Title
Compositionality and Systematicity in Connectionist Language Learning

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 14(0)

Author
Hadley, Robert F.

Publication Date
1992

Peer reviewed

Compositionality and Systematicity in Connectionist Language Learning

Robert F. Hadley
School of Computing Science
Simon Fraser University
Burnaby, B.C. V5A 1S6
hadley@cs.sfu.ca

Abstract

In a now famous paper, Fodor and Pylyshyn (1988) argue that connectionist networks, *as they are commonly constructed and trained*, are incapable of displaying certain crucial characteristics of human thought and language. These include the capacity to employ *compositionally structured representations* and to exhibit *systematicity* in thought and language production. Since the appearance of Fodor and Pylyshyn's paper, an number of connectionists have produced what seem to be *counter-examples* to the Fodor-Pylyshyn thesis. The present work examines two of these apparent counter-examples; one is due to Elman and the other to St. John and McClelland. It is argued that although Elman's and St. John & McClelland's networks discover a *degree* of compositionality, and display a degree of systematic behaviour, the degrees involved are substantially less than that found in humans, and (consequently) are less than what Fodor & Pylyshyn require (or presumably would require if the question were put to them).

1. Introduction

In a now famous paper, Fodor and Pylyshyn (1988) argue that connectionist networks, *as they are commonly constructed and trained*, are incapable of displaying certain crucial characteristics of human thought and language. These include the capacity to employ *compositionally structured representations* and to exhibit *systematicity* in thought and language production.[1] Since the appearance of Fodor and Pylyshyn's paper, an number of connectionists have produced what seem to be *counter-examples* to the Fodor-Pylyshyn thesis. In the present work I examine two of these apparent counter-examples; one is due to Elman (1990), the other to St. John and McClelland (1990). I have chosen these works because, on the face of it, both constitute *strong* counterexamples, and because both are directly concerned with language acquisition, which is a focal point of my discussion here. In (Hadley, 1992) I examine four other recent, apparent counter-examples (due to Pollack, Smolensky, Small, and Chalmers.) As will emerge, I argue that although Elman's and St. John & McClelland's networks discover a *degree* of compositionality, and display a degree of systematic behaviour, the degrees involved are substantially less than that found in humans, and (consequently) are less than what Fodor & Pylyshyn require (or presumably would require if the question were put to them).

2. Compositionality and Systematicity

In this section I examine two experiments which establish, to varying degrees, that connectionist networks (hereafter, c-nets) can discover the compositionality implicit in a training corpus of *sentences*. When describing the results of these learning experiments, researchers commonly argue from the fact that a network can correctly process *novel* sentences (not contained in the training corpus) to the conclusion that the network has indeed induced a compositional structure, and as a consequence is able to exhibit a degree of *systematicity*. As we examine the c-net experiments described below, it will be useful to distinguish different degrees of systematicity, according to the *degree of novelty* of sentences which a c-net is able to recognize (given the c-net's training regime). I shall distinguish three degrees of systematicity. No doubt, it would be possible to make even finer distinctions, but for our purposes the following should suffice. The degrees of systematicity are: weak, quasi, and strong.

1) *Weak Systematicity*. Networks exhibiting weak systematicity can perform at least the following kind of generalization: Suppose that a training corpus is "representative" in the sense that every *word* (noun, verb, etc.) that occurs in some sentence of the corpus also occurs (at some point in the training corpus) in every permissible syntactic position. Thus, although the training corpus omits some sentences permitted by the target grammar, any network trained on this corpus will have been trained to recognize every word in every syntactic position that the word will occupy in the set of novel test sentences which are used to demonstrate the network's generalization capacity. Assuming that this set of novel sentences contains only sentences which are syntactically isomorphic to sentences in the

[1] 'Compositionality' here presupposes that representations have a combinatorial syntax and semantics, whereas 'systematicity' refers to the systematic relationships which *result* when such combinatorially structured representations are employed.

training corpus, and that no new vocabulary is present, we shall say that a c-net exhibits at least *weak systematicity* if it is capable of successfully processing (by recognizing or interpreting) novel test sentences, once the c-net has been trained on a corpus of sentences which are *representative* in the sense described above. I describe such c-nets as (at least) weakly systematic in order to reflect the fact that their generalization capacity has only been tested upon sentences which are *weakly novel* with respect to the training corpus.

