TOWARD MORE SUBSTANTIAL THEORIES OF LANGUAGE ACQUISITION

by

Cinnamon Ann Jenson

A dissertation submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Philosophy

The University of Utah

December 2015

The University of Utah Graduate School

STATEMENT OF DISSERTATION APPROVAL

The dissertation of Cinnamon Ann Jenson

has been approved by the following supervisory committee members:

Patricia Hanna	, Chair	7/29/2015 Date Approved
Steven Downes	, Member	7/29/2015 Date Approved
James Tabery	, Member	7/29/2015 Date Approved
Dustin Stokes	, Member	7/29/2015 Date Approved
Edward Rubin	, Member	7/29/2015 Date Approved

and by Matthew Haber, Chair/Dean of the Department/College/School of Philosophy

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

Cognitive linguists argue that certain sets of knowledge of language are innate. However, critics have argued that the theoretical concept of "innateness" should be eliminated since it is ambiguous and insubstantial. In response, I aim to strengthen theories of language acquisition and identify ways to make them more substantial. I review the Poverty of Stimulus argument and separate it into four nonequivalent arguments: Deficiency of Stimulus, Corruption of Stimulus, Variety of Stimulus, and Poverty of Negative Evidence. Each argument uses a disparate set of empirical observations to support different conclusions about the traits that are claimed to be innate. Separating the Poverty of Stimulus arguments will aid in making each one more effective.

I offer three sets of considerations that scholars can use to strengthen linguistic theories. The Empirical Consideration urges scholars to address specific sets of empirical observations, thus ensuring that innateness theories are not used to explain dissimilar traits. The Developmental Consideration urges scholars to consider complex developmental processes of acquisition. The Interaction Consideration urges scholars to examine interactions between organisms and their environment during language acquisition. I support recent contributions to the approach of "biologicizing the mind" which encourages interdisciplinary collaboration between psychology and biology. I develop an account of language acquisition in terms of canalization, and use this account

to explain empirical observations used in Variety of Stimulus arguments. Finally, I argue that the conception of "innateness" can be understood in terms of canalization when it applies to traits that are canalized. Although the canalization conception of "innateness" is not generalizable, it can explain a certain set of empirical observations about language acquisition.

To Chris Jenson

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I owe my deepest gratitude to my advisor Patricia Hanna for her guidance, advice, and for our inspirational conversations on Chomsky. This dissertation would not have been possible without her unending patience through this long process.

I wish to thank my committee members who supplied helpful resources and feedback: Steve Downes, Jim Tabery, Dustin Stokes, and Ed Rubin. I am particularly grateful for the support I received over the years from Ed Rubin, who introduced me to the field of child language acquisition, and from Steve Downes, who introduced me to Cognitive Science. I would also like to thank Stephen Laurence from Sheffield University for inspiring me to read Chomsky in a new light.

I thank my family for their steady encouragement: My parents for their emotional support; Abbey Hepner for our lifelong friendship; and especially Chris Jenson for having confidence in me.

CHAPTER 1

INTRODUCTION

The central claim in cognitive linguistics is that humans are equipped with innate cognitive structures that aid in the ability to understand, parse, and produce language (Chomsky, 1966; Crain, 1994; Hauser, Chomsky, & Fitch, 2002; Jackendoff, 2003; Lightfoot, 1982; Pinker, 1994). The "innateness hypothesis" maintains that certain aspects of language including linguistic knowledge or rules of grammar are innate. This position has generated decades of important research and has aided in our understanding of the complex processes of language acquisition.

However, in important respects, scholars gloss over the complexity of language acquisition and oversimplify what is necessary to explain it. They do so in three ways. First, they rely on one argument, albeit a complex argument, called Poverty of Stimulus, to support the innateness hypothesis. Second, they offer one solution, namely "innateness" to explain a wide variety of empirical observations about linguistic behaviors. Third, they underestimate the role that the environment plays in acquisition. Traits are acquired by development between internal factors in an organism and environmental stimuli. Scholars limit the sources for theories of acquisition to cognitive science or psychology, thus cutting off other resources such as biological theories of innateness that respect the interactive nature of development. Each of these ways

contributes to offering a theory of language acquisition that underestimates the level of explanation required.

