Language Acquisition and the Argument from the Poverty of the Stimulus

Parissa Zohari

A Thesis

in

The Special Individualized Program

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Arts (The Special Individualized Program) at
Concordia University
Montreal, Quebec, Canada

April 2004

Your file Votre référence
ISBN: 0-612-91165-9
Our file Notre référence
ISBN: 0-612-91165-9

ABSTRACT

Language Acquisition and the Argument from the Poverty of the Stimulus

Parissa Zohari

This Thesis represents an attempt to illustrate the significance of two fundamental concepts related to First Language Acquisition (FLA): 1) The Primary Linguistic Data (the PLD), and, 2) The Argument from the Poverty of the Stimulus (POS). The central claim is that the notion of the PLD and the POS Argument, which have traditionally been presented as distinct, are intimately related. The PLD is viewed as the external basis for FLA as well as the POS Argument. This thesis proposes a new, tri-level definition for the Linguistic Data: a) the Available Linguistic Data (ALD), b) the Received [by the Acquirer's brain] Linguistic Data (RLD), and, c) the Perceived Linguistic Data (PLD). In addition, three issues related to the POS concept are discussed: Innateness, Negative Evidence, and The POS Argument. Chomsky proposed the POS concept as evidence that FLA cannot occur in the absence of innate language universals. The idea was that people attain knowledge of their language despite the impoverished Linguistic Data. This thesis explores Geoffrey Pullum and Barbara Scholz's (2002) recent challenges to the POS Argument which attempt to undermine the nativist view of FLA. Two POS exemplars studied and rejected by Pullum and Scholz are reexamined here. This thesis concludes that the claims made by Pullum and Scholz lack scientific evidence and, hence, the POS Argument still remains unchallenged. This thesis also suggests that the next crucial step in language acquisition research is to explore Second Language Acquisition (SLA) and its relation to language universals.

DEDICATIONS & ACKNOWLEDGEMENTS

I would like to dedicate this thesis to my beloved family. Though far away and in different parts of the world, their unwavering love and support has always been my guiding light.

I would like to thank Dr. Mark Robert Hale for providing me with invaluable insight and the much needed incentive to go farther than I ever thought possible, Dr. Charles Reiss whose intellectual challenges gave me the drive to never falter from further examining and improving my views, and Dr. Annette Teffeteller for offering the added value of an alternative perspective and for believing in me.

I would also like to thank Dr. Catherine Vallejo, my work supervisor, and my dear colleagues Ms. Marilyn Malofy and Ms. Suzanne Marcil for their support, understanding and enthusiasm throughout my studies and my thesis preparation process.

There is no need to suppose that anything material passes from objects to our eyes to make us see colors and light, nor that there is anything in these objects that is similar to the ideas or the sensations that we have of them: just as nothing moves from the objects that a blind man senses that must pass along his stick to his hand, and the resistance or the movement of these bodies which is the only cause of the sensations that he has of them, is in no way similar to the ideas that he conceives of them . And in this way, your mind will be freed of all those little images flying through the air, called *intentional species*, which so exercise the imagination of the philosophers.

Rene Decartes (Dioptrics 1637)

Having the capacity to do so-and-so is not the same as knowing how to do so-and-so; in particular, there is a crucial intellectual component in "knowing how".

Noam Chomsky (Rules and Representations 1980)

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

In the recent years, the field of cognitive science and the science of linguistics are undergoing major transformations. The very nature of scientific research is synonymous with transition and change. Nevertheless, the changes that linguistics and cognitive science, these fairly new fields of empirical and scientific exploration, are experiencing are comparatively much more radical than the evolutionary transitions that mark the usual course of research.

As exciting as this might sound, unfortunately, I will argue that this progression into practically uncharted scientific territory has led to many instances of confused terminological usage and a resulting misinterpretation of ideas in these fields. In my belief, a type of borrowing—not unlike cross-cultural linguistic borrowing[1]—is happening among related fields. Although the exchange of information between various sciences is common (and necessary) practice, at this point in time, it has led to major confusion about the accuracy of definitions and a failure to offer plausible interpretations for the borrowed ideas.

The need to step back and correctly identify the true meaning of fundamental idioms, expressions and ideas is undeniable at this crossroads where all sciences interested in uncovering the mysteries of the realm of cognition come head to head. After

[1] Borrowing is normally a socio linguistic aspect of language evolution. Different expressions or words are borrowed from one language to the other. Some borrowed words are interpreted correctly and others change in meaning through the course of borrowing and the passage of time.

all, the objective of all cognitive sciences (including linguistics) is to heighten our understanding of the mind and its ways of operation.

Consequently, instead of working in isolation and occasionally borrowing research results and scientific terminology, we must take the time to define and fully comprehend the nature of elements that constitute the foundation of our research. When it comes to proper methodical analysis, a lack of understanding will simply result in unscientific theorizing and baseless debate.[2]

The primary duty of all scientists should be to understand the overall nature of the elementary aspects of their research. What is being dealt with? What is its nature? How does it function? How does it affect the research results?

Consider, for example, the reconstruction of the Indo-European language (called Proto-Indo-European, P.I.E.). For Historical Linguists, this is an essential task since the actual language no longer exists. Moreover, there is no attested direct evidence for the original language which is believed to have been spoken by the Indo-European people.

In an interestingly similar manner, we are yet to attain direct access to the human brain or parts of it which are responsible for language acquisition and production. We must, therefore, carefully define (rather similar to reconstruction) the object of our theorizing based on the observable results of the processes that occur in the brain.[3]

In other words, we have to step back and look at some of the general characteristics of the object of our study. We must come to some type of agreement

---

[2] It is impossible to cover all instances of misuse. In this thesis, I will discuss some examples in relation to the Primary Linguistic Data (PLD) and the Poverty of the Stimulus (POS).

[3] That is, what is being produced—in this case the acquisition and production of the language of the environment (G1).

regarding key definitions. The goal must be to mainstream the interpretation of crucial basics before arguing the smaller details and the narrower, more refined aspects of research in all Cognitive Sciences.[4]

Since we cannot (yet) directly observe the neural networks of the human brain,[5] we must reconstruct a Proto-Cognitive-Map via empirical research and observing the general characteristics of the end result. That is to say, we must map cognitive processes through observing the outcome (i.e., cognitive behavior).

This thesis represents an attempt to illustrate the importance of understanding the meaning behind two central concepts of First Language Acquisition (FLA):

a) The Primary Linguistic Data (the PLD)

b) The Argument from the Poverty of the Stimulus (POS)

The central focus of this thesis is to argue that the notion of the PLD and the Argument from the POS, which have traditionally been presented as distinct, are intimately related.[6]

The Primary Linguistic Data is the external stimulant that triggers the Language Acquisition Device (the LAD) in the human brain to process and construct the grammar of the environment (also called the Target Grammar or G1 for abbreviation). I will argue that without a clear notion of what constitutes the PLD, we risk jeopardizing the entire basis for the theory of First Language Acquisition (FLA)— including our understanding

---

[4] To continue with the historical linguistic analogy, I think that defining the preliminary aspects of cognitive science could be compared to the discovery of the common elements among all descendents of the Indo-European language. This discovery initiated the hypothesis of a previously existing common language. The birth of this hypothesis, in turn, led to the reconstruction of this language for which no documented evidence exists (hence the addition of the word Proto to the reconstructed language: Proto Indo-European, P.I.E.).

[5] Of course, the work of many researchers is dedicated to discovering ways of directly observing brain activity. But, when it comes to Language Acquisition, we have a long way to go.

[6] I would like to thank Charles Reiss (Concordia) for pointing this out during our discussions regarding my thesis.

of issues surrounding the Poverty of the Stimulus (POS) Argument. This matter is discussed in more detail in Chapter 3.

The logic of the Argument from the Poverty of the Stimulus is quite simple, "...children come equipped with a priori knowledge of language just because it is unimaginable how they could otherwise acquire the complexities of adult language [to which they are not exposed due to the impoverished nature of the data; pz]" (Lasnik&Uriagereka 2002: p. 150).

In other words, "obviously, children are not born with the mental rules for any particular language [...] just as obviously, babies are not taught the rules of those languages" (Pinker 1995a: p.108). Yet, they end up with an accurate version of the grammar of their environment and, hence, it is a logical necessity "that children's mental learning mechanisms be constrained in some way, for otherwise they could not generalize correctly beyond their parents' sentences to the rest of the language" (Pinker 1995a: p. 109). This, in my view, is the very reason why the POS concept has evolved to its current prominent status in the Innatist view of Language Acquisition in such a short period of time.

Geoffrey Pullum and Barbara Scholz (also P&S, henceforth) have recently challenged the POS Argument (2002: p. 9-50). The central issue of this thesis is to examine and discuss P&S's arguments against the POS concept. These challenges are mainly explored in Chapter 3. However, as an introduction, a brief example will illustrate how (similar to many contemporary researchers) Pullum and Scholz seem to have misunderstood what, in my opinion, is a simple concept.

In the case of Auxiliary sequences, Pullum and Scholz (2002: p.27) oppose John Kimball's (1973: p. 73-75) claims[7] by stating that, "when Kimball was writing we did not have the tools we have today for searching corpora. Looking in the relatively small and in some ways unsuitable texts available to us (and we do not intend to suggest in any way that our preliminary testing represents a definitive result), we find hundreds of examples." (2002: p. 28).

It is hard to believe, however, that the 'hundreds of examples' Pullum and Scholz claim to have found come from a variety of literary novels, including Bram Stoker's *Dracula* and Herman Melville's *Moby Dick* (2002: p. 29).

Revealing the source of these examples should suffice to jeopardize the scientific validity of Pullum and Scholz's claims. Nonetheless, it is imperative to seriously consider this rather strong example of how easily researchers could misunderstand scientific concepts— in this case, what constitutes the Primary Linguistic Data (PLD), and how this lack of understanding could lead to radical complications.

Here is a list of possible problems which arise from using English novels (i.e., written data) as the equivalent of the Primary Linguistic Data:

a) The data in question are available to a limited number of people.

b) The data in question are limited.

c) It is unnecessary to go through records of electronic data to realize that the available data in the environment of every child are different and, yet, the end

[7] Kimball believes that the acquisition of complex auxiliary sequences by children who usually hear simpler forms of these sequences is evidence of acquisition despite impoverished data. This matter is further discussed in Chapter 3.

result is that all children learn the language of the environment in a relatively short period of time (between 0-5 years of age), and most important of all,

d) What a child hears is, in itself, not necessarily linguistically relevant. That is, data used as the PLD are not the same as data available in the immediate environment of the child.[8]

It is safe to say that the Primary Linguistic Data (PLD) is considered the external basis for acquisition in any language acquisition related theory— including the Argument from the Poverty of the Stimulus. Noam Chomsky first brought up the term "argument from the poverty of the stimulus" in 1980 (p. 34). The diachronic development of the POS concept is discussed in Chapter 2.

However, it is worth mentioning at this point that this concept was initially generated in support of the Innateness Hypothesis.

Figure 1.1 Legate & Yang 2002: p. 152

"If you know X, and X is underdetermined by learning experience, then the knowledge of X must be innate."

Many researchers agree that the way "a child masters a rich and highly unstructured system on the basis of degenerate and deficient data" (Hornstein and Lightfoot, 1981: p.9) is a dilemma that needs to be addressed. I will argue in favor of the Poverty of the Stimulus Argument in detail in Chapter 3. Interestingly, part of my discussions will deal

[8] This matter is discussed in great detail in Chapter 3.

with the fact that both the pro-POS and the anti-POS advocates have at one time or another misinterpreted this concept.

For instance, in 1982, Lightfoot points out that "there are no data available to the child which will suffice to establish some rule or principle. But the rule or principle is acquired anyway." (p.428). Pullum and Scholz oppose Lightfoot by stating that both "available to the child" and will "suffice to establish" need to be clarified (2002: p. 15).

My overall argument is that Lightfoot's claim is much more straightforward than what P&S interpret it to be. In my view, Lightfoot is simply saying that:

a) Not all children get the same data.

b) The data children get are highly impoverished and include speech errors.

c) There is no negative evidence available to children (i.e., children do not usually get corrected when making mistakes, and there is no evidence that they absorb corrections even if they *are* pointed out), and yet,

d) All children master the language of the environment regardless of the shape and form of the available data. Simply put, no child is exposed directly to structure (i.e., there are no grammar courses for 0-5 year-olds), and there are no strings of structural information attached to the linguistic data. Yet, all children achieve native mastery of the structure of their first language.

As mentioned above, this type of misconception is not exclusive to those who argue against the POS concept. For instance, Julie Legate and Charles Yang who are advocates of the Innatist Hypothesis and offer strong pro-POS arguments against Pullum and Scholz also fall into a similar trap and miss the point in some important respects when they mention the following:

Figure 1.2 Legate & Yang 2002: p. 157

> "The only realistic acquisition data they [P&S; pz] give, based on the Nina corpus in CHILDES, is curiously selective: they report counts from only one file, NINA05.CHA, which happens to be the file that has the most number of critical sentences, out of all 56 files. Even for this file alone, they don't give a denominator —how many adult sentences the file contains— to give us a sense of how robustly these critical sentences are attested."

