Lenguas Modernas 19 (1992), 41-55
Universidad de Chile

NOTES TOWARDS A CONNECTIONIST PROGRAM FOR THE STUDY OF SYNTAX

James W. Ney
Arizona State University

A connectionist program for the study of syntax would require a radical change in approaches to the subject matter. This change would be necessitated by the fact that grammarians and linguists have traditionally looked at language as though it is processed in a serial fashion. In most cases, the connectionist viewpoint would require that the linguist look at language as though it is processed in a parallel fashion, perhaps a massively parallel fashion. But users of language probably process the context of situation, the context of discourse and the linguistic elements of language in a parallel fashion, synthesizing information from a number of sources at once; that is, the users of language receive information from their senses, and from memory at the same time. Besides this, since connectionism is based on an associationist model with connection 'weights' and changes in these 'weights', the notion of 'rule of grammar' will have to be redefined.

1. Connectionism and language

1.1. *Connectionism as a paradigm shift*

According to some commentators, the new paradigm shift for linguistics (Sampson 1987:871) and psychology (Schneider 1987:73 and, perhaps, Hunt 1989 and Leahey 1992:308) has already arrived. This new paradigm is variously named connectionism or parallel distributed processing. As the latter name suggests, the new discipline comes out of the computing sciences, full of terms such as connection weights and parallel processing elements and under the leadership of two specialists in the field, James McClelland and David Rumelhart. At this point, it is impossible to say whether Sampson, Leahey and Schneider are correct in their assessment. But if they are, for linguists, the question arises: What would some of the characteristics be in a program for a linguistic study of syntax within the connectionist framework?

In the first place, it would ultimately be quite unlike anything ever seen to this point. The reason for this is that the Western intellectual tradition has simply led generations of grammarians and linguists to look at language as though it is processed in a serial fashion with certain kinds of entities. (See Yngve (1986, 1990).) In most cases, the connectionist viewpoint would require that the linguist look at language as though it is processed in a parallel fashion, perhaps a massively parallel fashion. This

in itself would require that language be viewed in a manner quite foreign to the conventional Western tradition.

1.2. *The nature of language and language studies*

Thus, it is possible to look at a phrase from natural language such as *The shooting of the hunters* and to parse the sentence so that in one reading the noun, *hunters,* is subject and in another reading the noun, *hunters*, is object. (The example is from Chomsky (1957:88).) But the question is this: Do native speakers when they encounter language in its natural state, that is, within the context of situation or within the context of a written or spoken discourse, parse the sentence prior to uttering or understanding it? The answer to the question is: Probably not. Experimentation by psychologists during the 60's could not demonstrate the existence of transformational processes for the processing of language or the fact that native speakers, in some sense or other, parsed sentences in their attempt to process the meaning in them. (See Newmeyer 1983: 131-136, Botha 1979: 313-317.) Users of language probably process the context of situation, the context of discourse and the linguistic elements of language in a parallel fashion, synthesizing information from a number of sources at once; that is, the user of language receives information from his eyes, sight, from his nose, smell, and from his ears, hearing, and processes the information from all three sources and from memory at the same time. Thus, information that is used by a language user comes from a number of sources and is processed in a parallel fashion.

Since parallel processing has generally been disregarded in linguistic analysis, the Western tradition has simply led generations of grammarians and linguists to look at language in an incorrect fashion. Although it is possible to analyze the phrase *The shooting of the hunters* so that in one reading the noun, *hunters,* is subject and in another reading the noun, *hunters*, is object, is it true to an analysis of language in its natural setting to do so? In any natural environment, the speaker/hearers know from information that is present in the situation who is doing the shooting and who is being shot. Perhaps, language simply provides rather inexact symbols to which language users attach meaning which they infer from the context of situation or from the context of a written or spoken discourse and the meanings 'subject of ', 'object of ', etc. are not to be found in the concatenations of the verbal symbols at all. Uhlenbeck (n. d. p. 28) is very close to formulating such a concept when he discusses this very topic and points out that *The shooting of the soldiers* is ambiguous in a 'reciprocal' reading versus a 'non-reciprocal' reading, whereas *The shooting of the soldier* is not. He also points out that the ambiguity of the reading cannot be attributed to the plural because the *The shooting of the gang* is ambiguous in the same way that *The shooting of the soldiers* is. He goes even further to suggest that *The broiling of the duck* would not normally be read with *duck* as the subject of *broil* but it could be read this way if the duck under discussion was a mythical hero in a fairy tale. Thus the meaning inherent in any collocation of language items would be attributable to the meaning derived from the linguistic elements themselves, and the context of situation or from the context of a written or spoken discourse in a parallel fashion and not from notions of 'subject', 'object', 'agent', 'instrument', 'noun', and 'verb' inherently attached to collocations of lexical items. In this sense also, the meanings of 'subject', 'object', 'agent', 'instrument', 'noun', and 'verb' would be derived indirectly from the context also.

1.3. *The nature of connectionism*

If the preceding discussion has a modicum of truth in it, the question then arises again: What would some of the characteristics in a connectionist program for a linguistic study of syntax be like? No answer to the question can be forthcoming until some of the relevant aspects of the discipline of connectionism (or associationism) are brought into focus. Fortunately, some of the terms in connectionism are not so new. In fact, 'connectionism' and its synonymously used term 'associationism' from Hull to the present derive from the behaviorist era as Lachter and Bever (1988: 234) have noted: "The learning models in McClelland and Rumelhart (1986) are a complex variant on traditional —or at least Hullean— s-r connections formed in time...". In addition, like the behaviorists, "The connectionist focuses on overt behavior ..." (Lachter and Bever, 1988: 89) Most important, however, connectionism is a term which can be used of machines that "learn." Thus, parallel processing provides for computer simulation through the scanning of input and the adjustment of 'weights' in 'networks' until the desired output is achieved. In the words of Fodor and Pylyshyn (1988: 6) connectionist "... networks can be made to learn; this is achieved by modifying the weights on the connections as a function of certain kinds of feedback..." The fact that machines can be made to "learn" is a startling discovery, which has caused some theoreticians to pay serious attention to associationism in its present form as a function of parallel distributed processing (PDP).