I aim to strengthen theories of language acquisition by identifying ways to make them more substantial. I do this by identifying ways in which theories of language acquisition in terms of "innateness" are unsatisfactory, and by offering solutions to create more substantial theories. Briefly, the structure of the discussion is follows: In Chapter 2, I critique the effectiveness of Poverty of Stimulus argument as it stands in the literature. I then I distinguish four distinct strands of arguments. Separating these strands will aid in making the arguments more effective. In Chapter 3, I discuss why the conception of "innateness" is unsatisfactory, and I offer a set of considerations that any theory of acquisition should address to strengthen the explanatory value. In Chapter 4, I develop a theory of language acquisition in terms of canalization as an example of a substantial theory that explains a subset of the empirical observations related to language acquisition. In Chapter 5, I argue that one particular conception of "innateness" in terms of canalization can explain a certain set of empirical observations.

Chapter 2 addresses the first way scholars underestimate the complexity of language acquisition, by using one argument, Poverty of Stimulus, to explain a host of diverse empirical observations. According to this argument, linguistic knowledge is acquired despite a deficiency in linguistic stimuli presumed to be important in learning certain aspects of grammar. This argument draws on "the observation that children's linguistic experience is quite limited" to support the theory that "syntactic knowledge is in large part innately specified" (Crain, 1994, p. 396). Once it is established that certain knowledge of language (including knowledge about phonological, morphological, or

semantic properties of language) is not learned, linguists conclude that it must be innate.

On one level, this is a plausible enough argument, but the conclusion that "knowledge of language is innate" is insufficient as it is missing an explanatory story. In this dissertation, instead of following tradition and viewing this as a single argument, I argue that there are actually four separate arguments. I use the term "Problems of Stimulus" (POS) to refer to this suite of arguments; these arguments are not equivalent and cannot be reduced to one single argument.

The four POS arguments can be summarized as follows: The first, Deficiency of Stimulus (DOS), argues that the linguistic stimuli in the environment is too impoverished to explain the fact that children acquire a rich grammar. The second argument, Corruption of Stimulus (COS), shows that the linguistic stimuli in the environment is degraded in that it contains ungrammatical and nonlinguistic sounds through which the child must sort to form her grammar. The third argument, Variety of Stimulus (VOS), points to a wide variety of linguistic stimuli in each individual's environment as creating a puzzle about why all children proceed through similar stages of development and converge on the same set of linguistic knowledge. The fourth argument, Poverty of Negative Evidence (PNE), argues that children acquire correct grammar in the face of poverty of negative evidence (indications in the environment that the child has acquired faulty grammar).

These arguments are not equivalent. Each argument relies on a different set of observations about the child's abilities and the environment and separating the arguments avoids the danger of lumping together empirical observations. Second, each argument has a different conclusion about what kinds of traits are innate. Separating the arguments is a

first step in formulating more effective theories of language acquisition that respect the complexity of the language acquisition process.

In Chapter 3, I address the second way in which scholars gloss over the complexity of language acquisition; viz., that of offering one solution, "innateness," to explain a diversity of empirical observations. Many theories of language acquisition conclude that traits are innate; however, without specifying what they mean it leaves the reader to struggle with a very general concept with disparate properties. "Innateness" is ambiguous, as it may refer to properties such as, "present at birth," "pancultural," "not learned," or "monomorphic." As a consequence, some critics argue that it is an insubstantial concept that impedes our efforts to advance research. Griffiths and Machery (2008) note, "the concept of innateness is an anti-heuristic which encourages researches to check the obvious sources of environmental input, and then to stop looking" (p. 405). Indeed, he argues, the theoretical term of "innateness" is so inherently confused that we should eliminate it from the sciences.

The charge that "innateness" is confused is a serious charge, as many theories of acquisition conclude that certain traits are innate. Scholars use the term "innateness" to refer to something about the organism (as yet unknown) that plays a causal role in the acquisition of traits. However, without further analysis, "innateness" can do no explanatory work on its own. What we need is a detailed account of how traits are acquired. This may involve identifying the kinds of traits proposed to be innate, specifying the cognitive structures that underlie the trait, and discovering the way in which the trait is acquired.

I propose three considerations that any theory of acquisition may address in order to

strengthen its explanatory value. These are the following:

> (1) *Empirical Consideration* (EC): An account of language acquisition must give careful attention to different sets of empirical observations. It should not oversimplify language acquisition by underestimating the diversity of observations to be explained.
>
> (2) *Developmental Consideration* (DC): A theory of language acquisition should describe cognitive mechanisms that underlie traits and the developmental processes that play a causal role in trait acquisition.
>
> (3) *Interaction Consideration* (IC): A theory of language acquisition should include information about the interaction between factors internal to the organism and environmental conditions.