On the one hand, the point Legate and Yang are trying to make seems scientifically legitimate. Any scientist would like to see numbers and statistics in order to prove one point or another. On the other hand, I believe that the study of language acquisition in particular needs to go beyond the scope of simple data gathering. The Innatists (including Legate and Yang) should avoid getting entangled in the web of statistical data when it comes to language acquisition.

That is to say, the number of times a certain subject (in this case a girl named Nina) has heard a certain structure is irrelevant to the overall homogeneous process of FLA. Not all children are exposed to the same set of data but all children master their native tongue. Regardless of the amount of gathered data, we cannot claim otherwise. Nor can we sensibly discuss theoretical issues in FLA using assumptions which would preclude this fact.

As previously mentioned, this thesis is an attempt to clarify the nature of the Primary Linguistic Data (PLD) and argue in support of the Poverty of the Stimulus Argument. The overall approach is to explain these two phenomena and their close

relation to the best of my ability and, at the same time, give examples of how researchers[9] often diverge from the correct path by borrowing and misusing scientific terminology of which they have little understanding.

Chapter 2 will give a brief historical background of the development of the POS concept. For the most part, Chapter 2 will be based on Margaret Thompson's article called "the Development of the concept of the poverty of the stimulus" (2002: p. 51-71).

The notion of the PLD will not be examined in Chapter 2 since I see the Primary Linguistic Data as a basic external tool for language acquisition and a foundation for the Poverty of the Stimulus Argument. The historical evolution of the PLD is, hence, of little relevance to this thesis.

The central focus of Chapter 3 will be the intertwined nature of the two key concepts under discussion: the PLD and the Argument from the POS. First, what I believe to be the real definition of the PLD, its location and its nature will be outlined. Then, I will draw some logical conclusions based on my definition of the Primary Linguistic Data. The focal point of the second part of Chapter 3 will be the Argument from the POS. Two of the four examples that nativists use as evidence in favor of the POS Argument and that Pullum and Scholz claim, "further undermine the linguistic nativist position" (P&S 2002: p.9) will be re-examined in the second part of Chapter 3.

In Chapter 4 I will conclude that more empirical evidence and further research is needed in relation to First Language Acquisition (FLA). Moreover, researchers will be warned against terminological misuse and amateurish borrowing of expressions since it

[9] This thesis is mostly focused on Pullum and Scholz's dispute against the POS Argument. Their work, in my view, is a good example of how a challenge with little scientific base could be the cause of endless debates.

could have serious effects on scientific speculations. My overall conclusion will be that the definition of the PLD needs to be further refined and that, so far, the POS Argument remains unchallenged. Furthermore, I will draw attention to the fact that these two key concepts are closely related and should be further studied in relation to one another.

Chapter 5 will outline what appear to be profitable directions for further research arising from this thesis. I will firmly suggest that the process of Second Language Acquisition (SLA) is similar to that of FLA for the most part. I will mention that scientific research on SLA has only recently begun and propose that both FLA and SLA should be treated in the same way when it comes to the Argument from Poverty of the Stimulus as well as the availability of access to Universal Grammar (UG) principles and constraints.

CHAPTER 2

HISTORICAL BACKGROUND

One of the key concepts of generative linguistics is that of the Poverty of the Stimulus. According to Margaret Thomas, "More than any other claim, the assertion that input to language learners is incommensurate with what they end up knowing distinguishes generative from non-generative linguistic literature." (2002: p. 51).

In order to have a better understanding of this chief concept (i.e., the POS), a brief historical background is mandatory. In this chapter, I have laid out a narrative history of the POS Argument and its diachronic development.

It is important to mention that the notion of Primary Linguistic Data (PLD) is not discussed in this chapter. As mentioned before in Chapter 1 (see above), the PLD is a basic linguistic component of language acquisition and its historical development is of little relevance to this thesis. Understanding the nature of the PLD, however, is crucial and key to comprehending language acquisition theories including the Argument from POS. The nature of the PLD is dealt with in detail in Chapter 3.

According to Steven Pinker (1995b), the scientific study of language acquisition and the birth of cognitive science were concurrent events that initiated in the late 1950s and "the historical catalyst was Noam Chomsky's review [1959; pz] of Skinner's *Verbal Behavior*" (p.137). Skinner's overall claim was that language is a form of verbal— and hence a learned —behavior. Chomsky argued that language acquisition falsified these beliefs in a single stroke.

Figure 2.1 Pinker 1995b: p. 137

> "Children learn languages that are governed by highly subtle and abstract principles, and they do so without explicit instruction or any other environmental clues to the nature of such principles. Hence, language acquisition depends on an innate, species-specific module that is distinct from general intelligence."

Chomsky's debates eventually laid the ground for the birth of the Poverty of the Stimulus concept. "The Argument from the Poverty of the Stimulus is Chomsky's most unique argument" (Wexler, 1991: p. 268). Since the POS Argument is the central focus of this thesis, it is necessary to recount its history and the way this fairly new concept has evolved in the past twenty four (24) years.[10]

The term "poverty of the stimulus" was first mentioned by Chomsky (1980: p.34) in Rules and Representations (P&S, Thomas 2002). The idea was to address the "logical problem of language acquisition" namely "how a child masters a rich and highly structured system on the basis of degenerate and deficient data" (Hornstein and Lightfoot 1981: p. 9). Chomsky believed this to be "a variant of a classical argument in the theory of knowledge" (1980: p.34). Chomsky cited Socrates and Descartes as pioneers of the belief that humans are born with innate knowledge about the world. That is "knowledge which cannot be derived from the environment..." (Thomas 2002: p.52):

[10] The historical background of the POS Theory mentioned in this chapter is mainly based on Margaret Thompson's article (2002: p. 51-62).

Figure 2.2 Chomsky 1980: p. 35

"...since the stimulus does not resemble what the mind produces on the occasion of stimulation."

Margaret Thomas correctly points out that, "the concept now referred to as 'the poverty of the stimulus' began to emerge before the term first appeared in 1980, and it has continued to evolve" (2002: p.52).

As mentioned before, the theory of Language Acquisition has only recently become a central theory in the nativist framework:

Figure 2.3 Thomas 2002: p. 52

"The acquisition of language has always had theoretical importance in the work of Chomsky and his colleagues, but in the late 1950s and early 1960s it did not have the central position that it now enjoys, nor was it represented in exactly the same manner as it is now."

Until the early 1970s, the concept of the POS, "a conceptualization of input to language learners and the role of that input in relation to the language faculty", was rarely noticed by Generativists and/or other linguists (Thomas 2002: p.53)

Nevertheless, the idea always seemed to be present in Chomsky's work. For instance, in 1957 Chomsky briefly mentioned that despite the "finite and accidental" (p.15) exposure to language humans are able to understand and produce an infinite number of utterances. However, when Chomsky discussed Skinner's views in *Verbal*

Behavior, he evidently discussed the concept of language acquisition, but did not mention the Poverty of the Stimulus (Thomas 2002: p.53).

The notion that the Primary Linguistic Data is insufficient and degenerate was brought forth in 1960s. For example, in 1962, Chomsky realized the importance of explaining the Competence/Performance discrepancy (Thomas 2002: p. 54). He put an emphasis on the existence in speech of "interrupted fragments, false starts, lapses, slurring, and other phenomena that can only be understood as distortions of the underlying idealized pattern" (Chomsky 1962: p. 531).

In 1965, Chomsky took an important step with in relation to the theory of language acquisition and the quality of the available data. Chomsky asserted that "the primary linguistic data that [the child: pz] uses as a basis for [...] theory construction may [...] be deficient in various respects" (1965: p.201). According to Thomas, in 1968, Chomsky considered the scattered and degenerate nature of available data as a good argument for the plausibility of Universal Grammar (2002: p. 55).

The poverty of the stimulus concept, however, did not take center stage until mid 1970s. John Lewis and Jeffrey Elman discuss Chomsky's view on "the innateness of the principle of structure dependence" via an example (2003: p.1):

1) *Is the man who is smoking crazy?*

2) **Is the man who smoking is crazy?*

The idea, brought forth by Chomsky, is that children never produce sentence (2) without ever hearing any evidence as to the ungrammaticality of (2).[11]

[11] Lewis and Elman mention that this claim was subsequently empirically tested and substantiated by Crain and Nakayama (1987).

That is, "children create only structure-dependent grammars although no feature of the input eliminates structure-independence (Thomas 2002: p.56). In 1975, in *Reflections on Language*, Chomsky posited that "the only reasonable conclusion is that UG constrains the principle that all such rules must be structure-dependent" (Lewis & Elman 2003: p.1).

The mid-1970s seem to represent the onset of a shift in the view of language acquisition. However, the term poverty of the stimulus is not mentioned until 1980. It was Stanley Peters, in 1972, who coined the term 'projection problem' (p.171). This was to label the task of deriving grammar from data (Thomas 2002: p.55). His idea would contribute to the formation of the POS concept.

Figure 2.4 Thomas 2002: p. 56

Peters' Projection Problem: "The surface features of language cannot fully communicate to learners the properties of the grammar which produces them."

In 1975, in *The Logical Structure of Linguistic Theory,* Chomsky noticed the importance of language acquisition. He moved the reference to language acquisition to the first chapter of his book[12] (Thomas 2002:p. 53). Following is Chomsky's view on the most important issue that a theory of language has to explain:

[12] Apparently, this reference and the whole discussion on language acquisition were originally in Chapter 10 (Chomsky 1975a: p. 3).

Figure 2.5 Chomsky 1975: p. 61

"A speaker of a language has observed a certain limited set of utterances in his language. On the basis of this finite linguistic experience he can produce an indefinite number of new utterances."

In 1979, Baker further developed the idea of the 'projection problem'. What follows is some of the famous structure-dependent English phrases (Thomas 2002: p. 56):

1) The ambiguity of a phrase like: *Flying planes can be dangerous.*
2) The grammaticality contrasts between phrases like: *The child seems sleepy* and **The child seems sleeping.*
3) The intuition that there are radical differences in structure between phrases like: *John seems eager to please* and *John seems easy to please.*

These examples were formerly used in order to illustrate native intuition. But Baker was the person who used them in the context of the learnability issue. Baker changed the 'projection problem' to a concept he called the 'deductive gap':

Figure 2.6 Thomas 2002: p. 57

Baker: "...the 'deductive gap' between input and competence— claiming that speakers cannot arrive at the relevant intuitions of grammaticality or ambiguity merely by exposure to primary linguistic data."

Thomas remarks that Baker's contributions to the formation of the Poverty of the Stimulus Theory were substantial (2002:p. 57). For instance, he not only examined and

improved upon Stanley Peters' 'projection problem' he also broke this theory down to its basic components. Here are some of Baker's contributions as listed by Thomas:

1) He questioned the relevance of Negative Evidence.
2) He proposed that the input data for language acquisition is intrinsically inadequate since:
    a) Input is finite and the learner's capacity to learn is infinite.
    b) Input is positive evidence about well-formed sentences and does not indicate matters of ungrammaticality and ambiguity.

Building upon such arguments, Chomsky finally coined the term "poverty of the stimulus" in 1980 (p. 34).

In 1981, David Lightfoot contributed to the development of the POS concept by creating the term 'the logical problem of language acquisition' (Thomas 2002: p.58). Hornstein and Lightfoot expanded on Baker's binary theory on the intrinsic inadequacy of the acquisition data. They proposed what has come to be known as the 'canonical tri-partite characterization' of the input data.[13]

Thomas mentions three other, and more recent, developments since the 1980s that are related to the evolution of the Poverty of the Stimulus concept (2002: p.59):

1) The works of Crain and Nakayama (1987): their focus was on the structure-dependence theory. They looked at the same question formations discussed by Chomsky— i.e., why do children choose the structure-dependent *Is the man who is wearing a yellow coat at home* over the structure-independent but

[13] This tri-partite theory will be discussed in more detail in Chapter 3.

ungrammatical phrase **Is the man who wearing a yellow coat is at home*. The importance of Crain and Nakayama's work is that they empirically demonstrated that "children do not produce the relevant ungrammatical question structure" (Thomas 2002: p.59). In other words, Crain and Nakayama provided some valuable empirical evidence in support of the Poverty of the Stimulus Argument.

2) The work of Wexler (1991) and Schwartz & Sprouse (2000) have asserted a methodological role for the poverty of the stimulus Argument (Thomas 2002: p.61). Wexler's ideas are discussed in more detail in Chapter 3, but it is important to discuss his general idea at this point:

Figure 2.7 Thomas 2002: p. 61

The logic behind Wexler's idea is that "if no model for feature X in the input can be found to be commensurate with the richness of speakers' demonstrated knowledge of X, then investigation of X is likely to lead to insight into the language faculty."

