism theory is a robot whether it has mathematical or mechanical form.

I doubt that any one would want to claim that the structures and processes of a behavioral isomorphism theory of man would constitute a theory of human cognitive processes, although there are clearly people who would be content with this type of theory. The famous Turing test, apart from missing the point of the controversy about thinking, implicitly hypothesizes just such a behavioral isomorphism model of man.

A concern for such models is a valid pursuit in its own right, but the cognitive psychologist is interested only in a subset of the behavioral isomorphism theories. This subset is defined by several further constraints which establish the class of psychological isomorphism theories. The functional components of this type of theory are not just any set which will generate the desired behavior, but are equivalent to the functions which are actually involved when people perform the behavior. Additionally, the functional components of the model are interrelated in the same way the functional components of people's cognitive processes are actually related.

These theories incorporated the insights and data from an abstract characterization theory as well as a behavioral isomorphism theory. If, for example, it were true that people's knowledge of the syntactic structure of sentences involved two levels of grammatical description, a psychological isomorphism theory of language might contain a functor which changed a deep syntactic structure into a surface syntactic structure. Notice, of course, that two robots, one constructed as only behaviorally isomorphic to man and the other as psychologically isomorphic would be indistinguishable to an observer who was not free to examine their internals workings.

Many of the psychological isomorphism models which have been attempted thus far have been computer simulation models. For example, the General Problem Solver of Newell, Shaw and Simon attempts to simulate aspects of certain types of human problem solving. GPS also exemplifies the use of an additional type of evidence characteristic of psychological isomorphism models. Namely, subject protocols of the thinking-out-loud variety, a type of data which has only recently (again) become respectable in psychology. What is gained from this type of data is information about the conscious mental experiences of someone engaged in some cognitive process. Aspects of problem solving which are not available to conscious awareness would, of course, not be derivable from protocol data and this sort of data cannot always be taken at face value. For example, someone with trick mathematical abilities might report that he solves complex problems by visualizing the problem on a mental blackboard and then letting a piece of chalk write down the answer. We would consider the matter carefully before we build a blackboard and an animate piece of chalk into our model of human problem solving.

Psychological isomorphism theories also construct functional interpretations of experimentally derived conclusions about covert states and events causally affecting behavior in order to construct its account of cognitive processes. In considering visual information processing, it seems reasonable to conclude that in order to get binocular depth cues, the input from one visual channel must be compared with the input from the other channel. Thus, the psychological isomorphism theorist might posit the existence of a single Binocular Comparison Component which operates to extract these binocular cues on the basis of input from the left visual field and the right visual field. Of course, the pure psychological isomorphism theorist ignores physiological data so he would not be aware or care how complex the actual situation is; nor would he find out that his model is not functionally equivalent to human visual information processing beyond the behavioral isomorphism type of functional equivalence. A really pure psychological isomorphism theorist might not even have noticed that man has two eyes. Of course, it would be rediculous to consider visual perception without taking into account all of the available physiological data. No one would do such a thing and expect his model to actually describe human visual information processing. One wonders, however, what there is about such things as problem solving and language which allows "materialist" scholars to feel comfortable with models of these processes having no basis in physiological fact.

Some people might be tempted to define cognitive psychology such that low level receptor functions are viewed as being merely physiological with no relevance to the psychologist. However, at what point in the merely physiological visual pathways do cognitive processes kick in? Does visual information processing become cognitive only after it has left the visual areas of the cortex and gone off into the mysterious "association areas" where mental processes are reputed to lurk? Unless we want to use the term cognitive process as a synonym for "conscious mental experience", there is no notion of "higher mental process" such that we can functionally model visual cognition while ignoring the neurophysiology of vision.

Psychological isomorphism theories are the first type of theory we have considered which in any sense deserves the label psychological and one which might reasonably be viewed as telling us something about cognitive processes. Still, not everyone would be satisfied that the internal processes underyling man's behavior had been fully explained. One problem is that the internal structure of the functional components of a psychological isomorphism theory have no constraints other than their capacity for performing the function. Even if the theory has one component corresponding to every one psychological component (assuming a one-to-one isomorphism were desired), the internal operations of the components could not hope to reflect the way people process information while ignoring the physiological facts entirely.

