# ELIMINATIVE CONNECTIONISM: ITS IMPLICATIONS FOR A RETURN TO AN EMPIRICIST/BEHAVIORIST LINGUISTICS[1]

[*Behavior and Philosophy* (1992) **20**, 21-35]

Ullin T. Place
Department of Psychology
*University of Wales, Bangor*

ABSTRACT: For the past three decades linguistic theory has been based on the assumption that sentences are interpreted and constructed by the brain by means of computational processes analogous to those of a serial-digital computer. The recent interest in devices based on the neural network or parallel distributed processor (PDP) principle raises the possibility ("eliminative connectionism") that such devices may ultimately replace the S-D computer as the model for the interpretation and generation of language by the brain. An analysis of the differences between the two models suggests that the effect of such a development would be to steer linguistic theory towards a return to the empiricism and behaviorism which prevailed before it was driven by Chomsky towards nativism and mentalism. Linguists, however, will not be persuaded to return to such a theory unless and until it can deal with the phenomenon of novel sentence construction as effectively as its nativist/mentalist rival.

## EMPIRICISM AND BEHAVIORISM IN LINGUISTIC THEORY

There was a time when linguists took for granted what the man and woman in the street still takes for granted, namely, that the ability to speak and to understand what is said is something that we learn to do. According to this common sense theory, linguistic competence is acquired, partly from verbal interactions between the child and its parents, partly from verbal interactions between the child and its peers. The evidence which has persuaded generations of both ordinary people and professionals is, of course, the sheer multiplicity and diversity of natural languages. Added to this is the evidence that the natural language and dialect which the child acquires is invariably that spoken in the linguistic environment within which it grows up. There is no apparent contribution in this from heredity.

Before Noam Chomsky (1958;1959) provoked the revolution which overtook linguistics in the late nineteen fifties, many professional linguists in the United States and elsewhere were looking to behaviorist "learning theory" to give scientific substance to this "common sense empiricism" in the theory of language acquisition. One thinks particularly of Leonard Bloomfield (1933) and the philosopher Charles Morris (1946), best known for the distinction (Morris 1938) between "pragmatics", "semantics" and "syntactics."

## CHOMSKY

This empiricist/behaviorist theory of the acquisition and maintenance of linguistic competence was stopped dead in its tracks by Chomsky's (1959) devastating review of Skinner's book *Verbal Behavior*. In its place Chomsky (1958; etc.) proposed a nativist/rationalist theory of linguistic competence according to which

(1) linguistic competence is defined as the ability both to construct and to construe an indefinite number of sentences which are both novel, in the sense that neither speaker nor listener have ever produced or heard them before, and well-formed, in the sense that they conform to the rules of a natural language and are intelligible to any competent speaker of that language,

(2) the acquisition and persistence of linguistic competence is explained by the child's possession of two kinds of knowledge:
    (a) syntactic knowledge, knowledge of the syntactic rules governing the way sentences are constructed out of words, and
    (b) semantic knowledge, knowledge of the meanings of the words themselves.

---

[1]Earlier versions of this paper were presented at the Easter Conference of the Experimental Analysis of Behaviour Group, University of York, April 1990, and at the Annual Convention of the Association for Behavior Analysis, Nashville, TN, May 1990.

(3) knowledge of the rules of syntax is of two kinds
   (a) learned knowledge of the "surface structure" which varies from language to language, and
   (b) innate knowledge of the "deep structure" which is invariant from language to language.

The ease with which this new paradigm was displaced the earlier empiricist/behaviorist formulation is to be explained, in my estimation, by the failure of existing behaviorist theory, including Skinner's (1957), to address the problem of novel sentence construction. Although Skinner attempts to address the problem of syntax in his account of the "autoclitic" in *Verbal Behavior* (Skinner 1957, Part IV, Chapters 12-14), nowhere in that book does he give evidence of having a concept of the sentence as a linguistic unit with a distinctive structure.[2]