3) The most recent development is, of course, Pullum and Scholz's claims regarding the Argument from the Poverty of the Stimulus. The discussion of their theory and my views on their proposals are the focal points of this thesis. It is important to mention that their work is scientifically significant since they have questioned the validity and empirical plausibility of the Poverty of the Stimulus Argument.

This chapter has been dedicated to a brief history of the origins, birth and the diachronic evolution of the Poverty of the Stimulus concept. As demonstrated, this concept has evolved to become firmly established as one of the most prominent arguments in support of generative linguistics and language universals.

We have seen, then, that over a relatively short period of time (i.e., the past 15-25 years) the Argument from the Poverty of the Stimulus has "gradually been knit more and more closely into both empirical study of language acquisition and theoretical work in the field" (Thomas 2002: p.62).

CHAPTER 3

The PLD & the ARGUMENT from the POS

In this chapter a debate regarding the current Chomskian view on the Poverty of the Stimulus theory initiated by Pullum and Scholz (*The Linguistic Review*, 2002, Vol. 19) will be examined. Pullum and Scholz's challenges to the Stimulus Poverty Argument will be explored.

Prior to discussing POS concept, what I propose to be a closely related notion—the Primary Linguistic Data (PLD) —will be examined. The idea is to show the interrelated nature of these two concepts. "Children most definitely *do* need to hear an existing language to learn that language [...] children with Japanese genes do not find Japanese any easier than English" (Pinker 1995b: p. 152).

Pullum and Scholz primarily base their arguments against the Poverty of the Stimulus concept on a rather peculiar definition of the Primary Linguistic Data (PLD). In fact, there seems to be a whole spectrum of views and a certain amount of inconsistency concerning the definition of the PLD. For instance, some researchers, including Pullum and Scholz, believe the PLD to comprise any (and all) linguistic data (uttered or in written form) that occur in the immediate environment of the Acquirer: "Given the application to language acquisition, it might be suggested that only material definitely addressed to children is relevant. But examples such as the following, from classic children's books, show that this challenge can be met..." (P&S 2002: p.29).

On the other hand, the majority of researchers treat 'all' speech directed at the acquirer as the PLD. "Not only do we know about the *output* of language acquisition, we

know a fair amount about the *input* to it, namely parents' speech to their children" (Pinker 1995b: p.137). In the following section, I will argue that only a very specific portion of the linguistic data available to the first language (L1) Acquirer forms the real Primary Linguistic Data.[14]

Considering the range and diversity of these views, it is crucial that a clear definition of the PLD, its location and, most importantly, its nature be set forward. Otherwise, there will be no solid foundation for the Stimulus Poverty debate.

3.1. WHAT IS THE PLD?

I will argue that the Primary Linguistic Data (PLD), like any other cognitive stimulus, cannot be something outside of (or apart from) the entity that is on the receiving end (the Receiver/ the Acquirer in the case of language acquisition). What we see (i.e., visual stimuli), for example, is not a literal copy of what exists in the environment, but rather a mental representation which results from our cognitive interpretation of the perceived stimulus.[15]

By the same token, speech utterances that are available in the environment can only be considered potential stimuli that, if perceived, will cause brain activity on the part of the Receiver (in this case, the Acquirer). The specific brain activity triggered by a given stimulus may be quite different based on issues such as attention, processing load, etc. The Acquirer, therefore, is not a mere recording machine. The Language Acquisition Device (LAD) does not simply hear, record, and allow the Acquirer to utter speech data.

[14] For a discussion regarding the nature of the PLD see also M. Hale, 1998, *Syntax*, p. 1-8
[15] The reception vs. perception analysis is the most fundamental aspect of Cognitive Science.

In this and all other cases of sensory perception, the data is not just received, but perceived and potentially interpreted via cognitive processes that take place in the Receiver's brain.

Consequently, it becomes rather obvious that in the process of language acquisition, the location of the PLD should be shifted from the immediate environment of the Acquirer to the part of the individual's brain responsible for language processing and acquisition. As a result, the definition of the PLD should also be modified. The PLD neither refers to 'all speech utterances in the immediate environment of the Acquirer' (P&S 2002) nor to 'all speech utterances directed to the Acquirer' (Pinker 1995). I suggest that instead we call such data the ALD —the Available Linguistic Data.

Moreover, in my attempt to pinpoint the real definition of the PLD, I would go one step further and assert that even the Linguistic Data received by the Acquirer is not the real PLD. A more appropriate name for this portion of the linguistic data, in my opinion, is the RLD —the Received Linguistic Data. Conclusively, a more accurate definition of the PLD should resemble the following:

The PLD refers to the portion of the Received Linguistic Data (RLD) that is selected, perceived and processed by the Acquirer's Language Acquisition Device (the LAD).

In other words, I view First Language Acquisition (FLA) as the cognitive processes involved in:

a) absorbing all the Received Linguistic Data (RLD) available in the immediate environment,

b) perceiving and processing a selective[16] portion of the data which could, at this point, be called the Primary Linguistic Data (PLD),

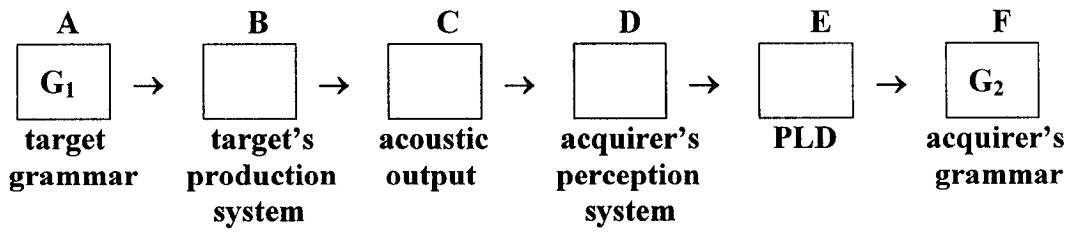
c) triggering the Language Acquisition Device (i.e., the part of the brain responsible for language acquisition) to apply the necessary algorithms and processes to the relevant portion of the perceived data (PLD),

d) constructing the Acquirer's version (i.e., G2) of the target grammar (i.e., G1; the grammar of the environment) based on the PLD and the Principles and Parameters of Universal Grammar genetically available to the Acquirer, and,

e) acquiring the knowledge and/or competence to produce speech and communicate via the target language (i.e., the Acquirer's version of the target language; G2).

The following figure will illustrate:

Figure 3.1 Mark Hale (1998: p. 8)



[16] The details of what part of the RLD is selected (and why) are too complex to be dealt with here, but it would be interesting to see the result of experiments based on these questions.

Figure 1 above shows how P&S (2002) as well as those who define PLD as any (and all) speech utterances in the environment are oversimplifying a very complex matter. A more in-depth look seems to reveal that the PLD should consist only of the portion of the available speech data that is perceived by the Acquirer's mind. And even then, it is only the combination of the perceived data and the application of the necessary algorithms and processes by the Language Acquisition Device (LAD) on the data that creates the final outcome; i.e., the Acquirer's version of the target grammar (G2). Hence, in my opinion, a more accurate definition of FLA should be:

The cognitive processes involved in constructing G2 (the Acquirer's perception of the Target Grammar) from the PLD which arises, indirectly, from the output of G1 (the Target Grammar).

Viewing FLA in this manner leads to certain logical assumptions about the way the RLD (Received Linguistic Data) is filtered by our cognitive processes to turn into the PLD and then used as a tool to trigger the LAD to construct the Acquirer's version of the Target Grammar (G2). What follows is a discussion of these logical assumptions.

3.1.1. The Acquirer *cannot* be born with a blank slate (*tabula rasa*)

If no biological predisposition to linguistic universals (i.e., structures and algorithms biologically available to the Acquirer) is presumed, it is impossible to explain

how the Acquirer is able to build a grammar simply by being exposed to linguistic data from the immediate environment. In assuming that humans are born with a blank slate, many cross-linguistically analogous patterns of First Language Acquisition would go unexplained.

For instance, we would be unable to account for the way members of the same speech community are able to cognitively create nearly identical versions of the target grammar (G1).

Unlike what Pullum and Scholz seem to imply in their arguments against the stimulus poverty exemplars (2002: p. 24-45), different children *are* exposed to different sets of linguistic data,[17] yet they learn essentially the same language and are able to successfully communicate with one another. That is, their version of the target grammar (G2; see figure 3.1 above) bears enough similarities to result in successful communication. How is it possible for different members of the same speech community to build relatively similar versions of the G1 on a "blank slate" when the ALD (Available Linguistic Data) of their individual environments could be (and most definitely is) different?

This dilemma becomes even more significant when we add to it the theory (proposed in this thesis) that the triggering data is neither the ALD, nor the RLD, but the

[17] In their attempt to undermine the four widely used exemplars of the POS (i.e., FLA in the absence of relevant and crucial data), Pullum and Scholz use children's novels as a means to prove the existence of exposure to certain syntactic aspects of the English language. "Given the application to language acquisition, it might be suggested that only material definitely addressed to children is relevant. But examples such as the following, from classic children's books, show that this challenge can be met..." (2002: p. 29). In their debate, P&S fail to account for one simple fact: NOT ALL CHILDREN ARE EXPOSED TO THOSE (OR ANY) BOOKS AND, YET, THEY ACQUIRE THE LANGUAGE OF THE ENVIRONMENT!

PLD which consists *only* of the portion of the linguistic data that is cognitively perceived. How could this dilemma be explained on the basis of the blank slate hypothesis?

Furthermore, if a biological blank slate is presumed, how do we explain the homogeneous ability among homo-sapiens to build a whole grammar structure based solely on the available data? Where do humans get the tools to do so? If the assumption is that there is no need for preexisting tools, how can we account for the fact that other species (e.g. cats, dogs, dolphins, etc.) or real world objects (e.g. Potatoes, House Plants, etc.) do not automatically acquire the language and/or the grammar of the environment.[18] In other words, if the data from the environment is the only necessary element, why is it that the Available Linguistic Data (ALD) does not trigger acquisition in other species?

Overall, I think it logical to assume that without linguistic biological endowment, language acquisition cannot begin. For example, a child (the Acquirer) exposed to a certain grammar structure (for instance, embedded clauses; e.g. *Mary said that she knew the man standing in the corner*), will not be capable of building the structure of the complex Noun Phrase (NP) if this structure (or the ability to linguistically analyze the structural relations of this complex NP) is not cognitively available. Simply put, to analyze a "complex NP" the Acquirer must be able to construct representations of NP's, which in turn requires that they have the capacity to represent N's, to build phrase structures for heads, etc. We thus can deduce a chain of necessary capacities of the acquirer, each building on some logically necessary antecedent capacity. The end of this chain cannot be without content. Thus, whatever content it has, *must* be innate.

---

[18] It has been shown that certain animals can learn how to communicate via words with humans. However, this is not an automatic process like that of a human infant, nor is word-usage "grammar".

A linguistic stimulus (in this case the PLD) can trigger a series of cognitive processes which help configure the embedded clause structure, if and only if this type of structural relation is cognitively available to the Acquirer prior to receiving the environmental stimulants (in this case the RLD). This brings us to the next logical assumption in discussing the nature of the PLD.

3.1.2. The grammar *is not* built on "all" existing data in the environment

As mentioned above (see section 3.1.), the Primary Linguistic Data (PLD) should only refer to all linguistic utterances perceived and linguistically processed by the Language Acquisition Device (LAD) of the Acquirer. I also discussed that the Received Linguistic Data (RLD) can be processed and analyzed if, and only if, the Acquirer (i.e. the human child[19]) is genetically equipped to process and analyze such data.

In my opinion, if the Acquirer were to build the structure of the language based on "all" the available linguistic data —which I called the ALD or Available Linguistic Data (see above) — the results would be chaotic and heterogeneous to say the least. In this hypothetical situation, there could be no entity called a speech community and successful communication among the members of the same species sharing the same environment would become extremely difficult —if not downright impossible— due to the inevitable diversity that would constitute their individually structured grammars.

An example will demonstrate the result of such a hypothetical state. In such a group of species, each neonate would absorb all the utterances in the immediate environment and use the accumulated data as a means to build the structure of the target

[19] I believe that Second Language Acquisition (SLA) follows the same rules as FLA. However, the topic of SLA should be extensively discussed in another paper.

Grammar. The RLD[20] would, hence, include all ALD which would in turn consist of all grammatical as well as erroneous speech utterances (e.g. speech errors, slips of tongue, ungrammatical sentences, unfinished utterances and all other examples of performance errors made by speaker in real-time speech on a regular basis). Several questions arise pertaining to such a situation:

1) How would the Acquirer make sense of such structural errors?
2) If resolved, what structure would the Acquired Grammar (G2) have?
3) How similar would G2 be to the target grammar (G1)?
4) Would the members of the same "speech community" be able to communicate with one anther via their individual grammars?
5) Most importantly, if this hypothetical community did indeed exist, could there possibly *be* a target grammar?

