C E N T E R   F O R   R E S E A R C H   I N   L A N G U A G E

June 1998 **Vol. 11 No. 4**

The Newsletter of the Center for Research in Language, University of California, San Diego, La Jolla CA 92093-0526
Tel: (619) 534-2536 • E-mail: crl@crl.ucsd.edu • WWW: http://crl.ucsd.edu/newsletter.html

• • •

FEATURE ARTICLE

ON THE COMPATIBILITY OF CONNECTIONISM AND COGNITIVE LINGUISTICS

Mark Collier
Department of Philosophy
University of California, San Diego

EDITOR'S NOTE

This newsletter is produced and distributed by the **CENTER FOR RESEARCH IN LANGUAGE**, a research center at the University of California, San Diego that unites the efforts of fields such as Cognitive Science, Linguistics, Psychology, Computer Science, Sociology, and Philosophy, all who share an interest in language. We feature papers related to language and cognition distributed via the World Wide Web) and welcome response from friends and colleagues at UCSD as well as other institutions. Please visit our web site: “http=://crl.ucsd.edu”.

SUBSCRIPTION INFORMATION

If you know of others who would be interested in receiving the newsletter, you may add them to our email subscription list by sending an email to listserv@crl.ucsd.edu with the line "`subscribe <email-address> newsletter`" in the body of the message (e.g. `subscribe jdoe@ucsd.edu newsletter`).

Please forward correspondence to:

Paul Rodriguez, Editor
Center for Research in Language, 0526
9500 Gilman Drive, University of California, San Diego 92093-0526
Telephone: (619) 534-2536 • E-mail: editor@crl.ucsd.edu

Back issues of this newsletter are available from CRL in hard copy as well as soft copy form. Papers featured in previous issues include the following:

PDP Learnability and Innate Knowledge of Language
David Kirsh
Department of Cognitive Science, UCSD
Vol. 6, No. 3, December 1991

Why It Might Pay to Assume That Languages Are Infinite
Walter J. Savitch
Department of Computer Science, UCSD
Vol. 6, No. 4, April 1992

Is Incest Best? The Role of Pragmatic Scales and Cultural Models in Abortion Rhetoric
Seana Coulson
Department of Cognitive Science, UCSD
Vol. 7, No. 2, January 1993

Marking Oppositions in Verbal and Nominal Collectives
Suzanne Kemmer
Department of Linguistics, UCSD
Vol. 7, No. 3, September 1993

Abstract Better Than Concrete: Implications for the Psychological and Neural Representation of Concrete Concepts
Sarah D. Breedin, Eleanor M. Saffran, and H. Branch Coslett
Center for Cognitive Neuroscience, Department of Neurology, Temple University
Vol. 8, No. 2, April 1994

Analogic and Metaphoric Mapping in Blended Spaces: Menendez Brothers Virus
Seana Coulson
Cognitive Science, UCSD
Vol. 9, No. 1, February 1995

In Search of the Statistical Brain
Javier Movellan
Department of Cognitive Science, UCSD
Vol. 9, No. 2, March 1995

Connectionist Modeling of the Fast Mapping Phenomenon
Jeanne Milostan
Computer Science and Engineering, UCSD
Vol. 9, No. 3, July 1995

Representing the Structure of a Simple Context-Free Language in a Recurrent Neural Network: A Dynamical Systems Approach
Paul Rodriguez
Department of Cognitive Science, UCSD
Vol. 10, No. 1, October 1995

A Brain Potential Whose Latency Indexes the Length and Frequency of Words
Jonathan W. King
Cognitive Science, UCSD
Marta Kutas
Cognitive Science and Neurosciences, UCSD
Vol. 10, No. 2, November 1995

Bilingual Memory: A Re-Revised Version of the Hierarchical Model of Bilingual Memory
Roberto R. Heredia
Center for Research in Language, La Jolla, CA
Vol. 10, No. 3, January 1996

Development in a Connectionist Framework: Rethinking the Nature-Nurture Debate
Kim Plunkett
Oxford University
Vol. 10, No. 4, February 1996

Rapid Word Learning by 15-Month-Olds under Tightly Controlled Conditions
Graham Schafer and Kim Plunkett
Experimental Psychology, Oxford University
Vol. 10, No. 5, March 1996

Learning and the Emergence of Coordinated Communication
Michael Oliphant and John Batali
Deptartment of Cognitive Science, UCSD
Vol. 11, No. 1, February, 1997

Contexts That Pack a Punch: Lexical Class Priming of Picture Naming
Kara Federmeier and Elizabeth Bates
Cognitive Science, UCSD
Vol. 11, No. 2, April, 1997

Lexicons in Contact: A Neural Network Model of Language Change
Lucy Hadden
Cognitive Science, UCSD
Vol. 11, No. 3, January, 1998

ON THE COMPATIBILITY OF CONNECTIONISM AND COGNITIVE LINGUISTICS

Mark Collier
Department of Philosophy
University of California, San Diego

Abstract

Is PDP Connectionism compatible with Cognitive Linguistics? It is unfortunate that this question has not received the attention it deserves, since at stake is the very possibility of a unified "West Coast Cognitive Science" approach to language. Part I of this paper argues that a systematic approach to the question of compatibility must involve an enumeration and analysis of the general principles used by each research program in their linguistic explanations. This approach is carried out in Parts II and III, and the conclusion is that the explanatory principles of PDP Connectionism are fundamentally data-driven, whereas those of Cognitive Linguistics rely essentially on structure that the mind contributes to language use. Part IV presents several computational models of metaphor and analogy as case studies of the practical difficulties involved in reconciling these two frameworks. Finally, Part V examines the philosophical foundations of their respective explanatory strategies in order to get a clearer view of the theoretical obstacles to a unified account.