## THE SERIAL-DIGITAL COMPUTER AS A MODEL FOR THE BRAIN

But it was not only Skinner's failure to address the problem of the novel sentence that led linguists to abandon behaviorism. Equally important was his refusal to ask what kind of brain mechanism might be involved. This meant that Skinner was cutting himself off from exciting developments in the new fields of cognitive science and artificial intelligence which were then emerging as a consequence of the adoption of the serial-digital (S-D) computer as a model for the control of behavior by the brain. Chomsky and his followers had no such inhibitions. Although I have no reason to suppose that Chomsky was influenced by the S-D computer model in constructing his theory in the first place, the congruence between the two is remarkable. The S-D computer is a device which manipulates symbols in a strictly mechanical way dictated by a set of symbolically formulated rules or instructions ("a program") which is stored in its "memory" and retrieved for use as required. Although programs can be constructed which change in the light of experience of success and failure, programs are typically fixed once and for all at their creation and inserted into the device from outside. Moreover every digital machine requires an operating program which is "hardwired" into it from its initial construction. It is not difficult to see the analogy between this and Chomsky's notion of novel sentences being generated by a set of inbuilt syntactic rules whose "deep structure" is innate.

## FODOR AND THE LANGUAGE OF THOUGHT HYPOTHESIS

There are other respects in which the parallel between Chomskean theory and the S-D computer model is less good; and these have led Jerry Fodor, who began as one of Chomsky's most loyal disciples, to diverge from his master in order to follow the serial-digital computer analogy to its ultimate logical conclusion. In *The Language of Thought* published in 1975, Fodor points out that the set of rules or instructions which constitute a computer program cannot operate as such, unless they are formulated in a language or code (in the sense of a system of symbols or "words" organized into strings or "sentences" in accordance with precise syntactic rules). An S-D computer responds to sentences written in the "programming language", either because it has been programmed so to do, or, as in the case of the so-called "machine language" or "machine code", because the propensity to respond appropriately to the sequence of digital pulses into which both software and data have to be converted, if they are to be "read" by the machine, has been hardwired into the device at its construction. Fodor's "language of thought" is the brain's machine language.

The notion that the construction and construal of novel sentences in natural language are performed by a device, the brain, which is innately pre-programmed to respond to a body of rules formulated in the symbolism of the language of thought, seems, at first glance, not very far removed from Chomsky's hypothesis whereby sentence construction and construal are generated by a set of syntactic and semantic rules whose "deep structure" is likewise innately pre-programmed. Consequently, linguists were able to incorporate the new idea without having to modify the "surface structure" of the theory of the rule-governed generation of well-formed sentences in natural language.

---

[2] In a paper published within a year of his death, Skinner (1989, p. 88) finally acknowledged both the sentence as a unit of verbal behavior and the problem of novel sentence construction. Sadly, he did not go on to explore the full implications of this concession.

At a deeper level, however, the change in the concept of language which is entailed by the language-of-thought hypothesis is profound. For if it is true that the brain is an S-D computer which has its own hard-wired machine language in the form of the language of thought, it follows that language is not, as we used to think, something that was evolved by human beings primarily in order to communicate with one another. On this view, it forms an intrinsic part of the thought (computation) process which runs, not just the human brain, but the brains of most, if not all, species of multi-cellular free-moving living organisms.

## THE PARALLEL DISTRIBUTED PROCESSOR

For most of the past twenty years, it has been possible to argue, as Fodor (1975;1987) has repeatedly done, that there is no plausible alternative to the S-D computer as a model of how the brain works. This has now changed. With the advent or, as it really is, the revival of the neural network or parallel distributed processor (Rumelhart, McClelland and the PDP Research Group 1986) as a model for the way the brain functions, the S-D computer now has a serious rival.

A parallel distributed processor (PDP) is a device in which neuron-like semi-conductor units or "nodes", as they are usually called, are linked together in the form of a network in the same way that neurons are linked together in the "grey matter" of the central nervous system. Like the S-D computer, a parallel distributed processor is an information-processor which generates an output appropriate to the current input. But whereas, in the S-D device, the output is computed by following a sequence of steps determined by a pre-ordained set of symbolically formulated instructions, in the case of a PDP, the nature of the output is determined by the way in which the pattern of stimulation is transformed as it passes through a network of synaptically connected "nodes". The way in which the input is transformed is determined partly by the size and complexity of the network and partly by the so-called "weights" of the individual synaptic connections between one node and another. The *weight* of a synaptic connection is a dispositional property of the connection whereby the firing of the node on the anterior or input side of the connection contributes either to the excitation or the inhibition of firing in the node on the posterior or output side of the connection.

