Lenguas Modernas 25 (1998), 51-70
Universidad de Chile

CONNECTIONIST MODELS OF LANGUAGE LEARNING: IMPLICATIONS FOR WRITING PEDAGOGY

JAMES W. NEY and JULIA C. ANGELICA
Arizona State University

Connectionism —an interdisciplinary approach that draws heavily from hard science— promises to be the new paradigm shift for linguistics and psychology, and has important implications for both composition studies and the teaching of writing. The models are innovative primarily because —in a manner extendable to neurobiological reality— they process in a parallel rather than a serial manner and address subsymbolic rather than symbolic representations. As neuroscientific knowledge expands, such models may be amended and developed to mirror learning of all types. Even at their current level of development, they provide several important insights into the nature of cognition. This investigation uses connectionist assumptions as analytical tools to explain much about past theoretical frameworks in written composition, and —more significantly— to suggest some important considerations for writing pedagogy.

As a result of technological breakthroughs that have revolutionized neurobiology, and in consequent anticipation of rapid developments in the understanding of cognition, neuroscientists are referring to the 1990s as "The Decade of the Brain" (Shepherd 1994). Because of developing knowledge and technology, they expect a "new emphasis on combining information from different levels of analysis into integrated models of brain function..." (p. 5). One aspect of this development has been the ability of neuroscientists to use positron emission tomography (PET) to analyze speech and language tasks too long unstudied because of their complexity and a lack of corresponding animal models (p. 498). As such neuroscientific understanding of cognition —and specifically language— expands, the resultant knowledge must necessarily have a major impact on related disciplines including composition studies. For example, current knowledge in neurobiology has already dealt with the traditional debate over whether brain areas are geographically specialized for local processing (modular) or whether information is processed as different areas interact (multimodular). The widely accepted answer for neuroscience is that both of the above apply: some processes occur within

Para correspondencia y solicitudes de separatas, dirigirse a: James W. Ney, Department of English, Arizona State University, Box 870302, Tempe, AZ 85287-0302, Estados Unidos.

local circuitries but many cognitive tasks require interaction between different brain areas. Further, much of the research, including the PET scan experiments on language tasks, suggests that the interaction between modalities "proceeds more by parallel pathways... than by the classical serial model" (Petersen et al. 1988, cited in Shepherd 1994, p. 684), thus laying to rest another traditional controversy. Even more important for those concerned with writing and creativity, studies of brain damaged —and particularly split brain— patients have provided solid evidence that, while the left brain hemisphere is "the language maven," it is the right brain hemisphere that "is superior to its counterpart in expressing and appreciating emotions" (Restak 1994, p. 127). As a result, those who study language must expect major theoretical rethinking and, indeed, according to some commentators, the new paradigm shift for linguistics (Sampson 1987) and psychology (Schneider 1987) arrived almost a decade ago.

While this new paradigm in its interdisciplinary entirety is categorized as cognitive science, the psychological and linguistic literature often uses the umbrella term *connectionism* to encompass its computational rather than neurological models. Admittedly, although these computational models are extendable to neurological reality, neuroscience feels that the "wealth of insights" they are currently providing only illustrate the ways "complex systems may process information" (Shepherd 1994, p. 10). A true understanding of real brain processing must be based on neuroscientific observation and experimentation. Nevertheless, existing connectionist models, particularly those that address widely distributed rather than local networks, promise and are already producing some intuitively sound principles that have made an impact on fields as diverse as physics, philosophy, and neurolinguistics. The question then arises: What relevance does connectionism have for the teaching of writing?

A direct answer to this question is perhaps difficult for, as Sampson (1987) pointed out, the consumers of the last linguistic paradigm shift were language teachers and literary students turned linguists. Neither these nor scholars of rhetoric are likely to greet or understand the new paradigm with enthusiasm since it is an outgrowth of information processing and artificial intelligence models. Thus, even at its most basic, this new discipline comes resplendent with such terms as connection weights, prototypical vectors, neural networks, and parallel distributed processing. Further, an understanding of the relationship between real brain function and the computational models developed to mimic them might require at least a basic knowledge of brain biology. Nevertheless, in light of the emerging body of knowledge about human cognition, those who are truly interested in language use and acquisition may find valuable some acquaintance with connectionism's major philosophical shift away from "atomic symbol structures as the basis of thought and . . . logical operations on symbol structures as the mechanism of reasoning" (Clark and Lutz 1992).

Characteristics of connectionist models

The basic assumption of most connectionist psychologists is that the brain contains layers of receptive fields consisting of neurons that react to different incoming stimuli. Thus, some fields respond to angle, some to line (vertical or horizontal), others to tone, and so forth through the realm of experience. Presentation of input arouses excitation in all potentially receptive neurons (that is, "it rings bells" throughout the appropriate receptive field) until, through a process of competition and inhibition, a particular configuration of neuron activation wins out as 'best fit' for the incoming data. This "best guess" can then be forwarded up or down to other layers/levels in a hierarchical cognitive system. While debates have been ongoing among computational modelers as to whether systems are local or hierarchical, neuroscience believes that

> the nervous system is organized in terms of functional units... [each] defined as a structural entity with a specific function... formed at different levels of organization, from genes and gene products through synapses, microcircuits, dendrites, neurons, and local circuits, to pathways and distributed systems. The nervous system is built of overlapping assemblies and hierarchies of such units of increasing extent and complexity. (Shepherd 1994, p. 8)

Thus, this analysis will concentrate on those connectionist models that are most highly compatible with neuroscientific assumptions about the hierarchical and distributed nature of the cognitive processing system.

The most widely cited connectionist model in the linguistic and psychological literature is the parallel distributed processing (PDP) model of Rumelhart, McClelland, and the PDP Research Group (1986a, 1986b). This model addresses the "microstructure of cognition" by computing the weight of connections between incoming stimuli (input), experience or patterns already present (hidden units), and the resulting state of excitation and its production (output). In other words, this model views cognitive processing as an associational dynamic process governed primarily by the firing of neural synapses; the strength (or 'weight') of connections formed between neurons as a result of this firing; and the patterns of excitation set up and reinforced as patterns are repeated, similarities recognized, and new information interpreted or amended according to existing expectations.