2) *Quasi-Systematicity.* We shall say that a system exhibits only quasi-systematicity if (a) the system *can* exhibit weak systematicity, (b) the system successfully processes novel sentences containing embedded sentences, such that both the larger containing sentence and the embedded sentence are (respectively) structurally isomorphic to various sentences in the training corpus, (c) for each successfully processed novel sentence containing a word in an embedded sentence (e.g., 'Bob knows that Mary saw *Tom*') there exists some *simple* sentence in the training corpus which contains that same word in the same syntactic position as it occurs within the embbeded sentence (e.g., 'Jane saw *Tom*'). A system would be *merely* quasi-systematic if 'Tom' needed to occur (in the training corpus) in the *object position* of a *simple sentence*, before the system could correctly process embedded occurrences of 'Tom' in object position. Analogous remarks apply to subject position, verb position, etc.

3) *Strong Systematicity.* We shall describe a system as strongly systematic if (i) it *can* exhibit weak systematicity, (ii) it can correctly process *simple* novel sentences containing words in positions where they *do not appear* in the training corpus (i.e., the word within the novel sentence does not appear *in that same syntactic position* within any *simple or embedded* sentence in the training corpus). Note that a system which has not been trained on embedded clauses may still exhibit strong systematicity, because neither condition (i) or (ii) requires that embedded sentences be present.

Having now distinguished three degrees of systematicity, I should emphasize that although these degrees are directly related to issues of learnability, their primary relevance to the Fodor-Pylyshyn controversy stems from the fact that degrees of novelty are at issue. That novelty is the central issue is underscored by the fact that Elman and others (cf. Hadley, 1992) base their claims to have undermined the Fodor-Pylyshyn thesis upon evidence that particular c-nets can process sentences which are novel with respect to training sets. Since these authors (and St. John and McClelland) take the ability to process novel input as evidence for generalization and systematicity, it seems fair turnabout and epistemically sensible to suppose that the ability to process various degrees of novelty should be taken as indicative of various degrees of systematicity having been induced. Moreover, quite apart from the Fodor-Pylyshyn controversy, I contend that the trifold distinction introduced here illuminates important differences between the respective abilities of humans vs. existing c-nets to process novel *kinds* of input. To establish this thesis, I shall first argue that humans exhibit the strongest of my three forms of systematicity.

To begin with, there is good reason to believe that even young children, who have not yet reached the stage of producing multi-word utterances, are frequently able to obey *simple* imperative sentences which contain words in syntactic positions where the child has never encountered the word before.

It is well known, for example, that in the few weeks which precede a child's first multi-word utterances, a "spurt" occurs in a child's acquisition of *nominals* (both common and proper nouns), and that during this period children are able *rapidly* to acquire the use of nominals by means of "what's that" games (cf. (Ingram, 1989; Dromi, 1987)). Once they have acquired nominals in this fashion, children are soon thereafter (i.e., within minutes) able to comprehend these words in sentences they encounter. This fact is established by Katz, Baker, and Macnamara (1974) who also present a strong case that the ability of young children to distinguish proper nouns from common nouns is much more a function of a child's *prior* ability to distinguish re-identifiable individuals from classes of objects than it is a function of some capacity to distinguish words which are syntactically preceded by an *article* from those which are not.

Moreover, as children begin to produce simple, multi-word utterances, they will often produce semantically reasonable, *albeit non-grammatical* combinations of the words they have acquired in previous contexts. In fact, children do this sufficiently often that some psycholinguists posit the existence of a "child grammar" (Ingram, 1989). Now, whether or not we accept the existence of a child grammar, the fact that children are able to recombine words in patterns that are not present in their training corpus strongly suggests that (at least in the early weeks of multi-word utterance) children have a much greater grasp of the *semantic* content of particular words than they do of their syntactic roles (in adult grammars). (The results of Katz, Baker, and Macnamara also reinforce this conclusion.) Furthermore, and more to the point, the ability of children to *sensibly* recombine words in patterns they have not been trained to produce clearly demonstrates that children are not nearly as dependent upon syntactic context as systems which are only weakly (or quasi) systematic.

