On Major Perspectives on Language Acquisition: Nativism, Connectionism, and Emergentism

Mohammad Khatib
English language and literature department
Allameh Tabataba’i University
South Allameh Street, Tehran, Iran
mkhatib27@yahoo.com

Somayyeh Sabah
College of Literature and Humanities
Khorramabad Branch
Islamic Azad University
Khorramabad, Iran
somayyehsabah@yahoo.com

Abstract
The phenomenon of language acquisition is a controversial issue within the field of language-related studies. Different approaches have been proposed to take account of this complicated matter. The present paper attempts at reviewing the current cognitive perspectives on language acquisition, i.e., Connectionism and Emergentism and exploring to what extent they are different from Nativism.

Keywords: Nativism, Connectionism, Emergentism, and language acquisition

1. Nativism
According to Ellis (1999), linguistics paves the way for the detailed descriptions of the patterns and relations in language. When language is analyzed out of context, intricate and intriguing structural systematicities emerge, and Generative Linguistics is scrupulous in its attempt to establish the set of rules that identifies the infinite number of sentences of a language. Putting it this way, Ellis argues that these careful descriptions are essential for a complete theory of language acquisition; nevertheless, they are not sufficient. Indeed, numerous cognitive scientists accept as true that linguistic descriptions differ from mental representations.

As said by Ellis, Generative Approaches, following Chomsky (1965, 1981, 1986, cited in, Ellis, 1999), have been directed by certain assumptions, which are as follows:

- Modularity: language is considered as a separate faculty of the mind;
- Grammar as a system of symbol-manipulating rules: knowledge about language represents a grammar, i.e., a complex set of rules and constraints that permits human beings to discriminate grammatical from ungrammatical sentences;
- Competence: the agenda of research ought to examine grammatical competence as an idealized hygienic abstraction rather than language use, which is despoiled by factors relating to performance;
- Poverty of the stimulus: given that learners approach the same grammar in more or less similar patterns of acquisition albeit the language input is degenerate, variable, and deficient in terms of reliable negative evidence, learnability arguments propose that there must exist strong constraints on the possible forms of grammars, the determination of which is the undertaking of Universal Grammar (UG);
- Language instinct: the necessary constraints of UG are innately represented in the brain, language is regarded as an instinct, linguistic universals are innate, and the faculty of language is modular by design;
- Acquisition as parameter setting: the acquisition of language is, thus, equivalent to the acquisition of the lexical items of a particular language along with the proper setting of parameters for that language. These assumptions guide the Generative Approach to the
As Ellis (1999) puts it, numerous cognitive scientists are suspicious of these assumptions, particularly modularity and language instinct together with the resultant analysis of the species-specific language faculty of human beings only, which is separated from semantics; the language functions; and the other social, biological, experiential, and cognitive aspects of humankind. As an alternative, the cognitive science proffers a different and more wide-ranging set of answers to the problem of mental representation than Generative Approaches.

In discussing the philosophical foundations of the mentalist paradigm, Bernat (2008) maintains that the Cartesian view places emphasis on the mentalist versus the behaviorist dimension of metacognitive knowledge, and the rationalist perspective that stands in opposition to the empiricist view accents the innate aspect of the mind in its accounts of learning. At this juncture, the emphasis is put on the individuality of mental knowledge as representations or schemata stored in the mind, and contextual influences are viewed as secondary. Thus, in this view, the properties of the mind are not necessarily dependent on the outside influences and, once established, are comparatively static.

Gregg (2003) argues that one of the merits of so-called nativist theories of language acquisition (first and second) lies in their capacity of provoking opposition. To Gregg, the very idea of an innate UG has from the outset been objectionable to numerous serious scholars, who have strived to demonstrate that language acquisition can be explicated without “appeal to an innate system of grammatical categories and principles” (p. 65) (e.g., Lieberman, 1984, 1991; O’Grady, 1987; Deane, 1992; Deacon, 1997; Sampson, 1997, 2001, cited in, Gregg, 2003).