One extremely important concept contained in the model is the ability to complete familiar patterns that are missing some elements (Rumelhart et al., 1986a). Such an ability obviously accounts for both the human tendency to convert nonwords into words, complete incomplete text as in Cloze tests/exercises, and to resolve lexical ambiguity based on context and world knowledge. A corollary to simple pattern completion is "content addressability," by which the human memory is able to correctly retrieve even extremely complex patterns from only a partial cue (p. 26). Even beyond the ability to fill in missing information, the human mind must also interpret totally new stimuli and experiences, and must, therefore, have predictive powers based on prior experience that allow mental modeling —or imagining— of possible

outcomes. Much of the above processing skill, therefore, is necessarily based on pattern matching or pattern approximation, with experience stored in patterns of excitation —or templates— to be used and modified as necessary. Further, the processor tends to look for measures of central tendency which allow it to make sense of incoming data.

Another extremely powerful connectionist model is the highly mathematical adaptive resonance theory of Stephen Grossberg (1987), which model recognizes that real-time synchronous processing is required for distributive spatial learning patterns to produce a self-organizing map whose adaptive weights can both increase and decrease. (For a more detailed technical discussion of the development of adaptive resonance models and their place in connectionism, see Grossberg 1998). This model is more encompassing than the PDP framework because it provides both a micro- and macrotheory based on a set number of basic theorems with explanatory value for ways in which information can be processed locally (within specific processing modules) and throughout a distributed network (between different processing levels/circuits in a cognitive hierarchy). One psychologist has coined an apt metaphor for adaptive resonance, that of a neurological aeolian harp over which the wind of incoming information plays (Peter Killeen, personal communication, March 25, 1996). As the wind touches the harp's elements, it sets off vibrations which resonate generating further vibrations among sympathetic neighboring elements whose vibrations in turn feed back and prolong the vibrations of the original element. Analogously, incoming stimuli set up patterns of excitation among clusters of neurons, which can in turn produce excitation in sympathetic groups of neurons in other receptive fields. The interactions between source units and receiving layers is nonlinear and competitive-cooperative (Stone 1994), mirroring neuronal interactions in which receptor sites are subject to both activating and inhibiting forces.

Grossberg and Stone (1987) emphasized the importance of recognizing that experience or familiarization changes the organism's internal state. Indeed, resonance theory is termed 'adaptive' because it recognizes the need for a balancing mechanism when different brain areas —often primed for different modalities— have to interact simultaneously with a wealth of incoming stimuli. Two unique features of Grossberg's theoretical framework are self-organization and 'masking fields'. Self-organization, a concept adapted from theoretical biology and physics, refers to the organism's ability to "adapt in real-time to environments whose rules can change unpredictably" (Grossberg and Stone 1987: 404). Obviously, this concept is crucial to human interaction within a linguistic environment. The second concept, the 'masking field,' is a regulatory neural network that, through parallel competitive/cooperative interactions within and between layers, can directly activate correct chunks or groupings without any prior search activity. Such immediate and massively parallel interaction may form the basis of heuristic (rather than algorithmic) cognition.

While the idea of aggregated (chunked) elements as individual units is not new to cognitive psychology, the neuronally inspired architecture within

which this dynamic process occurs is unique to connectionist models. In adaptive resonance, the network develops because of "simple rules of neuronal development: random growth of connections along spatial gradients, activity dependent self-similar cell growth, and competition for conserved synaptic sites" (Cohen and Grossberg 1987: 457). Unlike Rumelhart and McClelland's analysis of receptive fields in terms of levels equivalent to linguistic divisions (for example, phonemic levels), the more subsymbolic adaptive resonance theory proposes generalized representation units (Grossberg 1987). Thus, in the hierarchical processing system, levels of 'items' can be chunked into higher 'sublists' that in turn can be chunked into 'lists,' all designed to be retrieved with maximum processing economy. For example, selection for larger chunks/sublists (such as words) can be amplified so preventing smaller chunks (for example, individual letters) from masking them. The fields also allow for context-sensitive parsing of activity patterns, including expanding patterns and internal pattern changes. Most important, some masking fields "can anticipate, or predict, the larger groupings that may occur of which the item forms a part" (Cohen and Grossberg 1987: 473).

In sum, three intrinsic strengths of the connectionist paradigm have been identified by Stone (1994):

> It reflects variation across experience... variation is not collapsed to give a generic category label... Processing of stimulus is intrinsically *influenced by the context* in which it occurs... and connectionist systems pickup [sic] and *reflect degrees of statistical regularity* in experience through learning... which allows "rule-like" behavior. (p. 418)

Thus, the connectionist approach reemphazises the "perceptual flux" —continuous variations in the internal state of the system— rather than focusing on symbolic concepts as did earlier cognitive science (Stone 1994). To restate simply, the paradigm is revolutionary because it views cognition as produced through pattern processing, not symbol processing.

Theoretical applications for writing pedagogy

At this point the question can again be posed: What are the implications of connectionism for the teaching of writing? During the past two decades, the teaching of writing on a university level in the United States has become more and more the province of the developing field of composition and rhetoric. Nystrand and his colleagues (1993), prominent scholars in this area, have provided an excellent epistemological analysis of the development of composition studies in relation to academic needs and attitudes, and shifts in rhetorical, linguistic, and cognitive theory, thus emphasizing the potential for new paradigms to influence the study of writing. They proposed that the emergent body of composition scholars, who research in broader intellectual domains than traditional pedagogy, address the basic issues of "the nature and structure of composing processes, the context and course of writing development, the indirect effect of readers on writing, and most important, the problem of meaning in discourse" (p. 271). Specifically, they examined the formalist,

constructivist, social constructivist, and dialogist schools of thought, pointing out an aspect they feel atomistic approaches have ignored; namely, "important connections between evolving trends... responding to and conditioning the positions of those that come both before and after" (p. 271). In light of these interdisciplinary interactions, application of connectionist precepts to some of the questions raised and phenomena discussed can provide a practical example of the usefulness of these assumptions for the evaluation of writing pedagogy.

Besides the application of connectionist precepts to changing methods in writing pedagogy, seven representational composition texts, spanning the period of 1941 to the present (chronologically, Jones 1941; Warfel, Matthews and Bushman 1949; Guth 1969; Brooks and Warren 1979; D'Angelo 1984; Voss and Keene 1992; Axelrod and Cooper 1994) were examined to see how practices have conformed over time to principles derived from theoretical models. Approximate percentages of primary content were tabulated and attention paid to the authors' primary purpose(s) and their professed awareness of cognitive concerns. The examination suggested that, even though the material covered is similar in content, major shifts do mirror the trends analyzed by Nystrand et al. (1993). Moreover, attention in the discipline to cognitive theory has improved both pedagogical method and content organization in ways that are compatible with emerging understanding of neurocognition.