I would argue that the end result of a hypothesis where "all" linguistic data is assumed to be absorbed by the Acquirer would be, to say the least, impossible to predict. In real life and in all human speech communities, however, the homogeneous nature of FLA indicates that reality must be otherwise and the above hypothesis can, therefore, not be the case in the process of First Language Acquisition.

In fact, cross-linguistic studies show that the acquisition process of human infants is remarkably uniform. Children in all speech communities are able to construct very similar approximations of the target language in a relatively short period of time (about five years). A sensible way to account for the homogeneity of acquisition is to assume

[20] Notice that there would be no PLD since, in this imaginary world, there would be no Language Acquisition Device (LAD) to perceive and cognitively process the data.

that the mind of the Acquirer (i.e., the Language Acquisition Device), imposes structure on the input data. If we consider LAD to be innate and, hence, uniform, imposed structures will also be uniform.

So far, we have posited that the PLD triggers the LAD in the brain in order to construct the target grammar. We have also asserted that the PLD cannot refer to "all" speech utterances that occur in the environment, but only the portion perceived and processed by the Language Acquisition Device of the Acquirer. Moreover, we have stated that without genetic wiring and the existence of a LAD specialized in constructing the target grammar, language acquisition cannot begin.

This brings us to another important issue with regards to First Language Acquisition and the role human cognition plays in its process.

3.1.3. The human brain *must* be equipped with a Linguistic Faculty/Module

In opposing the Poverty of the Stimulus Argument (POS), Pullum and Scholz take a stance against the Nativist theory of linguistic knowledge, "...if the notion of innateness can be a plausible claim, then *general* nativism would certainly be a viable competing alternative to the concept of there being innate knowledge in our species, albeit concerning language" (2002: p.7).

Nevertheless, I will argue that if language acquisition is attributed to "General Nativism"— as opposed to a Language Faculty, it is imperative to find an answer to the following question: Why is this particular innate characteristic so much more reliable than other genetically-based human capabilities?

For example, how could we account for the fact that, cross-linguistically, every normal human child (i.e., one with no brain damage or damage to parts of the brain

responsible for language comprehension and/or production) born into a normal setting, will acquire the language of its immediate environment in a fairly rapid time span and in a fashion similar to that of other members of its species and the same is not true for the acquisition of the contents of fields such as Mathematics or Medicine?[21]

If all human abilities are attributed to General Nativism, then it must be concluded that they should have similar characteristics. However, we cannot predict that a human infant will learn how to play tennis if exposed to the game during the first few years of its life. But we *can* predict that all "normal" human infants will learn the language of the environment in a similar fashion and roughly within the same time span. In my opinion, this is the single most important characteristic that sets the process of first language acquisition (FLA) apart from other human traits.

In other words, the linguistic mastery of a 5-year-old child is highly predictable, but the same is not necessarily true, for instance, when it comes to achieving mastery at playing a game or being successful at school. Unfortunately, the advocates of General-Nativism and researchers like Pullum and Scholz don't seem to appreciate the significance of this key difference between general learning strategies and the process of First Language Acquisition (FLA). The reliable nature and the predictable outcome of the processes involved in FLA seem to provide solid proof for the specialized and innate nature of the language faculty. The assertion that there exists a specialized language faculty that is responsible for the construction of the grammar of the environment in the human brain logically leads to the next conclusion.

---

[21] Mathematics here does not refer to the general mathematical and/or intellectual abilities human infants seem to be born with. This example is in reference to mathematical formulas, rules and regulations that do not appear to be automatically available to humans (e.g. the formula for calculating the circumference of a circle).

3.1.4. The Acquirer *must* be open to all linguistic possibilities at the time of birth

"It seems natural and inevitable for human communication to depend on spoken words, even though this requires tens of thousands of arbitrary connections between noises and concepts [...] human children [...] learn these connections [...] even thought this requires analyzing the behavior of adults emitting concatenations of noises referring to logically structured combinations..." (Liberman 1995: p.55).

The cross-linguistic predictable nature of FLA requires that all human neonates be open to all possible structural (phonetic, phonological, morphological and syntactic) aspects of all human languages.

Otherwise, there would have to be cases of "normal" human children incapable of learning the language of the environment despite exposure to the Available Linguistic Data (ALD). If humans were not born linguistically open to all possibilities, there would be natives incapable of native-like speech. So, until we find examples of this peculiar phenomenon, we can logically assume that this is not the case.

Following is a recap of the four logical assumptions based on the cognitive nature of the Primary Linguistic Data:

a) The Acquirer *cannot* be born with a blank slate (*tabula rasa*)

b) The grammar *is not* built on "all" existing data in the environment

c) The human brain *must* be equipped with a Linguistic Faculty/Module

d) The Acquirer *must* be open to all linguistic possibilities at the time of birth

As mentioned above, this thesis is based on the elaborate relation between the PLD and the Argument from the Poverty of the Stimulus (POS). After having discussed the major aspects of the PLD, its nature and location, it is crucial to review the POS Argument based on the modified version of the PLD as discussed in this thesis.

3.2. WHAT IS THE ARGUMENT FROM THE POS?

The term Poverty of the Stimulus (POS) was first used by Chomsky (1980: p. 34). A good definition of the Argument from the POS is given by Norbert Hornstein and David Lightfoot (1981:p.9):

Figure 3.2 Hornstein & Lightfoot 1981: p. 9

> The POS Argument states that: "People attain knowledge of the structure of their language for which no evidence is available in the data to which they are exposed as children."

Since 1980 (Chomsky: p. 34), the POS concept has evolved to be considered one of the most powerful arguments in support of the Universal Grammar (UG). The idea is that the mental ability to build a whole language based on impoverished data (what is called the PLD in the literature but should really be called the ALD based on this thesis) is an indication that humans cannot be born with a blank slate. Chomsky (1980) proposed the Poverty of the Stimulus concept as evidence that FLA (or language acquisition in general; i.e., 2^{nd}, 3^{rd}...n^{th} language acquisition) cannot occur without innate language

universals that help the human brain construct an accurate and mutually comprehendible version of the grammar of the environment (= the Target Grammar; G1).

As mentioned above, the Poverty of the Stimulus (POS) Argument has recently been challenged by Geoffrey Pullum and Barbara Scholz (2002: p. 9-50). These challenges, in turn, have inspired a number of pro and con arguments by othe researchers in related fields.[22]

Although debates are crucial building blocks of scientific growth, I will argue that the points Pullum and Scholz bring forth in their attempt to challenge the Argument from the Poverty of the Stimulus lack scientific foundation and, therefore, fail to be a true challenge.

Pullum and Scholz do make some very good general points. For example, I agree with P&S when they mention the growing concern that, "linguists have not achieved what they are widely thought to have achieved. The empirical work that would be needed to support their claims has not been carried out" (2002: p. 10).

Linguists (and all other cognitive scientists) often base their speculations on statistical data and/or borrowed scientific terminology of which they have little understanding. As I have already pointed out in Chapter 1, there is a serious need to step back, and try to fully comprehend the overall implication of the claims we make. For instance, Pullum and Scholz are correct to mention that much empirical research is needed for all involved to reach a better understanding of the brain and its functions (2002:p. 9).

---

[22] For the Poverty of the Stimulus debate see *Linguistic Review* 2002, Vol 19.

Unfortunately, however, Pullum and Scholz seem to make the very same mistake they accuse others of making. In what follows, I will show how P&S do not appear to fully comprehend what needs to be done to improve the quality and coherence of linguistic research. Nor do they have a clear understanding of the Poverty of the Stimulus concept. I will also argue that their claims are based on little or no relevant empirical evidence. I will then conclude that the Poverty of the Stimulus Theory remains unchallenged and assert that the first step is to comprehend this concept in order to appreciate the Nativist claims regarding language acquisition.

In their article, Pullum and Scholz have argued against four exemplars often used by nativists as evidence in support of the Stimulus Poverty Argument. Their claim is that none of these exemplars hold up (2002: p. 9). They further assert that "the relevant kind of future work on this issue [the POS Argument; pz] is likely to further undermine the linguistic nativist position" (2002:p.9).

Since a plausible scientific theory must account for all aspects of the domain under investigation, I believe a careful look at Pullum and Scholz's claims is mandatory. In other words, the Argument from the Poverty of the Stimulus cannot be elucidated without a careful analysis of all opposing arguments.

What follows is an examination of a series of discussions by Pullum and Scholz against divergent views. If these arguments hold, then the whole concept of Poverty of the Stimulus Argument should indeed undergo major modifications. And if, as P&S assert, the nativist position on POS is undermined, then the entire generative conception of the nature of language acquisition will have to be reexamined.

3.2.1. INNATENESS:

Chomsky has proposed that language acquisition must rely on linguistically detailed knowledge. According to Chomsky (1975), Universal Grammar (UG) is posited to account for those "properties of language that can reasonably be supposed not to have been learned" (Lewis & Elman: 2001, p.1).

The idea of Innateness is the very core of the nativist theory. By questioning the scientific validity of the Poverty of the Stimulus Argument, Pullum and Scholz (2002) have challenged the plausibility of the linguistic innateness concept. In what follows, I have examined a series of examples in this relation.

3.2.1.1. Pullum and Scholz vs. Wexler:

The focal point of Pullum and Scholz's (2002) claims is that, on the whole, the nativist theory lacks empirical support and as well as a clear definition of its key concepts (e.g. the POS Argument). However—and despite the absolute necessity of constant re-evaluation of views and the need for empirical research, I will argue that the overall approach adapted by P&S also seems to be devoid of scientific support.

For instance, in their discussion against the Innateness Theory of language acquisition, P&S assert that the term "argument from the poverty of the stimulus" has never been clearly defined (2002; p. 12).

According to Pullum and Scholz, the term— first used by Chomsky (1980; p. 34), was merely picked up by other researchers without any effort to state the details of the Argument (P&S; p.12). As an example they mention the following excerpt from Wexler's article named 'On the Argument from the Poverty of the Stimulus':

Figure 3.3 Wexler 1991: p. 253

> "How does the child construct her grammar? In other words, why is the adult output grammar the one that it is? Chomsky's answer notes that the attained grammar goes orders of magnitude beyond the information provided by the input data and concludes that much linguistic knowledge must therefore be innate."

Pullum and Scholz state three opposing arguments against Wexler's view (above) and state that Wexler cannot conclude that linguistic knowledge must be innate based on this definition (2002: p.12). Nevertheless, they, themselves, appear to be more puzzled than those they accuse. Following is an illustration of the three arguments against Wexler's account followed by my discussion of how P&S seem to diverge from the real issues concerning the POS Argument.

a) P&S vs. Wexler: opposing argument #1:

> *"How does the child construct her grammar? In other words, why is the adult output grammar the one that it is?" (Wexler 1991: p. 253)*

In their argument against Wexler, Pullum and Scholz assert that "Surely, how a grammar is constructed is not the same question as why that particular grammar was constructed, yet Wexler appears to conflate the two questions" (P&S: p.12).

My disagreement with P&S is that the notion of Innateness in the context of Language Acquisition leaves no place for 'why'! Contrary to what Pullum and Scholz claim, Wexler is not conflating 'how' and 'why' in these two questions. My understanding is that by posing these questions, Wexler is merely confirming the unified nature of native linguistic competence in adults and the way children acquire the language of the environment. Wexler's questions are, therefore, not as trivial as 'why a language is constructed', but rather 'why it is constructed the way it is'. The point Wexler is trying to make is that the answer should be found within the genetic wiring of humans.

As previously mentioned (see above), the human mind constructs the grammar of the environment (i.e., the G1 to G2 process; see figure 3.1 above) and allows the Acquirer to achieve native ability in language usage simply because it is genetically wired to do so. Hence, for research purposes, it is imperative to acknowledge that the 'why' Pullum and Scholz are discussing (2002: p. 12) is fundamentally different from the 'why' Wexler has intended (1991: p. 253).

b) P&S vs. Wexler: opposing argument #2:

> *"...Chomsky's answer [to the first two questions – see (a) above; pz] notes that the attained grammar goes orders of magnitude beyond the information provided by the input data" (Wexler 1991: p. 253).*

Pullum and Scholz counter the above claim made by Wexler by stating that "...surely 'orders of magnitude' is not the right phrase (input data sets are finite, and generative

grammarians take grammars to generate not finite languages that are orders of magnitude bigger than input corpora, but infinite languages)" (2002: p. 12).

I think that Pullum and Scholz's opposing argument is a clear indication of what little understanding they have of the Poverty of the Stimulus concept. P&S might be right in suggesting that the mathematical term "orders of magnitude" is wrongly used by Wexler in this context, but in using the term correctly, they manage to misinterpret the concept of language acquisition. That is, Wexler's misapprehension of a mathematical term does not falsify his basic assertions about the process FLA.