By assigning a set of weights to the various synaptic connections within the network it is possible to give it the innate predisposition to respond in a particular way to inputs of a particular kind. But, unlike the S-D computer's predisposition to respond to rules and instructions formulated in the machine language, these predispositions are susceptible to modification through changes in the synaptic weights brought about by subsequent learning experiences.

This effect whereby weights are changed either up or down each time a particular connection is activated or inhibited gives the neural network its distinctive functional property, that of acting as *a pattern discrimination learning device*. But, in order for it to have that property, the changes that occur need to follow certain consistent patterns or "rules". According to McClelland and Rumelhart (1988) there are two such rules which have been followed in giving networks the capacity to learn, "the so-called Hebbian or correlational learning rule .... and the error-correcting or ‘delta’ learning rule." (McClelland and Rumelhart *op.cit.* p.83).

These learning rules differ from those which make up the programs which drive an S-D computer. Those rules *prescribe* what is to happen at each step in the process of computation where at each step a decision is made between two alternatives. By contrast, the learning rules of a PDP are formulae which *describe* a uniform change which occurs in the weights of *all* the synaptic connections within a network as a consequence of their previous activation. (The difference between these two kinds of rule is obscured by the fact that most PDPs in actual use are S-D computers programmed to operate as PDPs. In this case, of course, the learning rules become prescriptive rules incorporated in the program which sets up the device as a PDP. But if we consider a PDP which is assembled *ab initio* from electronic memory units linked together in the form of a network, and, still more clearly, if we consider its biological counterpart, the central nervous system, the descriptive nature of the rule which specifies changes in the process whereby excitation is transmitted across the synapse from one node or neuron to another becomes abundantly clear.)

In view of their descriptive nature we should not be surprised to find, as in fact we do, that the learning rules adopted by connectionists in order to give a network the ability to learn correspond rather precisely to principles of learning which have a long history within traditional associationist and behaviorist learning theory.

Thus McClelland and Rumelhart (1988) trace back their correlational learning rule to Hebb (1949) and James (1890). Hebb's statement of the principle links it to Thorndike's (1911) "Law of Exercise"; while James' formulation ties it in, on the one hand, to the principles of classical Pavlovian conditioning as interpreted by Rescorla and Wagner (1972) and, on the other, to the principle of association by contiguity which can be traced back through Aristotle's *De Memoria et Reminiscentia* to Plato's *Phaedo*.[3] In the case of the error-correcting or "delta" rule, the connection is obscured by the use of concepts such as `back-propagation', `gradient of descent' and the definition of `error' as the discrepancy between the actual and `target' outputs which have no immediately obvious counterparts in traditional learning theory. Nevertheless, the use that is made of this principle to generate learning according to the principle of trial and error-correction, first described by Thorndike (1898) in his classic study of cats learning to escape from a puzzle box, links the delta rule to Thorndike's (1911) "Law of Effect" and Skinner's (1981) "Selection by consequences." Through Thorndike's formulation of the Law of Effect in terms of 'satisfaction' and 'dissatisfaction', there is an obvious connection with the principle of psychological hedonism (Freud's, 1900/1913, "pleasure principle") which has a history going back to Epicurus.[4]

No one would now seriously doubt that a PDP is much more like the brain, both in its structure and in its functional properties, than is a conventionally constructed S-D computer. However, according to Pinker and Prince (1988), connectionism, as the movement which advocates the adoption of this model of brain functioning is called, comes in two forms:

(1) *implementational connectionism* which seeks to retain the S-D computer as a model for the way in which the brain or part of it operates, by proposing that the `classical architecture' of the S-D device is *implemented* by what is basically a naturally-occurring PDP, just as most artificially-constructed PDP's are implemented on S-D computers, and

(2) *eliminative connectionism* which proposes the total elimination of the S-D computer as a model for how the brain works and the substitution of the PDP as a guide to the way in which *all* the brain's functions are executed in all aspects and at all levels of complexity.

While it is not yet certain which of these versions of connectionism is going to prevail in the longer term, two things are clear:

(1) As long as eliminative connectionism remains an option, the claim that there is no alternative to the S-D computer model can no longer be sustained.

