Modularity and Non-modularity in Language Acquisition

CS784 Project Report

Submitted by:
Esha Tiwary (Y1131)

1 Introduction

"To what degree is language actually hard-wired into our brains? To what degree do we actually learn language?" This, in simple words, is the question that the scientific study of language acquisition attempts to answer. Chomsky declared that language acquisition depends on an innate, species-specific module that is distinct from general intelligence. This revolutionary and extremely controversial claim has sparked a heated debate that continues to this date – the debate between modularity and non-modularity in language acquisition.

When we think of language acquisition, we need to separate out its two major concerns: the study of language evolution (i.e. how humans acquired language), and the study of child language acquisition (i.e. how children acquire language). Several theories try to explain these phenomena and the debate between modularity and non-modularity is an ongoing one. In this report, we will look at some established researchers in this field, and explore their stands. We will also look at some upcoming ideas and try to examine their claims. Apart from theories, there have been several attempts to explain language acquisition in humans through synthetic modelling. In this technique, complex systems with artificial agents are created which try to simulate different aspects of language acquisition. These systems give us some insight into how the system might have worked. We will analyze some such experiments, and try to connect them with the existing theories.

2 Concepts

Before we go any further, we will briefly overview some of the existing concepts in language acquisition. These are the large camps in the field, and most prominent researchers fit into one of these categories.

2.1 Nativists

Nativists (also known as Rationalists and Innatists) believe that the language faculty is innate. It is something we are born with. How is it possible that native speakers come to know things about their language in conditions of sparse evidence? In particular, native speakers know things about their language which they have never been taught, and they have never experienced. The nativist's answer to these questions is that humans are endowed with an innate faculty of language.

There is a division in Nativism camp itself. There are two classes of believers:

2.1.1 Modularists

These people advocate the concept of a "dedicated potential" for language. They say that the language faculty is a dedicated one and is separate from other cognitive processes. They believe in what is known as FLN - the faculty of language "narrow". It is called I-language (Internalized Language) is the recursive linguistic component. It is this component that is innate and distinct from general intelligence.

2.1.2 Non-modularists

They contest the idea of a dedicated potential for language. There is no specific module dedicated to language. They believe that the language faculty is not modular; rather that it is FLB – the faculty of language "broad". It consists of the I-language, the sensory-motor system, and the conceptual-intentional system.

2.2 Empiricists

The empiricists claim that language is learned from experience. They oppose the idea of innate concepts. Empiricists claim that sense experience is the ultimate source of all our concepts and knowledge. They say that there is some primary linguistic data (PLD) which may be a rich linguistic state. Induction on this PLD gives rise to the rules of language.

2.3 Behaviourists

Behaviourists are at the other end of the spectrum from Nativists. They oppose any idea of an innate faculty, and even deny any sort of induction mechanism. They profess that the initial state of the brain is a clean slate, and all learning happens through associations. These are chains of stimulus-response pairs, and result in the knowledge of language.

Now that we have an idea, about the existing paradigms in the field, we will look at the views of some researchers and examine their claims.

3 Steven Pinker – "Rules of Language"

In this paper [1], Pinker analyses the existing theories about language acquisition by examining evidence from children with language impairments. One of the impairments he looks at is Specific Language Impairment (SLI). This is a syndrome of language deficits that is not attributable to auditory, cognitive, or social problems. The syndrome usually includes delayed onset of language, articulation difficulties in childhood, and problems in controlling grammatical features such as tense, number, gender, case and person.

The interesting thing about this defect is that it appears to have an inherited component. Pinker states that language impairments have been found in 3% of first-degree family members of normal probands, but in 23% of language-impaired probands. The impairment has been found to be 80% concordant in monozygotic twins, and 35% concordant in dizygotic twins. One case study mentioned by Pinker investigated a three generation, 30-member family, 16 of whom had SLI. The syndrome followed the pattern of a dominant, fully penetrant autosomal gene. This constitutes evidence that some aspects of use of grammar have a genetic basis.

Another impairment considered by Pinker is the Williams Syndrome (WS). This is associated with a defective gene expressed in the central nervous system, and causes an unusual kind of mental retardation. Older children and adolescents with WS show grammatical abilities close to normal in control testing, but their IQ is measured at around 50. Hence language seems to be preserved despite severe cognitive impairments, suggesting that the language system is autonomous of many other kinds of cognitive processing.

Pinker concludes that modern language research makes the idea of a general purpose learning mechanism being the sole complexity of the mind increasingly implausible. He claims that the evidence points to "a system that is modular, independent of real-world meaning, nonassociative, sensitive to abstract formal distinctions, more sophisticated than the kind of 'rules' that are explicitly taught, developing on a schedule not timed by environmental input, organized by principles that could not have been learnt, possibly with a distinct neural substrate and genetic basis.