Thus, one aspect of connectionism can be brought into focus from the works of theoreticians such as Fodor and Pylyshyn, who find that connectionism (or associationism) has two facets, both of which are important to language analysis: 1) the architecture of the systems with their analog in the human mind as a model; and 2) the implementation of these systems resulting in the weighting of associations and subsequent "learning" in the parallel computer networks. The first of these is comparable to the human neurological mechanisms that figure so prominently in Chomsky's system (Cellerier 1980: 83-89); the second is similar to classical associationism, which has developed within behavioristic psychology from Hull to the present. Within this system, Rumelhart and McClelland have ruled out language as a "rule-governed" phenomenon, a concept which figures prominently in transformational systems. They state (1987: 196):

> In our network models, the mechanisms that process language are constructed in such a way that there are no rules anywhere in them. Acquisition occurs by a simple process of adjusting connections between units. The behavior of the models is lawful (as lawful, we would argue, as the human behavior it simulates), but it is not based on the formulation or consultation of rules.

Thus, according to them, although language is not "rule-governed," it is "lawful" like other forms of human behavior. On this point, however, there might be some question about the meaning of "lawful" in Rumelhart and McClelland since they do use the term "rule," which seems to be a synonym to "law," in their system (1987: 206, 208, 226).

In any case, the difference between "rule-governed" and "lawful" is found in the fact that the latter is a statistical or probabilistic concept. That is, since the lawful

character of language is determined by connection weights, when the weights pass a threshold level they produce a lawful aspect of language behavior. When they fall below that level, the law is not in effect. This makes PDP or connectionism particularly useful in describing phenomena encountered in certain types of aphasia. As Clark (1989:169) states:

> ...imagine a kind of damage that decrements *all* the connectivity strengths by 10 per cent. This could move all the irregular words below the threshold while leaving the originally strong regular pattern functional.

It also provides for the production of regular and irregular patterns in language using the same mechanism. For example, on the morphological level, English requires that the past tense marker be /-t/ in voiceless environments as in *walk/walked;* however, in *buy/bought* the past tense marker /-t/ occurs in a voiced environment. Language behavior at this point is lawful; that is, the usual pattern requires /-t/ in voiceless environment. But that is not always the case. A connectionist model would produce both forms using the same mechanism. The connectionist weights would be altered by the processor to produce the /-t/ in the environment of *buy* ≈ *bough-*.

Similarly, on the level of syntax, as Bolinger (1968) has pointed out, reifying verbs usually require the gerundive complement as in *I enjoy going to the movies* (cf. *I want to go to the movies).* But in the case of *I managed to go to the movies yesterday,* a reifying verb, *manage,* requires the infinitival complement. Again, language behavior is lawful in that reifying verbs usually require the gerundive complement. But the law does not apply in every case; nevertheless, the same mechanism would be used to produce the usual case, reifying verb plus gerundive complement, and the less usual case, reifying verb plus infinitival complement, through an adjustment of connectionist weights. Actually, the case of *manage* is also lawful in that it requires a punctiliar aspectual marking in the complement which overrides the requirement of a reifying verb and dictates the use of an infinitival complement. (See Ney 1988:87 and the discussion below on the case of *happen* (section 2.2, example 4). Compare Wierzbicka 1988:24.) Viewed either way, mechanisms in a connectionist model can produce the required form. Since these models are parallel not serial processors, information coming from different sources can be used to produce the required form; that is, connections from one location in the processor can contribute to the weights in another location thus overriding a usual pattern.

2. Connectionism and the study of syntax

2.1. *Connectionist architecture*

It would seem, then, that the answer to the question about the implications of connectionism for the study of syntax can be found in the nature of the connectionist model which, in effect, makes it possible for computers to learn at least some facets of language. If, as Fodor and Pylyshyn maintain, the architecture of the networks determines in part the success of the model, then this architecture is similar to the neural networks in humans which are the basis for the notion "rule of language" or of "lawful language behavior." Fodor and Pylyshyn (1988:49) also complain that there is no formal structure in the connectionist architecture. When they do so they are

focusing on the "virtual machine" prior to operation, before it has made any of its many passes to develop an architecture which would be similar to the rules in the classical or transformational scheme. But to say that the connectionist model has no architecture is to misunderstand the nature of this model. The fact that connectionism employs parallel processors is in fact an architecture. How are these processors hooked together? What is their internal configuration? Do they operate like digital devices or are they more like analog devices? Answers to all of these questions imply a formal architecture of a given type —perhaps, an architecture very similar to that of a human brain. Furthermore, the fact that instructions to the machines are very often written in LISP (and can be described by algebraic or quasi-algebraic notations) also implies an architecture. In PDP, the machines and their architectures in no way constitute a tabula rasa. The crux of the matter in the definition of "formal architecture" is whether the architecture is that of a virtual machine that is based on a serial processor with complete instructions or a parallel processor which, in effect, develops many of its own instructions (though not all). The implication in the discussion above centering on the phrase *The shooting of the hunters* (section 1.2) is that the architecture of the connectionist model is in effect a more accurate, realistic architecture for the description of the syntax of language.