(2) If connectionism goes down the implementational road, linguistic theory will be unaffected; whereas if eliminative connectionism prevails, it will be faced with a revolution as dramatic as that brought about by Chomsky more than thirty years ago.

## THE IMPLICATIONS OF AN ELIMINATIVE CONNECTIONISM

The prediction that the adoption of the eliminative version of connectionism would lead to another scientific revolution in linguistic theory rests partly on the features which disappear once the S-D computer model is abandoned, and partly on the positive features of the alternative neural network model. In the event that the S-D computer is replaced by the PDP model, there are four major doctrines within contemporary linguistic theory whose disappearance can be confidently predicted:

(1) Fodor's notion that linguistically formulated thought precedes and makes possible interpersonal linguistic communication,

(2) the notion, common to both Chomsky and Fodor, that sentences are generated by formal semantic and

---

[3] "Now you know that when a lover sees a lyre or a garment or anything else his favourite is wont to use, this is his sensation: he knows the lyre and in his mind perceives the bodily form of the boy who owned it." (Plato, *Phaedo*, 73 D, translated by W.D.Woodhead, 1953)

[4] In tracing McClelland and Rumelhart's two learning rules back to their historical roots, I am deliberately sidestepping the complex and difficult issue which arises within both connectionism and traditional learning theory as to which principle most accurately describes the actual process taking place at synapses in the brain.

syntactic rules which are embodied in symbolic formulae inscribed on the brain's counterpart of a magnetic tape,

(3) the notion that the linguistic competence of the individual and the semantic and syntactic regularities on which the intelligibility of linguistic communication depends are made possible and depend primarily on an innate capacity, either Chomsky's innate knowledge of "deep structure" or Fodor's innate knowledge of "the language of thought", and

(4) the notion that learning is a matter of storing "information" in a localized memory store within the brain from which it can be "retrieved" as and when required (in the case of word-meanings, this memory store is identified by Katz and Fodor, 1963, as the brain's "lexicon").

With the adoption of eliminative connectionism we can expect these four doctrines to be replaced by

(1) the traditional common sense and behaviourist view that interpersonal linguistic communication precedes and makes possible linguistically formulated thought,

(2) the notion that sentences are generated by irrational associations which are shaped into conformity with the linguistic `rules' (`norms' or `conventions' would be more accurate) by the error-correcting practices of the linguistic community embodied in the `back-channels' or `response tokens' supplied by the listener in the course of ordinary conversation,

(3) the notion that the linguistic competence of the individual and the semantic and syntactic regularities on which the intelligibility of linguistic communication depends are acquired and maintained by the process of correct-confirmation and error-correction supplied by the listener's response, and

(4) the notion that learning takes place through the development of more or less permanent changes in the weights of a very large number of synaptic connections between the nodes or neurons which make up the network.

## THE CASE FOR A RETURN TO AN EMPIRICIST/BEHAVIORIST LINGUISTICS

In the light of these considerations, there can be little doubt that if eliminative connectionism eventually prevails and the PDP supersedes the S-D computer as the preferred model for the way language is interpreted and generated by the brain, linguistic theory will be compelled to revert to the empiricism and behaviorism which were repudiated thirty years ago in the wake of the Chomskean revolution. What is much more doubtful is whether linguistic theorists can be persuaded that that is the direction in which they ought to be moving *now*. The fact that eliminative connectionism is well on the way to becoming the predominant fashion in contemporary artificial intelligence will not be enough to persuade linguists to abandon the nativist/rationalist theories which have dominated the field for as long as the memory of most of them extends. Ultimately, the only argument that will persuade them to adopt such a radical shift in perspective will be a demonstration that, when exposed to the kinds of first language learning experience to which the average human child is exposed, a network can learn to construct and construe the sorts of novel sentence which a child exposed to similar learning experiences can handle. But that is an objective which even the most enthusiastic exponent of eliminative connectionism would have to concede is still a long way from being achieved.

Some of the barriers to progress have to do with the limitations of current connectionist models. Another and, perhaps, more important barrier, however, is a skeptical attitude on the part of linguists to anything that smacks of an empiricist/behaviorist approach to the linguistic competence.