4 J. Fodor – "Global Processes, Innate Concepts and a Few Words On Ontology"

This is extracted from an interview given by Fodor [2], in which he is talking about his new book "The mind doesn't work that way" and states his position about certain issues, particularly those related to language acquisition. He says that there are modular systems that are involved largely with perception and the articulation of action, and the design of most of the cognitive mind is not modularized. He further makes a distinction between global and local mental processes. The local processes can be perceived as modular. Thinking, reasoning, inference to the best explanation, problem solving, theory construction etc. (all the stuff that computers can't do) would be non-modular processes.

He also comments on the link between modularity and Darwinism. He believes that whatever parts of the mind are modular are probably adaptations. But he rejects the theory of the mind being 'massively' modular. The interesting part of the interview is his stand on what is innate. He says that for many concepts there are prototypes, prototypes being spectacularly good instances of the concept, typically also the high frequency instances. The prototype of a dog is a sort of a middle-size dog, the prototype of a chair is really a chair, not a stool, etc. The mind has got an inductive, essentially statistical, device for building prototypes, that being statistical and so on. The innate part comes in as linking the prototype to its concept. He talks of a kind of a mind, which, given a learned prototype then formulates a concept by exploiting some sort of innate mechanism. The condition on the concept that it formulates is what determines the right extension. So, the concept CHAIR, unlike the prototypical chair, has to apply to all chairs. Thus the mind is a mechanism that learns prototypes and, in some way or the other, pairs them with concepts.

Thus Fodor takes the stand that concepts are not innate. The innate faculty of language connecting concepts and prototypes implies that language is not separated from concepts; instead it is defined by them.

5 Paul Bloom - "How Children Learn the Meanings of Words"

In this book, Bloom analyses how children acquire the meanings of words. According to Bloom, children learn words through sophisticated cognitive abilities that exist for other purposes. These include the ability to infer others' intentions, the ability to acquire concepts, an appreciation of syntactic structure, and certain general learning and memory abilities.

Bloom contests the idea of a dedicated potential in word learning. He mentions several experiments done with children. In a particular experiment, novel objects are shown to children and referred to by several means including naming and linguistically and visually presented facts. For example, a linguistically presented fact would be "that which my uncle brought…". It is seen that children correctly identified the objects later, irrespective of the way they were referred to. This shows that fast mapping is not confined to naming, and gets extended to memory and learning. Hence, fast-mapping doesn't indicate a dedicated potential for word learning.

An interesting idea propagated by Bloom is that of the Theory of Mind. He says that word learning is basically an object mapping problem. And children solve the name-object mapping problem through inferring referential intentions of other people. He says that the mind reading ability used in language is the same as that used in intentional attribution more generally, and is not a product of a distinct module or sub-module. Hence there is no sub-module dedicated to communication. To substantiate this, he cites the "gaze" experiment. When an adult looks at an object and says its name, the child gaze follows the adult and looks at the object. If the adult shows some emotion instead of saying a name, like excitement or a frown, then too the child's gaze follows to the object. Hence a similar mechanism is at work – that of inferring intentions.

Bloom contests the idea of a dedicated potential, and claims that language is part of a larger cognitive system, and not an isolated module.

6 Experimental Evidence

We will now look at some experiments done on subjects with some brain damage, how it affects their behaviour, and the conclusions drawn from them.

6.1 "Beyond Modularity" – Emergent Modularity

In her book *Beyond Modularity*, Annette Karmiloff-Smith advances the notion of "emergent modularity". The primary data in support of her conclusion is that young children who suffer brain damage to the "language centres" of the brain are very often capable of learning language just as well as children without lesions. She claims that the MRI evidence shows fairly conclusively that they just use a different part of the brain to do language. This suggests that even if language is identified to some degree with one area of the brain (to what degree is still an open question), localization seems to be the result of learning a language, not its precondition. This undermines the idea of innate modularity in language.

Although she claims that children with brain lesions are able to use a different part of the brain for language, it isn't clear whether all language capabilities are achieved through the alternative areas. The point remains that it's possible that a single area of the brain is best suited to all the complexities of language, and when that part gets damaged, other parts may serve as substitutes for some functionality of language, but they don't encapsulate all its complexities. Hence these studies aren't very conclusive, and we need to probe deeper into the actual functionalities served by these alternate parts of the brain.

6.2 "Agrammatic but Numerate" - Dissociation between language and mathematical ability

In this paper [4], the authors examine the performance of three men with severe agrammatic aphasia. In addition to grammatical comprehension difficulties, all patients demonstrated severe limitations in grammatical production. Their performance is evaluated across a range of language, number, and calculation tasks. In particular, they examine behaviour on tasks that involved parallel operations across language and mathematics. In all tasks, the participants were required to use syntactic principles applied to mathematics that they were unable to use in language. The tasks included estimation tests, calculations with whole numbers and fractions, number infinity problems and bracket expressions.