The interpretation of what Pullum and Scholz suggest here is that based on the mathematical definition of 'orders of magnitude' a child should be able to construct an infinite number of languages from the 'Input Corpora' (i.e., what I have called the RLD or Received Linguistic Data; see section 3.1. above). And since this is quite obviously not the case (and nor what they want to imply), P&S should have simply pointed out that using the term 'orders of magnitude' is incorrect in this context.[23]

To be exact, if Pullum and Scholz aimed to heighten our understanding of the FLA process, they would focus more on what Wexler is trying to say; i.e., that Language Acquisition consists of the cognitive construction of a 'finite' set of rules (syntax) which govern the language of the environment with an 'infinite' number of sentences (lexicon). Evidently, children *do* learn the language of the environment and they *are* capable of creating an infinite number of sentences (for similar discussions see also Hornstein and Lightfoot 1981; Legate and Yang 2002; Fodor and Crowther 2002).

[23] Pullum and Scholz fail to address the important (and scientifically plausible) point Wexler is making. Instead they focus on his misuse of the mathematical term.

Overall, I believe that Pullum and Scholz should be more careful not to forget the focal point of their own debate. In other words, whatever mathematical (or other term) is used (or misused for that matter), the bottom line is that there is a significant difference between the number of speech data or elements of the 'input corpora' available to children and the 'magnitude' of their acquired lexical and syntactic inventory.

c) P&S vs. Wexler: opposing argument #3:

> *"...and concludes [Chomsky; pz] that much linguistic knowledge must therefore be innate" (Wexler, 1991; p. 253).*

Pullum and Scholz's next argument against the last part of Wexler's comments (see quote above) is even more interesting: "...Moreover, whatever the first two statements mean, it surely does not logically follow from them that 'much linguistic knowledge must...be innate'. Nevertheless, Wexler's claims are fairly typical" (P&S: 12).

I think it would be interesting to see how P&S resolve this dilemma! On the one hand we have the 'input corpora'— what I call the RLD —which is incomplete and impoverished for the most part. On the other hand we have cross-linguistic evidence that children acquire and construct the language of the environment based on the portion of this incomplete data perceived by their brain— what I call the PLD —in a relatively short period of time (normally during the first five years of their life). Other than presuming Innateness, how else could Pullum and Scholz possibly explain this phenomenon? Surely simply asserting that 'it does not follow' is not sufficient. I think the task of providing an

alternate (and plausible) account of how acquisition could possibly take place without invoking any (language specific) innateness is Pullum and Scholz's responsibility.

In my view, the most dangerous part of this type of debate is that, at the end, no practical solutions are offered. Moreover, the very fact that some researchers do not have a solid comprehension of mainstream fundamentals should undermine any possible suggestions they bring to the table. After all, having a solid grasp on a theory before one can argue against it is mandatory.

3.2.1.2. Pullum and Scholz vs. Hornstein and Lightfoot:

Another example of misapprehension can be found in Pullum and Scholz's debate (2002: p. 12-17) against a quote taken from Norbert Hornstein and David Lightfoot in an article that deals with 'the logical problem of language acquisition' (1981: p. 9).

In their debate counter the Argument from the Poverty of the Stimulus, Pullum and Scholz mention that they do not necessarily disagree with the POS notion, but they do disagree with one particular approach made by Hornstein and Lightfoot (2002: p. 14):

Figure 3.4 Hornstein & Lightfoot 1981: p. 9

"People attain knowledge of the structure of their languages for which no evidence is available in the data to which they are exposed as children."

According to P&S, "what Hornstein and Lightfoot claim is that some of the sentences children never hear are crucial evidence for learning from experience. That is, some aspects of languages are known to speakers despite the fact that the relevant positive evidence, *although it does exist*, is not accessible to learners during the acquisition

process, because of its rarity: linguists can in principle discover it, but children will not" (2002: p.14).

In my opinion, what is puzzling is how Pullum and Scholz draw the above conclusion from this particular quote. Hornstein and Lightfoot, along with other advocates of nativism (Chomsky 1975, 1980; Lasnik&Uriagereka 2002; Legate&Yang 2002, etc.) simply claim that adult knowledge of their L1 (= first language) is not a mere reflection of the Available/Potential Linguistic Data (neither structural nor lexical).

Contrary to what Pullum and Scholz declare, I believe Hornstein and Lightfoot have quite accurately pinpointed what has come to be called the "logical problem of language acquisition" in contemporary linguistic literature. According to Hornstein and Lightfoot, the fact that children acquire the complex structural and lexical system of the language despite the impoverished data requires explanation on three levels (=The Canonical Tri-Partite Characterization):

Figure 3.5 Hornstein & Lightfoot 1981: p. 9-10

a) The speech the child hears does not consist uniformly of complete grammatical sentences, but also of utterances with pauses, slips of the tongue, incomplete thought, etc.

b) The available data are finite but the child comes to be able to deal with an infinite range of novel sentences, going far beyond the utterances actually heard during childhood.

> c) People attain knowledge of the structure of their language for which no evidence is available in the data to which they are exposed as children [...] Children are not systematically informed that some hypothetical sentences are in fact ungrammatical, that a given sentence is ambiguous [...] the distinction between what is available to the linguist and the more limited data available to the child is of vital importance for our view of things.

To Hornstein and Lightfoot's comments I would also add that not all the available linguistic data is what constitutes the PLD. As discussed above (see section 3.1.), the relation between the available data and the Language Acquisition Device (LAD) of the Acquirer is multi-tiered. While all Available Linguistic Data (the ALD) could potentially act as a stimulus to trigger acquisition, the Acquirer receives only a portion of this data (the RLD). Moreover, only a part of the RLD is selectively perceived and processed (the PLD) by the Language Acquisition device (LAD). Considering this multi-tiered theory of the Linguistic Data could have dramatic effects on our understanding of the process of language acquisition.

3.2.2. NEGATIVE EVIDENCE:

Nancy Ritter[24] briefly discusses the notion of "Negative Evidence". With regard to language acquisition in children, she mentions that "...without relying upon negative evidence to discount some of these hypotheses [related to FLA; pz] and not others, the child would have no means for choosing one hypothesis over another" (2002: p. 4).

[24] Nancy Ritter is the Editor of The Linguistic Review (2002: Vol 19) which is entirely dedicated to the Argument from the POS debate.

Similarly, I believe that the question is: why should the child choose at all? If the child is born with a Language Acquisition Device that parses[25] the Perceived Language Data (PLD) into its structural components, then the LAD of the child is not really 'choosing', but simply decoding the string of sounds as it receives the data. An example would be a child born into a Spanish-speaking environment who repeatedly hears sentences with an overt subject, and then, a subject-less sentence.[26] By hearing sufficient evidence of sentences which appear to have no subject, the LAD of the Acquirer includes this particular parameter (i.e., the pro-drop parameter; Radford 1997) in the structure of the target grammar. On the other hand, the LAD of an Acquirer born into an English-speaking surrounding will never include such a particularity in the structure of English.

Other good examples would be bilingual children (who never mix the two languages they are exposed to). It would be interesting to look at, for instance, a Spanish/English or a Japanese/English bilingual child. Both these children are exposed to two systematically different languages (Pro-Drop in case of the former and headedness[27] in case of the latter).

With regard to headedness, for example, we could consider Japanese to be the "absolute" negative evidence for English. Yet, cross-linguistic studies support the claim that children natively exposed to Japanese and English —or any other two languages for that matter— do not confuse them and eventually acquire native competence in both

[25] Parse: To determine the syntactic structure of a sentence or other utterance (Hosting works dictionary 2004).
[26] Spanish is a Pro-Drop language in which phrases with covert subjects are grammatical (Radford 1997: p. 20).
[27] English is a left-branching/head-first language and Japanese a right-branching/head-last language (Radford 1997: p. 22).

languages.[28] Therefore, it becomes evident that there seems to be no need for negative evidence in the course of language acquisition. Many researchers believe that the lack of need for negative evidence provides support for the POS Argument. To cite an example, Janet D. Fodor and Carrie Crowther assert that, "the argument for linguist nativism appears to be solidly supported by the distinctive patterns of generalization that learners adopt in the absence of systematic negative evidence..." (2002:p. 105).

So far, I have argued that the only method to explain the homogeneous way in which all children learn the language of the environment is by positing nativism and genetic wiring. I have also suggested that negative evidence is *not* a necessary component of language acquisition. I agree with Fodor and Crowther who take the negative evidence discussion even one step further and posit that "POS includes *every* respect [i.e., poverty of positive evidence and lack of negative evidence; pz] in which learners' input underdetermines the adult grammar" (2002:p. 106).

As mentioned before, Pullum and Scholz have examined four examples often used by advocates of the Innateness Hypothesis as evidence for the POS theory. They have concluded that:

Figure 3.6 Pullum and Scholz 2002: p. 9

> "...linguists have some additional work to do if they wish to sustain their claims about having provided support for linguistic nativism, and we offer some reasons for thinking that the relevant kind of future work on this issue is likely to further undermine the linguistic nativist position."

[28] The study of bilingual children is too extensive and beyond the scope of this thesis.

In the next section, I examine two of the exemplars Pullum and Scholz have used in their dispute against the POS Theory. Like many others,[29] I find that their arguments do not hold up and conclude that the Poverty of the Stimulus Argument still remains unchallenged.

3.2.3. THE POVERTY OF THE STIMULUS ARGUMENT:

In the last two decades, The Poverty of the Stimulus (POS) Argument has become one of the most powerful tools used by nativists to support their belief that children "come equipped with a priori knowledge of language" (Lasnik & Uriagereka 2002: p.149):

Figure 3.7 Legate & Yang 2002: p. 151:

> "The so-called Innateness Hypothesis, which claims that crucial components of our tacit linguistic knowledge are not learned through experience but are given by our biological/genetic specification, is not really a hypothesis. Rather, it is an empirical conclusion mainly based on observations of child language acquisition, one of which is now known as the Argument from the Poverty of the Stimulus (APS) [APS is used by Pullum & Scholz as the equivalent of the POS; pz]."

When a central Theory is challenged, it is of utmost importance that the advocates of that theory carefully examine the opposing arguments and, if valid, modify the theory

[29] See also Lasnik&Uriagereka, Legate&Yang, and Fodor&Crowther 2002.

accordingly. Pullum and Scholz claim that they have brought such a challenge upon the POS Argument. To give an example, they argue that the Poverty of the Stimulus Argument does not provide strong enough evidence to eliminate Data-Driven Learning (DDL) or establish Innateness (2002: p. 20).

In proving their claim, Pullum and Scholz use the syntactically unique expression 'by and large' (2002: p. 20):

$$[_{PP}\,[_{P}\,\mathbf{by}]\,[_{Ad}\,[_{Cnj}\,\mathbf{and}]\,[_{Adj}\,\mathbf{large}\,_{Adj}]\,_{Adj}]\,_{PP}]$$

Their idea is that since this syntactic structure does not typically occur in English, it would provide important supportive evidence for the POS Argument if a speaker who never heard the expression before could judge its grammaticality correctly.[30] However, Pullum and Scholz move on to say, "...if the learner had definitely heard instances of the *by and large* construction that would support the idea that data-driven learning might have taken place" (2002: p. 20).

I would argue that, there are logical flaws in this type of reasoning:

a) This is not a reciprocal equation. That is, to prove that an acquirer has learnt an idiom via exposure does not necessarily (or logically) mean that data-driven learning (DDL) has occurred.

b) If we have learned an expression through exposure, we still have to have figured out its underlying syntactic structure in order to achieve correct usage and/or sound grammaticality judgement (see footnote 30). Therefore, even if the acquirer has learned the particular idiom through exposure, there *must*

[30] Judging correctly would mean that they would deem a sentence like *by and large, things are fine* to be grammatical and expressions like **by and small* or **of and large* to be ungrammatical (P&S 2002: p. 20).

exist a Language Acquisition Device capable of parsing the novel data correctly. And last but not least,

c) Hearing does not inevitably result into learning. That is, even if an acquirer must hear an expression like *by and large* in order to learn it, having heard the expression does not necessarily entail that acquisition has occurred.

I believe that this type of reasoning is Pullum and Scholz's biggest downfall in their challenges against the Poverty of the Stimulus Argument. In what follows, I will examine their disputes against two famous examples often used by advocates of nativism in support of the POS Argument, Auxiliary Inversion and Auxiliary-initial Clauses. I will conclude that the debates set forth by Pullum and Scholz do *not* undermine the POS Argument and, until there is a scientifically-based opposition, this concept remains unchallenged.

3.2.3.1. Auxiliary Sequences (Pullum & Scholz vs. Kimbal):

Before the term 'poverty of the stimulus' was coined by Chomsky (1980: p.34) John Kimbal (1973: 73-75) provided a clear example of the POS phenomenon. Kimbal stated that the following (Chomsky 1957) is the correct underlying phrase structure of an auxiliary sequence:

$$Aux \rightarrow T\ (M)\ (have + en)\ (be + ing)$$

Kimbal then concluded that a child should be able to correctly judge an Aux-sequence like *it may have been raining* to be grammatical based on the above phrase structure rule (P&S 2002: p. 27). According to Kimbal:

Figure 3.8 Kimbal 1973: p. 74

"sentences in which auxiliary is fully represented by a modal, perfect, and progressive are vanishingly rare[...]Thus, the evidence indicates that a great many English-speaking children will acquire the full auxiliary system [...] without having heard sentences directly illustrating each of the rules."