## THE PROBLEM OF NOVEL SENTENCE CONSTRUCTION AND CONSTRUAL

In order to persuade linguists, a case needs to be made for two propositions:

(1) There exists a viable empiricist/behaviorist theory of how novel sentences are understood and constructed.

(2) There are good reasons for thinking that such a theory is more likely to prevail in the longer term than are the existing alternatives.

Since the lack of such a theory in Skinner's book *Verbal Behavior* was first pointed out by Chomsky there have been, to my knowledge, two attempts to construct an empiricist/behaviorist theory of the construction and interpretation of novel sentences, One is a chapter entitled "Procedures for the acquisition of syntax" contributed by George Robinson (1977) to the Honig and Staddon *Handbook of Operant Behavior*. The other is an attempt by the writer (Place 1983; 1990; 1992; forthcoming) to explain the control exercised by novel sentences over the behavior of the listener by means of a version of Wittgenstein's "Picture Theory" of sentence-meaning, as developed in the *Tractatus* (Wittgenstein 1921/1971). On this theory, novel sentences act as signs or "discriminative stimuli", as Skinner (1938) calls them, for "contingencies" (i.e., antecedent-behavior-consequence relations) the like of which the listener may never have encountered, by virtue of an isomorphism between the structure and content of the sentence and the structure and content of the contingency to whose existence the listener is thereby alerted.

These two accounts complement one another. Robinson tries to show how the process of operant reinforcement ("error-correction", as the connectionist would say) can be invoked to explain how a child learns to construct syntactically well-formed sentences. "Contingency Semantics", as I call my theory, aims to explain how those sentences are able to control the behavior of the listener in the way they do.

## EXPLAINING THE EVOLUTION OF LINGUISTIC COMMUNICATION

Neither of these theories has, so far as I am aware, attracted the attention of linguists. Nor is it likely that they will do so, unless linguists can be provided with some good reasons for thinking that such a theory can account for the phenomena of language acquisition and use more effectively and more economically than any existing alternative. Are there any such reasons? I think there are.

One reason is that the empiricist/behaviorist theory provides a much more plausible explanation of why it is that only human beings can learn to communicate by means of language. It is a serious objection to Chomsky that he provides no account of how such a complex and sophisticated piece of genetic endowment as his theory postulates might have evolved. Fodor's theory avoids that difficulty by ascribing the language of thought to animals as well as humans; but this leaves us with no explanation of why animals are unable to exploit this innate linguistic ability for the purposes of interpersonal communication.

The empiricist/behaviorist theory assumes that the acquisition of linguistic competence depends on the same basic learning capacities as are available to other mammals. The evolution of language for the purpose of both interpersonal communication and symbolic thinking is seen to depend on the evolution of the vocal cords and the configuration of the mouth and palate required for vocal speech, coincidentally with the development of an enhanced manipulative ability, associated with the construction and use of tools. This combination allows the phonemes and phoneme combinations of natural language to acquire meaning in terms of the incipient evocation of a distinctive pattern of manipulative behavior. This contrasts with the circumstances of the parrot who can make the sounds but lacks the manipulative ability. The parrot's sounds remain just that, sounds without meaning.[5] This intimate connection between speech and manipulation explains both the use of gesture as an invariable accompaniment of speech and the use of sign language by the deaf as an alternative to vocal speech.

## EXPLAINING SYNTACTIC AND LOGICAL ERRORS

A second reason for thinking that the empiricist/behaviorist/ eliminative-connectionist theory of language is likely to prevail is that theories of linguistic competence which hold that sentences are *generated* by syntactic and logical rules make it very difficult to understand why human speakers and thinkers should make the syntactic and logical errors that they manifestly do. These errors are no problem for the empiricist/behaviorist/eliminative- connectionist theory. For on this view, it is only by making such errors, and receiving an "error message" from the listener as a consequence, that the organism or network can learn to conform to the relevant syntactic conventions and principles of logic.

[5] Ongoing research by Irene M. Pepperberg (1992) of the University of Arizona on the labelling behavior of an African Grey Parrot seems likely to show that, with appropriate training, even this time honored difference between humans and parrots can be made to disappear.

