Linguagem & Ensino, Vol. 8, No. 1, 2005 (161-183)

A new paradigm for learning language: Connectionist artificial inteligence*

(Um novo paradigma para a aprendizagem da linguagem: Inteligência Artificial Conexionista)

José Marcelino POERSCH
Pontifícia Universidade Católica do Rio Grande do Sul

ABSTRACT: Important improvements in the area of the connectionist paradigm were designed and executed in the last two decades. Connectionist modeling techniques have been used to help us understand how individuals acquire, maintain and, in some cases, lose mental functions. Connectionist models involve parallel distributed processing (PDP). Despite its successes, connectionism is far from presenting a final solution to all the problems of cognition. Our claim is that connectionism has better an explanatory power than constitutes a simulation of real brain processes.

RESUMO: Nas últimas duas décadas, principalmente a partir de 1986, significativos progressos na área do paradigma conexionista foram planejados e executados. As técnicas de simulação conexionista colaboraram para compreender melhor a maneira como as funções mentais

---

* This is an adapted version from the oral communication presented at the 17th International Congress of Linguists (Prague, 22 to 29 July 2003).

são adquiridas, armazenadas e, em certos casos, perdidas. Os modelos conexionistas baseiam-se num processamento distribuído em paralelo (PDP). Apesar de suas evidentes e valiosas contribuições, o conexionismo está longe de apresentar uma solução definitiva para todos os problemas da cognição. Somos da opinião de que esse paradigma corresponde mais a uma força explicativa do que corresponde a uma simulação perfeita dos verdadeiros processos cerebrais.

KEY-WORDS: modern artificial intelligence, language learning, connectionist modeling, neural networks, linguistic cognitive processes.

PALAVRAS-CHAVE: inteligência artificial moderna, aprendizado da língua, modelagem conexionista, redes neuroniais, processos cognitivos da linguagem.

INTRODUCTION

Connectionist modeling of language processing is a highly controversial activity. While some scholars argue that the modeling can be understood in connectionist terms, others argue that no aspects of language can be fully captured by connectionist methods. Recently, some of the limitations of connectionism have been overcome, re-opening the possibility that connectionism constitutes not an additional method but an alternative model of thought. Connectionism, based on a “neural inspiration”, means that the brain consists of a very large number of simple processors, neurons, which are densely interconnected into a complex network.

A large number of them operate simultaneously and co-operatively to process information. Furthermore, neurons appear to communicate numerical values rather than symbolic messages, and therefore neurons can be viewed as mapping numerical inputs onto numerical outputs. So, a neural network is a massively distributed processor that has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects: 1. Knowledge is acquired by the network through a learning process. 2. Interneural connection strengths known as synaptic weights are used to store the knowledge.

While symbolism processing, based on general purpose digital computers, aimed to model the mind as a symbol processor, connectionism (parallel distributed processing = PDP) has a different origin; it attempts to design computers inspired by the brain. The number in which the neurons of a neural network are structured is intimately linked with the learning algorithm used to train the network. These algorithms are structured in different layers: the input layer of source nodes projects onto an output layer of neurons. Between them are the hidden units responsible for the learning process of the machine network.

This article aims to present an alternative paradigm for language acquisition. So, an introductory debate on the theoretical bases of the connectionist paradigm is promoted. Furthermore, important aspects of modern artificial intelligence instanciated by connectionist modeling, is presented. First, we try to answer the question “Why a new paradigm for cognition?” Next, we present the characteristic features of connectionist. The central topic of the text is: What does connectionism modeling consist on? On this topic we analyze the meaning of neural networks, their architecture, and their capacity for learning. Finally, we present three simulations performed at the Center for Linguistic Research of the Pontifical Catholic University (RS – Brazil).

WHY A NEW PARADIGM FOR COGNITION?

The science of cognition is the area of knowledge that studies the input, the storage, the processing and the retrieving of knowledge, whether this knowledge is declarative or procedural, either natural or computer-simulated (artificial intelligence).