This also becomes more evident in an examination of the notions of regularity and irregularity in language. If excitation in the networks allows changes in a probabilistic fashion, this undoubtedly accounts for the irregularities in language. Thus, these two features of connectionism would account for the fact that some aspects of language are rule-governed, that is, they are "regular" while others are not, that is, they are "irregular." Rumelhart and McClelland themselves provide support for this view when they state (1987: 226): "... the performance of the model in Phase 1 essentially amounts to 'rote' learning. In Phase 2 it appears to be heavily rule-based, and in later phases its performance amounts to what might be described as a blend of rule-based, analogy-based, and rote-based." (This, of course, is reminiscent of the fact that analogy-based descriptions have long been a part of linguistic descriptions (Vincent 1978) and would find a place in connectionist studies.)

Thus, not only does parallel distributed processing account for the fact that language contains rule-based regularities and irregularities which can scarcely be captured by rules, but it goes through phases which accommodate these two facets of language: the first phase capturing the irregular forms through rote learning and the second phase concentrating on the rule-governed regularities. Finally, the model strikes a balance between the two poles of regularity and irregularity in language much as a child or an adult learning language would; that is, these models produce overgeneralized forms when they imitate human language learning much as a child or an adult produces *feets* in an attempt to master the irregular plural, *feet.* But, ultimately, the model, like the human language learner, strikes a balance where the irregularities and rule-based (lawful) regularities are produced correctly.

Not only does connectionism account for the regularities and irregularities of language but, as Clark (1989: 168-9) points out, phenomena such as developmental aphasia and surface dyslexia seem to support a parallel distributed processing approach to linguistic phenomena. As he states (1989: 169 see above): when a kind of "... damage ... decrements ... connectivity strengths ... " by a certain percentage, it moves "... the

irregular words below the threshold while leaving the originally strong regular pattern functional." Thus, connectionism can account for regular and irregular patterns of language as they occur in developmental aphasia and surface dyslexia. From both of these sources some of the relevant aspects of connectionism can be brought into focus for subsequent application to problems in the study of syntax.

2.2. *Mechanisms in the connectionist model*

The question then arises: What are the mechanisms by which this is accomplished? In the study of the machine learning of verbs, Rumelhart and McClelland used Wickelphones which are a highly redundant form of traditional phonemes and Wickelfeatures which are a highly redundant form of traditional Jakobsonian-type phonetic features. The Wickelphones, for instance, of the English word *pat* would be [#pae] [paet] and [aet#]. Thus, just as phonemes turn out to be "...complexes of features..." (Schane 1968: 712) of the Jakobsonian variety so too these same Jakobsonian features appear in clusters as Wickelfeatures in a connectionist program for generating verb forms (Rumelhart and McClelland 1987: 212). It is, of course, impossible at this point to predict what a connectionist program for the study of syntax might look like ultimately. It is, however, possible to make suggestions on the direction in which such a program might proceed starting with concepts already commonplace for the study of syntax. For this reason, in this study, the assumption will be made that "...syntactic and semantic criteria must go hand in hand in analysis as they do in everyone's perception and production of language." (Hall 1981: 223) Thus, it is expected that the contributions of linguistics to a connectionist program will come from studies which emphasize the relationship of semantics to syntax including those of Dirven (1980), Dixon (1979), Langacker (1987, 1988a, 1988b), Wierzbicka (1988), and Rudzka-Ostyn (1989) and including those which come out of the transformational generative tradition or are influenced by that tradition such as Allerton (1982), Jackendoff (1987), and Wilkins (1988). It is expected that the relationship of semantics to syntax will be expressed through the use of semantic features much like the Wickelfeatures used in the verb morpholygy studies.

The use of these features in the analysis of some well-known syntactic problems in a would-be connectionist program for the study of syntax permits some well-motivated solutions. For instance, in the attempts to formulate a suppletion rule for *some* and *any* in English, Lakoff (1969) has noticed some problems and despaired of ever formulating a rule. The problems are caused by sentences of the following types (among others):

1 a. I warn you that if you eat any candy, I'll whip you.
  b. I promise you that if you eat some candy, I'll give you ten dollars. (The examples are from Lakoff (1969: 611).)

Normally the *some/any* suppletion rule is written so that *some* occurs in affirmative sentences and *any* occurs in negative sentences and questions. But in the above, this is clearly not the case. Although instances such as these may be problematic for conventional grammatical analyses, they are of no problem for other types of grammatical systems. For instance, in a system similar to that suggested by Weinreich (1980) some

years ago, there would be a combination of features which assign a transfer feature, a negativeness factor of which could be given a 'weight' of [-10] in a connectionist system to a sentence with the negative marker *not* in it. The same negativeness factor would be assigned to sentences with verbs such as *warn* in them. Thus, any time a value of this nature ([-10]) is encountered, the use of *any* would be triggered. For the sentences at (1) there would be no problem, but the situation with *some/any* suppletion is even more complex than Lakoff imagined, however, because sentences such as (2) are possible.

2 I promise you that if you eat any candy, I'll whip you.

The solution with *warn* (above) would still apply, however, but this time all of the meaning bearing features of the sentence would be computed, coming up with the same value [-10] and thus triggering the use of *any* for this and all other sentences which imply a threat. That is, sentences such as (1a) or (2) are 'threat' sentences, and regardless of the lexical tokens which are in them these sentences would trigger a numerical value which would in turn trigger the use of *any*. Perhaps information from the context of situation or from the context of discourse, processed in a parallel fashion, would be needed to produce the correct weights. Such a scheme could be implemented quite easily within a connectionist system using parallel distributed processing and Weinreich-type features (or features very much like Wickelfeatures). (See Figure 1, below.)