Part I: Introduction.

1.1 The Nature of the Problem:

PDP (Parallel Distributed Processing) Connectionism and Cognitive Linguistics have each flourished since the 1980's, and both research programs have produced a series of important results and gained a number of adherents. Collaborative efforts have been rare, though, which raises concerns about the compatibility of the two frameworks. We feel that the question of compatibility has not received the attention it deserves, since at stake is the very possibility of a unified “West Coast Cognitive Science” approach to language. In this paper, we examine the practical and theoretical issues surrounding PDP Connectionist modeling of Cognitive Linguistics, and assess the philosophical grounds of the Compatibilist and Incompatibilist positions.

Motivation for a harmonious accommodation of these two research programs comes from both directions. From a PDP Connectionist point of view, perhaps Cognitive Linguistics can serve as an ally in their challenge to the Generative tradition in linguistics. This seems plausible, since Cognitive Linguistics, like PDP Connectionism, was developed as an alternative framework to mainstream linguistics.[1] Indeed, the PDP Connectionists might need Cognitive Linguistics as an ally in their defense as well. In a recent series of articles, Jerry Fodor, a partisan of the Generative approach, offers a “pre-emptive attack” on the possibility of a PDP Connectionist account of semantics, where meaning is cashed out in terms of the relative position of a representation in the network’s state space.[2] Replaying a theme from Fodor & Pylyshyn’s (1988) critique, PDP Connectionism is portrayed as nothing but old-fashioned associationism, and this characterization is taken to explain PDP Connectionism’s inability to account for aspects of language and thought deemed essential by the Generative Program. In the earlier debate, PDP Connectionists were not armed with a alternative theory of syntax and semantics, and thus were put in the defensive position of having to satisfy the Generative *desiderata*. To avoid repeating this reactive strategy, we suggest that PDP Connectionists look to Cognitive Linguistics for an alternative conception of language, and for the set of phenomena that stand in need of explanation. But this just brings us back to the original question of the compatibility of PDP Connectionism and Cognitive Linguistics.

The question of compatibility raises important issues for Cognitive Linguistics as well. One of the reasons that Cognitive Linguists ought to worry about the question of compatibility is that PDP Connectionist computational models can lend the theories of Cognitive Linguistics with a degree of biological and psychological plausibility. The development of computational models represents a prudent strategy, since direct inter-theoretic relations between Neurobiology and Cognitive Linguistics are currently unavailable. As Gilles Fauconnier notes, “[i]n spite of spectacular research in neurobiology, there is nothing remotely close to explanation of higher-level phenomena such as the ones discussed in the present work [on mental spaces]”.[3] Moreover, not only is the proof of the compatibility with lower-level explanations a desired goal, but, going in the other direction, any inconsistency with neurobiological data would raise problems for Cognitive Linguistics. In particular, it would create a tension between two of the main assumptions of Cognitive Linguistics:

(1) “The <u>Generalization Criterion</u>: linguistics is primarily concerned with the statement of general principles governing linguistic elements and structures at all levels.”[4]

(2) “The <u>Cognitive Criterion</u>: one’s analysis of natural language should be consistent with what is known about the mind and the brain generally.”[5]

The project of Cognitive Linguists is to uncover the high-level, general principles governing language use, and these principles must, according to these criteria, find explanations at a lower level. If there are no direct links with neuroscience then Cognitive Linguists are forced to couch their explanations in terms of some version of Connectionism; it is not clear that there is a plausible alternative by which Cognitive Linguistics can satisfy its own demands.

1.2 The State of the Debate: Langacker’s Argument for Compatibilism.

It is an unfortunate state of affairs that to date there has been very little written about the issue of the compatibility of these two approaches. The question was first broached by Ron Langacker in a previous CRL Newsletter (Vol.1, No.3), and it was taken up by William Bechtel and Adele Abrahamson in their book, Connectionism and the Philosophy of Mind. Bechtel and Abrahamson cite Langacker as the chief advocate of a Compatibilist position, since Langacker declares outright that "[C]ognitive Grammar (at least my formulation of it) is basically compatible with the Connectionist philosophy."[6] In the CRL Newsletter, Langacker offers four reasons why his approach to language is compatible with Connectionism:[7]

1. In Cognitive Grammar (CG), rules are nothing but schematized expressions; so, as with connectionism, "rules" are "immanent" in their instantiations and not discrete and separate structures.
2. CG, like Connectionism, avoids appeal to propositions; CG appeals only to phonological and semantical content.
3. In neither CG nor Connectionism is computation algorithmic.
4. In CG, "a linguistic system is viewed as simply an inventory of 'cognitive routines', which are interpretable as recurrent patterns of activation that are easily elicited by virtue of connection weights; the construction of complex expressions reduces to the coactivation of appropriate routines and "relaxation" into a pattern of activation that simultaneously satisfies all constraints."[8]

In a later paper, "A Dynamic Usage-Based Model," Langacker provides five additional reasons in support of his claim that the two frameworks are compatible; these involve a Connectionist "processing interpretation" of the central psychological processes appealed to by Cognitive Grammar:[9]

5. Entrenchment = an adjustment in connectionist weights that forms an attractor, which makes the re-occurrence of a particular pattern more likely.
6. Schema-extraction = a schema is a region in state space, and schemas are extracted when the commonalities between multiple expressions are represented by similar regions in state space.
7. Categorization = when an input (A) is captured by an attractor region (B) in state space.
8. Composition = co-activation of two patterns, to form a joint activation pattern.
9. Symbolization = the correlation between sets of patterns of activation.