Science consists of a constant search for truth, that is, a search for the theories that explain certain natural phenomena. The only theories that are truly scientific are the ones that present possibilities of assessment, theories whose veracity can be challenged. The existing theories must be continuously re-evaluated and tested; therefore new theories will come up due to the limitations of the preceding ones (Poersch, 1998). It is the scientist's responsibility to find out the positive aspects and the limitations of each theory.

Among the acquisition theories, there are two classic antagonist paradigms based on distinct philosophical backgrounds: behaviorism and mentalism (symbolism).

The behaviorist paradigm (Fig.1), based on the empiricist philosophy, emphasizes the senses and experience in order to approach the process of knowledge acquisition. It is a neuron-based paradigm; it denies the existence of the mind. Knowledge is learned through stimulus and response.

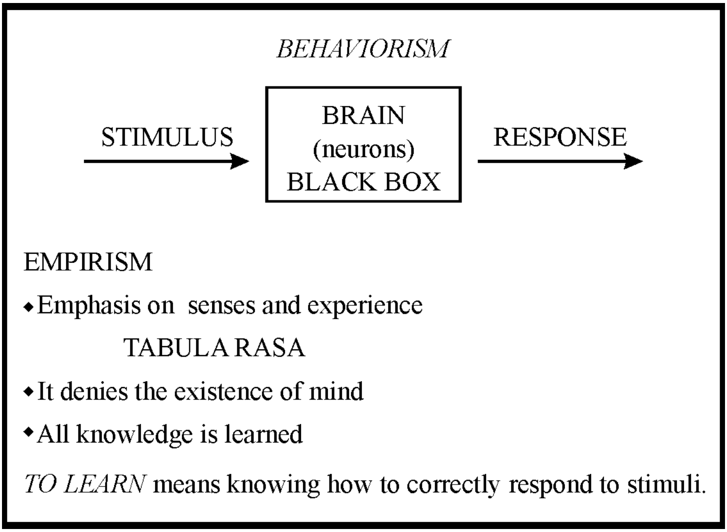

Figure 1 – Behaviorist Paradigm

The mentalist paradigm (Fig.2) emphasizes the role of the mind in the cognitive processes. Mind and brain are two realities of different substances, the first is spiritual, and the second one is physical. The higher-level cognitive processes take place in the mind where long-term memory is found. This paradigm postulates the existence of innate ideas (rules). Cognition is processed through the representation of the world in the mind by means of a serial processing of abstract and fixed symbols.

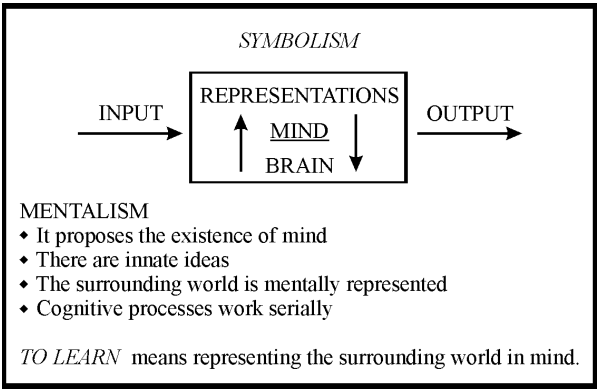

Figure 2 - Mentalist Paradigm

Once science never stops searching, it has found in this paradigm a number of limitations, of unexplainable aspects. The following ones are the most significant:

- How is knowledge, coded in the brain (a physical substance), filed-transferred into the mind (a spiritual substance)?
- Concepts are abstractions. How it is possible for an abstract reality that does not occupy a physical space to be stored in an abstract place, the mind?
- The mind stores the symbols that represent the reality of the world and are the object of declarative knowledge. How can procedural knowledge, not represented by symbols, be stored?
- What explanation can be given to the fact that speaking and writing, the product of a vast number of parallel problem-solving processes, turn into a serial sequence of sounds or letters?
- How does thought, an abstract and analogical reality, become language, a concrete and digital one?
- The verbal sign is made up of an external part, the symbol, and an internal one, the linguistic sign. The symbol and its

object are two distinct realities of a concrete nature; the linguistic sign results from the association of a concept and a sound representation, two abstract realities. How can abstract realities activate one another in the mind?