With regard to syntactic processes, the results reveal dissociation between the mathematical and language domains. They say that all patients were competent in mathematical syntactic functions that were not evident in their language syntactic performance. All were sensitive to reversibility and the role an element takes in a numerical expression: for example, as a divisor or dividend. Similarly, despite an inability to comprehend simple subject–verb–object sentences, all patients were sensitive to the embedded structure of bracket expressions and displayed capacity to solve such problems. Performance on tasks involving the productive use of syntactic principles also showed dissociation between language and mathematics. Although no patient was able to form productive clausal structures in language, all were able to display use of recursive principles in number infinity tasks.

The authors claim that their findings are incompatible with a claim (made by Hauser, Chomsky and Fitch in [5]) that mathematical expressions are translated into a language format to gain access to syntactic mechanisms specialized for language. Instead, grammar may be seen as a co-opted system that can support the expression of mathematical reasoning, but the possession of grammar neither guarantees nor jeopardizes successful performance on calculation problems.

Since these experiments looked at parallel operations between mathematics and language, these results allow consideration of two alternative interpretations regarding the syntactic mechanisms of the two faculties. One possibility is that a common and domain-general syntactic mechanism underpins both language and mathematics. However, mathematical expressions can gain direct access to this system without translation into a language format. In the case of patients with agrammatic aphasia, language representations are disconnected from the syntactic mechanism, but mathematical expressions can still gain access. The second possibility is that in the mature cognitive system, there are autonomous, domain-specific syntactic mechanisms for language and mathematics.

Hence we can’t say with conviction whether the faculties of language and mathematics have shared components, or whether they are independent of each other. However, the presence of dissociations between mathematics and language in people with developmental language impairments indicates the potential for autonomous mechanisms.

7 Genetic Assimilation

In *The Symbolic Species: The Co-Evolution of Language and the Brain* [6], Terrence W. Deacon tries to reconcile the conflicting ideas of modularists and non-modularists by proposing genetic assimilation as a solution to the problem of language acquisition.

Language, he argues, is not an instinct and there is no genetically installed linguistic black box in our brains. Language arose slowly through cognitive and cultural inventiveness. Ancient man tried to assemble an extremely crude symbolic system – one that we would not have recognized as language.

Language then improved by two means. First, invented linguistic forms were subjected to a long process of selection. Generation after generation, the newborn brain deflected linguistic inventions it found uncongenial. The guessing abilities and intricate nonlinguistic biases of the newborn brain acted as filters on the products of linguistic invention. Today's languages are systems of linguistic forms that have survived. The child's mind does not embody innate language structures. Rather, language has come to embody the predispositions of the child's mind. The second way by which language improved, in Deacon's view, had to do with changes in the brain. Crude and difficult language imposed the persistent cognitive burden of erecting and maintaining a relational network of symbols. In that demanding environment, genetic variations that rendered brains more adept at language were favored. Language began as a cognitive adaptation. Genetic assimilation then eased some of the burden. Cognitive effort and genetic assimilation interacted as language and brain co-evolved.

According to Deacon, Pinker and Bloom cannot propose that language is a cognitive invention that underwent genetic assimilation because they think genetic specialization for language must have begun the process ("There must have been a series of steps leading from no language at all to language as we now find it, each step small enough to have been produced by a random mutation or recombination"). He argues that language was a cognitive and cultural invention that underwent genetic assimilation. Language, he argues, was "acquired with the aid of flexible ape-learning abilities." It is not walled off from other cognitive functions such as interpreting and reasoning. Grammatical form is not independent of conceptual meaning.

Deacon claims that genetic assimilation built new wetware, largely in the area of prefrontal cortex that assisted attention, memory, and association, consequently easing the burden of language. These neurobiological changes were "a direct consequence of the use of words" "An idea," says Deacon, "changed the brain."

In theoretical linguistics, we have opposing camps that tend to dismiss each other’s points of view rather than confront them. To resolve this debate, we probably need to consider evidence from other human sciences as well. This is done quite effectively by Deacon. He is proposing a theory that is difficult to reject. If we accept what he days, i.e. language involved genetic assimilation, then the point of contention becomes – what was the initial structure of the linguistic faculty? Did we start from scratch (which he seems to suggest), and genetically assimilate features of language, or did we start with some initial state of the linguistic faculty (possibly an “exaptation”), that evolved over time?

8 Synthetic Modelling

There have been several attempts to try to understand language evolution in humans through artificial modelling. This is done through complex systems and there are three main approaches followed:

1. Genetic Evolution
In this approach, the linguistic structure is coded in the gene. It is a modular approach and assumes an innate Language Acquisition Device (LAD).
For example, McLennan built a complex system to simulate communication between agents. He tried simulation first using only genetic transmission. He found that the results improved by 50 times. But when he included adaptation in the agents' mechanism, the results improved by about 150 times. Although synthetic models don't serve as facts, they give some indication about the processes involved in language acquisition. McLennan's experiment gives an indication that adaptation improves survivability.