## THE ROLE OF LISTENER REINFORCEMENT IN THE ACQUISITION OF LINGUISTIC COMPETENCE

A third reason for expecting that the empiricist/behaviorist/ eliminative-connectionist theory of language will ultimately prevail is provided by E.L.Moerk's (1983) reworking of Roger Brown's (1973) study of the behavior of a mother as a first language teacher of her own child. Brown's study is the source which is commonly cited as evidence for the view that mothers do not correct the child's syntax and that consequently her subsequent compliance with the syntactic conventions of the language is not acquired from training she receives on the proverbial "mother's knee". Moerk's re-analysis of Brown's data clearly shows that Brown's principal conclusion, that the child's linguistic competence develops independently of the learning experiences provided, both deliberately and unconsciously, by the mother's response, does not stand up. It is true that the mother's reluctance to correct the child for emitting syntactically deviant forms (other than by modelling the correct form after an initial *Good girl!*), and her generosity in interpreting what the child is trying to say, leads to the prediction from reinforcement theory that the child would acquire many more deviant forms of utterance than are normally observed. But this assumes that the only relevant learning experiences in the child's acquisition of linguistic competence are those which come from its interactions with the mother. While the mother-child interaction no doubt provides an indispensable foundation, all the evidence suggests that it is the child's interaction with its peers which is the most important determinant of the pattern of speech that is ultimately acquired. The peer group may not selectively reinforce patterns of speech which conform to the rules laid down in the grammar books, but it is much less generous than the mother in its interpretation of what the child is saying. Utterances which are deviant by the syntactic conventions recognized by the group will receive powerful "error messages", at best bafflement and incomprehension, at worst ridicule.

## THE ROLE OF THE LISTENER IN MAINTAINING LINGUISTIC COMPETENCE

My fourth and final reason for thinking that the empiricist/behaviorist theory of language will eventually prevail relates to the maintenance of linguistic competence rather than its initial acquisition. It is provided by the linguistic phenomenon somewhat disparagingly referred to as the "back-channel". This is the phenomenon whereby, when two people engage in conversation, each time a speaker completes a sentence or adds a phrase which, taken together with the preceding sentence, makes a longer sentence, the listener acknowledges that sentence completion by an expression of reciprocal greeting, agreement, comprehension, compliance, surprise, amusement, concern depending on the nature of the utterance to which the listener is responding. This phenomenon is illustrated on Table 1 which is taken from the transcript of a conversation recorded in the Philosophy Departmental Office at Leeds University in 1985.[6]

On this table, the points where the current speaker completes a sentence are indicated by an upward pointing arrow, and you will see that with only four exceptions every time a sentence is completed it evokes an appropriate response or "response token" to use the term preferred by conversation analysts, from the listener. The four exceptions are:

(1) on lines 10-11 where there is no apparent response from Penny when Rose completes the sentence *the're all on that list* which echoes Penny's immediately preceding sentence *the're on that list*, and can, perhaps, be regarded simply as a response to that sentence rather than a sentence in its own right,

(2) on lines 11 and 12 where Penny's *yes* anticipates the completion of Rose's sentence at a stage when it is quite clear what the completion is going to be, and

(3&4) when Penny engages in the extraordinary piece of self-directed reasoning which makes up her final three sentence turn on lines 13-15 and which, apart from its conclusion, is ignored by Rose presumably because she is only interested in the conclusion and not in the tortuous mental process by which the conclusion is reached.

[6] A transcript of the complete conversation with an extensive analysis is given in Place (1991).

| | | |
|---|---|---|
| Penny: | it's *just* this bus'ness of (.) th' *par*ty [for the *fir*st y:e:ars. | 01 |
| Rose: | [ye:(s) yes= | 02 |
| Penny: | =*I* won't (.) be *i*:n tomorrow mo:rning. | 03 |
| Rose: | no= | 04 |
| Penny: | =I've *left* a *no*tice on the *board*. | 05 |
| Rose: | yeah.= | 06 |
| Penny: | =and there's a *no*te for them °of the money. | 07 |
| Rose: | wh*o* wants to pick it up?= | 08 |
| Penny: | =we:ll (.) th*e:*'re on that li:[st. | 09 |
| | [ | |
| Rose: | [oh the're °all on that *li*st.= | 10 |
| Rose: | = (.) and any-any of these people[can *have* it, (.) can they. = | 11 |
| | [ | |
| Penny: | [yes:: (.) | 12 |
| Penny: | =I *do*:: know John's girl friend *knows* about it. = | 13 |
| Penny: | =bu(t) *she's* not *free* at the same time as *them* tomorrow. = | 14 |
| Penny: | =so:th't *lots* of people *know* about it, = | 15 |
| Rose: | =anan *the:*'re goin(g) to *get* the shoppin(g) *ou*[t of it. (.) I *see*= | 16 |
| | [ | |
| Penny: | [yes (.) | 17 |