These unanswered questions force scientists to experiment with a new paradigm that makes up for these limitations. This paradigm is connectionism.

CHARACTERISTICS OF CONNECTIONISM

Connectionism (Fig.3) is a cognitive paradigm based on the findings of neuroscience and not on explanatory hypothesis. All cognitive processes take place in the brain; the mind is nothing more than the grouping of these processes. The mind is not an *"ens in se"*; it is a phenomenon that actually occurs, it is an *"ens in altero"*.

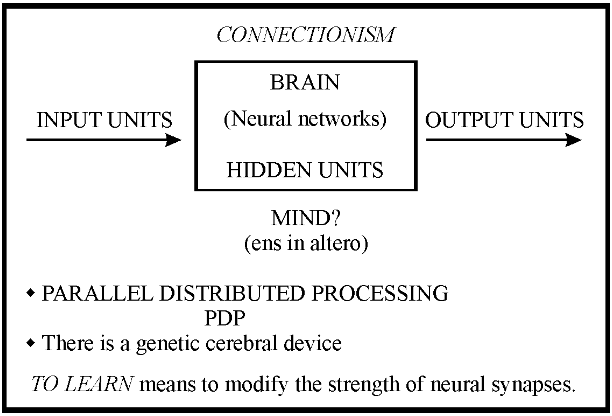

Figure 3 - Connectionist Paradigm

The brain contains thousands of neurons connected in parallel which form inter-neural nets. Each neuron (Fig.4) is constituted of a body and two kinds of filaments responsible for the net formation: the axons, electrical transmitters connecting a neuron body to a synapse, and the dendrites, electric impulses connecting the synapse to other neurons.

Where an axon reaches a dendrite there is a space in which chemical reactions are processed: the synapses. These reactions are responsible for learning. Learning means modifying the synaptic forces.

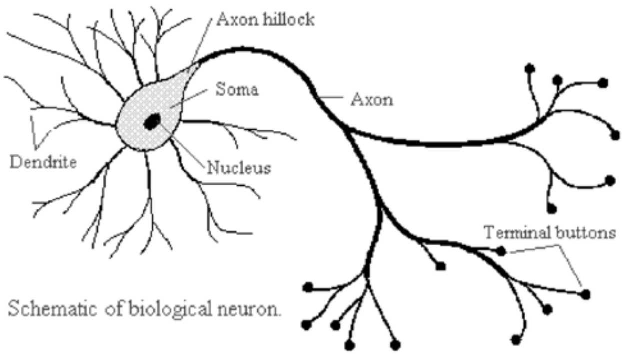

Figure 4 – Representation of a biological neuron

The brain possesses a genetic innate mechanism, a genetically encoded knowledge that makes its functioning possible. There are no innate rules for language processing (Rumelhart & Mc Clelland, 1986); the rules are inferred through a statistical data processing coming from experience (Seidenberg & Mac Donald, 1999). The declarative knowledge of the language and of the world, as well as the procedural knowledge from a variety of skills are coded in the brain not in the shape of fixed symbols that occupy designated places, but as fine-tuned elements distributed in different neurons connected between themselves. The processing does not occur serially as in the information theory, but in parallel, that is, many processes take place simultaneously.

There are important studies that simulate the functioning of the brain. The modeling is not done through algorithms that guide the functioning in a serial way, but through connectionist neural networks containing a device for learning from the input data. The meaning and the functioning of these networks are explained in the next section.

WHAT DOES CONNECTIONIST MODELING CONSIST ON?