The first major theoretical framework, formalism, was dominated by the ideas of "features" considered inherent to ideal texts (Nystrand et al. 1993) leading to a search within text of "a stable, singular, and universal core meaning" that could be explicated in the same manner as a math problem might be solved using "standardized reader response" (p. 276). Such formalism, influenced by the behaviorist precepts of the era, thus expressed itself linguistically as "the conditioned response of language users to repeated and highly patterned verbal stimuli" (p. 277). Obviously, from a connectionist point of view, the repeated presentation of highly and consistently patterned stimuli could be expected to build prototypes by increasing connection strengths for such patterns. The result would be a higher likelihood for decision-making competition to settle into equilibrium by accepting them as ideal regularities. Thus, such methods artificially imposed a statistical regularity that could indeed increase the potentiality of communal agreement on standards.

Specifically, formalism subscribed to the following four precepts:

1. Language is composed of objective elements organized into fixed systems.
2. The meaning of texts is encoded in "autonomous" texts themselves and is explicit to the extent that writers spell things out.
3. Written texts are more explicit than oral utterances.
4. Texts are properly interpreted only when readers avoid inference about the writer or the context in which the text was written. (Nystrand et al. 1993: 278)

Again, the first idea of identifiable organization of elements lends itself to the concept of templates; however, a major difference between a connectionist and a formalist model would lie in the idea of "fixed" patterns or "fixed" systems. Connectionism, rather, views the system, and its components, in a continual state of adaption to environmental stimuli which can at any time reweight the excitation caused by input. That the texts might be considered "autonomous" given that writers "spell things out" (according to the accepted "standards") is a feature of the communal agreement and the resulting similarity of recognition choices by the preprimed members of the formalist community. The probability of reduced explicitness for oral utterances as compared to text could be interpreted as a corollary of both processing modularity and pattern dependence. Not only is it more likely that the accepted standards be more fully presented in formal written language, but when such patterns are presented orally, the speaker employs somewhat different areas of the neuronal architecture than when processing written symbols. For it should be clarified that, while current neuroscience views the brain as being distinctly multimodular, such multimodularity is a function of the overlapping of motor, sensory, and linguistic elements needed for processing, not the interaction of a geographic 'language' area with 'motor' or 'sensory' areas. Therefore, language processed orally would be inextricably linked to the motor functions needed for oral processing and auditory sensing, while language processed visually and silently would produce excitation in different cortical sections. Contributing to the pattern dependence is the statistical likelihood that pattern recognition of previously stored templates would be facilitated when alternate choices/competing expectations in the feedback loop were kept to a minimum. Thus, formalism attempted to reduce "noise" and facilitate communal agreement on best choices.

Not surprisingly, the earlier texts (Jones 1941, Warfel et al. 1949) are highly product-oriented and focus heavily on error correction —from a connectionist viewpoint, the encoding of correct prototypical patterns. Of the two, the Jones text (1941) attempts a more cognitive approach with a progression from simple rhetorical task through error correction to advanced rhetorical task; however, the organization is rather perplexing as it intersperses sections on thinking habits (such as fidgetiness), separate sections on logic, reasoning, and "How the Mind Falls into Error," unrelated (and noncontiguous) sections on "Thinking with the Emotions" and "The Human, Subjective Element in Writing," and writing products as varied as short story, business letter, and social notes. Such a mix hardly provides cogent patterns for a pattern-oriented processor. After the introduction of elementary rhetorical tasks from paragraphs to short papers (14%), the handbook (33%) does at least limit content to the most common errors in college themes. The remaining 53% addresses more advanced rhetorical tasks of both narrative and expository types, with a professed goal of attacking "twin primary faults, thinness and *formlessness*" ([emphasis added] p. v). Rhetorical exercises include reading aloud (the value of which is questionable because of the modality shift), using probes for self-criticism, examination of emotional appeal, playing devil's

advocate, and practice writing assignments, of which the four latter should have metacognitive value. Grammar sections, however, are followed by pull-out pages that handle little full text correction but rather require answers of a Yes/No, Cloze, or correct word/mark type. Such exercises play only on simple pattern recognition tasks or minimal pattern completions, thus falling far short of facilitating correct production. Vocalization of corrections is also often required as is copying of correct elements. As a broad subsequent body of research in learning theory has shown, both copying and vocalization are ineffective encoding strategies. From a connectionist viewpoint, such failure may well stem from the different modalities/brain area interactions involved, most particularly because copying has been shown to be a function of the right hemisphere not the left (Restak 1994). Specifically, the pattern connections strengthened in the copying exercise demand very basic visual recognition and motor reproduction processes that make no demands on the higher cognitive functions needed to analyze pattern regularities and encode them as prototypes. A wide body of neurological research has shown that cognitive functioning fans outward in an increasingly more complex distribution from the primitive original midbrain to the highly developed cerebral cortex. Whereas the more primitive functions tend to be more localized, more complex cognitive tasks demand greater interaction across the distributed network and more conscious attention to processing. (For an excellent summary of the major neurological findings to date, see Deacon 1997.) Thus, it should be clearly understood that even though the brain is without doubt a highly distributed network with the capacity for parallel processing, it is a hierarchical system within whose layers the extent of localization or interaction with other brain areas differs according to the complexity of processing demands.

The highly rule-focused Warfel et al. text, following a short overview of specific college uses of English, devotes the second section (47% of content) to a copious coverage of the rules and conventions of usage in English, thus focusing heavily on template encoding. The remaining section (41%) covers composition as used in newspapers, magazines, and books: process —including six principles of composition and eight methods of developing the topic sentence; the four forms of discourse; library use and research; the narrative, descriptive, and argumentative modes; and writing products —including editorials, character sketches, and feature articles. Exercises consist of using analytic probes for sample themes or text readings, topic suggestions and writing assignments, and myriad grammar tasks such as recasting, correcting, underlining, evaluating, and explaining function. The grammar tasks, particularly, are highly formulaic and minimally productive; despite the motor involvement, such practice, again, may involve processing interactions/modularities other than those involved in grammar application while writing. Also, as the student's existing grammar patterns are deeply embedded in meaning, superficial practice is unlikely to alter the associations to a point where the amended version will be selected for as best fit. Also, provision of topic ideas (rather than invention prompts), while acting as a cue to

retrieval of stored patterns, may stifle creative potential currently suspected to result from a high level of interaction between brain hemispheres (Restak 1994) rather than facilitating selection from among choices as would occur from more open-ended heuristics.