2. Connectionism
As Gasser (1990) acknowledges, the recent connectionist models, also referred to as neural networks and Parallel Distributed Processing (PDP) models, are associated with the pioneering work by Neuroscientists and computer scientists in the 1940s and 1950s (McCulloch & Pitts, 1943; Rosenblatt, 1962, cited in, Gasser, 1990), who are said to be concerned with the computational power of networks consisting of simple neuron-like processing units. The current interest in these models has been prompted by the discovery of new-fangled learning algorithms as well as by the dissatisfaction with the achievements of classical symbolic models of cognition. To be precise, Gasser argues that the majority of connectionist models adhere to the subsequent fundamental characteristics:

- The system’s memory is composed of a network of simple processing units, which are attached by means of weighted connections. Each weight is a quantity that determines the extent to which the unit at the source end of the connection either activates or inhibits the unit at the destination end of the connection.
- The behavior of units is rooted loosely in that of neurons. They gauge the inputs they get on connections and work out an activation, which is considered as a function of the entire input, and an output, which is regarded as an activation function. The output of a unit is passed along its output connections on the way to other units. The existing pattern of activation on the units in the system is equivalent to short-term memory in further traditional models, and inputs and outputs to the system take the form of patterns of activation over groups of input and output units.
- The analogue of long-term memory in other models is the set of weights on the network connections. In learning models, these weights are attuned as a result of processing.
- Processing is parallel. In a good number of traditional models, as well as in conventional computers, decisions and actions are made one at a time. Similar to the brain, there is activity in several places concurrently.
Control is distributed. In contrast to the traditional cognitive models, connectionist systems possess no central executive whose job is to decide which rule or rules are presently applicable and to perform them. In effect, there exist no rules to be implemented.

According to Smith (1999), the last twenty years has witnessed the emergence or re-emergence of a drastically different approach to the study of mind. This approach that is identified as connectionism, neural networks, or PDP by any means rejects the need for symbolic representations. The entire complexities concerning the human thought and language can emerge from interactions amid a set of processing units, which are capable of undertaking divergent activation values. A connectionist network consists of a set of nodes that are responsible for collecting inputs from a range of sources that are both inside and outside the system, transmitting inputs to other nodes, and, thus, activating them alternatively. The connections may be unidirectional or bidirectional and are divergently weighted so that the next node along may be either restrained or stimulated. Putting it this way, learning is said to stem from training a network by continually exposing it to enormous numbers of examples and instances of the patterns to be acquired. What is more, it is not necessary to assume any kind of initial domain-specific structure to the network. That is to say, the linguists’ and psychologists’ appeal to modularity, particularly any form of genetically determined modularity, appears to be unessential. The complex structure of the modular mind is conceived of as an emergent property dependent solely on the input, especially the number of times a particular stimulus appears in that input. In other words, the statistical frequency of the input tokens is deemed to be vital to a network’s learning success, a property which enables it to capture the sensitivity to such things as word frequency effects.

In explicating the characteristics of neural networks, Poersch (2005) discusses that the work on neural networks has been based on the point that the brain computes in a way that wholly differs from the conventional digital computer, wherein the symbols are combined serially. The brain has a great number of neurons, which are enormously interconnected between each other. As a consequence of this, the brain comprises a tremendously well-organized structure. As Poersch puts it, “the brain is a highly complex, non-linear, and parallel computer” (p. 170). It has the ability of organizing neurons so as to execute certain computations many times more rapidly than the fastest digital computers. It encompasses a specific structure and possesses the capacity of constructing its own rules through experience, which is built over the years. The most remarkable development of the human brain is said to take place during the first years, generating millions of synapses per second.

As said by Haykin (1994, p. 2, cited in, Poersch, 2005, p. 171), “synapses are elementary structural and functional units that mediate the interaction between neurons,” and “a neural network is a massively parallel distributed processor that has a natural propensity for storing experiential knowledge and making it available for use.” Poersch (2005) states that “a presynaptic process liberates a transmitter substance that diffuses across the synaptic junction between neurons and then acts on a postsynaptic process” (p. 171). Thus, a synapse turns a presynaptical electric signal into a chemical signal and then converts it back to a postsynaptic electrical signal. It is presumed that synapses are simple connections that have the capacity of enforcing mutual activations between neurons. A significant characteristic of the brain is the plasticity provided by synapses, which allows the developing neuron system to adjust to the environment surrounding it. Synapses are acted upon via two cell filaments, i.e., the axon and the dendrite. The procedure utilized to carry out the processes of learning is called a “learning algorithm” whose function is “to modify the synaptic weights of the network in order to attain a desired design objective” (p. 171).

2.1. Major characteristics of connectionist models

According to Elman (2001), there are more than a few fundamental characteristics that are central to the way these networks operate. First, the response or the activation function of the units is frequently non-linear. This implies that the units may be particularly sensitive under certain conditions but remain rather insensitive under other circumstances. This non-linearity entails remarkably significant consequences for processing, chief amongst which is that networks can
occasionally function in a discrete, binary-like manner, and, thus, capitulate the crisp categorical behavior. In other circumstances, the system is able to yield graded, incessant responses. Second, what the system makes out is, to a great extent, accomplished by means of the pattern of connections, and the weights that are assigned to every connection, with the weights functioning as multipliers. Third, rather than making use of symbolic representations, the vocabulary of connectionist systems encompasses patterns of activations across various units.