There are a number of reasons, however, why one ought to be cautious about allowing such a strong Compatibilist position as a resolution of the debate. First, the nine "processing interpretations" do not constitute a full enumeration of the terms of Cognitive Grammar (e.g. there is no interpretation for Comparison, which is the Cognitive Grammar term for mappings between domains), nor is Cognitive Grammar a complete representative of all of Cognitive Linguistics. Even if this *were* a complete enumeration, however, the argument rests on the dubious assumption that a translation of idiom entails a translation of theoretical framework. The proof that the frameworks are in fact compatible is, as they say, in the pudding. Bechtel and Abrahamson are skeptical of Langacker's Compatibilism for this reason, since "to date there have been limited attempts to directly implement Cognitive Linguistics in a connectionist framework... Hence, it is premature to judge how fruitful the link will be."[10]

The lack of successful, available models might be interpreted in two very different ways, depending on one's prior assumptions about the issue of compatibility. From a Compatibilist point of view, for example, it might only imply a lack of sufficient attention, or simply a lack of current sophistication; nonetheless, the Compatibilist might maintain, a solution is around the corner. From an Incompatibilist point of view, of course, the scarcity of actual simulations serves as evidence of the incommensurability of the two research programs.

1.3 Overview of the paper:

Because of the ambiguous nature of the current evidence, the most fruitful approach to the issue of compatibility will involve conceptual analysis and comparison of the central principles involved in each research program. Part II begins this project with a look at two paradigmatic examples of the PDP Connectionist approach to language, and isolates some of the main principles which underlie their explanations of language acquisition. Part III continues it by surveying some of the central principles posited by Cognitive Linguists to underlie language use. Part IV examines two computational models of metaphor and analogy as case studies that demonstrate some of the *practical* difficulties involved in building PDP Connectionist models of Cognitive Linguistics. Finally, Part V looks at some of the *theoretical* obstacles to a unified account of PDP Connectionism and Cognitive Linguistics, and assesses the philosophical grounds of the Incompatibilist position.

Part II: Some Main Principles of the PDP Connectionist Approach to Language.

2.1 Rumelhart and McClelland's model of past-tense:

Rumelhart and McClelland's model of the acquisition of the English past tense is an important landmark in the PDP Connectionist approach to language. In their chapter of the PDP Volumes, they challenge the dogma in linguistic research that a domain-specific language acquisition device is necessary in order to account for the observable stages of children's acquisition of past-tense morphology.[11] Their simulation demonstrates that a single associative mechanism can learn both irregular and regular English past-tense forms. Their model relies on a simple "Pattern Associator" with modifiable connections between two pools of units. The input pool represents the root form pattern (both input and output were represented in terms of 360 phonological Wickelfeatures), and the output pool represents the network's best guess at the past-tense form of the input.[12] (e.g. "Walk - Walked") In the simulations, the network is repeatedly provided with such pairs of base/ past-tense forms, and the matrix of connections between the pools are adjusted according to a Perceptron Convergence Procedure, which is a discrete version of the Delta Rule.[13] Since a Pattern Associator can store a number of mappings between inputs and outputs, it is able to store exceptional and regular patterns in the same matrix. It is also able to generalize to novel base form patterns, by exploiting the regularity that exists in the previous mappings.[14]

Although the Rumelhart and McClelland paper spawned a huge debate in cognitive science, we are here only interested in the general principles they invoke in their explanations, and so we can steer clear of the controversy that surrounds particularities of their model, such as the nature of the corpus and the phonological representations that were used. This will allow us to isolate two of the basic principles that are invoked in their explanation of language acquisition. First, and very generally, the network learns to associate the base forms and past-tense forms by virtue of repeated exposure to their pairing in the data; as the researchers write: "[t]he pattern associator can teach itself the right set of interconnections through experience processing the patterns in conjunction with each other."[15] Second, the network associates patterns based on their respective similarities:

"Pattern associator models have the property that uncorrelated patterns do not interact with each other, but more similar ones do. Thus, to the extent that a new pattern of activation... is similar to one of the old ones, it will tend to have similar effects."[16]

The principle of similarity explains the capacity of the network to generalize to novel past-tense forms.

2.2 Elman's SRN account of lexical category learning:

Whereas Rumelhart and McClelland's used a simple, two layer architecture, later PDP Connectionist models introduced a variety of alternative architectures and learning techniques. One important line of development is the use of recurrent architectures, which introduce time as an element into processing. In one class of recurrent architectures, the network's activation at time t is copied onto a set of context units, which are then fed back, along with new input, onto the input units at time t+1, thus providing the network with short-term memory. A very influential Connectionist model of language learning that relies on a Simple Recurrent Network (SRN) is Jeff Elman's (1991) model of lexical category learning. The SRN is trained on a corpus composed of 10,000 two- and three-word sentences composed from a lexicon of 29 nouns and verbs.[17] Backpropagation is used to train the network to predict the next word in a sentence. The striking finding is that, although the network does not learn to accurately predict the next word in a sentence, it does manage to predict the lexical category of the next item.[18]