The move to a more cognitive approach to human thought necessarily generated a more constructivist interpretation of the act of writing which urged less focus on codified strategies and greater attention to the "rhetorical considerations" inherent in the writing process. This process "entails making choices and decisions about a given rhetorical problem, audience, and possible constraints . . ." (Nystrand et al. 1993: 278), and so inherently addresses the need for self-organization and predictive power. Further, Emig (1971) specifically used a computer analogy to criticize formalist tenets by suggesting that, as taught in the American system, they were "algorithmic" (p. 52). Thus, this shift to a more cognitive analysis is analogous to a shift in emphasis from the repetition and strengthening of prototypical pattern connections to the cooperative-competitive dynamics by which a selection is made from among several potentialities. Also of interest for a connectionist interpretation is the observation by Emig that "processes of writing do not proceed in *a linear sequence*: rather, they are recursive" ([emphasis added] Emig 1981: 26). Nonlinearity is an inherent feature of distributed adaptive models. (Admittedly, neural networks are restricted in the limit to which they recognize recursive generalization; however, so is the human mind [Fetzer 1992]).

Another important development, subsequent to the work of Emig, was the work of Flower and Hayes (1981) whose challenge against traditionalism Nystrand et al. paraphrased as follows: "In structuralist terms [writing] was described as the writer's translation of an underlying, *hierarchically organized* cognitive representation into text" ([emphasis added] p. 281). The intermodal and adaptive ideas taken from Flower and Hayes are seminal to connectionist interpretation:

> Readers and writers do not simply "find" meaning, [Flower and Hayes] argued; rather, they "construct" it by organizing, selecting, and connecting information in terms of mental structures. Nor is the resulting mental representation necessarily linguistic; it may also be imagistic or kinetic. Furthermore, it was said to evolve as people read situations, revise their goals, write and revise their texts. (Nystrand et al. 1993: 282)

Obviously, models such as the adaptive resonance theory of Grossberg offer great potential for analysis of both the interplay between language and imagistic or kinetic modalities and the ongoing revisional strategies. One of the advantages of the network models over earlier one-concept-one-mode models, especially those that proposed the dominance of either top-down or bottom-up processing, is that they allow for dynamic interactivity in which there is both bottom-up and top-down processing, often, as in the Grossberg model, between specialized processing units that form sublayers of a hierarchy.

Further, the work on writing heuristics originated by Flower and Hayes (1979, 1981) —and developed by Flower (1989, 1993, 1994) as cognitive rhetoric— evolved out of AI (artificial intelligence) models that were precur-

sors of connectionist architectures. Such models, however, used classical serial symbolic architectures (although in fairness it must be pointed out that the Flower and Hayes studies identified a parallel processing potential that prefigured PDP models). In support of the new networks, Elman (1990) suggested that distributed representations allow far greater flexibility than models where "there is strict mapping between concept and individual nodes" (p. 377). Even more important for teaching pedagogy, networks have the capacity to be sensitive to context, "the mechanism underlying the ability to abstract and generalize" (p. 377). Elman also suggested that connectionism offers a unique new way to approach cognition because, whereas in classical models words and similar representational elements act as building blocks from which structures are built even though the blocks themselves remain unaltered by the construction process, in neural networks "there is no separate stage of lexical retrieval. . . no representations of words in isolation. . . words are not building blocks as much as they are cues that guide the network through different grammatical states. . . " (p. 378). It follows, then, that in a multilayered architecture, such a dynamic could occur on any linguistic level —from the subsymbolic to the allophonemic to the semantic/semiotic— with input from other modalities contributing to both lexical decisions and adaptive modifications. Thus, connectionism, because of its parallel rather than serial nature and its shift into subsymbolic dimensions, represents "a Copernican revolution" (Clark and Lutz 1992: 9-10).

A shift away from pure rule generation to a greater emphasis on process —"the means by which *good writing is produced*" (p. xiii)— is observable as early as the Guth (1969) text. While some chapters of the grammar/mechanics handbook are designed for possible class use, this section (39% of text) is primarily intended as a reference source. Guth designed the text to focus on the "Whole Theme"— specifically introducing process, thesis and support, classification, and comparison and contrast, while incorporating outlining, introductions, and conclusions. It also addresses writing about literature and draws on a "greater diversity of concerns . . . from a wider range of both professional and student writing" (p. xiii). Individual chapters include observation and description, personal experience, opinion, definition, logic, persuasion, and tone and style. Specific sections on expository product cover the research paper, writing about literature, summaries, letters, and essay examinations. Guth also strived for a "cumulative program" (p. xi) —providing "less exhortation, more demonstration" (p. xiii)— that presented major principles in logical sequence; a presentation far more compatible with the expectations of a processor that seeks for regularities and measures of central tendency. Exercises include *generating* topic lists, *formulating* key questions, *inventorying* subject materials (all promoting more cognitive production/hemispheric interplay than the exercises in the earlier text), evaluating and analyzing sample texts and examples, and writing assignments. These exercises provide input and practice for various levels of processing layers —word, sentence, paragraph, and whole theme— and requests for ranking effectiveness do ask the students to supply reasons. While there is no doubt that during language

acquisition skills are aggregated through these developing levels, it is questionable whether the adult thinker —capable of item chunking— would benefit from practice in the separate levels unless some actual learning deficiency exists on one of them. All levels need attention, but in a hierarchical system they do not exist autonomously. Although the text provides possible topic themes for each section, it does ask for some class discussion and/or consensus so incorporating a collaborative dimension for cognitive bridging and negotiation of meaning. Ironically, despite the intention of demonstrating, there is not one instance of a sample text revision.