The successful grammaticality judgment on the part of native speakers of English makes the Aux-sequence rule a good attestation of the Argument from the Poverty of the Stimulus.

Pullum and Scholz first tackle the inaccessibility claim. "Looking in the relatively small and in some ways unsuitable texts available to us [...] we find hundreds of examples." (2002: p.28):

i. *I must have been falling asleep* (Bram Stoker, Dracula)

ii. *Your sister would have been living now* (Emily Bronte, Wuthering Heights)

iii. *We stand together again at the steps of this symbol of our democracy – or we would have been standing at the steps if it hadn't gotten so cold* (Ronald Reagan's inaugural address)

Having realized how far-fetched these examples are, Pullum and Scholz say that they could refer to many examples from classic children's books and, therefore, meet the inaccessibility challenge (2002: p. 29).

Although P&S's examples of speech data are already self-discrediting, I believe a few issues need to be pointed out. In my view, to count written work as the Primary Linguistic Data— as suggested by Pullum and Scholz —is to misunderstand the whole concept of language acquisition. As far as I am concerned, an instance of an Aux-sequence in a book proves nothing since:

a) The data in question are available to a limited number of people (i.e., those who do have access to these specific, or indeed, any books).
b) The data are limited.
c) All acquirers are exposed to different sets of data yet they learn the language of the environment in a similar manner and are able to communicate with one another in the given language.

The second point Pullum and Scholz mention with regard to Aux-sequences is that "there are significant consequences for learnability. If auxiliaries are complement-taking verbs there is no need to assume that strings like *may have been writing* must be heard before their grammaticality can be known [...] all of this can be learned from examples containing one item acting as head of the complement of another." (2002: p. 30-31).

Hence, even Pullum and Scholz are admitting that a structure such as *is able to avoid appearing to be drunk* is acquired through understanding the relationship between its syntactic components (= parsing). However, they fail to explain their view on where this knowledge stems from in a 3-year-old. That is, the only way a child can understand the above structure is through parsing it into its underlying structural components. Obviously, parsing can occur if— and only if —the acquirer is cognitively equipped to

do the parsing. There is no evidence in Pullum and Scholz's claims that this language-specific parsing can be done using general cognitive strategies. Nor do they provide any empirical evidence that notions such as 'head' or 'complement' can arise from any such strategies.

Pullum and Scholz's third opposing argument against Aux-sequence acquisition as an instance of POS is based on an analogy on how humans learn facts through their general learning abilities and challenging the necessity of domain-specific innate abilities in relation to language acquisition.

Figure 3.9 Pullum &Scholz 2002: p. 31:

> "To posit learning from experience for facts of complement type selection is no more than would be assumed for learning any other structure of parts recombining to make particular types of wholes. We learn – at a very young age – that houses contain rooms, and rooms are different types such as kitchens and bedrooms [...] Our experience does not establish for us definitively that dressers with faucets are impossible [...] yet we come to believe these things too [...] and in a way that is dependent on our environment (we could not have learned the same generalizations growing up in a rural African village), but without negative evidence. It would be rather radical to claim that learning facts of this sort from experience is impossible, so that there must be innate domain-specific knowledge about architecture, furnishings, and appliances."

In my view, there are many flaws in this type of reasoning, especially when it comes to language acquisition:

1) The knowledge of architecture might not necessarily be internal or domain specific (either way it is outside the scope of this thesis). Nevertheless, when we point to a room and tell an acquirer "this is a room" and the acquirer does not conclude that only the part of the room we were pointing at is called 'a room', is internal.
2) Language is not the mere concatenation of its smaller parts. Language acquisition is about how a series of sound waves (the Available Linguistic Data – ALD) are perceived and parsed into meaningful structural and syntactic units by the mind of the acquirer.
3) Speakers do not see language as a simple collection of smaller pieces but as a relational network with underlying structures. That is, the Acquirer's manipulation of language is structure-dependent (structure-dependence was briefly discussed in Chapter 2 and will be further discussed in the following section).
4) The fact that we have knowledge of, for example, the grammaticality of infinitely long sentences or that we know the speed or the pitch with which sentences are uttered has no grammatical relevance shows that we must be born with some pre-wired understanding of the concept of language.
5) Language is a communication system that is live, active and generative. A house, on the other hand, is an immobile, static and non-generative entity. This

significant difference in nature makes it impossible to compare our perception of a house to the way our mind processes and produces language.

3.2.3.2. Auxiliary-initial clauses/Auxiliary Inversion:

Pullum and Scholz state that, "the apparently strongest case of alleged learning from crucially inadequate evidence [for the POS Argument; pz] discussed in the literature, and certainly the most celebrated, concerns auxiliary-initial positioning in polar interrogatives [i.e., yes/no questions; pz] in languages like English and Spanish" (2002: p.36).

This notion was discussed briefly in Chapter 2 (see above) and is normally used by advocates of the Innateness Theory as evidence for structure-dependence. Consider the following examples:

a. The dog in the corner is hungry.

b. Is the dog in the corner hungry?

Compare to:

c. The dog that is in the corner is hungry.

d. Is the dog that is in the corner hungry?

e. *Is the dog that in the corner is hungry?

The idea here is that the "generalization concerning the formation of such sentence types is structure-dependent: it is based on structural relations (dominance among constituents), not just temporal sequence (precedence among words)" (P&S 2002: p.36). Since this case is purely syntactic and children are not instructed on such structural relations but still

manage to acquire them,[31] the Aux-inversion phenomenon could be considered the most crucial evidence in favor of the Poverty of the Stimulus Argument.

In a debate with Jean Piaget, Chomsky has mentioned that "a person might go through much or all his life without ever having been exposed to relevant evidence, but he will nevertheless unerringly employ [the relevant structure-dependent rule; pz], on the first relevant occasion" (Chomsky, in Piattelli-Palmarini 1980: p.40).

Pullum and Scholz focus on the inaccessibility criterion mentioned by Chomsky in Piattelli-Palmarini (1980). Specifically the claim that "the child is deprived of crucial positive evidence that supports the main-clause-auxiliary generalization over the incorrect [but much simpler; pz] first-auxiliary generalization" (2002: p.39).

Their first claim is that if people so rarely produce or come across such structures, "there could be speakers around who have acquired an 'incorrect' structure-independent generalization instead, but who are never detected because of the rarity of the crucial situations in which they would give themselves away" (P&S 2002: p.40).

In my opinion, Chomsky is merely stating that since this structure is rare children have little or no evidence to support the structure-dependent hypothesis and, yet, they learn to use it correctly despite the impoverished data. Moreover, I do not believe Pullum and Scholz would want to open Pandora's Box by claiming that there are actual 'native' speakers who are walking around with the 'wrong' underlying grammar.[32]

---

[31] See Crain and Nakayama (1987) for experimental confirmation that children *do* correctly learn Aux-Inversion.

[32] The notion of 'wrong' here refers to 'UG-illicit' and not just say a non-standard dialect. Pullum and Scholz are actually referring to people with types of grammar that are prohibited by UG. I would like to thank Mark Hale (Concordia) for pointing this out.

Moreover, since Pullum and Scholz mainly argue based on what they deem to be empirical evidence, I strongly suggest that they show at least one such case (i.e., a native speaker with the 'wrong' underlying representation of the grammar of the environment) in order to validate such reasoning against Nativist claims.

Pullum and Scholz's second argument is based on the empirical issue. They mention Robert Freidin's claims that the POS instance based on Aux-initial clauses "is based on the empirical assumption that children encounter (or pay attention to) simple sentences prior to those with subjects containing finite relative clauses [...] Whether this is actually the case has not been established" (1991: p.618).

Pullum and Scholz then start their usual (and scientifically dubious) style of empirical evidence presentation. For instance they claim Sampson (1989) suggests that "it is probably not true that children are deprived of access to questions with auxiliary-containing subjects [...] during the acquisition process" (P&S 2002: p.40).

Even though it seems that Pullum and Scholz might have found a flaw in the POS Arguments via Sampson's discoveries, it turns out that these so-called examples, as always, include sentences like the following:

Figure 3.10 Pullum and Scholz 2002: p. 40

> "He [Sampson; pz] takes William Blake's well-known poem '*Tiger*' to be typical of language encountered by almost every English Child during their early years, and notes that the line *Did He who made the lamb make thee?* offers crucial positive evidence for the structure-dependent rule [...]

He finds another relevant example by scanning a list of questions in an encyclopedia for young children."

Once again, I believe that these examples are questionable in and of themselves. But for the sake of argument, I would like to ask the following questions:

1) Do scientists like Sampson, or Pullum and Scholz really believe that every single English-speaking child has been exposed to the '*Tiger*' poem?
2) How would they explain the proper acquisition of the structure-dependent rule in those children who have never been exposed to said poem?
3) How many 0-5-year olds do we know who use '*the encyclopedia for young children*'?

Unfortunately, examples like this are plentiful in what Pullum and Scholz see as their star evidence against the Poverty of the Stimulus Theory. These examples range from the *'Tiger'* poem to the script of *The Importance of being Earnest* (P&S 2002: p.43) to the seemingly more scientific results of empirical research done by Betty Hart and Todd Risley (1995: p.132):

Figure 3.11 Pullum & Scholz 2002: p. 44-45

"One rather startling result of Hart and Risley's concerns the magnitude of the difference in language experience between classes: it is calculated that a child in a working-class family will have heard 20 million word tokens by the age of 3, and a child being raised in a family on welfare will have heard only 10 million."

In my opinion, the statistics accumulated by Hart and Risley work in favor of the Poverty of the Stimulus Argument. It is evident that all these different classes (the Working-Class Child and the Welfare Child and even the Wealthy Upper-Class Child) *will* perceive the Grammar of the Environment (G1) via the Available Linguistic Data (the ALD) and *will* cognitively construct their version of the Target Grammar (G2).

A big percentage of language-related differences between these classes are related to the idioms, expressions and the type of vocabulary and are, therefore, lexical and not syntactic. That is, the lexicon of the different classes might be different but not the underlying structural representations of the syntactic relations. Though, even if there are syntactic and structural differences, the key is not that they can be different, but that they can be the same (i.e., all these children are equipped to acquire— and will acquire —the language of the environment).

After having looked at Pullum and Scholz's arguments against two of the exemplars of the Poverty of the Stimulus Argument (Auxiliary Sequences and Auxiliary-Initial Sentences/Auxiliary-Inversion), I conclude that this type of reasoning does not create a scientifically sound challenge against the Argument from the POS. In my opinion, Pullum and Scholz have done nothing but trap themselves in data analysis and empirical statistics. In their quest for a better understanding of the matter at hand, they have provided nothing but evidence of language usage— mostly in the form of written words —in isolated environments.

Hence, it is my belief that the Poverty of the Stimulus Theory has certainly not been challenged by Pullum and Scholz and still remains as one of the most important examples in support of the Innateness theory of language acquisition.[33]

[33] For scientific accounts against Pullum and Scholz's views in the case of Auxiliary-Inversion see also Lasnik & Uriagereka (2002: p.147-150) and Legate & Yang (2002: p.151-162).

CHAPTER 4

CONCLUSION

"Intelligence requires construction and manipulation of symbolic representations. Interacting intelligently with the world requires the ability to *parse* input (assign it a representation). Learning is a form of intelligent interaction with the world, thus learning requires parsing into representations. Without an innate set of representational primitives, learning cannot begin" (Hale&Reiss 1998: p. 1).

This thesis was based on the Poverty of the Stimulus Theory and the challenges that were recently brought against this Theory by Geoffrey Pullum and Barbara Scholz (2002: p.9-51). The thesis has discussed two main points with regard to the theory.

First, it was pointed out that cognitive sciences, including the science of language acquisition, are undergoing major transitions and researchers were warned against irresponsible borrowing and misuse of idioms and scientific expressions.

Second (and in keeping with the first proposition), a) the nature of the Primary Linguistic Data (PLD), and, b) the Argument from the Poverty of the Stimulus (POS) in relation to First Language Acquisition (FLA) were explored. It was proposed that the definition of the nature and the location of the PLD need to be modified. Moreover, Pullum and Scholz's challenges against the POS Argument were examined for their validity. It was concluded that these allegations did not stand the test of validity and that the Poverty of the Stimulus Argument still remains unchallenged.

The focus of Chapter 1 was on two things. First, the fact that Linguists and Cognitive Scientists need to be careful when using or borrowing scientific expressions

from other sciences was discussed. At this point, I think it appropriate to give, yet, another example. In viewing Sampson's (1989, 1999) arguments against Hornstein and Lightfoot's claims,[34] Pullum and Scholz get tangled up in the following (seemingly plausible) discussion.