As with the Rumelhart and McClelland model, our interest is in the general principles which underlie how the network managed to learn its task. According to Elman, the SRN learns to "identify inputs as belonging to classes of words based on distributional properties and *co-occurrence information*,"[19] and is able to "induce lexical category structure from statistical regularities in usage."[20] The sequences of words that constitute the training data include a certain probability density function, and the network learns the conditional probabilities that a particular lexical class will follow a given word of the sequence. This structure is implicit in the data, and is extracted as a result of the pressures of the learning task to accurately predict the next word in the stream. The crucial point for our purposes is that Elman's explanation of the induction of lexical classes appeals to principles that are similar in kind to those of Rumelhart and McClelland's. First, as is the case with Pattern Associators, the learning laws of a SRN are, in broad terms, associationist, since learning is driven by the frequency with which items co-occur in the input, and SRNs learn by treating items which follow one another in time as being closely located together in state space. As Tim van Gelder notes:

"[M]uch of connectionist work is clearly associationist... The immediate goal of Elman's SRN models, for example, is to get the network to absorb the statistical regularities in the training set, and successful performance for the network is even defined as behavior that is perfectly in accord with those regularities. The model is behaving correctly if its prediction of the next word in the sentence is in exact agreement with the statistics of the training set." [21]

Second, as with the Pattern Associator, the Elman net operates according to the principle of similarity. As cluster analysis reveals, the hidden layer come to embody what Elman calls a "similarity structure", in which similar items cluster together in closer regions of state space.[22]

Part III: Some Main Principles of the Cognitive Linguistic Approach.

This section presents a rough and ready overview of some of the central principles invoked by Cognitive Linguists to account for language use. Obviously, this analysis makes no claim to comprehensiveness; rather, the focus will be on two important areas of Cognitive Linguistics, which Fauconnier entitles "projection mappings" and "mental-space mappings."[23] This research includes some of the major findings of the field to date, however, and provides a sufficient sample of the principles involved in the cognitive apprehension of language to allow comparison with PDP Connectionist accounts in the later sections.

3.1 Metaphor and the "Invariance Hypothesis":

One of the central discoveries of Cognitive Linguistics has been the pervasive role that metaphor plays in everyday discourse and reasoning. (Lakoff and Johnson 1980, Lakoff 1987, Johnson 1987). A classic example of this type of projection mapping, so common to everyday language and thought, is the "time is space" metaphor (e.g.. "Summer is around the corner").[24] Metaphors are considered to be projection mappings, because the production and comprehension of metaphors involve the projection of information from a source domain onto a target domain. The source domains are usually more concrete, and invoke structure from experiential and conceptual domains, image-schemas, mental images, and generic knowledge.[25] The structure of the source domain is used to set up asymmetrical, partial correspondences with a target domain, which is typically more abstract than the source. These mappings have varying degrees of conventionality; on one end of the spectrum are metaphors that are completely conventional and automatic, and on the other end are novel metaphors that must be processed on-line. The most important point about these mappings for our purposes is that they are not arbitrary, but are constrained by very general principles. George Lakoff supplies an example of one such principle:

Invariance Principle: "the portion of the source domain structure that is mapped preserves cognitive topology (though, of course, not all the cognitive topology of the source domain need be mapped)".[26]

3.2 The Principles of Analogical Mappings:

Analogy is another area in which Cognitive Linguists and Cognitive Psychologists have found evidence for projection mappings. (Gentner 1983, Holyoak and Thagard 1989) A classic example of analogy is the comparison of the structure of an atom to the solar system. As with metaphor, analogy involves projections between two domains, and the task is to find coherent correspondences between highly abstract and relational structures.[27] As with metaphor, analogical reasoning is constrained by very general principles, which are "soft" and may compete with one another in a given context. Although there is no consensus about the precise nature of these principles, some leading candidates involve:

Structural Consistency Principle (Gentner 1983) or Isomorphism Principle (Holyoak and Thagard 1989): "the process of comparison is one of structural alignment between two mental representations to find the maximal structurally consistent match between them."[28]

Parallel Connectivity Principle: (Gentner 1983, Holyoak and Thagard 1989) "if two predicates are matched, then their arguments must also match."[29]

One-to-one mapping Principle: (Gentner 1983, Holyoak and Thagard 1989) "requires that each element in one representation correspond to at most one element in the other representation".[30]

Systematicity Principle: (Gentner 1983) "when there are multiple interpretations of a pair, all else being equal, the one that preserves the maximal (i.e. largest and deepest) connected relational structure is preferred)".[31]

Semantic Similarity Principle: (Holyoak and Thagard 1989) "elements with some prior semantic similarity (e.g. by virtue of their membership in a taxonomic category) should tend to map to each other."[32]

Pragmatic Centrality: (Holyoak and Thagard 1989) "a mapping structure should give preference to elements that are deemed especially important to goal attainment, and should try to maintain correspondences that can be presumed on the basis of prior knowledge."[33]

3.3 Mental Space Mappings:

A final example of the kinds of mappings in everyday language and thought involve the mappings between mental spaces that occur during on-line discourse. (Fauconnier 1985) Mental space mappings have been used to explain many aspects of language and thought, such as opacity and counterfactual reasoning. One of the key insights of this research has been to re-conceptualize the nature of sentence meanings. In this framework, sentences don't contain meaning, but invoke meaning construction. During communication, a speaker's utterances act as "prompts" or "cues" for the listener to construct mental spaces in which to connect the appropriate elements of discourse, including background knowledge and frames, local deictic knowledge, roles, relations, etc. As with projection mappings, mental space mappings are governed by general principles: "The principles governing the operations are, in themselves, simple and general. They appear to be universal across languages and cultures."[34] One of the central principles isolated by the work on mental spaces is the "Identification Principle":

Identification Principle or Access Principle: (Fauconnier 1985) "allows elements in mental spaces to be accessed in terms of elements connected to them, and situated in other mental spaces."[35]

Part IV: The Practical Difficulties of Modeling Cognitive Linguistics.