One of the main goals of connectionism is to provide an account of the mechanisms that support cognitive processing (Poersch, 2001). Connectionists are interested in describing the internal states of brain activity even though they may view them as fundamentally associative in nature. Connectionist models are getting more and more complicated. In the early days of the connectionist revival, researchers tried to impress upon their audience that much could be achieved with rather simple models that did away with a lot of the excessive baggage of classical cognitive theories. "Today, we see increasingly sophisticated connectionist models in use as researchers attempt to explain a greater range of facts and exploit our rapidly deepening knowledge of real neural systems in the brain (Plunkett, 2000, p.111). It remains to be seen whether these additional sources of constraints on connectionist model building amount to a reinvention of the principles proposed by cognitive psychology decades before – even though in associative clothing. Nevertheless, it is apparent that connectionism has established itself as a major player in cognitive science.

What are neural networks?

The work on neural networks has been grounded on the recognition that the brain computes in an entirely different way from the conventional digital computer (symbols serially combined). The brain possesses a tremendous big number of neurons, massively interconnected between each other. The result is that the brain constitutes an enormously efficient structure.

The brain is a highly complex, nonlinear, and parallel computer. It has the capacity of organizing neurons so as to perform certain computations many times faster than the fastest digital computer. It has a specific structure and the capacity for constructing its own rules through experience. This experience is constructed over the years, with the most dramatic development of the human brain in the first years, producing millions of synapses per second.

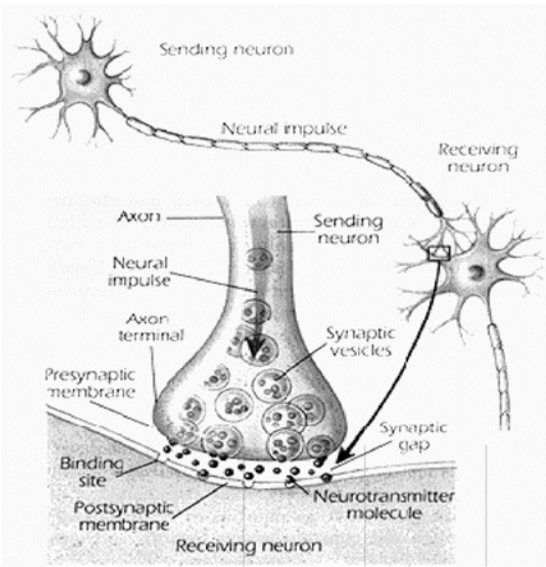

Figure 5 – How neurons communicate (From users.rcn.com/ jkimball.ma.ultranet/ BiologyPages/~/Neurons.html)

"Synapses (Fig.5) are elementary structural and functional units that mediate the interaction between neurons" (Haykin, 1994, p.2)". A presynaptic process liberates a transmitter substance that diffuses across the synaptic junction between neurons and then acts on a postsynaptic process. Thus a synapse converts a presynaptical electric signal into a chemical signal and then back to a postsynaptic electrical signal. It is assumed that synapses are simple connections that can impose reciprocal activations between neurons. An important feature of brain is the plasticity that synapses provide it with. This plasticity permits the developing neuron system to adapt to its surrounding environment. Synapses are performed by means of two cell filaments: the axon and the dendrite.

Just as plasticity appears to be essential to the functioning of neurons in the human brain, so it is with neural networks made up of artificial neurons. A neural network is a machine designed to model the way in which brain performs a particular task or function of interest. The network is usually implemented using electronic components or simulated in software (algorithmic program) capable of performing useful computations through a process of learning using a massive interconnection of simple "processing units".

"A neural network is a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use" (Haykin, 1994, p.2). The procedure used to perform learning processes is called a "learning algorithm"; the function of this algorithm is to modify the synaptic weights of the network in order to attain a desired design objective.

How are neural networks designed?