Lakoff (1969:610) also documents other instances where *some* and *any* are in what the structuralist called contrastive distribution:

3 a. If you eat some spinach, I'll give you ten dollars.
b. If you eat any spinach, I'll give you ten dollars.

In this case, it is quite likely that *any* specifies a smaller quantity than *some* since *any* could be glossed as "any at all" whereas *some* perhaps signifies "a reasonable quantity" from the speaker's viewpoint —differences in meaning selected by the speaker. Thus, the contrast in (3a,b) would be encoded in a semantic level with information about the context of situation or the context of discourse and processed in a parallel fashion. Within a connectionist system, this kind of phenomena would most likely be handled in what, to autonomous syntacticians, is a hybrid model with weights coming in from a semantic area of a processor to be included in the weightings of a network in the syntactic area, the distributed network, where the *some/any* suppletion rule would be located as in the diagram at Figure 1. Such a model fits well within the functionalist (Givón 1984, 1991) or cognitivist mode (Langacker 1987, 1988a, 1988b).

Another area in the study of syntax which can be used to illustrate the utility of a connectionist approach for the study of syntax is the distribution of the gerundive and infinitival complements in English. Following Bolinger (1968) verbs marked with a feature for reification require the gerundive complement. Those marked with a feature for hypothetical action require the infinitival complement. There are, however, apparent exceptions to this general scheme. For example, *happen* is a reifying verb but requires an infinitival complement:

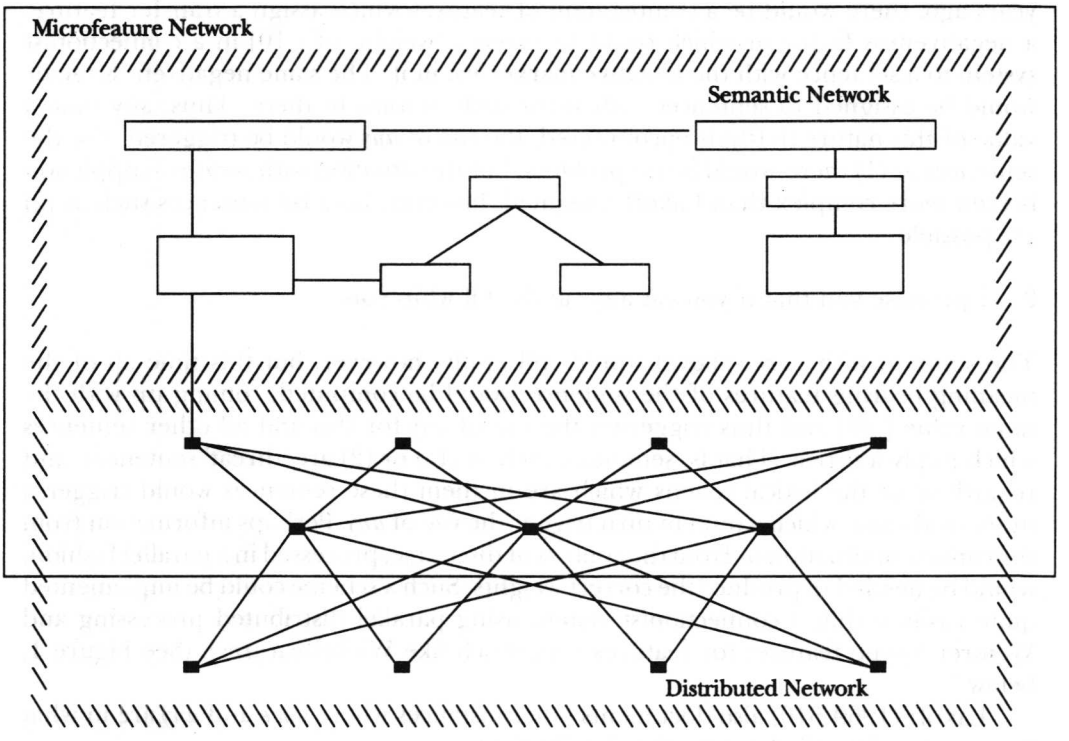

Figure 1: A Hybrid Network Model
(Adapted from Hendler 1989: 233)

4 He happened to find her near the carousel.

(See Wierzbicka (1988:35-6) and Celce-Murcia and Larsen-Freeman (1983:434-436) for a discussion of these.) For these and similar sentences, it is apparent that the complement is chosen on the basis of aspectual markers following Rosenbaum's suggestion (1967). In particular, *happen* requires that the complement be marked for the punctiliar aspect. So it is that when a lexical item such as *like* can take either the gerundive or the infinitival complement, the choice is made on the basis of whether a durative or a punctiliar marker is required by the higher verb of the complement as in the sentences at (5):

5 a. He likes going downtown.
b. He likes to go dowtown.

Since, in this case, the choice of the gerundive or infinitival complement depends on native speaker intention as in the case of *some* and *any* in sentences (3a,b) a similar strategy would have to be used for their production or description, that is, for sentences such as those at (4) and (5a,b). Within a connectionist system, this kind of phenomena

would be described with weights coming in from a semantic area of a processor (or an entirely different processor) which would also be dedicated to information about the context of situation or the context of discourse. These weights would override the weightings in the area where the hypothesis / reification distinction is located, perhaps a distributed network dedicated to syntax as in Figure 1. In a linguistic program for the study of syntax within a connectionist mode some such system could be easily devised since connectionist systems operate in a parallel fashion.