4.1 Thomas and Mareschal's (1996) "Metaphor by Pattern Completion" model:

Unfortunately, PDP Connectionists have tended to pay little attention to metaphor. Luckily, Thomas and Mareschal counter this trend with a simulation of simple "A is B" metaphors. Their model is designed to capture the on-line comprehension of mappings between concepts, which "accounts for how the semantic features of a target word (A) are transformed by the semantic properties of a knowledge base (B)."[36] The model consists of a three layer, feed-forward architecture, in which three concepts (APPLE, BALL, AND FORK) are represented in terms of 13 semantic features. The researchers train the network on one of the concepts, and then input an exemplar from another of the concepts. The network attempts to recognize the exemplar as an instance of the concept on which it has been trained, and subsumes the exemplar under the learned prototype. For example, they train the network on exemplars of the concept BALL, and then present the network with an exemplar of the concept APPLE. The network attempts to recognize the input, and in so doing, it transforms the APPLE pattern to make it consistent with its knowledge base. The interesting feature of such transformations is that it produces what Thomas and Mareschal call "meaning enhancement".[37] For example, in the "apple is a ball" example, the network automatically reduces the "edible" semantic feature of the apple, while it increases its "thrown", "hardness", and "roundness" features. According to Thomas and Mareschal, this captures the interactive nature of "A is B" metaphors.[38]

Although Thomas and Mareschal's model is an interesting first step for PDP Connectionist modeling of what Lakoff and Johnson entitle "image metaphors", it fails to capture an essential aspect of projection mappings. The reason is that the model has trouble explaining the *partiality* of mappings which is fundamental to Lakoff's Invariance Principle. The pattern completion mechanism in their model superimposes the *whole* vector of prototypical features of the source domain onto the target domain. The subsumption of the target image under the source prototype is unselective, as it merely averages the vectors of each domain. As a result, the model will produce too many undesired correspondences; for example, if the model were given the metaphor "time is money", it would produce correspondences such as "time is green", or "time consists of coins and paper", as well as the appropriate correspondences.

4.2 Computational Models of Analogy:

Modeling analogy is notoriously hard. As Gilles Fauconnier puts it:

"Analogical mapping is so commonplace that we take it for granted. But it is one of the great mysteries of cognition. Given the richness of the domains and their complexity, how are the 'right' schemas consistently extracted, elaborated, and applied to further mappings?"[39]

According to Hummel, Burns, and Holyoak, the difficulty of building a computational model of analogy stems from the requirements that the analogical correspondences involve disparate domains, invoke partial correspondences, and involve relational structure rather than direct similarity; in addition, analogical mappings are difficult to search, since the number of possible correspondences between two domains scales exponentially with the number of elements in the analogs.[40] Analogical reasoning provides an additional set of problems for PDP Connectionism. According to Holyoak and Barnden, "[t]he complexity of analogical processing as traditionally conceived presents a considerable challenge to connectionism."[41] One reason such modeling has proven difficult is that:

"In contrast to the typical connectionist learning algorithms, which require statistical generalization from many repetitions of large numbers of examples, analogy allows rapid 'two-trial' learning, in which a single previous case can generate a rich set of inferences about a second case."[42]

The computational complexity of modeling analogy has pushed a number of researchers towards the use of hybrid models, which incorporate symbolic representations. For example, two influential models of analogy, ACME (Analogy Constraint Mapping Engine, Holyoak and Thagard 1989), and LISA (Learning and Inference with Schemas and Analogies, Hummel and Holyoak, 1995), invoke either a simplified predicate calculus (ACME) or semantic units (LISA). The benefit of using a hybrid model is that symbolic representations make the mapping problem easier to solve. In the LISA model, and recent modifications of ACME, the researchers use temporal synchrony to bind correspondences between representations. For example, if the system is given the two analogs, "love (Jim Mary)" and "love (Bill Susan)", both the Jim and Bill units fire synchronously, since they decompose onto the same semantics features of "male" and "person".[43]

These hybrid models do not help establish the compatibility of PDP Connectionism and Cognitive Linguistics, since they solve the problem of analogy only by shifting away from Connectionist principles. A truly PDP Connectionist approach to analogy is still an open research question, although one that has been called into doubt for the reasons mentioned above. It is an open question whether it is possible to construct a model along the lines of ACME, but which uses solely distributed representations, rather than a localist or predicate calculus representations, to set up the analogical correspondences. The trade-off, however, is that symbols were originally introduced to simplify the account of the representations that are to be mapped, and without them the modeling task becomes more difficult. In LISA, for example, the use of a localist, semantically-labeled nodes allows the modelers to feed it simple propositional analogs with pre-specified common features, and let temporal synchrony bind the appropriate correspondences. But with a distributed architecture, there are no clear, prestructured, semantically-labeled units, and so there is no guarantee that temporally synchronous firing between such representations would pick out the right analogical correspondences.[44]