A theory of language acquisition should explain specific empirical observations about traits (fulfilling EC) by describing the cognitive mechanisms or processes responsible for the acquisition of traits (fulfilling DC). EC address one concern about the treatment of the use of the term "innateness," namely, that it is used as a panacea to explain a large array of empirical observations about linguistic traits. This requires that scholars give more careful consideration to the diversity of empirical observations and ensure that their theory of acquisition explains them. DC, by contrast, requires an adequate description of the cognitive mechanism or developmental process that underlie specific traits, and play a causal role in the acquisition of those traits.

The third consideration, IC, addresses the concern about how scholars underestimate the complexity of language acquisition process. I explain how cognitive linguists limit the sources for theories of acquisition to cognitive science or psychology, thus cutting off other resources such as biological theories of innateness. They do this by discussing two approaches used in the study of language. The first approach is "psychologizing the mind." This approach stipulates that cognitive traits are special and explanations should be given in terms of psychology. This approach encourages scholars to ignore the

developmental story and instead rely on biologists to fill in the details. It also creates a false dichotomy between nature vs. nurture, and genes vs. environment.

I favor a second approach, "biologicizing the mind," which was introduced by Ariew (1996, 1999, 2006) and supported by many cognitive linguists (Birch, 2006; Berwick, Friederici, Chomsky, & Bolhuis, 2013; Chapman, 2000; Chomsky, 2005, 2007, 2012; Collins, 2005, 2011; Fitch, 2009; Hauser, et al., 2002; Hauser & Bever, 2008). This approach seeks to integrate biological categories and concepts with psychological ones. It recognizes that cognitive traits are subject to the same constraints in development and acquisition as other biological traits. A greater emphasis must be placed on the interaction between the genome and environment, thus preventing artificial distinctions between "innate" and "learned." IC requires an interactionist approach to language, given that a full explanation of language will involve multilevel explanations from both psychology and biology.

In Chapter 4, I give a detailed account of VOS, an argument that is undervalued by scholars. The VOS argument is unique because it involves several sets of empirical observations about children's abilities. An explanation of these abilities requires a theory about development and acquisition. I develop a theory of language acquisition in terms of developmental canalization in order to provide an example of a more substantial theory that addresses the three considerations I suggested. This theory uses Chomsky's Principles and Parameters account (Chomsky 1986; 1988; Chomsky & Lasnik 1995). I show how this account can explain certain features of language acquisition (that are involved in VOS arguments), thus indicating a way in which linguistic knowledge can be acquired via canalization. Canalization can initially be understood as the capacity to

acquire a particular trait in spite of variability in the environment during the course of development (Ariew, 1996, 2006; Waddington, 1957). I focus on environmental canalization (as opposed to genetic canalization) where a trait is more or less canalized depending on the degree to which it is insensitive to environmental perturbations.

Though a theory of canalization seems to explain language acquisition, few scholars have developed an account. Chomsky (2012) notes that:

> Every normally developed learner goes through uniform stages of acquisition, where they acquire rules of syntax in order, despite environmental differences. Canalization seems like a useful explanation in context with language acquisition, yet nothing beyond a general discussion has been given. (p. 239)

I draw on Ariew's work (1996, 1999, 2006), and other scholars contribute to his account (Birch, 2009; Collins, 2011; Jenkins, 2004; Khalidi, 2009). I do not claim that canalization can explain the acquisition of *any* trait; my claim is that it explains a range of important traits, namely, those involved in VOS arguments.

Chapter 5 returns to a discussion about the conception of "innateness." I examine Ariew's (2006) definition of "innateness" in terms of canalization, which I call "c-innateness." I argue that it is a useful concept for VOS arguments, as it explains empirical observations used in VOS arguments.

In identifying ways in which theories of language acquisition can give more substantial explanations, I hope to contribute to the effort in understanding the complexity of language acquisition. I hope that my discussion of the four arguments that make up the Problems of Stimulus will clarify these roles so that nativists and empiricists alike can identify what needs to be explained and fill in the details required for more substantial theories of language acquisition.

CHAPTER 2

PROBLEM OF STIMULUS ARGUMENTS

The best-known argument for innateness in language acquisition is the "Poverty of Stimulus" argument (Berwick, Chomsky, & Piattelli-Palmarini, 2012; Chomsky 1980, 1986; Clark & Lappin, 2011, 2013; Cowie 1999; Crain 1994; Lightfoot, 1982; Margolis & Laurence, 2013; Pullum & Scholtz, 2002). The argument is used to support "linguistic nativism," the view that humans have innate knowledge of certain linguistic features.