As we consider somewhat older children, who have acquired a rudimentary syntax (but not necessarily the use of prepositional phrases or relative clauses), it becomes transparently clear that humans can learn to use nominals long before they have encountered them in all possible positions. For example, a child visiting a zoo with her parents may hear her mother exclaim, "Susie, look at the otter". Susie may reply, "What's an otter?" The mother, pointing, replies "Here, this is an otter". If Susie is adept at language, she may learn the (approximate) meaning of 'otter' rapidly, by this ostensive means, and may soon utter, "Look, Mommy, this otter is chasing the other one". Although the child has never encountered the word in *subject* position, she is able to use it in that position once its meaning has been surmised. Of course, most children will require a few repetitions before an ostensively introduced word enters long term memory, but these repetitions need not present the word in all legal positions. With adults, new words may enter the vocabulary even more rapidly, as when one surmises a word's meaning during the course of conversation, or when listening to a brief exchange during a meeting. Once a word's meaning has been surmised, most adults can use it freely in embedded sentences and simple sentences, although they may only have heard the word used in a single syntactic position.

We turn now to consider connectionist systems which, *prima facie*, challenge Fodor and Pylyshyn's view on the limitations of c-nets vis-a-vis compositionality and systematicity. In considering these systems we should bear in mind that Fodor and Pylyshyn are concerned with the kind of full-fledged compositionality and systematicity that human thought and language exhibit.

2.1 St. John and McClelland

St. John and McClelland (1990) present a connectionist model which learns to assign "semantic representations" to English sentences which are presented as input. Although the details of their model are somewhat complex, the overall gist is that, via backpropagation, the network is trained to produce a correct semantic representation of the situation *described* by each input sentence. Situations (or events) described by input sentences consist of relationships, and the objects involved in those relationships. Input sentences are fed into the network in presegmented constituents. As each constituent is processed, an inspection is made to see whether the network has output the desired, complete representation of the target situation. Backpropagation is performed after each such inspection. Because it is usually not possible to predict the *entire* target representation on the basis of isolated sentence constituents, the network is forced to learn associations between individual constituents and particular objects or relations in the target situation.

Sentences which serve as input constitute a highly simplified version of English, in that all articles are deleted and only singular nouns are present. However, certain prepositional phrases are permitted. Each target semantic representation consists of an *ordered* series of role/filler pairs. Roles are *agent, action, patient,* etc., and fillers are "concepts" (*my* scare quotes) corresponding to individual nouns and verbs. Thus, each semantic representation is a *structured, concatenated* sequence of pairs. By itself, this aspect of the model would seem to undermine any potential the model might possess for deposing Fodor and Pylyshyn's thesis that human thought requires structured, internal representations. For the experimental design presupposes the existence of such representations (at the point where backpropagation is employed). Moreover, the ability to form such representations *presupposes* that the learner has *already discovered* a compositional, systematic method of representing situations. Thus, from Fodor and Pylyshyn's standpoint, the model's design concedes one of their major contentions. However, the question still remains whether the model acquires knowledge of the compositional, systematic nature of its *input* sentences. St. John and McClelland (hereafter, St.J&Mc) clearly claim that it does (p. 250, 1990), and it is this claim we now consider.