**Table 1**

As you will see from this excerpt the overwhelming majority of these response tokens are "correct" messages or reinforcers, i.e., expressions of agreement or comprehension, surprise, sympathy, etc. which serve to maintain the flow of the speaker's ongoing verbal behavior from sentence to sentence. This flow is occasionally interrupted by "error" messages or "disinforcers", to use the term proposed by Harzem and Miles (1978) in the form of expressions of disagreement, incomprehension, disbelief, etc. which require the speaker to repeat, amplify or justify what has just been said. In this excerpt the nearest we get to such an error message are the two requests for clarification and confirmation which Rose makes on lines 08 and 11.

Three conclusions would seem to follow from this evidence:

(1) Response tokens play a significant role in ensuring that effective interpersonal linguistic communication takes place.

(2) They do so by acting as "correct messages" or reinforcers when communication is proceeding smoothly, and as "error messages" or disinforcers when communication breaks down.

(3) By ensuring the speaker's conformity to the semantic and syntactic conventions of the language, they contribute to the process whereby effective linguistic communication is maintained within the linguistic community constituted by all competent speakers of that language.

## CONCLUSION

In a lecture delivered and published as recently as 1987, Chomsky gives an account of his approach to the problem of the acquisition of language which includes the following statement:

> The mind/brain is considered to be an information processing system, which forms abstract representations and carries out computations that use and modify them. This approach stands in sharp contrast to the study of the shaping and control of behavior that systematically avoided consideration of the states of the mind/brain that enter into behavior, and sought to establish direct relations between stimulus situations, contingencies and behavior. This behaviorist approach has proven almost entirely barren, in my view, a fact that is not at all surprising since it refuses in principle to consider the major and essential component of all behavior, namely, the states of the mind/brain.

It is ironic that at the very moment when those words were being pronounced, a new way of looking at how the mind/brain operates was already beginning to undermine Chomsky's confident dismissal of the behaviorist approach to language acquisition. Why, nearly thirty years after his review of *Verbal Behavior*, does Chomsky still find it necessary to reiterate the inadequacies of a behaviorism whose death and burial he and his confederates have been proclaiming ever since. Is there, perhaps, a lurking suspicion that the corpse may yet rise up and strike down its would-be assassins?

I have some sympathy, of course, with Chomsky's criticism of Skinner's refusal to concern himself with the workings of the brain. I am sure that Skinner was entirely right to insist that we can and should study the relation of behavior to environmental contingencies without any preconceptions as to how that relationship is managed by the brain. But to claim, as he sometimes did, that the behavioral psychologist has no business concerning him or herself with such matters is plainly absurd. Nevertheless Skinner's refusal to countenance neurophysiological speculation has had one beneficial consequence. If nothing else, it has prevented behavior analysts from following the crowd down the blind alley of the serial-digital computer as a model for the functioning of the brain. What behavior analysts need to appreciate is the advantage this gives them in coming to terms with the implications of the connectionist revolution in artificial intelligence, as compared whose thinking has been dominated for years by what has come to be known as the "classical cognitive architecture", and who are having painfully to unlearn the thought habits of a generation, if not a lifetime.

## REFERENCES

Bloomfield, L. (1933) *Language*. New York: Holt.

Brown, R. (1973) *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.

Chomsky, N. (1958) *Syntactic Structures*. 's Gravenhage: Mouton.

Chomsky, N. (1959) Review of B. F. Skinner's *Verbal Behavior*. *Language*, 35, 26-58.

Chomsky, N. (1987) On the nature, use and acquisition of language. In N. Chomsky *Generative Grammar: Its Basis, Development and Prospects*. Kyoto: Kyoto University of Foreign Studies.

Fodor, J. A. (1975) *The Language of Thought*. New York: Crowell.

Fodor, J. A. (1987) *Psychosemantics*. Cambridge, MA: MIT Press.