A decade later, the Brooks and Warren (1979) text significantly omits the handbook altogether devoting 9% of content to beginning to write; 46% to forms of discourse and the methods of their exposition (modes of invention); 23% to special problems of discourse such as diction, sentence, paragraph structure, metaphor, and tone/style; and 22% to research and critical writing on literary text. Thus, formalistic rules are deemphasized in favor of basic patterns/prototypes for text selection and organization on a variety of processing levels. Being convinced "that good writing is a natural expression of necessary modes of thought and not at all a matter of rules or tricks," the authors' stated purpose is to make students aware of their own individual cognition: "the student learns to write by coming to a deeper realization of the workings of his own mind and feelings, and of the way in which those workings are related to language" (p. vii). Thus, assignment topics are left to the discretion of the student or instructor (although few heuristics are provided for invention). Rather than lists of exercises, each presentation of principles is followed by an application of the elements taught, so involving brain areas needed for the skills rather than those implicated in visually and orally processing the instructions. The tasks are varied —rearrangement of sentences in logical order, a log for observations of logical fallacies, editing of jargon— and often imaginative —for example, using a Cloze technique to fill in emotional appeal in a Faulkner passage. The latter obviously provides a creative dimension to an otherwise minimal task. Nevertheless, most tasks are primarily individual, thus failing to address the communal aspect of discourse.

When analyzed from the view of cognitive processing, the concepts of social constructionism and membership in discourse communities, while indicative of a major sociological shift, contain some of the same processing implications as did formalism. As formalist pedagogy focused on the building of acceptable templates whose correct regurgitation was easily quantifiable, so must the discourse community teach its members to recognize and assimilate the "discursive practices [of the community that] constrain the ways [writers] structure meaning" (Nystrand et al.: 289). The major difference lies in whether the cognitive processor is regurgitating prototypes or using predictive (imaginative) powers to make felicitous choices from multiple possibilities; in other words, between prescription and convention.

Gardner (1972) analyzed the key characteristics of structuralism (both constructivist and social constructivist) as follows:

1. a *strategic aspect* concerning the identification of universal patterns in the flux of everyday experience,
2. a *formal aspect* involving the derivation of rules and general laws informing human behavior and institutions, and
3. an *organismic aspect* concerned with identifying the dynamics of whole organisms, behaviors, and institutions especially as such transformations affect the parts.

(Summarized in Nystrand et al. 1993: 292-293)

Within a connectionist framework, the identification of universal patterns is a property of pattern recognition/completion in an organism that looks for measures of central tendency, and the rules and laws are regularities both within the organism's environment and its reaction to that environment, all occurring through dynamic cognitive processes —some of which are necessarily predictive— designed to assist adaptation to a continual flux. Historically, a variety of underlying cognitive pattern structures have been suggested, including "frames" (Minsky 1975) and "scripts" (Schank 1976). Rumelhart (1975) also subscribed early to "schemata," and similar "knowledge structures... assumed to be the basis of comprehension," but has subsequently pointed out that they can only illustrate human cognition fully when presented in mutual interaction that "capture[s] the generative capacity of human understanding in novel situations" (Rumelhart et al., 1986a: 9).

Awareness of schemata informs the D'Angelo (1984) text, which also specifically revives with even greater emphasis the classical rhetorical tradition, much of which is itself pattern oriented. Thus, after devoting approximately 11% of text to the process of invention and the use of *paradigms* for arrangement, the author applied such principles in a *consistent pattern* to each of the classical rhetorical modes. About 22% of the text covers the *patterns* of thought —addressing purpose, audience, discourse type, invention, and paradigm within each section; 11% covers persuasion and inherent logical considerations; and 23% analyzes paragraph development, sentence- and word-level stylistic choices, and (briefly) revising and editing. Unlike Brooks and Warren (1979), D'Angelo has addended a handbook for reference (17%). The remainder of the text addresses the research paper. The social aspect of the discourse community is addressed more than in earlier texts as more exercises are designed for classroom discussion. The text does provide suggestions for writing assignment topics —in a matrix, however, that emphasizes invention as a goal— and exercise sections include topic ideas for practicing skills on primary processing levels —word, phrase, sentence, paragraph, and essay. Both the paradigms used for invention and arrangement, and the levels of analysis used for imitation and modeling are pattern-based. Nevertheless, despite the consistency of organization within each chapter, interpreted from an ideal of aggregation or chunking, the exercise sections introduce skills in an insular (independent) and somewhat random manner, and, because different topics are suggested for each skill, the exercise sets are not cumulative.

As outlined above, the basis of connectionist models is the use of pattern-based and pattern-oriented processing —dynamic cognitive processes, some of which are necessarily predictive— designed to make sense of the perceptual flux of human experience. A vital part of survival within this flux is analysis and identification of other organic entities in terms of their patterns of behavior and amendments to these patterns (transformations). For connectionists, rules are interpreted as regularities —thus being more dynamic and having the potential for amendment— and general laws are accepted regularities sanctioned and perpetuated as conventions by the related community. Such regularities are necessarily recognized, constructed, reinterpreted, and amended in terms of patterns —both micro- and macrocognitive, and internally and externally constrained— and are reflective of the pattern propensity of the human species. Such an interpretation implies less differentiation between the strategic, formal, and organismic aspects because the underlying dynamics are functions of interrelated hierarchies of pattern processing governing the human organism's need to self-organize in a perpetually changing environment. Such a framework necessarily subscribes to the postmodernist view that meaning is constructed in response to situational needs. Indeed, connectionist concepts of dynamic internal process(es) by which the highly unique organism engages in continual adaptation to the external are in no way antithetical to the "*dynamic, temporal process of negotiation,*" the "*intersubjective...* interaction between users," the "*semiotically mediate[d]* interactions, " the "*dialectical* constitutive relationship," and the focus on "*situated discourse*" that are presented as important themes in dialogism (Nystrand et al. 1993: 300-301). Further, developing comprehension of the nature of human neuronal interaction highlights ever more drastically the total uniqueness of each human brain (Posner and Raichle 1994). In the light of such singularity, any communicative language medium must be both examined and fostered in the cultural and social milieu to which it is a mediator.