The argument was introduced by Chomsky as "a variant of a classical argument in the theory of knowledge, what we might call 'the argument from poverty of the stimulus'" (1980, p. 34). He argues that language acquisition is "better understood as the growth of cognitive structures along an internally directed course under the triggering and partially shaping effect of the environment" (1980, p. 35). To support his conclusion, he gives a number of linguistic examples that demonstrate the complexity of language, thus dispelling any intuition that children could somehow learn grammar just from hearing sentences. In the decades that followed, nativists cited a range of empirical observations to support the Poverty of Stimulus argument.

What is clear is that the argument should include empirical observations about children's abilities and the nature of the linguistic stimuli in the environment in which children are raised (i.e., the "primary linguistic data," or *pld*). Depending on the particular

combination, these observations raise various puzzles about how it is that children can acquire linguistic competence in the face of linguistic stimuli that are deficient, degenerate, or idiosyncratic. These observations include:

> (1) *Speed*: Children learn languages fast.
>
> (2) *Reliability*: Children succeed at acquiring language.
>
> (3) *Universality*: All normally developed children acquire language. Language is ubiquitous among all societies.
>
> (4) *Selectivity*: Children pick their grammar from among an enormous number of seductive but incorrect alternatives.
>
> (5) *Underdetermination*: Children acquire knowledge of principles of grammar that are underdetermined by linguistic stimuli.
>
> (6) *Degeneracy*: Linguistic stimuli in the environment include degenerate samples (ungrammatical sentences, false-starts, pauses, etc. as well as nonlinguistic noises).

It is no easy task to outline the Poverty of Stimulus argument. As Pullum and Scholtz note, they mix together an "idiosyncratic cocktail of claims," although "no one attempts to state the argument" (Pullum & Scholtz, 2002, pp. 12, 11). Depending on which scholar is consulted, these claims are used to support a wide range of conclusions. Laurence and Margolis (2001) write that it is "more a collection of related, mutually supporting claims than a neatly circumscribed argument" (p. 221).

Take, for example, the following description of the Poverty of Stimulus argument from Garfield (1994):

> The examples of the target language to which the learner is exposed are always jointly compatible with an infinite number of alternative grammars, and so vastly underdetermine the grammar of the language. The corpus always contains many examples of ungrammatical sentences…There is, in general, no explicit reinforcement of correct utterances…Since it is impossible to explain the learning of the correct grammar – a task accomplished by all normal children within a few years—on the basis of

> any available data or known learning algorithms, it must be that the grammar is innately specified, it is merely 'triggered' by relevant environmental cues. (p. 369)

Garfield first gives a complicated argument about underdetermination and selectivity but follows it with a quick summary of another problem, viz., that of degenerate environments (i.e., "the corpus always contains many examples of ungrammatical sentences"), and poverty of negative evidence in the environment (i.e., there is "no explicit reinforcement"). He explains universality and speed ("a task accomplished by all normal children within a few years") by concluding that, "the grammar is innately specified" and merely triggered. In one short paragraph, he appeals to "innateness" to explain a variety of observations, each of which are independent of one another. These empirical observations are diverse, and it is not clear how some of them support any version of a conclusion that grammar is innate, let alone how they all support a single concept of "innateness." For example, it is easier to see how an innate grammar can explain universality since all children would possess the same grammar. It is more difficult to see how having a set of innate principles of grammar explains how children can sort through degenerate stimuli that include nonlinguistic noises. This is a problem about how children distinguish relevant linguistic stimuli from ungrammatical sentences or nonlinguistic sounds.

In this chapter, I aim to strengthen the Poverty of Stimulus argument by distinguishing four arguments. Separating the arguments will clarify the kinds of empirical observations that should be used in each argument, and identifies the proper conclusions that can be drawn by particular sets of empirical observations. I begin by discussing the essential elements common to any Poverty of Stimulus argument, and

motivate a need to distinguish four nonequivalent arguments (Section 2.1). I then discuss the Deficiency of Stimulus argument (2.2), Corruption of Stimulus argument (2.3), Variety of Stimulus argument (2.4), and Poverty of Negative Evidence argument (2.5). Each section provides examples of the argument in the literature, outlines the essential argument, and identifies the proper sets of empirical observations appropriate to support the particular conclusion that is drawn in each argument. In Chapter 4, I highlight the Variety of Stimulus argument and argue that it draws on unique empirical evidence about acquisition and development.

2.1 The structure of the Poverty of Stimulus argument

The Poverty of Stimulus argument is an inference to the best explanation. Nativists infer that the best explanation for linguistic competence is that some linguistic knowledge is innate.