4.5 Modeling Mental Spaces:

As far as I know, there are no PDP Connectionist models of mental space mappings. Is this a simple sin of omission on the part of PDP Connectionist modelers, or does this lacuna represent a potential incompatibility of the two frameworks? One explanation is that mental space mappings, like projection mappings, involve partial mappings across distinct and dissimilar domains, and so present similar difficulties as projection mappings. But are there any problems for modelers particular to mental space mappings? One potential problem presented by the mental space data is that the unit of analysis for studying meaning across mental space mappings is not the sentence, but temporally extended discourse.[45] Although this raises a host of difficulties for linguistic models which make individual sentences the primary unit of analysis, though, it should not raise any particular worries for PDP Connectionists. Intersentential meaning propagation poses no problem for PDP Connectionists, since recurrent networks offer a suitable architecture for temporally extended discourse. As Elman points out, PDP Connectionists are also moving in the direction of going "beyond sentences."[46] Whereas traditional schemes have difficulty keeping information from previous sentences available to discourse purposes, processing in a Simple Recurrent Network allows later sentences to be effected by prior sentences:[47]

"This approach to preserving information suggests that such networks would readily adapt to processing multiple sentences in discourse, since there is no particular reanalysis or re-representation of information required at sentence boundaries and no reason why some information cannot be preserved across sentences."[48]

Another potential difficulty for PDP Connectionist models of mental space mappings is that speakers not only have a facility for fluently accessing "background knowledge", but can also access "locally introduced" domains in the construction of mental space mappings.[49] Whereas background frames or schemas translate smoothly into PDP Connectionist weights or prototypes, accessing locally introduced domains strains the resources of PDP Connectionism. The reasons are similar to the one-trial learning problem that Hummel et al. pointed out with regards to PDP Connectionist accounts of analogy. In this case, locally introduced domains are by definition those for which the network has had little or no previous experience. The only resources available to a PDP Connectionist network to process novel discourse situations, then, is to extrapolate from background knowledge, but there is no *a priori* guarantee, without actual simulations being run, that restricting the possible mappings in ways that conform to past experience will allow the appropriate mental space mappings to be set up.

Part V: The Theoretical Difficulties of PDP Connectionist Modeling of Cognitive Linguistics.

5.1 The principle of similarity revisited:

In Part IV, we saw that metaphorical, analogical, and mental space mappings present PDP Connectionists with the difficult engineering problem of simulating novel mappings across partial domains. I now want to argue that the problem is more than a mere practical difficulty of current modeling techniques, but represents a theoretical challenge that goes to the heart of the PDP Connectionist explanatory strategy. One source of the problem is the PDP Connectionist commitment to the principle of similarity as part of their linguistic explanations. Since the similarity between two representations is defined in terms of their relative distance in state space, and distance is itself a function of co-occurrence information, representations which infrequently occur together, such as "time" and "money", will end up in remote regions of state space, and so will be treated as dissimilar by the network. Nonetheless, English speakers can fluently comprehend novel metaphors that recruit structure from the domain of money to elicit inferences about time. In order for a Connectionist network to model structured mappings between superficially dissimilar representations, the pre-existing similarity relation must be overridden. It is not clear, though, how this can be done without the recruitment of additional principles or domain-specific mechanisms. Now PDP Connectionist networks are capable of overriding perceptual similarity and treating two representations as conceptually similar, as in the case of the solution to the XOR problem. In this case, however, the network is being asked to override the conceptual similarity relation itself.[50]

This is not to say that past co-occurrence information plays no constitutive role in the similarity relation that allows comprehension of metaphorical mappings. As Lakoff and Johnson point out:

"Conventional metaphors (orientational, ontological, and structural) are often based on correlations we perceive in our experience. For example, in an industrial culture such as ours there is a correlation between the amount of time a task takes and the amount of labor it takes to accomplish the task. This correlation is part of what allows us to view TIME and LABOR metaphorically as RESOURCES and hence to see a similarity between them."[51]

Nonetheless, correlations can play at best a limited explanatory role, since a crucial aspect of metaphorical language and thought is the creation of *new* similarities. Once a metaphorical mapping between domains is set up, albeit on the basis of experienced correlations, metaphorical extensions establish novel structured similarity relations for which there has been no co-occurrence information. For example, once the metaphor PROBLEMS ARE OBJECTS is established, it can be extended to PROBLEMS ARE PRECIPITATES, and can support novel inferences, such as "problems can dissolve".[52]

5.2 The co-occurrence principle and the philosophical grounds of the Incompatibilist position:

We saw in Part IV that the reliance on frequency information lies behind the practical problem of one-trial learning. This reliance on co-occurrence information in fact raises a fundamental theoretical difficulty for PDP Connectionist modeling of Cognitive Linguistics as well. Both PDP Connectionist models surveyed in Part II learn to reproduce the statistical properties of the training data as either sets of connections in a Pattern Associator, or as structure in the state space of a SRN, and both form generalizations according to the regularities it has extracted. The explanatory principles on which learning relies in these cases are essentially *reproductive* in nature. In direct contrast, an important aspect of all of the Cognitive Linguistic principles is their *productive* nature.[53] These principles are productive in the sense that they organize experience, and extend existing categories by supporting novel mappings.