More focus on discourse in the community, social awareness, increased cognitive development, and an ongoing tendency for consistent patterns of chapter organization all accompany the next primary shift away from the rhetorical modes as an organizational focus. One alternative is the Voss and Keene (1992) text, which views writing as "characterized by the writer's primary purpose" (p. viii). These authors wished to combine the product-oriented and process-oriented approaches and, while admitting that the "traditional 'modes' have value as *recognizable structures for patterning thought* or developing lines of writing, [they did] not endorse them as models for whole written discourse" ([emphasis added] p. vii). Reasons, purposes, and getting started comprise the first 19% of text, followed by subjects for writing —people, places, things, facts, and ideas (18%); *patterns* —narration, description, definition, comparison, analysis, and argumentation (26%); special applications —critical literary essays, essay tests, research papers, reports, proposals, letters and oral presentations (24%); and a handbook (12%). The text stresses the physical (motor) involvement in the writing process and gives activities for getting to know the writing community, so addressing both

intermodularity and social interaction. To facilitate processing and connection, activities are arranged in parallel fashion throughout the chapters, including highly connective bridges between readings and issues. Analytic probes examine a wide variety of elements —diction, gender bias, imagery, style, tone, audience, organization and so forth. Topic suggestions are provided. Technical aspects such as stasis theory and Toulmin's argument scheme, that might seem daunting if presented in a formal deductive manner, are gently "slipped in" inductively. Such a method might hope that the ideas become linked incidentally (by simultaneous presentation) to accompanying materials.

The Axelrod and Cooper (1994, 1997) texts focus even more specifically on the writing situation. Primary writing activities (55%) are organized around genres of increasing rhetorical difficulty. (Interestingly, the genres themselves are defined in terms of specific conventions that attest to the pattern/paradigm focus of cognitive processes). "Toolbox" sections outline skills such as critical thinking —invention and reading (5%), writing strategies— cues and modes (11%), and research strategies —investigation and methods of citation (8%). The text also provides an assessment section —examinations and portfolios (18%), as well as a handbook (3%). Thus, the writing activities constitute an identifiable body of writing practice for which the other "tools" can be used. The entire writing process as a conscious recursive dynamic is stressed as students work through invention, planning, drafting, reevaluation, and revising, following heuristics for each activity often in collaboration with teacher or peers. Content continues the classical rhetorical tradition, but real-world scenarios, general topic areas, and heuristic guidelines are explicity designed to help students elicit their own situations and topics (rather than providing them with a list of specific topics). The authors wished not only that students examine their own thought processes, but that they "learn to use writing to think critically and communicate effectively with others" (p. iii). Thus, group inquiry is assigned for both reading analysis and the writing process, and critical self-assessment follows completion of each assignment. To facilitate processing/understanding, each genre chapter also follows a parallel plan —scenarios, group activity, readings, rhetorical situation, process, revision guidelines and sample, and metacognitive self-reflection. In its effort to provide a wide variety of prompts for invention, however, the text leans toward a verbosity that visual learners may find frustrating. While the later edition is much improved visually, the text might still profit from more paradigmatic presentation of prompts and major principles. Also, students produce autonomous genres —the argumentative papers take a position, propose a solution, evaluate a subject, speculate about causes, and interpret stories (for those whose curricula include literature). More important, the tasks are not cognitively linked by such devices as the bridging activities provided by Voss and Keene (1992), thus giving less attention to the connective properties of learning dynamics.

Discussion

The question remains, therefore, as to whether the developing pedagogical focus in writing texts maximizes what is beginning to be understood about the human cognitive process. The above outline suggests clear shifts in pedagogical emphasis from a highly rule-based production of error-free, appropriately organized text to the cognitive and pragmatic process of writing, initially with a primary focus on the individual, then with a growing concern for the writer as part of a community. Gradually, organization of content has become clearer and more consistent, and idea presentation has become more concise. Such presentation subscribes to the pattern-based nature of the human processor; especially, when rhetorical and mechanical skills are connective, build upon each other, and are elicited in ways that require students to formulate their own metacognitive patterns. Notwithstanding, as discussed below, the medium of information presentation —mostly in the form of expository text— may not maximize the parallel processing of multiple modalities of which the processor is capable.

Connectionist assumptions suggest several implications for the teaching of writing and the tools used therein. Primary considerations are the human processor as a pattern completer and pattern extrapolator who seeks similarities and measures of central tendency, the hierarchical nature of the processing system, and the assumed potential for interactivity between processing modules. An overall implication is that providing heuristics to facilitate cognitive processing will not produce a felicitous "best fit" if insufficient pattern connections exist between which the parallel processor may choose. Therefore, the question of what, when, and how much is far from answered, particularly as an understanding of how language is actually processed in the brain is still in its infancy. Nevertheless, these precepts intuitively suggest some potential benefits and drawbacks of current practices.

Because of the role of pattern extrapolation and completion in cognitive processing, the classical modes of invention should not be undervalued. (Indeed, much of the rhetorical practice that emerged from the observations and experience of the ancient oral tradition is highly compatible with connectionist precepts [Angelica 1997]). These traditional guidelines that enable students to formulate their own metacognitive patterns should be encouraged —but not codified. Also of obvious cognitive value as retrieval cues are analytical probes for eliciting student understanding or existing knowledge of a subject —they constitute the initial breath of wind that sets the strings of the Aeolian harp in motion. Subsequently, the existing experiential knowledge of each individual can be modified or "adapted" in response to group or class discussion in which meaning is "negotiated" with others in the writing community. Such interaction can bridge the experiential differences unique to each individual. Ideally, such probes should address a metacognitive level above simple memory and recall. Thus, at minimum, they should develop what Rumelhart et al. refer to as 'content addressability' —the skill of retrieving complex patterns from even a minimal cue.

Student comprehension of the purpose and logic of a text is probably

greatly facilitated by consistent organization/presentation. Also an aid to processing are clear patterns such as checklists, models, and paradigms that provide readily encodable templates, and schematics for explication of text. Such patterns can also be viewed as "containers" to be filled so offering a variety of heuristic (Coe 1987). They should not, however, become ideal forms that place limits on creativity. Deconstruction and reconstruction of text to exhibit the structural parts, while initially not leading to original creativity, can contribute to the building of a repertoire of connective patterns from which the experienced writer can choose. (Analogously, a chess master's ability to select a best strategy from among hundreds of potential moves may involve some genetic predetermination, but this potential does not activate until the possible moves have been learned.) Similarly, on a syntactic level, pattern formation techniques such as sentence combination in which correct syntactic patterns are practiced may have value, although for such patterns to become habit may require more repetition than is realistic for the typical college writing classroom. The benefit of exercises providing a maximal cue to produce only minimal pattern completion, however, is unlikely.

