Journal of Applied Linguistics and Language Research
Volume 1, Issue 2, 2014, pp. 27-33
Available online at www.jallr.ir
ISSN: 2376-760X

The Localist and the Distributed Models of Connectionism

Morteza Montazeri
Department of English Language, Sari Branch, Islamic Azad University, Sari, Iran

Hadi Hamidi
Department of English Language, Sari Branch, Islamic Azad University, Sari, Iran

Bahman Hamidi
Department of English Language, Tonekabon Branch, Islamic Azad University, Tonekabon, Iran

Javad Yaghoobi
Department of English Language, University of Mazandaran, Babolsar, Iran

Abstract
Connectionism is the theory that sees brain in terms of neural or parallel distributed processing networks of interconnected units. The present paper reviewed the basic assumptions of connectionism and two main types of connectionist models were explained; the localist model and the distributed model. The drawbacks of the localist connectionism were mentioned. Properties of distributed connectionist networks were delineated. In the end, general problems with connectionist models were discussed. It was mentioned that the major drawback of connectionism that would cast doubt on the usefulness of a connectionist approach was that this approach had its basis on the sciences of math and physics, while the brains of human beings, or language learners, are biological entities. This seems to mar the usefulness of this approach to language learning, since it can be hardly assumed that the mathematical principles can be extended to biological ones. Language learners, language teachers as well as neurologists and psychologists may find the discussions of the present study useful in the process of language acquisition.
Keywords: connectionism, distributed model, localist model

INTRODUCTION

Connectionism is the theory that sees brain in terms of neural or parallel distributed processing networks of interconnected units (Piske & Young-Scholten, 2009). These connections, as Piske and Young-Scholten (2009) maintain, are either strengthened or weakened through activation or nonactivation. They further assert that connectionist approaches to language acquisition argue that language is learned by learning rules from input alone, with no LAD involvement. According to connectionism, the mind

Correspondence: Hadi Hamidi, Email: hamidi_tefl@yahoo.com

makes links or connections between information nodes and the creation of new links when new input is received means that the network becomes progressively larger and more complex (De Angelis, 2007).

Connectionism, according to Jordan (2004), rejects the assumption made by nativists that the brain is a symbol processing device similar to a digital computer. He further argues that the brain relies on a type of computation that emphasizes patterns of connectivity and activation.

BASIC ASSUMPTIONS OF CONNECTIONISM

The basic assumptions of connectionism are as follow:

1- Information processing takes place through the interaction of a large number of simple units which ate organized into networks and operate in parallel,
2- Learning occurs through strengthening and weakening of interconnections in a particular network in response to examples faced with in the network, and
3- The result of learning is often a network of simple units acting as though it knows abstract rules, in spite of the fact that rules already exist only in the form of association strengths distributed across the entire network (Richards & Schmidt, 1985)

TWO TYPES OF CONNECTIONIST MODELS

Connectionism, according to Rast (2008) involves the use of computer processing in order to stimulate the functions of the mind and predict how human will act under different conditions. A connectionist architecture consists of network of a large number of interconnected elements called nodes, and the knowledge of the networks lies in the information given to these nodes and the strength of connection between nodes (Rast, 2008), and the associations between these nodes are, according to Mitchell and Myles (2004), named connection strength or pattern activation, the strength of which changes with the frequency of input and nature of feedback. Atkinson (2011) states that the connectionist models of cognition are self-organizing systems-their structure emerges in direct response to environmental input. Connectionist networks, hence, show nontrivial environmental engagement and their complexity results from environmental complexity, rather than being pre-built into the cognitive system (Ellis, 1998, cited in Atkinson, 2011).

As Rast (2008) maintains, generally we can distinguish between two major types of connectionist models: localist symbolic models and distributed subsymbolic models. In the localist tradition, as he explains, representations are coded for distinct pieces of information, such as visual information for all letters in the alphabet or in the entire word. In distributed subsymbolic, however, information or knowledge is coded as pattern activation across many processing units which contribute to a number of different representations. “The focus of these models is on the emergence of skilled human performance through learning” (Rast, 2008, p. 6).

The localist model

Elman (1990) contrasts between localist schemes of connectionism and distributed schemes of connectionism in that in localist schemes each node stands for a separate concept, and acquiring new concepts requires adding new nodes, whereasin distributed schemes concepts are expressed as activation patterns over a fixed number of nodes.In localist approach, as Elman later in 1999 maintains, nodes are assigned discrete interpretations. In such models, nodes may represent grammatical roles or relations, and these may be bound to other nodes which represent the word-tokens which instantiate them either by concurrent activation or other techniques (Elman, 1991).