Freud, S. (1900/1913) *Die Traumdeutung*. Leipzig und Wien: Franz Deuticke. English translation as *The Interpretation of Dreams* by A. A. Brill, London: George Allen & Co.
Harzem, P. and Miles, T. R. (1978) *Conceptual Issues in Operant Psychology*. Chichester: Wiley.
Hebb, D. O. (1949) *The Organization of Behavior*. New York: Wiley.
James, W. (1890) *Principles of Psychology*, Vol.1. New York: Holt.
Katz, J. J. and Fodor, J. A. (1963) The structure of a semantic theory. *Language*, 39, 170-210.
McClelland, J. L. and Rumelhart, D. E. (1988) *Explorations in Parallel Distributed Processing*. Cambridge, Mass.: MIT Press.
Moerk, E. L. (1983) *The Mother of Eve - As a First Language Teacher*. Norwood, NJ: Ablex.
Morris, C. W. (1938) Foundations of the theory of signs. *International Encyclopedia of Unified Science*, Vol.1, No.2. (pp.78-137). Chicago: University of Chicago Press.
Morris, C. W. (1946) *Signs, Language and Behavior*. New York: Prentice-Hall.
Pepperberg, I. M. (1992) A search for equivalence classes in an African Grey Parrot: Equivalence involving objects and auditory and visual labels. Paper presented to the 15th Symposium on Quantitative Analyses of Behavior (Stimulus Relations), Harvard University, June 14th 1992.
Pinker, S. and Prince, A. (1988) On language and connectionism: analysis of a parallel distributed model of language acquisition. *Cognition*, 28, 73-193.
Place, U. T. (1983) Skinner's *Verbal Behavior* IV - how to improve Part IV, Skinner's account of syntax. *Behaviorism*, 11, 163-186.
Place, U. T. (1990) Intensionalism, connectionism and the picture theory of meaning. *Acta Analytica* **6**, 47-63.
Place, U. T. (1991) Conversation analysis and the analysis of verbal behavior. In L.J.Hayes and P.N.Chase (Eds.) *Dialogues on Verbal Behavior: Proceedings of the First International Institute on Verbal Relations*. Reno, NV: Context Press, Chapter 5, pp. 85-109.
Place, U. T. (1992) Behavioral contingency semantics and the correspondence theory of truth. In L. J. and S. C. Hayes (Eds.) *Understanding Verbal Relations*. Reno, Nevada: Context Press, Chapter 9, pp. 135-151.
Place, U. T. (forthcoming) Holism and cognitive dissonance in the discrimination of correspondence between sentences and situations. *Acta Analytica*, **8**
Rescorla, R. A. and Wagner, A. R. (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black and W. F. Prokasy (eds.) *Classical Conditioning*, Vol. 2: *Current Research and Theory*. Englewood Cliffs, NJ: Prentice-Hall.
Robinson, G. M. (1977) Procedures for the acquisition of syntax. In W. K. Honig and J. E. R.Staddon (Eds.) *Handbook of Operant Behavior*. Englewood Cliffs, NJ: Prentice-Hall.
Rumelhart, D. E., McClelland, J. L and the PDP Research Group (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Volumes 1 and 2. Cambridge, MA: MIT Press.
Skinner, B. F. (1938) *The Behavior of Organisms*. New York: Appleton-Century.
Skinner, B. F. (1957) *Verbal Behavior*. New York: Appleton-Century-Crofts.
Skinner, B. F. (1981) Selection by consequences. *Science*, **213**, 501-504.
Skinner, B. F. (1989) The behavior of the listener. In S. C. Hayes (Ed.) *Rule-Governed Behavior: Cognition, Contingencies and Instructional Control*. New York: Plenum, Chapter 2, pp. 85-96.
Thorndike, E. L. (1898) Animal intelligence: an experimental study of the associative process in animals. *Psychological Monographs*, **2**, No.8.
Thorndike, E. L. (1911) *Animal Intelligence*. New York: Macmillan.
Wittgenstein, L. (1921/1971) Tractatus logico-philosophicus. *Annalen der Naturphilosophie*. *Tractatus Logico-philosophicus*. With second English translation by D. F. Pears and B. F. McGuiness. London: Routledge and Kegan Paul.