A major reason people are willing to accept models of such things as problem solving which ignore the brain has to do with a real world problem for real world investigators that doesn't trouble idealized, imaginary scholars. It is clear that a cognitive process such as problem solving is intimately tied up to very many other cognitive processes. An explanation of the structure of information in memory, the form of incoming perceptual data the limitations of short-term memory, learning, motivation, language, and almost everything else would be required before anyone could make any reasonable claims about how human beings actually solve problems. All of these processes, in their turn, are intertwined with problem solving and the rest.

The most common solution to this problem has been to attack only a small part of one of the cognitive processes and beg the rest. This is called making "simplifying assumptions". Usually the small part which is chosen to be explained is the data from a particular experiment. One finds an article entitled something like, "A formal model of the structure of linguistic concepts" which turns out to be a computer simulation theory which accounts for the learning of six pairs of three-letter nonsense syllables under specified conditions of presentation and rehearsal by five male English-speaking college sophomores from Middlebury, Connecticut. If another investigator comes along and varies one of the parameters, one thing leads to another and a new **sub**field of psychology is born.

An additional reason for this over-specialized treatment is the desire to avoid claims which do not arise directly from experimental results in the standard format. Since the models developed amount to a summary of particular data, their applicability always depends on the particular way the problem was presented to the subjects, the amount and type of rehearsal, backgrounds of the subjects, and so on. There is an excellent chance that cognitive psychology, as behavioral psychology before it, will end up consisting of an unstructured heap of superficial, unrelatable experimental data. It does not seem reasonable to allow this type of scientific investigation to continue to be the dominant activity in the field of psychology. If psychology is ever going to get off the ground, some way must be found to incorporate data into a general framework where it can interact with all the other data being gathered. One way to begin to tie together the data which has been gathered is to broaden the scope of the mainstream of psychology to include physiological facts.

One may decide as a matter of practical fact that it is unreasonable to go beyond the psychological isomorphism level in describing cognitive processes or that one doesn't care to try. Nevertheless, it is not unreasonable for someone to wish for an explanation of cognitive processes which did explain exactly how human beings perform these functional operations. A type of theory which would fully explain human cognitive processes is the theory I have called psycho-physical isomorphism. This theory contains the insights of physical

description theories <u>and</u> psychological isomorphism theories. It presents an explicit answer to the question of the relationship between the mind and the brain. It explains how neurophysiological processes in human beings embody psychological functions. The development of such a theory is, in my opinion, the prime goal of cognitive psychology.

Some psychologists might be tempted to avoid a psycho-physical orientation on the grounds that neurophysiology has not progressed far enough to say anything of significance about cognitive processes. I seriously doubt that this is true. There is an enormous amount of relevant information about the brain and its functioning. Even if one constructed a psycho-physical theory on the basis of physiological data which turned out to be wrong, the theory would be closer to an adequate theory than one which ignores the physiological data entirely. Neural net theories can be constructed at a level of generality which leaves them almost certain to be correct models of information processing in the brain. Further specificity could await further data.

Some physiologists might be tempted to avoid a psycho-physical orientation on the grounds that <u>psychology</u> has not progressed far enough to say anything of significance about cognitive processes. However, the behaviorists have amassed a prodigious quantity of data and over the past few years information processing models have begun to develop very impressive computer simulations of aspects of cognitive processes.

The best thing which could happen to both fields would be the development of a stronger interest in psycho-physical isomorphism theories. The best thing which could happen to an individual investigator would be to have a clear idea of what he is really interested in explaining: the brain, behavior, the mind, robots, or cognitive processes. He should know whether the data he has used is the kind of evidence which leads to models of the sort he is interested in. And finally, he should carefully examine his complete model to see what he has modeled.