Gasser (1990) contends that in localist approaches units represent particular concepts like blue, Elvis Presley, and transitive clauses and in distributed approaches complex concepts are distributed over many units, and each unit participates in the representation of many concepts. The localist model is argued by Ingram (2007) to be of limited learning capabilities. He contends that this model has its functional architecture hard-wired, in which every unit has a designated task. It is, for example, incapable of learning to recognize new words, which is according to Elman (1995), intrinsically context free, and in order to introduce each new word, one has to rewire the system (Ingram, 2007). Ingram (2007) further maintains that improvements in the spatial and temporal resolution of functional neural imaging have tipped the evidence in favor of a modular and localist account of sentence processing. Furthermore, Ellis (1999) contends that frequency of chunks in the input, and regularity and consistency of associative mappings with other representational systems, is conducive to the emergence of effectively localist units, especially at the lexical level.

TRACE model

The TRACE model of spoken word recognition is an elaboration of the localist network of word reading (McClelland &Rumelhart, 1981, cited in Ingram, 2007).This model according to Ingram (2007), was first the connectionist model of recognition/retrieval in order to demonstrate the possibility of dispensing with a separate retrieval mechanism, within the simple and integrated architecture of a localist neural network which is asserted by Gass and Selinker (2008) to be at the heart of connectionism. This model, as explained, was the first to take the form of a computer simulation which successfully modeled a range of pre-lexical and lexical effects. “Although its localist network architecture has been superseded by distributed networks with more powerful learning capabilities, TRACE remains one of the most comprehensive and successful simulation of a broad range of known perceptual effects” (Ingram, 2007, p. 144). It is also emphasizedby deBot, Lowie, and Verspoor (2005) that computer simulation of neural networkshave shown that neural networks can learn skills such as recognizing face, reading, and discovering simple grammatical structures, and can extend these simple structures into more complex ones.

Architecture of TRACE

The TRACE model is different from its predecessors in two respects. Firstly, the feature detector nodes have been redesigned for TRACE in order to extract phonetically relevant acoustic parameters from the speech signal, and, secondly, in order to accommodate the temporal sequential nature of speech signals a major complication was introduced (Ingram, 2007). In this model, words are represented as patterns of activation feature, phoneme and word nodes, or traces, that build and decay during time. In this model, according to Ingram (2007), time is modeled as a sequence of time frames, where each frame replicates the entire set of network nodes and interconnections.

TRACE 1 vs. TRACE 2

TRACE 1 developed by Elman and McClelland (1986) was applied to stimulate phonetic processing and pre-lexical effects in speech perception (Ingram, 2007). This model, As Ingram (2007) explains, operated on spoken word input, extracting in parallel from the speech eleven acoustic phonetic features, each coded for eight distinct levels of activation. Each distinctive feature consisted of a mini-network of 'level' nodes; one for each of the eight activation levels of a feature. The level nodes in a feature network were each connected in an excitatory manner to a transducer tuned to respond to a particular level of the acoustic property that the feature was designed to detect.

TRACE 2 model was developed by McClelland and Elman (986)primarily used to simulate word recognition and lexical influences on phoneme recognition (Ingram, 2007). For these recognitions, as Ingram (2007) states, the phoneme units were fed predetermined patterns of appropriate feature-level activation, to reduce computational overhead so that interactions between word and phoneme levels could be better explored using test vocabularies of a reasonable size.

Drawbacks of the localist connectionism

Elamn (1991) states the localist approach has a number of important drawbacks. First, the localist dictum, "one node one concept", taken together with the fact that networks typically have fixed resources, appears to be at variance with the open-ended nature of language (Elman, 1991). He maintains that if needs are preallocated to define roles subject or agent, then so as to process sentences with multiple subjects or agent there must exist appropriate number and types of nodes. "But how one how is one to know just which types will be needed or how many to provide?" (Elman, 1991, p. 196).

The second shortcoming to the use of localist representation is argued by Elman (1991) to be the fact that they often underestimate the actual richness of linguistic structure. Even the basic notion "word," which can be assumed to be a straightforward linguistic primitive, turns out to be more difficult to define than one might have thought (Elman, 1991). He further maintains that there exist "dramatic differences in terms of what counts as a word across languages; and even within English, there morphological and syntactic processes which yield entities which are word-like in some but not all respects (e.g. apple pie, man-in-the-street, man for all reasons)" (p. 92). The third shortcoming a connectionist network from a localist perspective is suggested by Elman (2004). Elman (2004) states that a localist representation of a connectionist network deliberately deprives the network of any information about grammatical category, meaning, inflection, etc.

The distributed model

The distributed networks, according to Ingram (2007) are interpreted in two distinct senses. In the usual sense, distributed network is described as a neural network whose cells are distributed across a wide region in the brain and in a more technical senses it is described as a network that distributes its information storage or representational states as patterns of activation over the nodes of the network as a whole, distinct from the localist network where the activation level of each node in the network represents the status of a distinct object. The distributed models are comprised of one or more internal layers of nodes, in addition to an input and output layer, and the internal or the hidden layer is the locus of input's being processed before becoming output.