As mentioned, St.J&Mc's training corpus includes sentences containing prepositional phrases. Unfortunately, when testing their network for the acquisition of compositional knowledge (which is manifested as *systematicity*) St.J.&Mc used simpler training corpora, which lacked prepositional phrases. Two experiments were conducted to test for systematicity of behaviour - one syntactic, the other semantic. Tests for syntactic systematicity involved only 10 objects and 10 reversible actions. Each object (action) uniquely corresponds to a particular noun (verb) in the training corpus. Both active and passive verb forms were permitted, and each input sentence had the general form: [noun verb-form noun]. Given that both active and passive forms are possible, a total of 2000 sentences are possible. All 2000 sentences were generated. Of these, 1750 comprised the training corpus, and the remaining 250 were set aside for later testing. Although St.J&Mc do not explicitly say so, their remarks elsewhere (p. 243, 1990) suggest that these 250 sentences were randomly selected. Assuming they were, it is highly probable that the remaining 1750 sentences contained occurrences of every word in every legal syntactic position. (Otherwise, 80% of the 250 test sentences would have to contain the *same* particular noun or verb in the same syntactic position. Given that there are 10 nouns and 10 verbs, this is extremely unlikely.[2]) Moreover, St.J&Mc give *no indication* that the training corpus (of 1750 sentences) *does not* include every possible word in every possible position. On the available evidence, therefore, it is reasonable to believe that the training corpus does present every word in every possible position. This conclusion is reinforced by St.J&Mc's remark that "What makes this a generalization task is that some of the sentences were set aside and not trained: *some agents were never paired with certain objects*" (my emphasis). The fact that sentences in the *test* corpus describe novel agent-object combinations does present convincing evidence of generalization, but does not suggest that anything stronger than weak systematicity and compositionality were tested for. To be sure, the network does display some degree of systematicity. The network assigns the correct semantic representation to 97% of the novel 250 sentences. However, given the above considerations, it seems entirely likely that the network displays only *weak syntactic* systematicity.

The test for *semantic* generalization is analogous, in relevant respects, to the one just described. The se-

[2] Note that 10% of the 2000 original sentences contain a given noun or verb in a given position. So, if a given noun or verb does not occur in a given position within the 1750 training sentences, then 200 of the 250 test sentences must contain that given word in the given position.

mantic test involved a set of 400 possible sentences, of which 350 were used for training and the remaining 50 were used for testing. St.J&Mc explicitly note that the 50 test sentences were randomly selected from the set of 400. As before, the set of 400 sentences exhausts the space of possible sentences. Now, since the 50 test sentences were randomly chosen, it is extremely probable (by analogy with the reasoning given in the previous footnote) that each word occurred in a syntactic position within the *test* corpus that it also occupied within the training corpus. Thus, it is virtually certain that the test for *semantic* generalization established only weak systematicity. Certainly, we are given no reason to suppose otherwise. Also, it is clear that St.J&Mc's model was not even intended to display the kind of strongly systematic behaviour and rapid integration of semantic knowledge which our example of the child at the zoo illustrates (involving the word 'otter').

It should be acknowledged, however, that despite the weaknesses mentioned above, the network we have considered yields some impressive results, including the ability to learn "to disambiguate ambiguous words; instantiate vague words; assign thematic roles; and immediately adjust its interpretation as each constitutent is processed" (p. 220, 1990). Even the ability to demonstrate weak systematicity is no small feat. However, it should be remembered that humans appear to exhibit a much stronger form of systematicity than this.

2.2 Elman

We turn now to the work of Elman (1989, 1990) on connectionist learning of syntactic structure. Elman contends that "the sensitivity to context which is characteristic of many connectionist models, and which is built-in to the architecture of the networks used here, does not preclude the ability to capture generalizations which are at a higher level of abstraction." In addition, Elman clearly opposes his results (and those of others, including St.J&Mc) to the conclusions advanced by Fodor and Pylyshyn (1988), and to Fodor's (1976) Language of Thought thesis. Yet, while it is clear that Elman's networks do generalize and acquire a degree of systematicity, it is by no means clear that they display the degree of systematicity that humans exhibit. Moreover, since Elman's research does not address issues of *semantic* systematicity and compositionality, it is unclear whether this work actually threatens Fodor's views on the Language of Thought. After all, we saw that St.J&Mc were able to train their network to discover *semantic* compositionality only when they assumed the prior existence of a concatenative, structured set of internal representations. However, let us consider Elman's results in some detail.