A more extensive example of the same principles can be provided from an on-going study (Ney 1990). In this study, features, developed on the basis of dictionary entries in the OED, are assigned to a verb such as *break* in an attempt to predict the syntactic patterning of the verb from these features. In the chart below, features are assigned to different meanings of the verb, *break,* on the basis of the nouns which occur with this verb in focus roles (usually preverbal) or patient roles (usually postverbal) positions. The absence of any of these features is assigned a negative number; the presence, a positive number.

Chart I: Features and Feature Values for Meanings of *break.*

Feature	Value		Feature	Value	
	Pres.	Abs.		Pres.	Abs.
Focus Role Second Rank			Patient Role Second Rank		
[SELFACT]*	+1	−1	[LONG]	+5	−5
[MACHINE]	+2	−2	[ROTUND]	+6	−6
	+2	−2	[NONSPEC]	+7	−7
Focus Role First Rank [ANIMATE]	+3	−3	Patient Role First Rank [PHYSICAL]	+8	−8
			[CONCRETE]	+9	−9
Patient Role Third Rank [ACTIVITY]	+4	−4	[FRAGIBLE]	+10	−10

* [SELFACT] = self actualizing as in *The storm was breaking.* [NONSPEC] - the patient role is not specified as to its shape. The positive value of a feature indicates its presence; the negative value of a feature indicates its absence. Thus, in *He broke his promise, He* is animate [+3], but it is not at all self-actualizing [-1], and *promise* is not breakable in a physical sense [-FRAGIBLE] [-9], not concrete [-8], not physical [-7]; neither is it an activity [-4]. As per the above, the most important features, impressionistically, have the greatest weights. (See Chart II.)

As per Chart II, the literal meanings of *break,* 'snap,' 'burst' and 'shatter', have positive scores; the figurative meanings have negative scores with some highly figurative, *break a promise* (-33) and *break a conspiracy* (-27) and others not so highly figurative, *break a hold* (-19) (in wrestling) and *break a trail* (-19). Here the term figurative follows the use established by Ruhl (1979) where the entities behind lexical items such as *promise* or *conspiracy* are not physical in the sense of the entities behind lexical items such as *glass* or *twig* and thus are figurative in some sense of the word or another. Somewhere in

Chart II: Representative Feature Matrices and their Values for Meanings of *break*.

Break 'snap' She broke a pencil.		Break 'burst' She broke a balloon.		Break 'shatter' She broke a wine glass.		Break a trail He broke a trail in the snow.	
[-SELFACT]	−1	[-SELFACT]	−1	[-SELFACT]	−1	[-SELFACT]	−1
[-MACHINE]	−2	[-MACHINE]	−2	[-MACHINE]	−2	[-MACHINE]	−2
[+ANIMATE]	+3	[+ANIMATE]	+3	[+ANIMATE]	+3	[+ANIMATE]	+3
[-ACTIVITY]	−4	[-ACTIVITY]	−4	[-ACTIVITY]	−4	[-ACTIVITY]	−4
[+LONG]	+5	[-LONG]	−5	[-LONG]	−5	[-LONG]	−5
[-ROTUND]	−6	[+ROTUND]	+6	[-ROTUND]	−6	[-ROTUND]	−6
[-NONSPEC]	−7	[-NONSPEC]	−7	[+NONSPEC]	+7	[+NONSPEC]	+7
[+PHYSICAL]	+8	[+PHYSICAL]	+8	[+PHYSICAL]	+8	[+PHYSICAL]	+8
[+CONCRETE]	+9	[+CONCRETE]	+9	[+CONCRETE]	+9	[-CONCRETE]	−9
[+FRAGIBLE]	+10	[+FRAGIBLE]	+10	[+FRAGIBLE]	+10	[-FRAGIBLE]	−10
Total	24	Total	18	Total	19	Total	−19
Break a conspiracy.		**Break a promise.**		**Break 'inoperable' It broke.***		**Break a hold in wrestling.**	
[-SELFACT]	−1	[-SELFACT]	−1	[+SELFACT]	+1	[-SELFACT]	−1
[-MACHINE]	−2	[-MACHINE]	−2	[+MACHINE]	−2	[-MACHINE]	−2
[+ANIMATE]	+3	[+ANIMATE]	−3	[-ANIMATE]	−3	[+ANIMATE]	+3
[-ACTIVITY]	+4	[-ACTIVITY]	+4	[-ACTIVITY]	−4	[-ACTIVITY]	−4
[+LONG]	−5	[-LONG]	−5	[+LONG]	−5	[-LONG]	−5
[-ROTUND]	−6	[-ROTUND]	−6	[-ROTUND]	−6	[-ROTUND]	−6
[-NONSPEC]	+7	[-NONSPEC]	+7	[-NONSPEC]	+7	[+NONSPEC]	+7
[-PHYSICAL]	−8	[-PHYSICAL]	−8	[+PHYSICAL]	+8	[+PHYSICAL]	+8
[-CONCRETE]	−9	[-CONCRETE]	−9	[+CONCRETE]	+9	[-CONCRETE]	−9
[-FRAGIBLE]	−10	[-FRAGIBLE]	−10	[+FRAGIBLE]	+10	[-FRAGIBLE]	−10
Total	−27	Total	−33	Total	15	Total	−19

*Technically, although *it* in the sentence *It broke* (e.g. *The electric canopener broke*) is in a focus role, it still functions in a patient role in this as in related sentences (e.g. *Jack broke a canopener*) and thus is marked for patient role features.

between these two extremes are the entities behind such lexical items as *a hold* (in wrestling) and *a trail* which although they are not physical in the sense of *glass* or *twig* have instantiation in the physical world in a way that *promise* and *conspiracy* do not. The values (weights) assigned to their particular features show this with the more figurative receiving a more negative weight.