The pattern extrapolation tendency of the language processor also explains why conventions such as topic sentences at the beginning of technical paragraphs/sections, consistent subject focus, or parallel rhetorical and organizational presentation are helpful to readers of English. (As illustrated by Bander [1978], logical and rhetorical thought patterns characterize all languages.) Making writers aware of such conventions strengthens both analytical and communicative skills. Nevertheless, such attention to reinforcing patterns and pattern awareness may seem to some reminiscent of formalism. However, it should not be construed with such rigidity —such conventions are guidelines to successful communication according to situation, audience and purpose, not formulaic laws for production. To illustrate, Axelrod and Cooper (1994, 1997), whose cognitive emphasis has been demonstrated above, categorize theses statements, topic sentences, paragraph structure, and cohesive devices such as transitions under the rubric of "Cueing the Reader," which purpose these devices indeed serve in linear English logic.

By the same token, pattern storage by connective weight is a double-edged sword. The rehearsal/repetition of patterns leads to more heavily weighted connections and a greater likelihood that those patterns be chosen as "best fit". Unlawful regularities or "errors", then, rather than signaling a deviation from ideal competence, are stored patterns that are unacceptable to the writing community. Allowing such "unlawful" regularities to be perpetuated may simply strengthen the connections that produce them. As a result, the basic pedagogical task becomes to strengthen appropriate templates and modify inappropriate ones, while the creative responsibility remains to provide sufficient choices from which the expert system may choose. Accordingly, novice writers should not sacrifice the creative process in order to stop and correct every error at the expense of creative strategies, for different brain areas are employed for differing levels of attention, tasks, and experiential memory banks. That the writer realize unfortunate choices at

some point means that the appropriate connections have been satisfactorily weighted but that, during the creative process, different neuronal clusters are in a state of excitation. (A series of PET scan studies by Posner and Raichle (1994) have already definitely identified separate anterior and posterior attentional systems, each involving several brain areas and specialized functions.) Nevertheless, any pedagogical view that stresses creativity to the point of simply having students write "a lot" with no attention to accepted conventions may well be strengthening connection weights for pattern regularities that academic discourse communities may find inappropriate.

Connectionist architectures are not only hierarchical but highly connective, and therefore may be well served by cumulative programs such as the bridging activities presented by Voss and Keene (1992). In such programs, new ideas are introduced in relationship to material already covered, so maximizing similarity matching capabilities. Approaches that work progressively from elementary to more complex tasks (as in Jones 1941, Axelrod and Cooper 1994) do address both the aggregational and the hierarchical aspect; however, autonomy or insularity of assignment presentation may fail to maximize connective potential. For example, in curricula based on modes of invention, students produce(d) autonomous essays of description, exemplification, analysis, classification, comparison/contrast, causal analysis, argument and so forth (few of which modes occur autonomously anywhere except in essay exams). However, such patterns can be presented in a related way: analysis may produce elements that can be classified according to similar characteristics, or they can be divided according to similarities and differences; analysis of causes can often lead to a suggested solution which can be argued; examples and descriptions help the reader during any exposition. Admittedly, such relationships do not hold true in all situations; nevertheless, pointing them out or asking students to find them can facilitate comprehension of how the patterns work and relate to each other.

One final connectionist issue —connections between different processing modularities— also raises pedagogical concerns. As already mentioned, the later texts (Voss and Keene 1992, Axelrod and Cooper 1994) pay particular attention to the link between the physical (motor) involvement in the writing process; and most of the texts surveyed promote discussion, so introducing an oral/aural component that increases intermodularity. Also, ideas are presented in manageable chunks of text (particularly, D'Angelo 1984, Voss and Keene 1992, Axelrod and Cooper 1994) giving some consideration to the maximal attention span which is estimated in the learning literature as somewhere between 12 and 15 minutes (Johnson, Johnson and Smith 1991). Nevertheless, the preferred method of content delivery is still prose text with little exploration of visual aids (trees, matrices, pyramids, and similar analytic schematics much used in tutorials and workshops) or imaginative prompts such as illustrations, cartoons, and color plates. Admittedly, current print technology allows better use of visual cues such as color, special type face, bulleted block indent, shaded boxing, and so forth to break up text blocks. Also, some sporadic attempts at graphics have been made. Jones (1941)

—believing that "one seeing is worth a hundred tellings" (p. v)— did experiment with both graphic devices and imaginative typeface variations: the result, however, is not felicitous. The Brooks and Warren (1979) text does occasionally include schematics to illustrate certain modes, text logic, and argument structure. Voss and Keene (1992) have reproduced graphics that accompany reading samples but themselves only use symbols (keys, checkmarks, magnifying glass) as dingbats to signal particular sections. They have also prefaced each chapter with a full color plate: none, however, is connected to text. Not only can visual variation prevent terminal glaze, but educators know that learning styles differ. Yet —whether because of tradition, a predilection for verbal processing, production costs, or the pressures of publish or perish— prose is the predominant tool in composition texts. The connectionist might well ask why, if student learning styles differ and human experience and language processing is multimodal, writing textbooks are limited primarily to the modality of the discipline and not to the parallel capabilities of the learner.

In sum, connectionism is a highly interdisciplinary paradigm which explains phenomena already noted in cognitive psychology in terms of connection strengths within and between both local and distributed neural networks. Because of its extendability to neurological reality, the connectionist paradigm promises to deliver far more critical insights into human cognitive processing than have been provided by earlier models in either psychology or linguistics. Its technology makes possible both the testing of hypotheses about brain functions and the replication of observed dynamics. While not yet providing a complete theory of language production, the paradigm suggests certain assumptions about the language processor that hold implications for the teaching of writing. Most important is that future pedagogy must recognize the learner as a *pattern* not a symbol processor, and consciously design teaching materials and methods that maximize these pattern-processing abilities. Further, while the progressively complex layers of the hierarchical cognitive system do contain localized as well as distributive functional units, current investigation suggests that the individual psyche is capable of a far more massively parallel dynamism than implied in earlier learning models. Consequently, while responding in real time to the shifting demands of external discourse communities in a complex social setting, the language user is also both motivated and constrained in discourse and strategy selection by a far more highly complex and uniquely individual internal cognitive architecture than previously imagined.

References

ANGELICA, J.C. (1997). A connectionist reinterpretation of the classical *topoi*. Manuscript in revision.

AXELROD, R.B. and C.R. COOPER. (1994). *The St. Martin's guide to writing*. (4th ed.). New York: St. Martin's Press.