Elman (1989, 1990) describes two experiments, both employing recurrent networks with a context layer feeding back into the hidden layer. The training procedure for both networks is essentially the same. Simplified English sentences (articles are absent) are fed into the network one word at a time, and backpropagation is used in a (*prima facie*) attempt to train the network to *predict* the next word it will receive as input. However, since a large training corpus is employed (10,000 sentences in each experiment), the network cannot learn to predict the next input word, but does learn (in essence) to predict the *syntactic category* of the following word. The first of the two experiments is designed, in fact, to demonstrate that the network does indeed develop a set of syntactic categories which correspond to the traditional grammatical categories. Cluster analysis on the network's hidden-layer activation values reveals that the network acquires *approximately* traditional categories, as well as (approximate) subcategories corresponding to animate noun, inanimate noun, transitive verb, etc.[3] The syntactic corpus for this experiment consists entirely of simple 2 and 3 word sentences. Both singular and plural nouns are included, and the network does learn to detect number agreement.

The second experiment is designed to test whether a somewhat more complex recurrent network can discover syntactic structure. In this experiment the training corpus includes relative clauses, embedded to a maximum depth of two (judging by examples provided). Now, although the acquisition of *approximate* syntactic categories in the first experiment seems to indicate that a *degree* of systematicity has been discovered, only in this latter experiment is a test for systematicity explicitly performed. We therefore concentrate our attention upon the latter experiment.[4]

The training regime for the second experiment consisted of four phases, the first of which presented the network with a continuous stream of 10,000 sentences, containing no relative clauses. The three remaining phases each built upon the preceding phases, and involved increasingly high percentages of relative clauses. This controlled, graduated exposure to relative clauses raises questions about the psychological plausibility of the design, which I shall explore in section 3. However, our present concern is with systematicity. Given that the initial training phase involved 10,000 sentences, comprised only of 8 common nouns, 2 proper nouns, and 12 verbs, we have good reason to suppose that the *initial* training corpus presented every word in every syntactically legal position.[5] Assuming this is so, it is clear that the first phase could induce only weak systematicity. Moreover, there is no reason to suspect that any stronger form of systematicity is established by the later training phases, since every possible arrangement

[3]These categories only approximate traditional categories because (for example) the representations developed for subject and object tokens of the same noun are not identical, though they do cluster together.

[4]Also, it is quite clear that the training corpus for the first experiment presented every word in every syntactically legal position. This can readily be established on the basis of the number of nouns and verbs available. It follows that the first experiment establishes only weak systematicity at best.

[5]Note that even if we assumed that every verb optionally takes a direct object, the total number of possible *simple* sentences is: [10 nouns × 12 verbs × 10 nouns = 1200] plus [10 nouns × 12 verbs = 120].

of nouns and verbs that could occur as the complement of a relative clause appears to have been present within simple sentences in the first training corpus.

In passing, it is worth noting that Elman does not say whether his test corpus included greater *depths* of embedding than were present in the training corpus. This is unfortunate, since the ability to generalize to greater depths is an important component of human thought.

3. Plausibility of Training Regimes

In the preceding pages I have occasionally commented upon psychologically problematic aspects of certain of the training regimes invloved. Although none of the authors considered here make strong claims for the psychological plausibility of their methods, it is important to consider whether the results obtained actually *require* learning conditions which are truly implausible. For, even *competence* models of cognitive behaviour (as well as performance models) are normally expected to preserve (or at least approximate) extensional relationships between an agent's *real* input and *real* output. If a particular c-net training regime *requires* the existence of input copora or external error feedback which simply do not occur in human conditions, then serious doubts arise as to whether the c-net model can even provide insight into human cognition. This is especially true when there appears to be no way to *modify* the c-nets involved such that more realistic sets of input and output can be accommodated.

In what follows I examine aspects of the work of St.J&Mc (1990) and Elman (1989, 1990) which *prima facie* (at least) involve seriously unrealistic assumptions about certain learning and/or biological conditions involved in human language acquisition.

3.1 St. John and McClelland (1990)

Recall that the training regime of St.J&Mc presupposes that the learner has *already* apprehended, at the time a given input *sentence* is processed, the particular external state of affairs that the sentence describes.