At the heart of any connectionist model is an interconnected web of processing units. It is helpful to think of each processing unit as a neuron that receives activity from other neurons through synaptic connections (Fig.6). Like real neurons in the brain, the activity of a connectionist neuron depends on the amount of activity reaching it. Synapses between neurons can be excitatory or inhibitory, strong or weak. The pattern of connectivity in a connectionist network determines how it will respond to sensory input or information from other networks with which it communicates. In fact, the pattern of connectivity defines what the network knows about the problem it has been designed to work on.7

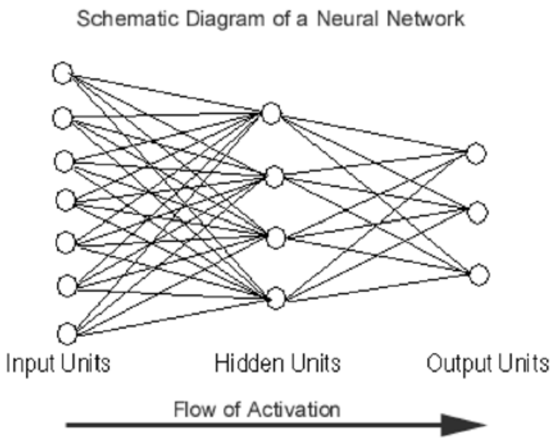

Figure 6 – Representation of a neural network (From www.citations.neural/networks/Haykin.html)

An important aspect of connectionist networks is their ability to learn. Most connectionist models come equipped with a built-in learning algorithm that enables them to learn from their experiences. There are a wide variety of learning algorithms in use today. These algorithms alter the strength of the connections in the network in response to neuronal activity evoked by sensory input or information from other networks. By changing the connections between neurons, the network encodes information about its environment (Fig.7).

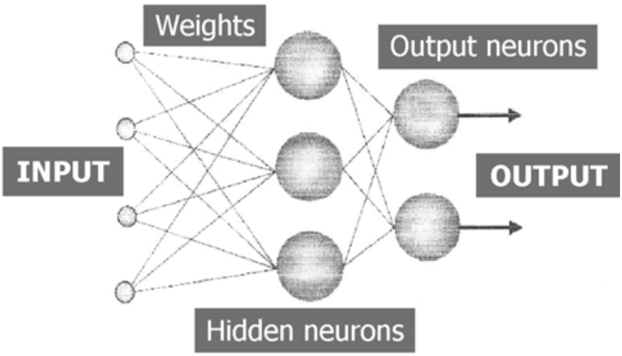

Figure 7 – The weight of connections is responsible by coding the information (From www.citations.neural /networks/Haykin.html)

Connectionist models come in a range of flavours, each with their own architectural constraints, set of learning rules and assumptions about how the environment is presented to the model. All these factors conspire to constrain the performance of the model and its ability to learn about the environment. A judicious choice of network architecture and learning rule may be all that is required to ensure a particular outcome given a particular set of experiences. The problem is to identify the characteristics of the system that yield these constraints. These characteristics offer a comprehensive strategy for investigating a range of connectionist models and their application to different types of complex cognitive and linguistic domains. A common strategy is to search for the simplest type of network architecture (consistent with what we know about brain structure) that is able to match the behavioural data when exposed to a structured environment.

An important insight is to recognise the trade-off between environmental resources and computational/ architectural complexity, and the timing of the interaction between them. Connectionist modellers explore this trade-off by investigating a wide range of conceptual assumptions associated with the nature of the environment, the computational machinery and its appropriate application (Plunkett, 2000).

The learning algorithms

Connectionist models can be trained to perform a wide variety of tasks, e.g. predict the reappearance of an object from behind a screen, change a verb into its past form, predict the next word in a sentence, categorize objects, categorize speech sounds, pronounce written text or, catch a ball. In each case, the learning algorithm (Fig.8) fine-tunes the strength of the connections in the network until adult-like performance is achieved. The network can then be analysed to see how it performs the task, providing a source of hypotheses as to how adults may perform the task.