AXELROD, R.B. and C.R. COOPER. (1997). *The St. Martin's guide to writing*. (5th ed.). New York: St. Martin's Press.

BANDER, R.G. (1978). *American English rhetoric: A writing program in English as a second language.* (2nd ed.). New York: Holt, Rinehart and Winston.

BROOKS, C. and R.P. WARREN. (1979). *Modern rhetoric.* (4th ed.). New York: Harcourt Brace & Jovanovich.

CLARK, A. and R. LUTZ. (Eds.). (1992). Introduction. *Connectionism in context.* Pp. 1-15. London: Springer-Verlag.

COE, R.M. (1987). An apology for form: or, Who took the form out of process? *College English* 49: 13-28.

COHEN, M. and S. GROSSBERG. (1987). Neural dynamics of speech and language coding: Developmental programs, perceptual grouping, and competition for short term memory. In S. Grossberg (Ed.), *The adaptive brain II: Vision, speech, language, and motor control.* Pp. 457-498. New York: Elsevier Science.

D'ANGELO, F.J. (1984). *Process and thought in composition.* (3rd ed.). Boston: Little, Brown & Company.

DEACON, T.W. (1997). *The symbolic species: The co-evolution of language and the brain.* New York: W.W. Norton.

ELMAN, J.L. (1990). Representation and structure in connectionist models. In G T.M. Altmann (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives.* Pp. 345-382. ACL-MIT press series in natural language processing. Cambridge, MA: MIT Press.

EMIG, J. (1971). *The composing process of twelfth graders.* Urbana, IL: National Council of Teachers of English.

EMIG, J. (1981). Non-magical thinking: Presenting writing developmentally in schools. In C. Frederiksen and J. Dominic (Eds.), *Writing: Process, development, and communication.* Vol. 2, pp. 21-30. Hillsdale, NJ: Erlbaum.

FETZER, J.H. (1992). Connectionism and cognition: Why Fodor and Pylyshyn are wrong. In A. Clark and R. Lutz (Eds.), *Connectionism in context.* Pp. 38-56. London: Springer-Verlag.

FLOWER, L. (1989). *Problem solving strategies for writing.* (3rd ed.). San Diego: Harcourt, Brace & Jovanovich.

FLOWER, L. (1993). Cognitive rhetoric: Inquiry into the art of inquiry. In T. Enos and S. C. Brown (Eds.), *Defining the new rhetorics.* Pp. 171-190. Sage series in written communication. Vol. 7. Newbury Park: Sage.

FLOWER, L. (1994). *The construction of negotiated meaning: A social cognitive theory of writing.* Carbondale, IL: Southern Illinois University Press.

FLOWER, L.S. and J.R. HAYES. (1979). *A process model of composition.* Technical Report No. 1. Washington, DC: Document Design Project.

FLOWER, L.S. and J.R. HAYES. (1981). A cognitive process theory of writing. *College Composition and Communication* 32: 365-387.

GARDNER, H. (1972). *The quest for mind: Piaget, Levi-Strauss, and the structuralist movement.* New York: Vintage.

GROSSBERG, S. (Ed.). (1987). *The adaptive brain II: Vision, speech, language, and motor control.* New York: Elsevier Science.

GROSSBERG, S. (1998). Birth of a learning law. *INNS/ENNS/JNNS Newsletter* 21: 1-4.

GROSSBERG, S. and G. STONE. (1987). Neural dynamics of word recognition and recall: Attentional priming, learning, and resonance. In S. Grossberg (Ed.), *The adaptive brain II: Vision, speech, language, and motor control.* Pp. 403-455. New York: Elsevier Science.

GUTH, H.P. (1969). *Words and ideas.* (3rd ed.). Belmont, CA: Wadsworth.

JOHNSON, D.W., R.T. JOHNSON, and K.A. SMITH. (1991). *Active learning: Cooperation in the college classroom.* Edine, MN: Interaction Book Company.

JONES, E.S. (1941). *Practical English composition.* (3rd ed.). New York: Appleton-Century-Crofts.

MINSKY, M. (1975). A framework for representing knowledge. In P.H. Winston (Ed.), *The psychology of computer vision.* Pp. 211-277. New York: McGraw-Hill.

NYSTRAND, M., S. GREENE, and J. WIEMELT. (1993) Where did composition studies come from? An intellectual history. *Written Communication* 10(3): 267-333.

PETERSEN, S.E., T.T. FOX, M.I. POSNER, M. MINTON, and M.E. RAICHLE . (1988). Positron emission tomographic studies of the cortical anatomy of single word processing. *Nature* 331: 585-589.

POSNER, M.I. and M.E. RAICHLE. (1994). *Images of mind.* New York: Scientific American Library.

RESTAK, R.M. (1994). *The modular brain.* New York: Charles Scribner's Sons.

RUMELHART, D.E. (1975). Notes on a schema for stories. In D.G. Bobrow and A. Collins (Eds.), *Representation and understanding.* Pp. 211-236. New York: Academic Press.

RUMELHART, D.E., J.L. MCCLELLAND, and the PDP Research Group. (1986a). *Parallel distributed processing: Explorations in the microstructure of cognition.* Vol. 1. Cambridge, MA: MIT Press.

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

SAMPSON, G. (1987). [Review of the article Parallel distributed processing: Explorations in the microstructures of cognition]. *Language* 63(4): 871-886.

SCHANK, R.C. (1976). The role of memory in language processing. In C.N. Cofer (Ed.), *The structure of human memory.* Pp. 162-189. San Francisco: Freeman.

SCHNEIDER, W. (1987). Connectionism: Is it a paradigm shift for psychology? *Behavior Research Methods, Instruments and Computers* 19(2): 73-83.

SHEPHERD, G.M. (1994). *Neurobiology.* (3rd ed.). Oxford: Oxford University Press.

STONE, G.O. (1994). Combining connectionist and symbolic properties in a single process. In S.D. Lima, R.L. Corrigan, and G.K. Iverson (Eds.), *The reality of linguistic rules.* Pp. 417-444. Amsterdam: John Benjamins.

VOSS, R.F. and M.L. KEENE. (1992). *The Heath guide to college writing.* Lexington, MA: D.C. Heath & Company.

WARFEL, H.R., E.G. MATTHEWS, and J.C. BUSHMAN. (1949). *American college English.* New York: American Book Co.