Elman (2001) presents the debate that because of the significance of the weighted connections in these models, a basic question that is posed is associated with the factors that determine the values of these weights and program the networks. The connectivity in the early models has been set by hand, and this is still the case for “structured” connectionist models. Nevertheless, one of the stimulating advancements that has made connectionism enormously interesting to many is the development of algorithms by means of which the weights on the connections yield themselves to learning. To put it in plain words, the networks are capable of learning the values determined for the weights on their own; that is, they can be self-programming. What is more, the learning style is inductive; that is to say, examples of a target behavior are presented to the networks. As an example, the network gets exposed to the proper responses that are presented to a set of diverse stimuli. The network modifies the weights in small incremental steps through learning in such a way that eventually the accuracy of the network’s responses gets improvement. The network is also said to possess the capacity of generalizing its performance to the new stimuli and, thus, signifying that it has learned the essential generalization that connects outputs to inputs instead of only memorizing the training examples. This type of learning is known as the so-called “supervised learning,” which is merely one of many miscellaneous types of learning possible in connectionist networks. Other learning procedures are not concerned with any prior notion of “correct behavior” at all. Instead, the network might learn, for example, the correlational structure that lies behind a set of patterns.

**2.2 Connectionism and item-based learning**

As said by Hulstijn (2003), the acquisition of a lot of forms of cognition, say, language, happens in more than a few phases, namely, the accumulation of several information units, which are frequently referred to as “items, instances, or exemplars;” the construction of a network of these units with different “association strengths” between them; revealing “frequency” and “regularity” effects of the input; and, ultimately, the building of the “abstract categories” and “combinational rules.” For example, at a certain stage, the language learners may be concurrently developing the accretion of the phonological, morphological, lexical, and syntactic knowledge, “such that, in each domain, (1) some knowledge has the form of an associated network, (2) in which some common patterns are emerging, (3) some of which are said to result in the construction of the rules” (p. 418).

In this respect, Ellis (2002, cited in, Hulstijn, 2003) contends that the acquisition of the common patterns and the category construction are frequency-driven. According to Hulstijn (2003), it is through getting exposed to a great deal of input that the learners implicitly learn that certain patterns are very much common whereas others are not. For instance, the phoneme combination *sm is considered to be common while *ms is not, and it is through exposure to large amounts of language that native English speakers learn that verbs such as *give and offer are different from verbs like denote and present for the reason that the former category permits dative alternation while the latter does not. That is to say, the sentence He gave a present to his sister and the sentence He gave his sister a present are both correct whereas only the sentence She donated some money to the university is regarded to be correct, and the sentence *She donated the university some money [italics in the original] is not correct (p. 41).

As Saville-Troike (2005) also puts it, connectionism is another cognitive framework for concentrating on learning processes. It differs from other current frameworks for the study of SLA in not considering language learning to engross either innate knowledge or the abstraction of rules.
and principles. Rather, it takes account of the language learning as a process that emanates from escalating the strength of associations or connections between the stimuli and responses.

2.3 Criticisms against connectionist models
As Carroll and Meisel (1990, cited in, Gregg, 1996) point out, the connectionist accounts fail to explicate that human beings possess knowledge that goes beyond the input, a point that is the very heart of the logical problem. The spreading activation can possibly be exploited in the establishment of certain connections between, say, irregular verbs and their past-tense endings; however, one cannot fall back on a lack of activation for her or his knowledge that one sentence (e.g., *She may have been being misled*) is a possible sentence of English whereas another sentence (e.g., *She may been have being misled*) [italics in the original] is not possible. Likewise, as Fodor and Pylyshyn (1988, cited in, Gregg, 1996) put it, it appears to be difficult to understand how connectionism can tackle such inferential capacities that are capable of engendering an indefinitely large number of negative beliefs, e.g., earthworms cannot tapdance.