The contrast between the reproductive and productive nature of the respective explanatory principles lies at the heart of the tension between PDP Connectionism and Cognitive Linguistics. The reproductive/productive dichotomy, of course, is not new. It was drawn by Immanuel Kant in order to criticize David Hume, and it was later used by Gestalt Psychologists to beat up on Behaviorism. Kant's critique of Hume, in fact, brings into clear view some of the difficulties of unifying PDP Connectionism and Cognitive Linguistics, and points towards the philosophical grounds of the Incompatibilist position. It can do so because both PDP Connectionism and Hume put forth regularity-associationist theories of cognition, in which patterns are learned and generalized on the basis of the principles of frequency and similarity. Now, as every student of philosophy has been taught, Kant showed the insufficiency of associationist accounts of cognition, and demonstrated that reproductive principles must be supplemented by principles that are grounded in the structure of the mind. Hume rejected concepts such as substance and cause because he could not find them *in* the perceptions of experience. Kant's reply is that Hume was looking in the wrong place; he agrees with Hume that these concepts are not derived from experience, but argues that they must be supplied by the cognizing agent.

One might draw a parallel moral for those who wish to extend current PDP Connectionist models of syntax and morphology to semantics in all of its complexity and richness. PDP Connectionism, from a Neo-Kantian perspective, looks for semantics in the wrong place — in the sentences — rather than in the general principles which make sentence processing possible in the first place.[54] An important aspect of all the Cognitive Linguistic principles is that they enable the language user to *go beyond* or *transcend* the received linguistic input in order to construct meaning. Remember, one of the key insights of Cognitive Linguistics is the realization that language serves as a prompt, or clue, to the construction of meaning.[55] The principles that govern language use are not extracted from the linguistic corpus, but are supplied by the cognitive system.

There are two avenues of response to this line of criticism open to friends of Connectionism. The first, taken by the Berkeley Structured Connectionists, involves the decision to build biologically motivated structures into the architecture of the network in order to enable productivity. According to Terry Regier, for example, the question is:

"[T]o what degree are the regularities in human behavior due to the structure of the organism, and to what degree are they due to the structure of the environment? An implicit claim of those who advocate fully unstructured PDP models is that these regularities are due almost entirely to the structure of the environment; an implicit claim of structured or constrained connectionism is that they are at least partially due to the makeup of the organism."[56]

The second avenue is to object to the criticism on the grounds that it relies on an overly narrow construal of the data available to PDP Connectionist networks.[57] Perhaps what is needed to supplement the linguistic data are not prewired structures, but just more data, including a full panoply of bodily, emotional, and social-cultural information. While this response escapes the Kantian-inspired objection, however, it does so at the cost of threatening to make the computational modeling intractable.

The moral of this paper is that if PDP Connectionism is going to provide a cognitively realistic model of language, then it must pay more attention to the productive nature of the imaginative processes that underlie it. An integrated account of Connectionism and Cognitive Linguistics will need to show how productivity can be incorporated into an inherently reproductive framework. This is not intended as an *a priori* proof of the insufficiency of PDP Connectionism — an argumentative strategy that failed many times in the past. Rather, we expose some of the obstacles to a unified "West Coast" approach to language at the service of those who, we hope, will begin to tackle an area of computational modeling into which few have ventured.

References

Elman, J. (1990), Finding Structure in Time. Cognitive Science, 14, 179-211.

Elman, J. (1995) Language as a Dynamical System. In *Mind as Motion: Dynamical Perspective on Behavior and Cognition*. Port, R. and van Gelder, T. (Eds.) Cambridge, MA: MIT Press.

Fauconnier, G. (1985) *Mental Spaces*. Cambridge, MA: MIT Press.

Fauconnier, G. (1996) *Mappings in Language and Thought.* Cambridge University Press.

Fodor, J. and LePore, E. (1995) "Paul Churchland and State Space Semantics" in McCauley, R. (Ed.) *The Churchlands and Their Critics.* Blackwell Press.

Fodor, J., and Pylyshyn, Z. (1988) Connectionism and cognitive architecture: a critical analysis. In S. Pinker and J. Mueller (Eds.) Connections and symbols. Cambridge, MA: MIT Press.

Gentner, D. (1983) Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science*, 7, 155-170.

Gentner, D. and Markman (1994) Structural Alignment in Comparison: No Difference Without Similarity". *Psychological Science*, Vol. 5, No. 3.

Holyoak, K. and Thagard, P. (1989) Analogical Mapping by Constraint Satisfaction. Cognitive Science, 13, 295-355.

Holyoak, K., and Barnden, J. (1994) Introduction. In Advances in Connectionist and Neural Computation Theory. Vol 2: Analogical Connections. Norwood, NJ: Ablex Publishers.

Hummel, J., Burns, B., and Holyoak, K. (1994) Analogical Mapping by Dynamic Binding: Preliminary Investigations. Advances in Connectionist and Neural Computation Theory. Vol 2: Analogical Connections. Norwood, NJ: Ablex Publishers.

Hummel, J. and Holyoak, K. (1995) LISA: A Computional Model of Analogical Inference and Schema Induction. *Seventeenth Annual Conference of the Cognitive Science Society*. Erlbaum Press.

Johnson, M. (1987) *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and reason.* Chicago: University of Chicago Press.

Lakoff, G. (1987) *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind.* Chicago: University of Chicago Press.

Lakoff, G. (1989) Philosophical Speculation and Cognitive Science. *Philosophical Psychology*, Vol.2, No. 1.

Lakoff, G. (1990) Invariance Hypothesis: Is Abstract Reason Based on Image-Schemas? *Cognitive Linguistics* 1-1, 39-74.

Lakoff, G. and Johnson, M. (1980) *Metaphors We Live By.* Chicago: University of Chicago Press.