Now, it turns out that these values (or weights) have important consequences in the syntax of English as traditionally construed. Thus it is interesting to note that the three meanings of the surface verb *break* having the greatest positive weights have a peculiar distribution with the progressive tense marker. Thus, in most environments, **I'm breaking a glass* would be ungrammatical since the verb, *break,* is perfective in its meaning. In these environments, the term **halfbroken* would be very unusual, to say the least. If something is broken, that's that — it's broken. In this respect *break* is different from verbs such as *cut* and *tear* as Morreal has pointed out "... *cut* and *tear*

imply an ongoing, progressive division, while *break* does not..." (Morreal as cited in Ruhl (1979:203)). Similarly, the action in the meaning of the verb, *break,* is punctiliar. The act of breaking happens in an instant. As a result, in most environments, the surface verb *break* is not normally compatible with the progressive tenses.

6 a. *John is breaking a wine glass.
b. *The wine glass is being broken.

This is perhaps less true of the figurative meanings of *break* where the following sentences seem to be acceptable:

7 a. John is breaking his promise again.
(Cf. ?*John is breaking a glass again.)
b. The conspiracy is breaking up.

From the first of the above (7a), it should be apparent that if the action is repeated (iterative) then the progressive occurs more readily in the sentence. Thus, the progressive tenses can be used with *break* as in the following:

8 a. Since John is going on the wagon, he is right now breaking every wine glass in the house.
b. Believe me; the wine glasses are being broken.
(This would be used in the context of (8a) above.)

Thus if the action in the verb is iterative, the progressive tense marker can be used with the verb *break.*

9 a. The waves are breaking against a rock.
b. The otters are swimming along underwater, breaking the surface every once in a while.
c. Their ranks are breaking like thin clouds before a Biscay gale. OED 1824 Macaulay Ivry 43
d. John is breaking every doll in the house.
e. He has a reputation for breaking mechanical things.
f. The glacier was evidently breaking beneath our feet. OED 1860 Tyndall Glac. ii. ^ 17. 317
g. Like hounds breaking up a fox. OED 1875 Buckland Log-bk. 155
h. The Army breaking, My husband hies him home. OED 1601 Shaks. All's Well iv. iv. 11
i. A Ring of Cudgel-Players..breaking one another's Heads. OED 1711 Budgell Spect. No. 161 <page> 3
j. That tumult in the Icarian sea, dashing and breaking among its crowd of islands. OED<ante> 1744 Pope (J.)

This is especially noticeable in sentences dealing with liquids; where the punctiliar nature of the action in *break* is still present, the iterative nature of the action permits the use of the progressive tenses with these verbs. In (9 a,b,j), the act of breaking is viewed as a series of acts and the underlying verb is marked [+ITERATIVE] so that the progressive marker is permissible with the surface verb *break.* Even if the object of the verb is singular, iterative or continuous action in the verb can permit or require

the progressive as in (9c,f,h). Another way to look at this phenomenon is to see it as dealing with underlying plural (9c,d) or massive (9f,h) semantic entities as in (9d,f) even though none of the nouns are marked syntactically for the plural. Thus, when the surface verb *break* is marked with an iterative aspect or associated with plural of massive semantic entities, it can occur with the progressive aspect.

For phenomena such as this linguists such as Weinreich have recommended the use of transfer features, as previously mentioned, which, in a connectionist scheme, would be interpreted as various weights in the differing matrices. As a result, a value of [-15] could be assessed for 'plural object', 'iterative action' or 'massive entity' (such as army or glacier). For this reason, verb matrix values associated with the surface verb *break* less than [9] would be more inclined to occur with the progressive than those with values greater than [10]. Transfer features would thus be applied to the literal meanings of *break* with a value of [20] or higher to account for the occurrences of the progressive in the sentences at (9). The non-literal meanings of *break* with matrix values of [-19] or less would not need the application of such features but would more readily occur with the progressive as a theory formed on the basis of this reasoning would predict. Thus, *break*, 'shatter', normally having a weight of [19], would, in the context of *glacier,* a massive entity, have a weight of [5] accounting for the sentence at (9f). Similar statements could be made about *army* (9h) and *sea* (9j). In the model suggested here, these values would be introduced into the syntactic network from a semantic network with information from the context of situation or the context of discourse or produced in the syntactic network by scanning all of the nodes in that network and computing their various weights in a parallel fashion (as in Figure 1).

To this point, the discussion has focused on the architecture of connectionist systems which utilize parallel distributed processing and how this architecture is readily adaptable to the description of a solution to one type of problem in syntax. Another problem in syntactic description that is frequently referred to in discussions of connectionism is the problem of recursion (Fodor and Pylyshyn 1988: 33 and 14, note 9, Gelder 1990: 358). Most often, these discussions arise out of the assumption that current versions of transformational generative grammar can account for recursion better than other grammars, an assumption which may or may not be supported in principle in the absence of any criteria to state what constitutes 'proof' in intra-paradigmatic debates. Thus, for instance, Lockwood devotes a page to recursion in his *Introduction to Stratificational Linguistics* and thus demonstrated that recursion is not problematic for this type of linguistic analysis. Quite apart from the fact that Chomsky has recently renounced generativity as a goal for transformational generative grammar (Chomsky 1986:3. See also McCawley 1988: 355) and quite apart from the fact that the recursive properties of language (and subsequently, linguistic description) are used to account for the fact that the "...sentences of a language constitute an infinite set...", other grammars account for the recursive properties of language quite well. For instance, Longacre (1964: 26-31) described this characteristic of tagmemic grammars some time ago. The Longacre argument was picked up by Ney (1968) and expanded. In essence, tagmemic grammars account for recursiveness by developing a string such as A_1 B_1 C_1 and repeating that string with different slot fillers in a slot D_1 where D_1 is filled by a string A_2 B_2 C_2 so that A_1 B_1 C_1 D_1 is realized as A_1 B_1 C_1 A_2 B_2 C_2 (A_1