3. Emergentism
According to Ellis (1999), emergentists draw on the debate that the complexity of language emerges from rather simple developmental processes through exposure to an enormous and enormously complex environment. To O’Grady (2008), emergentism has its roots in the work of John Stuart Mill (1930 [1843], cited in, O’Grady, 2008), who suggests that a whole system can have properties that bring about results more than the sum of its individual parts. For O’Grady, Mill’s insight is also conducive to the scrutiny of the “so-called ‘Complex Systems’ – ranging from atoms to flocks of birds in flight to the weather – whose dynamic, non-linear behavior involves many interacting and interconnected parts” (p. 448). Putting it this way, O’Grady (2008) argues that the proponents of emergentism within linguistics are generally committed to the thesis, which is as follows:

The phenomena of language are best explained by reference to more basic non-linguistic (i.e., ‘non-grammatical’) factors and their interaction—physiology, perception, processing, working memory, pragmatics, social interaction, properties of the input, the learning mechanisms, and so on. (p. 448)

As said by O’Grady (2008), a great deal of emergentist studies within linguistics make use of the techniques of connectionism, which represents a perspective on the study of mind that attempts at modeling the learning process and cognition in terms of networks of neuron-like units whose relationship with respect to each other is characteristically graded and probabilistic (e.g., Elman, 1999; Christiansen & Chater, 2001; Palmer-Brown, Tepper, & Powell, 2002, cited in, O’Grady, 2008). A number of the varieties of connectionism rebuff the subsistence of the kinds of symbolic representations, say, the syntactic structure, which have played an indispensable role in the explanatory work resting on the human language. Regardless of one’s views on the Symbolist/Eliminativist disagreement, one point seems to be apparent; that is, the connectionist modeling paves the way for trying out a range of predictions on the subject matter of language acquisition, processing, change, and evolution.

In this sense, Gass and Selinker (2008) present the debate that in the field of language acquisition, emergentists maintain that certain simple mechanisms of learning, which are of the kind demonstrated elsewhere in cognition, are enough to amount to the emergence of complex language representations. With regard to SLA, it is debated that emergentism presumes that the process of learning a second language takes place on the basis of the extraction of regularities from the input.

3.1 Types of emergentism
In principle, Gregg (2003) refers to emergentists as “a fairly heterogeneous group” (p. 95), although they have much in common in rejecting the “nativist accounts of language that appeal to something like UG” (p. 96). Nevertheless, Gregg makes a distinction between two subsets of emergentism, i.e., “O’Gradian nativist emergentism” and “empiricist emergentists,” a term that, to
draw on Gregg’s terminology, accurately consists of each and every one of other “self-proclaimed emergentists” (p. 96).

O’Grady, Lee, and Kwak (2009) debate that despite the enormously substantial miscellany of the emergentist thought, there appears to be at least one essential thesis to which every one of its diverse proponents adhere; that is, the complexity of language needs to be identified with respect to the interaction of simpler and more fundamental non-linguistic factors. Nevertheless, O’Grady et al., state that there exist two types of emergentist approaches to language acquisition in terms of the dominant strategy, which is adopted. On the one hand, there is a particularly leading and inspiring body of research that concentrates on the significance of the input (or usage) for making sense of how language acquisition works. Ellis (2002, 2006, cited in, O’Grady et al., 2009) is said to present an extensive debate regarding this approach. On the other hand, a smaller body of research investigates the role of the processor-working memory interface at work in language acquisition and deals with the issues of learnability and development that have typically been the exclusive domain of the UG-based work.

As said by O’Grady et al., (2009), one of the first examples of a systematic input-based approach to language learning is the Competition Model (MacWhinney, 1987; Bates & MacWhinney, 1987, cited in, O’Grady et al., 2009). This approach, which continues to be very impressive, proffers a theory of how language learners recognize and give priority to a variety of competing cues (word order, animacy, case, agreement, etc) that are pertinent to the comprehension of the sentence. The basic variables, as MacWhinney puts it, are to be identified in the input; that is, how often the cue exists while a specific pattern is being interpreted (cue availability), and how often it informs on a specific interpretation (cue reliability). In exploring the role of the input frequency in language acquisition (first or second), it is necessary to keep in mind a principal problem based on which what counts does not refer to how many times learners come to hear a particular form; rather, the important issue is that how many times they come across mappings between a form and its meaning.

O’Grady et al., (2009) present the argument that the foundation of processor-based emergentism is the standpoint offered by Hawkins (2004, cited in, O’Grady et al., 2009) and O’Grady (2005, cited in, O’Grady et al., 2009), which entails that the basic properties of the syntactic phenomena that have been utilized to support the UG-based approaches to language for a long time are better elucidated with respect to the processing factors. Hawkins develops this proposal for numerous phenomena, which are essential to typology while O’Grady’s work concentrates more straightforwardly on the problem of language acquisition whose essential argument entails that a simple processor that is determined to the task of decreasing the burden and load on working memory lies at the core of the human language faculty. Even though such a processor does not use grammatical principles, its operation is central to explicating the properties of numerous core syntactic phenomena, i.e., binding, control, agreement, island constraints, scope, etc. What is more, it plays a significant role in taking account of the way those properties can be acquired in response to the limited kinds of experience accessible in the early years of life.