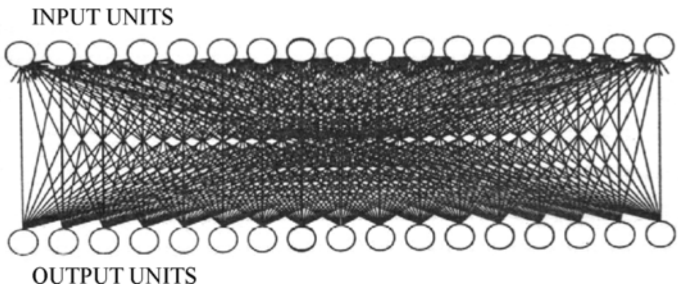

Figure 8 – The learning algorithms (From Plunkett, 1997, p.44)

It is also possible to examine the developmental profile of the network going to its mature state by taking snapshots of the network at regular intervals. If the behaviour of the network during training mimics the behaviour of the child during development, perhaps the snapshots can tell us something about the state of the child at different points in development. Likewise, if damaging the network produces unusual patterns of performance that resemble behaviours in disordered populations, then we might be able to gain some insights into the causes of these disorders.

Networks learn by changing the strength of the connections in response to neuronal activity. These changes usually take place gradually (often determined by a parameter called the learning rate). Usually consecutive learning experiences reinforce each other.

The success of connectionist networks in mimicking behavioural, cognitive and linguistic development lies in their sensitivity to the statistical regularities inherent of the environment. It is important to choose the right kind of network to reach these statistics. Once chosen, the network (or system of networks) can integrate information from multiple sources and modalities to construct cognitive representations that could not have emerged from isolated domains. The whole is greater than the sum of the parts. Connectionist modeling offers the psychologist with a powerful tool to investigate interactionist, epigenetic accounts of development, encompassing not just general developmental profiles but individual differences in learning and critical period effects.

SAMPLES OF SIMULATIONS

Here are three simulations realized at the Centre for

Linguistic Research (Pontifical Catholic University – Brazil) under the advisory of Poersch and the helpful assistance of internationally recognized centers for connectionist researches.

a) The acquisition of passive constructions: a cross-linguistic study.

Rosangela Gabriel designed and built the connectionist neural network of her simulation at the Department of Experimental Psychology (Oxford University) under the assistance of Kim Plunkett and the use of the T-learn program.

This investigation attempts to contribute for the understanding of the nature of language and mind. The passive construction is a focus of considerable interest in psycholinguistic research in the last decades. Located in the field of language acquisition, this investigation aims to shed light on the following question: How do children learn the passive constructions? To answer this question, two techniques have been used: analysis of empirical data and computer simulation of neural processing. The empirical data come from four studies that differ from other similar studies. The crosslinguistic results provided evidence for modeling the acquisition and processing of passive constructions in a computer neural network. A connectionist neural network was built, relying on the assumption that learning is based on associative processes involving modifiable synaptic weights and connections between networks of simple computing units. The research stresses the need for constant interchange between empirical data, neurological findings and computational techniques.

b) The learning of inferencing strategies in reading.

Ana Elisa Sigot had the assistance of Walter Kintsch and Eillen Kintsch and designed the architecture of her simulation with the help of Randall O'Reilly using the LEABRA++ software, at the Institute of Cognitive Science of the University of Colorado at Boulder.

It is generally assumed that the environment where students learn a language influences their reading performance, since they are exposed to different linguistic input and contexts of the utterances. Based on this assumption, the differences in the construction of the mental representation of texts produced by Brazilian students learning English as a foreign language both in Brazil and in the United States was investigated. An analysis of the reading inferencing of these students was realized, based on two different approaches: The construction of the textbase and the situation model representations and a much lower level of analysis using connectionist networks. We collected and analysed some empirical data from undergraduate students learning English as a foreign language both in Brazil and in the United States was collected and analysed. Afterwards a simulation was run to model some of the relatedness data and observe the differences in the data obtained empirically and the results from the simulations. The results showed that there is a clear difference between both groups of students in terms of their text representations. The differences found in the performance of the networks reflected the results obtained from the empirical data, that is, there were differences with respect to the generation of inferences produced by adults learning English as a foreign language both in Brazil and in the USA. In the reading process there is an interaction of both explicit and implicit information. This study on the processing of inferences approaching different text representations constitutes an effort to explain the interplay between the explicit and the implicit aspects of text processing.

c) The cross-linguistc tranfer of reading processes in L2 reading.