The learner apprehends this state of affairs by having a *structured*, sequentially ordered representation of this state of affairs in mind. I have already remarked that these structured representations resemble, in spirit at least, those of Fodor's Language of Thought. However, our present concern is with a different problem, *viz.*, is it legitimate to assume that among all the various states of affairs perceptually available to the agent at the time the sentence is presented, the agent's attention is drawn to the particular state of affairs described by the sentence?

St.J&Mc briefly address the above difficulty when they say (p. 249, 1990) "The problem of discovering which event in the world a sentence describes when multiple events are present would be handled in a similar way, though we have not modelled it. Again, the aspects of the world that the sentence actually describes would be discovered gradually over repeated trials, while those aspects that spuriously co-occur with these described aspects would wash out". However, there may be a serious problem with St.J&Mc's suggestion. For, given their experimental design, if spurious states of affairs were frequently presented to the backpropagation algorithm as the *intended* target state of affairs, the *number of iterations* required to wash out the spurious information may well be utterly implausible. Even without spurious information, the network requires over 300,000 iterations *before* it begins to master sentences in passive voice. Were a substantial percentage of spurious states of affairs to be presented, the complexity of the learning task would certainly increase, and we have no reason to suppose the number of iterations involved would fall within anything resembling a plausible range. Even the existing figure of roughly 300,000 raises doubts. These doubts would not be so unsettling if we had reason to believe that alternative architectures would dramatically decrease the iterations involved, but the authors present no arguments to that effect. Moreover, we must bear in mind that the learning task has already been dramatically oversimplified by (a) the absence of articles in the input copora and (b) the fact that the agent's internal representation of the target state of affairs contains a *marker* indicating whether the input sentence is in active or passive voice. It is very difficult to see how a *perception* of an external state of the world could yield an indication as to whether the given *sentence* was active or passive. Also, it seems implausible that the agent would represent voice information *before* the active-passive distinction had been at least partially mastered. Note that *that* distinction is discoverable in the relationships between the input sentences and the internally represented states of affairs, not in the latter representations alone.

Another difficulty concerns the way in which St.J&Mc employ the backpropagation algorithm. During the training procedure, as each constituent of a sentence is processed in turn, the resulting pattern on the output layer is compared to the *entire target* state of affairs. Differences are noted, and backpropagation of error is employed *after each constituent is processed* (although weight changes are accumulated and adjusted after each 60 trials). Now the lack of a biological correlate for the standard backpropagation algorithm is a well known problem, but the defense is commonly made that there may exist some unknown biological process whose effects are roughly analogous to those of this algorithm. This defense is reminiscent of the kind of hand-waving that some connectionist find lamentable in classical AI. However, even if this hand-waving response is accepted, and even if we accept a suggestion of Smolensky that connectionist processing occurs at a more abstract level than the neural level, still there must be *some* biological process which is presumed to support the more abstract process which is supposed to (roughly) correspond to the backpropagation algorithm. Moreover, this biological process would presumably occur each time the backpropagation algorithm is executed in the training regime, and this biological process requires time. Given the complexity of backpropagation, it is difficult to believe that a biological process supporting the algorithm's abstract analogue could occur during the interval between the uttered constituents in a sentence. In light of this, serious doubts arise as to the legitimacy of invoking the backpropagation algorithm *each time* a sentence constituent is heard.[6] At best, the burden of proof rests upon St.J&Mc to show that this application of the backpropagation algorithm has even a *rough* physiological basis.

3.2 Elman (1989, 1990)

Like St.J&Mc, Elman employs backpropagation, but he does attempt a justification for doing so. Recall that, in an effort to teach his networks the syntactic categories of lexical items, Elman trains the network, via backpropagation, to *attempt* to predict the *next* word in a successive stream of words. In defense of this "error-feedback" strategy, Elman remarks that "it does seem to be the case that much of what listeners do involves anticipation of future input". Presumably, Elman takes this as evidence that listeners are constantly attempting to predict the next word they hear. This strikes me as a dubious extrapolation needing empirical support. However, a more serious objection is that Elman's invocation of backpropagation after each word is processed is subject to the same criticism as St.J&Mc's usage. It is difficult to believe that a biological process supporting anything analogous to backpropagation could occur between succeeding words in an utterance.