B_1 C_1 = John loves Mary. A_2 B_2 C_2 = Mary loves Jack. A_1 B_1 C_1 A_2 B_2 C_2 = John loves Mary who loves Jack.)

Now if a connectionist model were built on a tagmemic organization, it could easily produce strings requiring recursion by reserving a particular area in memory for strings of a given nature and addressing that area repeatedly until no further need for recursion was reached. This kind of procedure is implicit in the architecture of connectionist systems and would not require deviation from normal connectionist models. In the type of system suggested above, strings of (syntactic/ semantic) features developed in a given memory location would be addressed repeatedly until the requirement for even multiple embeddings were fulfilled.

One other aspect of connectionism which would facilitate the description of syntactic information is the probabilistic fashion in which connectionist systems operate. That is, since the experimentation of Quirk and his colleagues (Greenbaum and Quirk 1970) and since the work of Labov and his colleagues, it is quite widely known that some aspects of language can best be described in a probabilistic fashion. Hence, Labov (1966, 1968, 1969) has advocated the notion of the variable rule which for any given syntactic phenomenon might mean that a rule is in effect for 80% of the time while for 20% of the instances it is not in effect. This is true also for any individual's rule use. It is this kind of phenomenon that has prompted Rumelhart and McClelland (1987) to describe some aspects of human behavior as 'lawful' rather than rule-governed and it is this kind of phenomenon for which connectionist systems are so well adapted. Quite apart from the controversy that variable rules have caused within the linguistic establishment (Romaine 1981, 1985, Kroch 1985), the phenomena described by these rules are intrinsic to the nature of these systems. The form of knowledge within these systems is of the substance of variable rules; in short, knowledge is represented in a probabilistic fashion. In fact, any information in the networks suggested above would be in this form. As a result, any linguistic program which is designed to fit into connectionism could easily be cast in the form suggested by sociolinguistic research.

In conclusion, two trends in contemporary linguistics are peculiarly adaptable to a connectionist system. One is the probabilistic nature of certain aspects of language from the sociolinguistic (Labovian) position or from the empirical research of Quirk and his colleagues. The other is the structure of certain aspects of language which are peculiarly amenable to description within a connectionist system such as the difference between regular and irregular syntactic forms. In this manner, it is the architecture of such systems which contributes to a ready solution to problems of the description of language and thus to notes towards a connectionist program for the study of syntax.

References

ALLERTON, D.J. (1982). *Valency and the English verb.* New York: Academic Press.

BOLINGER, D. (1968). Entailment and the meaning of structures. *Glossa* 2: 119-127.

BOTHA, R.P. (1979). External evidence in the validation of mentalistic theories: A Chomskyan paradox. *Lingua* 48: 299-328.

CELCE-MURCIA, M. and D. Larsen-Freeman. (1983). *The grammar book.* Cambridge, MA: Newbury House.

CELLERIER, G. (1980). Some clarifications on innatism and constructivism. In Piattelli-Palmarini (Ed.), 83-89.

CHOMSKY, N. (1957). *Syntactic structures.* 'S-Gravenhage: Mouton.

CHOMSKY, N. (1986). *Knowledge of language: Its nature, origin and use.* New York: Praeger.

CLARK, A. (1989). *Microcognition, cognitive science and parallel distributed processing.* Cambridge, MA: MIT Press.

DIRVEN, R. (1980). Towards a semantic syntax of English: The case of the modals *can* and *may*. *Papers and studies in contrastive linguistics* 12: 99-147.

DIXON, R.M.W. (1979). Ergativity. *Language* 55: 59-138.

FODOR, J.A. and Z.W. PYLYSHYN. (1988). Connectionist and cognitive architecture: A critical analysis. *Cognition* 29. 3-71.

GELDER, T. VAN. (1990). Compositionality: A connectionist variation on a clasical theme. *Cognitive Science* 14: 355-384.

GIVÓN, T. (1984). *Syntax: A functional-typological introduction.* Volume I. Amsterdam: John Benjamins.

GIVÓN, T. (1991). Isomorphism in the grammatical code: Cognitive and biological considerations. *Studies in language* 15(1): 85-114.

GREENBAUM, S. and R. QUIRK. (1970). *Elicitation experiments in English: Linguistic studies in use and attitude.* London: Longman.

HALL, R.A. (1981). Can linguistics be a science? *Lingua:* 53. 221-226.

HENDLER, J.A. (1989). Problem solving and reasoning: A connectionist perspective. In Pfeifer and others (Eds.), 229-244.

HULL, C.L. (1943). *Principles of behavior: An introduction to behavior theory.* Boston: Scribner.

HUNT, E. (1989). Connectionist and rule-based representations of expert knowledge. *Behavior research methods, instruments and computers* 21: 88-95.

JACKENDOFF, R. (1987). The status of thematic relations in linguistic theory. *Linguistic Inquiry* 18: 369-411.