Marcia Cristina Zimmer designed the neural network at the Department of Psychology of Carnegie Mellon University under the supervision of David Plaut using the LENS simulator.

One of the main aspects in which second language acquisition (SLA) differs from first language acquisition is that L1 patterns are usually transferred to L2. Grapheme-phoneme knowledge seems to be transferred when Brazilian Portuguese (BP) speakers name English words. As most graphemes are common to both languages, but the corresponding phonemes are not, many English words may be pronounced with a Brazilian accent because of a grapheme-phoneme conversion bias, that is, the tendency to assign L2 graphemes the same or a similar phoneme they would activate in their L1 phonological system. Although naming English words has been studied in connectionist accounts during the last 15 years, it has not been studied in English as a Second Language. The aim of the present study was to look into the Portuguese-English grapheme-phoneme transfer processes among 157 adult Brazilian ESL students - divided into four groups of ESL proficiency (beginner, intermediate, upper intermediate and advanced) – during phonological recoding sessions of high and low frequency regular words, exception words and non-words. The sessions have been recorded and phonetically transcribed and the transfer processes in which the participants engaged were listed. After the empirical research was completed, the computational modeling of reading aloud in BP has been undertaken.

The findings of this investigation lead us to the conclusion that adult learners may normally show an accent when reading orally in a second language because their cognitive system has largely been used to solve a lot of oral reading problems in their mother tongue; the perception of the phonetic categories of the mother tongue models the learner's phonetic space. The child, by its turn, probably attains a better performance because his cognitive system has not yet been totally framed by his mother tongue. So, we can formulate a connectionist explanation of linguistic transference: the phonetic knowledge from the mother tongue and possibly from other foreign languages which the learner may have been exposed to influences the perception and the production of the phones in a specific foreign language.

CONCLUSION

Connectionist modeling of language processing has been highly controversial. While some scholars argue that no aspects of language can be fully captured by connectionist methods, others state exactly the contrary. "And the controversy is particularly heated because, for many, connectionism is not just an additional method for studying language processing, but an alternative to the traditional symbolic accounts. Indeed, the degree to which connectionism supplants, rather than complements, existing approaches to language is itself a matter of debate (Christiansen and Chater, 1999, p.417).

Connectionism, differently from traditional digital computers that follow symbolic rules, designed computers inspired by the brain, computers that learn from input data, from the experience. Differenly from the group including Fodor & Pylyshyn (1988), Pinker & Prince (1988), and Smolensky (1988) who typically assume that connectionist modeling should start with symbol processing models and be implemented by connectionist nets, and differently from Chater and Oaksford (1990) that argue in favor of a two-way influence between symbolic and connectionist theories, *radical connectionist* in the field of language processing state that the new paradigm substitutes, rather than implements, the symbolic approach.

Seidenberg and MacDonald (1999) as well as O'Brien and Opie (2002) also argue that connectionist models will be able to replace the currently dominant symbolic models of language structure and language processing, throughout the cognitive science of language. They suggest that connectionist models exemplify a probabilistic, rather than a rigid rule guided, view of language, that requires the foundations of linguistics as well as the cognitive science of language more generally to be *radically* rethought.

O'Brien and Opie (2002, p.327) have defended radical connectionism. They state that "Radical connectionism claims, as against both classicism and ecumenical connectionism, that cognition never involves an internal symbolic medium, not even when natural language plays a part in our thought processes. On the face of it, this renders the human capacity for abstract thought quite mysterious. However, we've argued that connectionism, because it adopts an analog conception of neural computation, is committed to a structural resemblance theory of representational content. Representation of the abstract is no more problematic for a system of analog vehicles that structurally resemble their target domain, than for a symbol system. Natural language is therefore not required as a representational medium for abstract thought. Indeed, since natural language is arguably not a representational medium at all, but a conventionally governed scheme of communication, the role of internalized (i.e. self-directed) language is best conceived in terms of the coordination and control of cognitive activities within the brain.