Another difficulty with Elman's approach is that, when relative clauses are involved, training occurs in 4 distinct phases. Phase 1 presents the network with a concatenated string of 10,000 *simple* grammatical sentences (no relative clauses are included). This string of 10,000 sentences is presented to the network 5 times over. Now, not by the wildest stretch of the imagination is this a psychologically plausible regime. Normally, a child would encounter many breaks even during a series of 20 sentences. During some of these breaks the child may hear sentence fragments, or even simple names. Almost certainly, the child would be exposed to a substantial percentage of unfinished and ungrammatical sentences. The question naturally arises, would Elman's networks be able to induce systematic regularities under these conditions? Not likely, but if not, what are the real implications of this research?

Returning to the succeeding phases of Elman's regime, phase 2 modifies phase 1 by having 25% of the 10,000 sentences contain relative clauses. Phase 3 contains 50% relative clause sentences, and phase 4 contains 75% relative clauses. Clearly, this training regime is highly contrived. Children are not exposed to anything like this artificial partitioning of the input copora.

In fairness, I should note that the artificiality of Elman's training regime is certainly not unique to his work, and he would no doubt readily concede its artifice. Somewhat analogous remarks would apply to St.J&Mc (who also train their c-nets with implausibly long strings of sentences). It is doubtful whether these authors would attempt serious defenses of the size and presentation of their input copora. However, in the absence of such defense we must ask whether these networks could discover even the moderate degree of compositionality they do discover if they were subjected to the erratic, mixed, and often ungrammatical input that humans receive.

[6] Note that even if weights are modified only after every N invocations of the algorithm, the strategy described requires equally as many invocations just to enable the information to be gathered for later weight modification.

4. Summary

In the foregoing I have examined c-net experiments which arguably establish that c-nets can be trained to discover compositionality and exhibit systematicity. In neither of the cases examined does there appear to be any reason to suppose that the c-nets involved exhibit anything stronger than weak systematicity. I have also argued that humans exhibit a much stronger form of systematicity than these c-nets, and thus there is no reason to suppose that the results of Elman and St.J&Mc defeat the Fodor-Pylyshyn thesis. Moreover, I have argued that the experiments considered here involve seriously unrealistic training regimes , and this in turn casts doubt upon the cognitive significance of the experiments. I do not suggest that these experiments are uninteresting; it may be that they will ultimately illuminate an important aspect of the overall puzzle. However, as it stands, it is difficult to see what the cognitive implications of these experiments are.

References

Dromi, E. (1987) *Early Lexical Development*, Cambridge University Press, Cambridge, UK.

Elman, J.L. (1989) "Representation and Structure in Connectionist Models", CRL Technical Report 8903, University of California at San Diego.

Elman, J.L. (1990) "Finding Structure in Time", *Cognitive Science*, Vol. 14, pp. 179-212.

Fodor, J.A. (1976) *The Language of Thought*, Harvard University Press, Cambridge, Mass.

Fodor, J.A. & Pylyshyn, Z.W. (1988) "Connectionism and Cognitive Architecture: A Critical Analysis", *Cognition*, Vol. 28, pp. 3-71.

Hadley, R.F. (1990) "Connectionism, Rule Following, and Symbolic Manipulation", Proceedings of AAAI-90, MIT Press, Boston.

Hadley, R.F. (1992) "Connectionism and Systematicity in Connectionist Language Learning", Technical Report, CSS-IS TR 92-03, Centre for Systems Science, Simon Fraser University.

Ingram, D. (1989) *First Language Acquisition*, Cambridge University Press, Cambridge, UK.

Katz, N., Baker, E. & Macnamara, J. (1974) "What's in a Name? A Study of How Children Learn Common and Proper Names", *Child Development*, Vol. 45, pp. 469-473.

St. John, M.F. & McClelland, J.L. (1990) "Learning and Applying Contextual Constraints in Sentence Comprehension", *Artificial Intelligence*, Vol. 46, pp. 217-257.