2. Adaptation
In this approach, only the cognitive system (Perceptual Motor System + Learning System) is genetically transmitted. The linguistic structure is not coded in the gene. It is a non-modular approach (i.e. language acquired and stored in memory).
For example, de Boer conducted an experiment to simulate phonology in humans [7]. He tried to simulate the vowel systems. The results showed that despite starting from a clean slate, realistic vowel systems emerged in the agents.

3. Genetic Assimilation
This method is the reconciliation of modular and non-modular principles. It works on the Baldwin principle, introduced in 1896, which basically says that if weak biases are better adapted to the environment, over a period of time they become strong biases. This means that if some property is acquired by an organism, and it is well suited to the environment (i.e. it stabilizes), then over generations it gets coded into the genetic structure of the organism.

8.1 Genetic Assimilation

We'll look at an example of a simulation which works on the principle of genetic assimilation. It gives some interesting results, and forces us to consider genetic assimilation as an increasingly possible solution to the question of language acquisition.

This work is titled *Cultural transmission, learning cost and the Baldwin effect in language evolution*, authored by Steve Munroe and Angelo Cangelosi [8]. According to them, the Baldwin effect has always been invoked to explain the evolution of an innate language acquisition device (LAD) (Pinker & Bloom, 1990). However, in this paper, they propose an alternative role for the Baldwin effect in language evolution. Rather than playing a role in the evolution of linguistically-specialized structures such as the LAD, the Baldwin effect can explain the assimilation of neural substrates that favour the evolution of general cognitive abilities. These also favour the evolution of linguistic abilities through the co-evolution of language, brain and cognition. This links up with the theory proposed by Deacon, and their results lend some weight to Deacon's claim.

They use a multi-agent model to simulate the evolution of shared compositional languages, extending the work of Cangelosi (2001). Neural networks simulate the process of language learning and cultural transmission. A genetic algorithm models some of the mechanisms of natural selection. Two major parameters were varied to ascertain their effects on the type and strength of the Baldwin effect. A first parameter controls the variability of the language environment, i.e. the level of stochastic noise in the process of cultural transmission. The second parameter determines the cost of language learning, i.e. its fitness cost for the individual.

They report some interesting results about how variations in these parameters show changes in the Baldwin effect seen within the system. In the first case, the language environment is varied during cultural transmission and there is an associated high learning cost. Then the agents develop an increased predisposition to learn the language quickly and efficiently. No actual linguistic structures are assimilated in the agents' genome. At each generation organisms have to learn the prevailing language, and in later generations they are able to do this in a very efficient way.

In the second case, when the language environment remains static and there exist high learning costs, the agents incorporate aspects of the structure of language into their genome. Before cultural transmission starts, agents already have some knowledge of the language to be learned. The analysis of the internal representations of the agents' neural networks reveals that these Baldwin effects consist in the evolution of adaptive neural structures. These favour cognitive abilities and linguistic abilities as an indirect consequence of it. However, in the third case, if the learning costs are low the effect is much reduced, as there is little evolutionary pressure to translate the lifetime learning task into genetic structures.

Their experiment illustrates the Baldwin effect observed in the system depending on different parameters. What emerges is that with a stable language, and high cost for learning, Baldwin effect is quite high. This has clear analogies with human systems, since language is necessary for survival and hence the high learning costs (rewards). These results support Deacon’s theory, and give us a new way of understanding language acquisition.

References

[1] S. Pinker, *Rules of Language*, Science 253 (1991): 530-535

[2] *Global Processes, Innate Concepts and a Few Words on Ontology*: *An Interview with Jerry Fodor*
http://www.netn.fi/201/netn_201_fodoeng.html

[3] P. Bloom, *How children learn the meaning of words,*
http://cognet.mit.edu/library/books/view?isbn=0262024691

[4] Varley, Klessinger, Romanowski, Siegal, *Agrammatic but Numerate,* 2005

[5] Hauser, M. D., Chomsky, N. & Fitch, W. T. (2002) *Science* 298**,** 1569–1579.

[6] Terrence W. Deacon, *The Symbolic Species: The Co-Evolution of Language and the Brain* (W.W. Norton & Company, 1997)

[7] Bart de Boer, *Self-organization in vowel systems*, Journal of Phonetics, 2000
http://www.idealibrary.com

[8] S. Munroe, A. Cangelosi, *Cultural transmission, learning cost and the Baldwin effect in language evolution*
http://www.ling.ed.ac.uk/evolang2002/ABSTRACTS/munroe.txt