### 3.2 On emergentism and nativism

To put in plain words, O’Grady (2008) puts forward the debate that emergentism does not stand in opposition to nativism in its own right given the point that the brain is innately structured in a variety of ways. Nevertheless, emergentists reject the idea that there exist innate linguistic constraints on the computational system for language, a point that serves as a fundamental tenet of the grammatical nativism, i.e., UG.

In line with the foregoing argument, Misyak and Christiansen (2011) debate that the dialogue in the sciences concerning the issue of language has by tradition foregrounded oppositions between linguistic-nativist or modularist perspectives, on the one hand, and the views of emergentist, connectionist, or neo/neuro-constructivist positions on the other. According to the former stance, the syntactic ability is regarded to be made available by a specialized neural substrate
that universally develops across individuals when appropriate and rather minimal environmental inputs are provided. This kind of inborn or genetic endowment is embodied in Chomsky’s (1965, 1981, cited in, Misyak & Christiansen, 2011) proposal of a UG, which represents the formalization of a set of universal language-specific constraints. On the contrary, the emergentist, connectionist, neo/neuro-constructivist, and similar views (e.g., Elman, Bates, Johnson, Karmiloff-Smith, Parisi, & Plunkett, 1996; Mareschal, Johnson, Sirois, Spratling, Thomas, & Westermann, 2007; Tomasello, 2003, cited in, Misyak & Christiansen, 2011) highlight the experiential processes that, according to Christiansen and Chater (2008, cited in, Misyak & Christiansen, 2011), act together with the individual’s general learning mechanisms over the developmental time as well as over the evolutionary time. Putting it this way, Misyak and Christiansen present the argument that the complex, species-typical patterns of behavior identified as language can come to pass without involving certain language-specific constraints or predetermined, domain-specific, and neurobiological circuits. Although such positions are intermittently misunderstood as the tabula rasa empiricism, there exists no theoretical necessity for giving up certain genetic biases. Such theories are capable of allocating a convincing role to small initial biases of the learning system in the midst of an interactive, developmental, and ecological milieu. In essence, the linguistic-nativist theories fundamentally advocate a strong structurally and/or functionally specific biological foundation for language whereas the emergentist position proposes small, germinal biases or broad intrinsic constraints impinging upon further general, low-level biological systems, which are of relevance to language. These are, therefore, varied suggestions given for the ways wherein language may have genetic bases and effects.

3.3 Criticisms against emergentism

Eubank and Gregg (2002, p. 238, cited in, Jordan, 2004, p. 249) challenge emergentism and pose the question regarding the way “children know which form-function pairings are possible in human-language grammars and which are not, regardless of exposure.” What is more, Eubank and Gregg (2002, p. 238, cited in, Jordan, 2004, p. 249) argue that how emergentists can explicate cases of instantaneous learning or “knowledge that comes about in the absence of exposure (i.e., a frequency of zero) including knowledge of what is not possible.”

Accordingly, Jordan (2004) acknowledges that the poverty of the stimulus argument lies at the heart of the problem of any empiricist approach. Emergentists, by adhering to an associative learning model along with an empiricist epistemology, wherein the existence of some kind of innate architecture is permitted while innate knowledge and, indeed, innate linguistic representations are not taken into account, must deal with the extremely difficult task of explicating how children come to possess their linguistic knowledge. To put it in plain words, they need to explain how “general conceptual representations” that operate on the environmental stimuli clarify the “representational system of language” that children reveal (p. 249).

4. Conclusion

In due course, the present paper has attempted at presenting a brief overview regarding the major cognitive perspectives on language acquisition and reviewing some of the differences between nativist and emergentist approaches. However, as Sinha (1999) puts it, further cognitive-linguistic inspired investigations of language acquisition are required. Besides, one needs to include the developmental perspective in the heart of making sense of the human language faculty not in terms of an innate module or a subset of modules amongst others, but as one dimension of an integrated yet complexity differentiated embodied neuro-cognitive system, which is functionally coupled and co-evolve with its socio-cultural surrounding. What is more, Jordan’s (2004) conclusion is worth considering on the basis of which one is required to “avoid pushing epistemological positions to extremes” (p. 250). The innatist-emergentist debate does not need to be a confrontational conflict, and it is merely by adopting an extreme stance in either camp that one comes to incompatible discrepancies and, undoubtedly, indefensible positions.
References