Properties of distributed connectionist networks

Gasser further mentions the following properties of distributed connectionist networks:

1- Robustness, graceful degradation
The systems do not breakdown when inputs are incomplete or even a part of the network is destroyed.
2- Graded representations
The concepts that the systems acquire hardly resemble the discrete categories of traditional models.
3- Fixed memory size
Because knowledge is shared in the system's connections, the addition of new knowledge does not appear to require new units and connections.
4- Automatic generalization, rule-like behavior
As connectionist systems learn about specific patterns, they are also constructing the knowledge that allows for handling a range of similar patterns. In contrast to rules of traditional models, these generalizations are not seen explicitly in the network. Rather they appear as needed during process.
5- Interaction of multiple sources of knowledge

Connectionist systems work by integrating information in the form of parallel spread of activation in various parts of the network.

THE BASIC FEATURES THE CONNECTIONIST MODELS SHARE

Most connectionist models are asserted by Gasser 1990 to share the following features:

1- The systems memory comprises a network of simple processing units joined by weighted connections. Each weight is a quantity which determines the degree to which the unit at the source end of connection activates or inhibits the unit at the destination end of the connection.
2- The behavior the units depict has loosely its basis on that of neurons. They gather input they receive on connections and compute activation. A unit's output is passed along its output connections to other units.
3- The analogue of long-term memory in other models is the set of weights on the network connections. In learning models these weights are adjusted as a result of processing.
4- Processing is parallel.
5- Control is distributed. Unlike traditional cognitive models, connectionist systems have no central executives whose job it is to determine which rule or rules are applicable and to execute them.

PROBLEMS WITH CONNECTIONIST MODELS

Edelman (1988, cited in Atkinson, 2012) argues that connectionist models are implausible due to drawing their inspiration from statistical physics and engineering, not from biology.

Furthermore, according to Elman (1998, cited in Atkinson, 2011) connectionist models are problematic due to:

1- Their disembodied nature
2- Their environmental passivity, and
3- Their modeling of cognition as a brain-bound phenomenon

FINAL REMARKS

Connectionism is the theory that considers brain in terms of neural or parallel distributed processing networks of interconnected units and argues that language is learned by learning rules from input alone, with no LAD involvement. This theory argues that mind makes links or connections between information nodes and the creation of new links when new input is received means that the network becomes progressively larger and more complex. It seems that connectionism rejects the assumption made by nativists that the brain is a symbol processing device similar to a digital computer. Here, the brain relies on a type of computation that emphasizes patterns of connectivity and activation.

However, it should be further mentioned that different groups can benefit from the connectionist models which are designed by software. These models can be used by language learners, language teachers as well as neurologists or psychologists (Hamidi & Montazeri, 2014). It seems that the major drawback of connectionism that casts doubt on the usefulness of a connectionist approach is that this approach has its basis on the sciences of math and physics, while the brains of human beings, or language learners, are biological entities. This seems to mar the usefulness of this approach to language learning, since it can be hardly assumed that the mathematical principles can be extended to biological ones.

REFERENCES

Atkinson, D. (2011). A sociocognitive approach to second language acquisition: How mind, body, and world work together in learning additional languages. In D. Atkinson (Ed.), *Alternative approaches to second language acquisition*. (pp.143-166). London: Routledge.

De Angelis, G. (2007). *Third or additional language acquisition*. Buffalo: Multilingual Matters Ltd.

DeBot, K., Lowie, W., & Verspoor, M. (2005). *Second language acquisition: An advanced resource book*. London: Routledge.

Ellis, N. (1999). Cognitive approaches to SLA. *Annual Review of Applied Linguistics, 19*, 22-42.

Elman, J. L. (1990).Finding structure in time. *Cognitive Science, 14*, 179-211.

Elman, J. L. (1991). Distributed representations, simple recurrent networks, and grammatical structures*. Machine Learning, 7*, 195-225.

Elman, J. L. (1995). Language as a dynamical system. Retrieved November 20, 2014, from: www.serch-document.com

Elman, J. L. (2004).An alternative view of the mental lexicon. *TRENDS in Cognitive Sciences, 8* (7), 301-306.

Gass, S. M., & Selinker, L. (2008).*Second language acquisition: An introductory course.* New York: Routledge.

Gasser, M. (1990).Connectionism and universals of second language acquisition. *Studies in Second Language Acquisition, 12*, 179-199.

Hamidi, H., & Montazeri, M. (2014). *Dictionary of second language acquisition*. Retrieved December 5, 2014, from http://www.iranelt.com/index.php/introduction-to-sla.

Ingram, J. C. L. (2007). *Neurolinguistics: An introduction to spoken language processing and its disorders*. Cambridge: Cambridge University Press.

Jordan, G. (2004). *Theory construction in second language acquisition*. Amsterdam: John Benjamins Publishing Company.

Piske, T., & Young-Scholten, M. (2009). *Input matters in SLA*. Bristol: Multilingual Matters Ltd.

Rast, R. (2008). *Foreign language input: Initial processing*. Toronto: Multilingual Matters Ltd.

Richards, J. C., & Schmidt, R. (1985). *Longman dictionary of language teaching and applied linguistics.* London: Pearson Education.