Sampson states that Hornstein and Lightfoot's claims regarding the 'logical problem of language acquisition' are self-contradictory. His argument is as follows. If a Linguist (*X*) claims that a Grammatical Fact (*F*) about a Language (*L*) has been learned by the speaker (*S*) the question is: How does the linguist know that *F* is a fact about *L*? Sampson claims that there are two ways:

a) *X* knows that *S* has learned *F* through evidence, which means there is positive evidence.

b) *X* knows *L* natively and claims that *S* has learned *F* just like *X* during the course of FLA.

Sampson then concludes— and Pullum and Scholz agree — that the argument set forth by Hornstein and Lightfoot is viciously circular and since "there are no other cases to consider, the argument refutes itself" (P&S 2002: p.15).

At the first glance, this form of reasoning may sound scientifically and logically sound. However, Sampson, Pullum and Scholz all forget one simple point:

There are always *F*'s that *S* is not exposed to but manages to produce.

[34] See Chapter 3, Section 3.2.1.2. for a brief discussion of 'the logical problem of language acquisition'.

Therefore, as logical as Sampson's argument (or P&S's interpretation of this argument) seems to be, it does not refute the simple fact that the linguistic competence of every single individual is indefinitely larger than the input they receive and perceive from the environment.

Chapter 2 dealt with the historical background and the evolution of the Poverty of the Stimulus concept. The history was recounted in a narrative style similar to that of Margaret Thomas (2002: p. 51-72).

The evolution of the Poverty of the Stimulus was the main point of this chapter. The fact that this theory is relatively young and has come to life only in the past 15-20 years was discussed. It was also pointed out that this theory has constantly evolved to taking center stage as one of the most celebrated arguments that enforce the validity of the Innateness Theory.

Pullum and Scholz were mentioned as the scientists who have recently questioned the legitimacy of the POS Argument. It was pointed out that these allegations should not go unnoticed. Moreover, it was mentioned that if Pullum and Scholz's challenges turn out to be legitimate, the whole idea of the Argument from the Poverty of the Stimulus would need to be reexamined and modified accordingly.

Chapter 3 was dedicated to two topics: a) the Primary Linguistic Data (PLD) which was discussed as the external fuel for First Language Acquisition (FLA), and, b) the Poverty of the Stimulus Theory and its impact on the nativist view of FLA, as well as a study of the challenges brought upon the POS Argument.

The Primary Linguistic Data was examined in the first part of Chapter 3. The main point of this examination was to determine the true location and nature of the PLD.

It was concluded that the definition of the PLD (i.e., all the linguistic data in the environment of the acquirer) would have to be modified. It was suggested that the linguistic data should be regarded in a tri-level fashion: The Available Linguistic Data (ALD), the Received Linguistic Data (RLD), and the Perceived Linguistic Data (PLD).

Four logical conclusions were drawn based on this tri-leveled analysis of the Linguistic Data:

1) The acquirer cannot be born with a clean slate. In other words, the *tabula* cannot be *rasa* (Hale & Reiss 2003: p. 219).
2) The acquirer does not build the target grammar on all existing data in the environment.
3) The human brain must be equipped with a Linguistic Faculty/Module and there must be some innate constraints that allow the mind to recreate the target grammar.
4) The acquirer must be open to all linguistic possibilities at the time of birth. Otherwise, there would have to be examples of normal children incapable of learning the language of the environment.

The second part of Chapter 3 was dedicated to the examination and exploration of the Argument from the Poverty of the Stimulus. The importance of the POS Argument and, therefore, the necessity to address the challenges brought upon the POS Argument were pointed out. It was argued that Pullum and Scholz (2002) have not been successful in challenging the POS Theory.

At this point it is appropriate to give another example of how other researchers view the Pullum and Scholz article. The following quote will illustrate:

Figure 4.1 Fodor & Crowther 2002: p. 105

> "The argument from the poverty of the stimulus as Pullum and Scholz define it (their APS) is undeniably true; given that all language learners acquire the ability to generate more sentences of the target language than they have heard. Uniformity across learners with respect to the additional sentences they project suggests that grammar induction is guided by general principles, which must be innate."

To the points made by Fodor and Crowther I would also like to add that natural languages seem to follow certain rules. That is to say, certain structures do not occur in any language. The sounds of the world, the phonology of all languages as well as their syntax seem to follow certain rules (e.g. see the structure-dependence theory, discussed briefly in Chapter 2 and 3; see also the discussion on UG parameters: Headedness and Pro-drop in Chapter 3).

With regard to the POS Argument, three related issues (including the Argument itself) disputed by Pullum and Scholz were discussed: 1) Innateness, 2) Negative Evidence, and, 3) The POS Argument.

Innateness: P&S's arguments against Wexler (1991) and Hornstein & Lightfoot (1981) were examined. In both cases, it was asserted that Pullum and Scholz seem to have misinterpreted the claims made by Wexler and Hornstein & Lightfoot.

Negative Evidence: The fact that many researchers misapprehend the notion of Negative Evidence was briefly discussed. The conclusion was that negative evidence does not appear to be a prerequisite for First Language Acquisition.

The Poverty of the Stimulus Argument: The POS concept was discussed as one of the principal theories that support the Innateness Hypothesis. Pullum and Scholz have claimed that they have undermined this theory by weakening the four chief exemplars of the POS phenomenon widely used by the advocates of the Innateness Hypothesis.

Pullum and Scholz's arguments against two of these exemplars were studied: a) Auxiliary Sequences, and, b) Auxiliary-Inversion/Auxiliary Initial Sentences. A careful examination revealed that in both cases, Pullum and Scholz have resorted to the same type of reasoning. Their emphasis is on empirical evidence. However, it was argued that their so-called empirical evidence does not mesh well with a phenomenon like language acquisition.

CHAPTER 5

SUGGESTIONS FOR FURTHER RESEARCH

In my opinion, Second Language Acquisition[35] (SLA) is similar to First Language Acquisition in many respects:

1) It is processed and acquired via direct access to Universal Grammar.
2) Linguistic Competence: The cognitive processes involved in creating a representation (G2) of the target grammar (G1) in the mind of the Acquirer.
3) Linguistic Performance: The ability to build novel sentences based on Primary Linguistic Data (PLD) and the UG constraints directly available to the Acquirer.

There is a whole spectrum of speculations among Linguists and Cognitive Scientists with respect to Second Language Acquisition (SLA). In my opinion, research (in its scientific sense[36]on SLA is at its starting stages.

Scientists have only recently started to fully appreciate the importance of determining the source and developmental process of a second language in the mind of non-children (= adolescents and adults). Similarities between First Language (L1) and Second Language (L2) acquisition are becoming more and more apparent. Researchers now agree that answers to questions like, "how an adult acquires a new language", "how

[35] In this thesis, L2 acquisition refers to adolescents and adults since children, if exposed to a second language early in life, are universally said to be capable of achieving native-like abilities in that language.

[36] There is an extensive body of research on SLA within the realm of Applied Linguistics and in relation to second language teaching and learning. However, the interest in understanding the cognitive processes of SLA is fairly new.

native-like this acquisition will be", or "how UG-related this acquisition process is" will have scientific significance in the overall study of Language Acquisition.

As mentioned before, I strongly believe that many similarities exist between L1 and L2 acquisition. My view of the Language Acquisition phenomenon (both L1 and L2) is based on the belief that under normal conditions (i.e., in the absence of conditions like aphasia, language and/or genetic disorders) full-access to UG is available to all humans through the course of their life. That is, humans are biologically wired for learning languages and born with a set of constraints implemented by Universal Grammar (UG) and a set of parameters (binary ways that the surface grammar could be set; e.g. the pro-drop parameter; i.e., whether a language allows covert subjects or not; see Chapter 3 for discussion).

The overall point is that those parts of UG that are not used in a given language (e.g., English does not allow covert subjects so this aspect of UG is not used in acquiring English) are not lost to humans. The idea is that:

A novel set of available data (= a new language) will trigger the Language Acquisition Device (LAD) in the same manner as the L1.[37]

The general point is that if an adult Second Language Acquirer (L2'er) is capable of understanding the language of the environment and is able to create novel sentences

[37] The Subset Principle (SP) offers a different view on this subject. It postulates that we are born with a highly restrictive set of primitives and that these restrictions are relaxed based on positive evidence through the course of acquisition. But one could adopt a version of the subset principle that says that this relaxation is Grammar-specific, rather than a property of the individual acquirer. I would like to thank Mark Hale (Concordia) for pointing this out.

based on exposure to the limited and impoverished available linguistic data, then direct access to the LAD and the UG is available during SLA.

In other words, the ability to communicate abstract thought through a new language which one is exposed to during adolescence or adulthood is evidence of direct access to Universal Grammar (UG) and another example of the Poverty of the Stimulus concept.

Further research might reveal that most L2 errors are caused by extra-linguistic factors. Here are a list of what I think could count as extra-linguistic or, at least, extra-syntactic:

a) Affective factors: e.g. being too shy or intimidated and hence not expressing one-self as well as one could.

b) The first/native language (L1) interfering with L2 during the early stages of acquisition acting as a hindrance to the SLA process.[38]

c) The existence of an impoverished lexical database; i.e., weak vocabulary.

When it comes to SLA, I believe the biggest difference between children and adults is that adults come equipped with a first language. Nevertheless, if direct access to UG is accepted as the foundation of L2 acquisition, the first language interference should not have much effect on L2 acquisition in the long run. Taking this fact into consideration, most adult speech errors would be judged as performance errors and an impoverished lexical inventory and not necessarily as lack of syntactic or structural competence.

[38] The nature of this 'interference' is the centerpiece of most debates on SLA.

Moreover, by assuming full-access to UG during the course of SLA, we are positing that the source or foundation of the L2 Acquirer's lexical representation, phonological rules and output forms must be their language faculty and, hence, similar to that of the L1 acquirer.

Of course, all languages seem different on the surface, but they are still restricted to the innate primitives, which are a part of the genetic foundation of *Homo sapiens*. In other words, there are no un-learnable languages in the world since they are all restricted to what our biological nature allows. Therefore, even though lexical representations and phonological and structural rules of each language might seem different on the surface, they are the same since they abide by phonological and structural universals and constraints and are, therefore, acquirable.

I believe that output forms constitute the biggest difference between L1 and L2 acquirers. As we know, under normal circumstances, all children will achieve native ability in the language of the environment roughly by the age of five (5). SL Acquisition, however, is far from uniform across the board. After all, it seems like most L2'ers never achieve native or even near-native abilities especially when it comes to output forms (= 'Accent' in lay terms).

Looking at different communities with the same L1 who learn English as a second language, for example, seems to confirm that there is a lot of L1 influence on L2 when it comes to accents. For instance, it is easy to determine whether the second language learner's first language is Spanish, Arabic, Farsi, Chinese, etc. That is why I believe that the ability to attain mastery in output forms in a given language is the most

significant distinguishing factor between child (L1) and adult (L2) acquisition. Nevertheless, a few points are worth considering in this matter:

a) Given the opportunity, L2'ers can distinguish the difference between their own and the 'standard' output forms. For example, L2'ers would know if someone is imitating them. The question is: Do these second language acquirers have the correct underlying representations? Could their inability to produce the correct output form be due to performance errors and not competence in their knowledge of the L2?[39]

b) Extra-linguistic factors should be taken into consideration. For example, the vocal tract may not be used to making the sound(s) required in the new language: e.g., Farsi does not have the voiceless dental fricative 'theta' [θ] like in 'both' or voiced dental fricative 'edh' [ð] like in 'there'. I can personally vouch that Iranians have a lot of difficulty producing these sounds when learning English. But I can also testify that the production of these sounds can be achieved through practice. In my opinion, this indicates that the innate ability to distinguish and produce these phonetic sounds is not lost to Iranians; they are merely out of practice.

As mentioned above, other extra-linguistic factors could be shyness, over-involvement with the L1 speech community which lengthens the process of L2

[39] This characteristic is very similar to that of children who make output errors, but correctly recognize the same output forms as *errors* when made by adults.

acquisition, and also enforces similar performance errors, and last but not least, self-doubt.[40]

I strongly believe that cognitive science research on second language acquisition is of utmost importance since it will provide:

a) A better understanding of how the human mind works when acquiring a language other than their native tongue,

b) A different angle in the analysis of scientific findings related to language acquisition, and most important of all,

c) The much needed empirical evidence arising from the collection of L2 language data and the scientific analysis of this data.

Simply put, by focusing on SLA research, cognitive scientists will be able to close the gap between philosophical and/or scientific speculation and hard-core empirical evidence to back up these speculations. In other words, cognitive scientists and especially linguists should focus on collecting and analyzing adult L2 data which will help them revise old-fashioned ideas and speculations with regards to SLA and will refine the overall understanding of the processes involved in language acquisition.

What follows are some suggestions for further research and some ideas that need to be taken into consideration when researching SLA.