KROCH, A.S. (1985). Review of S. Romaine *Socio-historical linguistics: Its status and methodology.* (Cambridge: University Press, 1982). *Language* 61: 698-704.

LABOV, W. (1966). *The social stratification of English in New York City.* Washington, DC: Center for Applied Linguistics.

LABOV, W. (1968). *A study of the non-standard English of negro and Puerto Rican speakers in New York City.* Volume I. ERIC ED 028 423.

LABOV, W. (1969). Contraction, deletion and inherent variability of the English copula. *Language* 45: 712-762.

LACHTER, JOEL and T.G. BEVER. (1988). The relation between linguistic structure and associative theories of language learning: A constructive critique of some connectionist learning models. *Cognition* 29: 195-247.

LAKOFF, R. (1969). Why there can't be any *some / any* rule. *Language* 45: 608-615.

LANGACKER, R.W. (1987). The cognitive perspective? *Newsletter of the Center for Research in Language* 1(3) (Feb. 1987): 1-7.

LANGACKER, R.W. (1986a). An overview of cognitive grammar. In Rudzka-Ostyn (Ed.), 3-48.

LANGACKER, R.W. (1988b). A view of linguistic semantics. In Rudzka-Ostyn (Ed.), 48-90.

LEAHEY, T.H. (1992). The mythical revolutions of American psychology. *American Psychologist* 47(2): 308-318.

LONGACRE, R. (1964). *Grammar discovery procedures.* The Hague: Mouton.

LOCKWOOD, D. (1972). *Introduction to stratificational linguistics.* New York: Harcourt Brace and Jovanovich.

MACWHINNEY, B. (Ed.). (1987). *Mechanisms of language acquisition.* Hillsdale, NJ: Lawrence Erlbaum Associates.

MACCAWLEY, J. (1988). Rev. of Noam Chomsky, *Knowledge of language: Its nature, origin and use. Language* 64: 355-365.

MCCLELLAND, J.L. and D.E. RUMELHART. (Eds.). (1986). Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 2. Cambridge, MA: MIT Press.

MCCORNMACK, W.C. and H.J. IZZO (Eds.). (1979). The *Sixth LACUS Forum.* Columbia SC: Hornbeam Press.

NEWMEYER, F.J. (1983). *Grammatical theory.* Chicago: The University of Chicago Press.

NEY, J.W. (1968). Transformational analysis. *Linguistics* 45: 36-49.

NEY, J.W. (1988). *Transformational grammar: Essays for the left hand.* Lake Bluff, Illinois: The Jupiter Press.

NEY, J.W. (1990). Polysemy and syntactic variability in the surface verb *break.* Paper read at the *Seventeenth LACUS Forum,* California State University, Fullerton, Aug. 10, 1990.

PFEIFER, R., Z. SCHRETER, F. FOGELMAN-SOULIE, and L. STEELS. (Eds.). (1989). *Connectionism in perspective.* Amsterdam: North Holland.

PIATELLI-PALMARINI, M. (Ed.). (1980). *Language and learning: The debate between Jean Piaget and Noam Chomsky.* Cambridge, MA: Harvard University Press.

ROMAINE, S. (1981). On the status of variable rules in sociolinguistic theory. *Journal of Linguistics* 17: 1-119.
ROMAINE, S. (1985). Variable rules, O.K.? Or can there be sociolinguistic grammars? *Language and Society* 5: 53-67.
ROSEMBAUM, P. (1967). *The grammar of English predicate complement constructions.* Cambridge, MA: MIT Press.
RUDZKA-OSTYN, B. (Ed.). (1988). *Topics in cognitive linguistics.* Amsterdam: John Benjamins.
RUDZKA-OSTYN, B. (1989). Prototypes, schemas, and cross-category correspondences: The case of *ask*. *Linguistics* 27: 613-661.
RUHL, CH. (1979). The semantic field of *break, cut* and *tear*. In W.C. McCornmack and H.J. Izzo (Eds.), 200-214.
RUMELHART, D.E. and J.L. MCCLELLAND. (1987). Learning the past tense of English verbs: Implicit rules or parallel distributed processing. In B. MacWhinney (Eds.), 195-248.
SAMPSON, G. (1987). Rev. article: *Parallel distributed processing: Explorations in the microstructures of cognition.* Ed. by D.E. Rummelhart, J.L. McClelland and the PDP Research Group. *Language* 63: 871-886.
SCHANE, S. (1968). On the non-uniqueness of phonological representations. *Language* 44: 709-716.
SCHNEIDER, W. (1987). Connectionism: Is it a paradigm shift for psychology? *Behavior Research Methods, Instruments and Computers* 19: 73-83.
UHLENBECK, E.M. (n.d.). *Critical comments on transformational generative grammar.* The Hague: Smits, Drukkers-Uitgevers B.V.
VINCENT, N. (1978). Towards an explanation of some analogies in perfect and past participle stems in Latin and Italian. *Semasia* 5: 47-64.
WEINREICH, U. (1980). *On semantics.* Ed. by W. Labov and B. Weinreich. Philadelphia: University of Pennsylvania Press.
WIERZBICKA, A. (1988). *The semantics of grammar.* Amsterdam: John Benjamins.
WILKINS, W.W. (Ed.). (1988). *Thematic relations: Syntax and semantics,* Vol. 21. San Diego: Academic Press.
YNGVE, V. (1986). *Linguistics as a science.* Bloomington: Indiana University Press.
YNGVE, V. (1990). Green cheese and the burden of proof in science. *Language Sciences* 12(2/3): 285-289.