Connectionism is beginning to have a considerable influence on the science of psycholinguistics. The final extend of this influence depends on the degree to which practical connectionist models can be developed and extended to deal with complex aspects of language processing in a psychological realistic way. If realistic connectionist models of language processing can be provided, then the possibility of a *radical* rethinking not just of the nature of language processing, but of the structure of language itself, may be required.

REFERENCES

CHATER, N. & OAKSFORD, M. Autonomy, implementation, and cognitive architecture. A reply to Fodor and Pylyshyn. *Cognition*, n.34, p.93-107, 1990.

CHRISTIANSEN, M.H. & CHATER, N. Connectionist natural language processing: the state of the art. *Cognitive Science*, v.4, n.3, p.417-437, 1999.

FODOR, J.A. & PYLYSHYN, Z.W. Connectionism and cognitive architecture: A critical analyses. *Cognition*, n.28, p.3-71, 1988.

GABRIEL, Rosângela. *A aquisição das construções passivas em português e inglês: um estudo translingüístico.* 212f. Tese de doutorado orientada por José Marcelino Poersch, Programa de Pós-Graduação em Lingüística, PUCRS, Porto Alegre, 2001.

HAYKIN, S. *Neural networks: A comprehensive* *foundation*. New York: Macmillan College Publishing Company, 1994.

O´BRIEN, G. and OPIE, J. Radical connectionism: thinking with (not in) language. *Language & Communication*, n.22, p.313-329, 2002.

PINKER, S. & PRINCE, A. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition*, n.28, p.73-193, 1988.

PLUNKETT, K. O conexionismo hoje [connectionism today]. In: Poersch, J.M. (Eds.) *Psicolingüística, ciência e arte.* [Psycholinguistics, science and art], p.109-122. Porto Alegre: EDIPUCRS, 2000.

POERSCH, José Marcelino. Contribuições do paradigma conexionista para a aquisição do saber lingüístico. *Anais do 4º Encontro Nacional de Aquisição da linguagem*, p.37-72. Porto Alegre: CEALL, 1998.

POERSCH, José Marcelino. A apropriação do conhecimento lingüístico: uma abordagem conexionista. *Anais do 1º Encontro Internacional de Aquisição da linguagem*, p.399-432. Porto Alegre: CEALL, 2001.

RUMELHART, D.E. & MCCLELLAND, J.L. On learning the past tenses of English verbs. In: McClelland, J.L., Rumelhart, D.E. (Eds.). *Parallel distributed processing*, Cambridge, MA: MIT Press, v.2, p.216-271, 1986.

SEIDENBERG, M.S. & MACDONALDS, M.C. A probalistic constraints approach to language acquisition and processing. *Cognitive Science*, n.23, p.569-588, 1999.

SIGOT, Ana Elisa. *The processing of inferences in reading in English as a foreing language according to a* *connectionist approach*. 159f. Tese de doutorado orientada por José Marcelino Poersch, Programa de Pós-Graduação em Lingüística, PUCRS, Porto Alegre, 2002.

SMOLENSKY, P. On the proper treatment of connectionism. *Behavioral and Brain Sciences*, n.23, p.589-613, 1988.

ZIMMER, Márcia Cristina. *A transferência do conhecimento grafema-fonema do português brasileiro (L1) para o inglês (L2) na recodificação leitora: uma abordagem conexionista*. 220f. Tese de doutorado orientada por José Marcelino Poersch, Programa de Pós-Graduação em Lingüística, PUCRS, Porto Alegre, 2003.

Recebido: Setembro de 2003.
Aceito: Dezembro de 2003.

Endereço para correspondência:

José Marcelino Poersh
Av. Parobé, 3435 Cx. Postal 359
93001-970 – São Leopoldo – RS
poerschjm@pucrs.br