Langacker, R. (1987) *Foundations of Cognitive Grammar: theoretical perspectives,* Vol. 1. Stanford, CA: Stanford University Press.

Langacker, R. (1987) The Cognitive Perspective. Center for Research on Language Newsletter 1(3). UCSD: CRL.

Langacker, R. (To appear), A Dynamic Usage Based Model.

Regier, T. (1996) *The Human Semantic Potential*. Cambridge, MA: MIT Press.

Rumelhart, D. & McClelland, J. (1986) On learning the past tense of English verbs. In J. McClelland and D. Rumelhart (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition.* Cambridge: MIT Press.

Thomas, M., and Mareschal, D. (1996) A Connectionist Model of Metaphor by Pattern Completion. *Proceedings of the 18th Annual Meeting of the Cognitive Science Society*. Erlbaum Press.

Van Gelder, T. (1992) Comment on Elman's Grammatical Structure and Distributed Representations. In *Connectionism: Theory and Practice.* Stephen Davis (ed.) Oxford: Oxford Univerisity Press.

1.Bechtel and Abrahamson, p.295. "[Chomsky's] framework has been challenged by a number of theorists (e.g. Fauconnier, 1985, Lakoff, 1987, Langacker, 1987a), who deny that we can understand either the syntax or semantics of language without understanding the psychological processing that underlies it."

2.See Fodor's Articles in The Churchlands and Their Critics.

3. Fauconnier (1985, 2nd edition), p. xxxiii.

4.Lakoff (1989) p 55.

5. Ibid, p.59.

6.Bechtel and Abrahamson, p.295.

7.Ibid, p.295.

8. Ibid, p.295.

9.Langacker, (to appear), pp.4-5.

10.Bechtel and Abrahamson, p.296.

11.Rumelhart and McClelland (1986), p.219.

12.Ibid, p. 222-3.

13.Ibid, p.226.

14.Ibid, p.233.

15.Ibid, p.36.

16.Ibid, p.38

17.Elman (1995), p.348.

18.Ibid, p.350.

19.Ibid, p.221 (my emphasis).

20.Ibid, p.221.

21.van Gelder (1992), p.188.

22.Elman, p.207. "Conceptual similarity is realized through position in state space; words that are conceptually distant produce hidden unit activation patterns that are spatially far apart."

23.Fauconnier (1996) pp. 9-11.

24.Ibid, p. 9

25.Mark Johnson (1987).

26.Lakoff,(1990), pp. 39-74.

27.Hummel, Burns, and Holyoak, (1994), p. 417.

28.Gentner and Markman, (1994), p. 152.
29.Ibid, p.152.
30.Ibid, p.152.
31.Ibid, p.152.
32.Hummel et al. p. 418.
33.Ibid, p.418.
34.Fauconnier (1985), p.xviii.
35.Ibid, p.xxi.
36.Thomas and Mareschal (1996), pp.696-697.
37.Ibid, p.698.
38.Ibid, p.697.
39.Fauconnier (1996), p.20.
40.Hummel, et al., p.417.
41.Holyoak and Barnden, (1994), p. 1.
42.Hummel et al., p.417.
43.Hummel and Holyoak (1995), p. 355.
44.A similar point was brought to my attention by Gilles Fauconnier in discussion.
45.Fauconnier (1985), pp. xix-xx.
46.Elman (1995), p.215.
47.Ibid, p.215.
48.Ibid, p.216.
49.Fauconnier (1996), p.1.
50.Which additional principles would be needed? Perhaps PDP Connectionists can account for the data accumulated by Cognitive Linguistics simply by employing more architecture. Perhaps what is needed are more hidden units above the level of the conceptual prototypes, which would allow for either temporal binding or weighted connections between areas at far ends of the state space. The problem with this, as we saw with the computational accounts of analogy in Part 4, is that it isn't clear how the use of temporal binding would allow the network to synchronize the appropriate elements, if the units that are bound have not occurred together in previous experience. Or, perhaps what PDP Connectionism simply needs is a new architecture that includes a separate module for on-line novel comparisons. Perhaps this module will include a mechanism for overriding superficial dissimilarities, perhaps by bending the state space along the lines of the solution of feedforward networks to the Either/Or problem. Nevertheless, this just re-raises the question of what the governing principles of such a module will be. If they will be frequency and similarity, then we end up back where we started.
51.Lakoff & Johnson (1980), pp.151-152.
52.Ibid, p.152.
53.Note that "productivity" here must be distinguished from the use to which Generative grammarians put the term in their critique of Connectionism. (e.g. one who understands the sentence 'John loves Mary' must also be capable of understanding the sentence 'Mary loves John'.) See Fodor and Pylyshyn (1988).
54.Now of course the nature of the supplementative principles differs drastically between Kant and Cognitive Linguists. Kant was interested in image-schemas that are formed from the Pure Categories; Cognitive Linguists are more naturalistically minded, and look to the role of embodiment and motor actions to organize experience.
55.PDP Connectionists might respond that, at least to a first approximation, they share this conception of sentences as "prompts", rather than "containers" of meaning. It is open for a PDP Connectionist to construe linguistic understanding as the response (e.g. spreading activation in state space) of the network to linguistic input. The crucial difference, though, is that this response is inherently reproductive and is hostage to previous training. The problem, as Cognitive Linguists point out, is that the principles which allow meaning-constructions do not show up in the overt form of the linguistic data, nor are they implicit in the way that past-tense rules or lexical word categories might be.
56.Regier (1996), p.49.
57.This was pointed out to me by Will Peterman.