5.1. Comparative study of similarities in L2 errors by speakers with a common L1

As mentioned above, there are common characteristics could be observed in the speakers who share the same L1 (say Farsi) when acquiring the same L2 (say English). The study of these common characteristics could open a window to the way the mind

[40] Unfortunately, it is commonly believed that L2'ers can *never* learn L2 as well as their L1. This belief, in

works while processing a second language by revealing, for example, where they originate in divergent competence (= knowledge) or performance effects of the type outlined above. This, in turn, could shed light on the source and nature of language acquisition in general.

For instance, statistical analysis of L1-initiated performance errors could give us an idea of the extent of L1 interference in L2 acquisition. It might also show us whether (and to what level) this influence tapers off as the acquirer gains better mastery of the second language.

By doing this type of comparative study, we might be able to understand how and to what extent this common background is influencing the acquirer's connection with their Language Acquisition Device (LAD). Subsequently, this could lead us to a better understanding of the way LAD works and the way Universal Grammar is employed while learning a 2^{nd}, 3^{rd}, n^{th}, language.

Most importantly, by studying the influence of a common L1 on L2 acquisition, we might be able to find ways to bypass these influences and facilitate the often cumbersome task of SLA for the Acquirer.[41]

5.2. Extra-linguistic influences (the Need/Necessity Hypothesis)

Unfortunately, I believe that scientists take extra-linguistic factors (see above) that effect SLA too lightly. I even go as far as suggesting that a number of researchers have never questioned the (in my view) outdated belief that reaching native-like ability is impossible in SLA.

---

my opinion, could be a major hindrance for adult L2'ers which can stop them from making the attempt.

[41] This kind of research could have great influences on second language teaching.

If we give weight to the depth of the influence that extra-linguistic factors have on SLA, our whole approach to the matter will have to be readjusted. As an example, let us contemplate SLA from a 'need/necessity' point of view. Consider the following speculation:

It is in the nature of humans to want to communicate abstract thought.[42] Once an individual acquires a first language, this need is fulfilled and the acquisition of a second language becomes a very low priority for the brain.

This idea might not sound very scientific at first, but if we posit that UG is available throughout our life span, then the acquisition of any language (2^{nd}, 3^{rd}, n^{th}...) should be as automatic and uniform as that of a first language. Unfortunately, L2 evidence seems to show otherwise. For example, there are individuals who live in a foreign country for years and still perform poorly when it comes to production of the language of the environment.

Nevertheless, I would like to argue that extra-linguistic factors might be the cause of this apparent L1/L2 discrepancy. My question is: what if the sole reason for language acquisition is to fulfill this 'need/necessity' to communicate? And, what if once this need

[42] This need for communication also includes "psychic integration" (=the ability to talk to oneself) as well as communicating with other fellow humans. I would like to thank Mark Hale (Concordia) and Annette Teffeteller (Concordia) for pointing this out.

is fulfilled, going through the whole process again is not the highest priority for the brain?

If we consider this hypothesis, we realize that L2 Acquisition will not occur (or will occur only haphazardly; e.g. speaking enough of the 2nd language to be able to 'get by') unless there is an absolute 'need' for it. If we view the world in this manner, we will see that in today's world, second language acquirers are usually a part of a community who speak their first language. Therefore, they either never learn the language of the environment (like so many older parents of immigrants who never or hardly ever communicate with the society around them), or learn enough of it to be able to communicate their so-called 'needs' to the world around them.

To emphasize the importance of this matter, here is an imaginary experiment that I think would clarify my point:

Consider a random individual (John Doe!) who speaks a native language *Lx* (the name or type of the language is of no importance). John Doe is a young adult, with a University Degree. He is an ambitious individual who has finished at the top of his class, and is now ready and eager to climb the social ladder!

Imagine that John Doe is informed that he can have his dream job only if he is willing to move to another country. This country shares one common language *Ly* which is completely foreign to John Doe. Moreover, no one in this country speaks *Lx* and can only communicate in their native *Ly*. John Doe is told that the only condition for him to get this job is to become a fluent speaker of *Ly*. He is also assured that his financial and personal needs will be provided while he is in the process of acquiring *Ly*.

John Doe is ecstatic and fully prepared to do all that it takes to get this job. In order to get a full exposure to *Ly*, he is transferred to live with a family in the new country. To summarize, one minute John Doe was in familiar surroundings where he communicated with every one with ease via his native *Lx*, and the next minute, he finds himself in a foreign setting where he does not understand the spoken language. Nevertheless, John Doe is fully motivated to get on with the one and only task at hand; i.e., fluently acquiring the language of the environment.

Given these conditions, I truly believe that John Doe *will* learn to speak *Ly*, and he will achieve native or near-native ability in *Ly*. As far as I am concerned, this imaginary example puts an emphasis on the importance of the Need/Necessity Hypothesis in L2 acquisition. I would also like to draw attention to the similarities between John Doe's situation and that of a newborn child. I assert that like all children born into normal surroundings, if every adult was given the sole task of learning the ways of communication with the immediate environment, the outcome of L2 acquisition would be very different from what it is in reality.

Another good example of the Need/Necessity Hypothesis is the well-known stereotype that English speakers are the worst L2 acquirers. If this happens to be true, we have two ways of looking at this apparent dilemma:

a) speakers of English are biologically different from the rest of the world, hence, their version of UG does not allow them to acquire a 2nd language, or,

b) English speakers hardly have a need for a 2nd language since they can communicate in English almost everywhere in the world.

Basically, it is my belief that the time has come for all involved to open their minds to new ideas about the SLA phenomenon. Linguists might want to pay a lot more attention to extra-linguistic matters like the Need/Necessity Hypothesis.

5.3. The case of the forgotten high-end!

In the study of second language acquisition (similar to other areas of research) scientists rely heavily on statistical data. It is safe to say that in every L2 related linguistic experiment, a percentage of L2-speaking subjects perform with native or near-native ability on all linguistic tasks. Unfortunately, based on the rules of statistical and empirical data analysis, the high- and the low-end of the results are always omitted in order to calculate a more accurate medium.

In my view, this is the worst way of statistical analysis when it comes to linguistics and more specifically SLA research. I strongly suggest that the high-end individuals should be the focal point of SLA research. Scientists must ask themselves: Why did these particular L2'ers perform correctly? What is their background? How fluently do they speak the L2 in real time (performance)?

I postulate that the results of studies done on these individuals could change the way researchers look at SLA or at least provide us with very different results when it comes to L2 acquisition.

5.4. CONCLUSION

Research on Second Language Acquisition (SLA) is only in its preliminary stages. There are many gaps in the related research and disagreements across the board.

More importantly, many myths still exist regarding learning a 2nd language. Old-fashioned beliefs that could (and for the most parts have) severely biased even the best

researchers still are commonly used. For example, some scientists still firmly believe that there is no-access to UG during the course of SLA; others believe there is partial-access. Even the advocates of the full-access theory think that second language acquirers can never achieve native-like mastery in the second language.[43]

It is my belief that the first and most important duty of L2 researchers is to establish whether there is direct-access to UG during the course of SLA. Authenticating the source and the nature of SLA will serve as a crucial stepping stone leading to a better comprehension of this phenomenon which in turn will move us forward in our quest to fully understand the workings of the human mind.

[43] For an in-depth discussion of UG-access during the course of L2 see Epstein et al (1996).

BIBLIOGRAPHY

Baker, C.L. (1979). Syntactic theory and the projection problem. *Linguistic Inquiry* 10: 233-280.

Chomsky, Noam (1957). *Syntactic Structures*. The Hague: Mouton.

— (1959) A review of B. F. Skinner's *Verbal behavior*. *Language* 35, 26-58.

— (1962). Explanatory models in linguistics. In *Logic, Methodology and Philosopy of science*, Ernest Nagel, Patrick Suppers, and Alfred Tarski (eds.), 528-550. Stanford, CA: Stanford University Press.

— (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

— (1975a). *The Logical Structure of Linguistic Theory*. Chicago: University of Chicago Press.

— (1975b). *Reflections on Languages*. New York: Pantheon Books.

— (1980). *Rules and Representations*. New York: Columbia University Press.

Crain, Stephen, and Mineharu Nakayama (1987). Structure dependence in grammar formation. *Language* 63: 522-542.

Epstein, Samuel D., Suzanne Flynn, and Gita Martohardjono (1996). Second language acquisition: Theoretical and experimental issues in contemporary research. BBS.

Fodor, Janet D., and Carrie Crowther (2002). Understanding stimulus poverty arguments. *The Linguistic Review* 19: 105-145.

Freidin, Robert (1991). Linguistic theory and language acquisition: a note on structure-dependence. *Behavioral and Brain Sciences* 14: 618-619.

Gleitman, Lila R., and Mark Liberman (1995). The cognitive science of language: Introduction. In *Language: An Invitation to Cognitive Science*, Lila R. Gleitman, and Mark Liberman (eds.), xix-xxxviii. Cambridge: The MIT Press.

Gleitman, Lila R., and Elissa L. Newport (1995). The invention of language by children: Environmental and biological influences on the acquisition of language. In *Language: An Invitation to Cognitive Science*, Lila R. Gleitman, and Mark Liberman (eds.), 1-24. Cambridge: The MIT Press.

Hale, M.R. (1998). Diachronic syntax. In *Syntax* 1, 1-18. Blackwell: Oxford.

Hale, M.R., and Charles Reiss (2003). The subset principle in phonology: Why the tabula can't be rasa. In *Journal of Linguistics* 39: 219-244.

Hart, Betty, and Todd R. Risley (1995). *Meaningful Differences in the Everyday Experiences of Young Children.* Baltimore: Paul H. Brookes.

Hornstein, Norbert, and David Lightfoot (1981). Introduction. In *Explanation in Linguistics: The Logical Problem of Language Acquisition*, 9-31. London: Longman.

Hosting Works Dictionary. Online, Internet, March 06/2004. Source: http://hostingworks.com/support/dict.phtml?jargon=parse

Kimbal, John P. (1973). *The Formal theory of Grammar.* Englewood Cliffs, NJ: Prentice-Hall

Lasnik, Howard, and Juan Uriagereka (2002). On the poverty of the challenge. *The Linguistic Review* 19: 147-150.

Legate, Julie, and Charles Yang (2002). Empirical re-assessment of stimulus poverty arguments. *The Linguistic Review* 19: 151-162.

Lewis, John D., and Jeffrey L. Elman. *A Connectionist Investigation of Linguistic Arguments from the Poverty of the Stimulus: Learning the Unlearnable.* Online, Internet, March 07/2003. Source: crl.ucsd.edu/~elman/Courses/cog202/PS/lewis-elman-cogsci2001.pdf.

Liberman, Mark (1995). The sound structure of Mawu words: A case study in the cognitive science of speech. In *Language: An Invitation to Cognitive Science*, Lila R. Gleitman, and Mark Liberman (eds.), 55-85. Cambridge: The MIT Press.

Lightfoot, David (1982). Review of Geoffrey Sampson, *Making Sense. Journal of Linguistics* 18: 426-431.

Peters, Stanely (1972). The projection problem: How is a grammar to be selected? In *Goals of Linguistic Theory*, Stanley Peters (ed.), 171-188. Enlewood Cliffs, NJ: Prentice Hall.

Piattelli-Palmarini, Massimo (ed.) (1980). *Language and Learning: The Debate Between Jean Piaget and Noam Chomksy.* Cambridge: Harvard University Press.

Pinker, Steven (1995a). Why the child holded the baby rabbits: A case study in language acquisition. In *Language: An Invitation to Cognitive Science*, Lila R. Gleitman, and Mark Liberman (eds.), 107-133. Cambridge: The MIT Press.

— (1995b). Language acquisition. In *Language: An Invitation to Cognitive Science*, Lila R. Gleitman, and Mark Liberman (eds.), 135-182.Cambridge:The MIT Press.

— (1996). *Language Learnability and Language Development*. Cambridge: Harvard University Press.dff

Pullum, Geoffrey, and Barbara Scholz (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review* 19: 9-50.

Radford, Andrew (1997). *Syntax: A Minimalist Introduction*. New York: Cambridge University Press.

Sampson, Geoffrey (1989). Language acquisition: growth or learning? *Philosophical Papers* 18: 203-240.

— (1999). *Educating Eve*. London: Cassell.

Scholz, Barbara, and Geoffrey Pullum (2002). Searching for arguments to support linguistic nativism. *The Linguistic Review* 19: 185-223.

Schwartz, Bonnie D., and Rex A. Sprouse (2000). When syntactic theories evolve: Consequences for L2 acquisition theory. In *Second Language Acquisition and Linguistic Theory*, John Archibald (ed.), 156-186. Oxford: Blackwell.

Thomas, Margaret (2002). Development of the concept of "the poverty of the stimulus. *The Linguistic Review* 19: 51-71.

Werker, Janet F. (1995). Exploring developmental changes in cross-language speech perception. In *Language: An Invitation to Cognitive Science*, Lila R. Gleitman, and Mark Liberman (eds.), 87-106. Cambridge: The MIT Press.

Wexler, Kenneth (1991). On the argument from the poverty of the stimulus. In *The Chomskyan Turn*, Asa Kasher (ed.), 252-270. Oxford: Basil Blackwell.