The general form of the Poverty of Stimulus argument is:

> (P1) Children have an ability to understand, interpret, and produce language. This ability requires "knowledge of language," or knowledge about certain aspects of language.
>
> (P2) Knowledge of language is either acquired by a process of learning, or some set of linguistic knowledge is innate.
>
> (P3) If knowledge of language is acquired by learning, then the environment must contain linguistic stimuli sufficient for the children to infer knowledge of language.
>
> (P4) The linguistic stimuli in the environment are insufficient to acquire knowledge of language.
>
> (C) Some set of knowledge of language is innate.

The argument can be separated into two parts: One part consists of a collection of

empirical observations about children's abilities (P1) and about the environment in which children are raised (P4). The other consists in a series of inferences (P2 and P3) to the conclusion that children could not learn knowledge of language if there is a poverty of stimuli.

Premise (1) states that children have the ability to understand, parse, and produce a natural language such as English or Japanese. Part of this ability works by taking the linguistic stimuli (e.g., sentences spoken in the environment) and assigning structural descriptions to that input (i.e., "parsing" or "interpreting" linguistic stimuli), and then producing sentences. The ability relies on "knowledge of language," the tacit knowledge involved in interpreting linguistic stimuli.

In this context, "knowledge of language" is treated as whatever knowledge is required in the ability for competent speakers to understand, interpret, and parse linguistic stimuli into syntactical properties (such as verb phrases or noun phrases). This definition is used by both empiricists and nativist theories, and does not favor one over the other. The idea for both nativists and empiricists is to discover (by a bit of reverse-engineering) what the *initial* knowledge of language must be that would explain linguistic competence,[i] i.e., what kinds of tacit knowledge people must have when they acquire language.

Knowledge of language may contain principles about phonology (the study of the sound patterns of language), morphology (the study of the structure and meaning of words), syntax (the study of the structure of sentences), and semantics (the study of linguistic meaning). An example from phonology is knowing that the [*ny*] sound is possible in the middle of words, such as in "canyon" or "onion," but it is not acceptable at the beginnings of words. "Grammar" is a formal system that describes rules about how

to manipulate words and phrases in well-formed sentences. People obviously acquire some knowledge of language by a process of learning such as lexicon or principles of grammar particular to natural languages such as English or French. However, there may be a set of knowledge of language where it is not obvious how people come to know those principles.

Taken together, Premises (1) and (4) set up a puzzle: If there are insufficient linguistic stimuli in the environment to demonstrate some linguistic principle (e.g., the principle of subjacency), how is it that children can produce sentences that follow that principle? From what source did they acquire knowledge of the principle? Premise (4) is a key premise as it provides empirical evidence that the environment does not contain the necessary linguistic properties from which children can learn all aspects of knowledge of language. Scholars offer set of empirical observations about the impoverished quantity or quality of linguistic stimuli. They do so by citing cases, both common and rare, that demonstrate how the environment does not contain enough linguistic stimuli, or the right sorts of linguistic stimuli, from which the child could learn knowledge of language.

Premises (2) and (3) set up a conditional where knowledge of language is acquired in two ways: It is either learned, or it is innate. If knowledge of language is learned, then it is by process of learning such as induction and this requires exposure to the right kinds of linguistic stimuli in the environment (called "primary linguistic data," or *pld*.) The *pld* must be rich enough, and must contain the right sort of linguistic stimuli, to allow children to infer principles of grammar. Premise (3) states that in order for an empiricist theory to succeed, the environment must contain linguistic information rich enough for the child to learn from it and if there is a poverty of such linguistic information, then any

such attempts at learning from the impoverished environment will fail.

Poverty of Stimulus arguments are always offered as an attempt to show how it is difficult (or impossible) to learn all aspects of language; as a consequence, they take empiricism as their target.[ii] Empiricists argue that grammar is acquired by learning, where learning may include any psychological process of acquisition, including conditioning, induction, or abduction (Elman, et al., 1999; Karmiloff-Smith, 2009; Prinz, 2012; Reali & Christiansen 2005; Saffran, 2002; Seidenberg, 1997). They argue that humans are equipped with domain general mechanism used to "notice patterns, recognize familiar objects and make decisions based on prior decisions" (Prinz, 2012, p. 168). Empiricists explain acquisition of linguistic knowledge by saying that children come equipped with domain general systems that "extract" or "pick up" patterns in the linguistic stimuli. These patterns correlate with rules of grammar, phonology, morphology, etc. For example, children can infer that the sound [*ny*] occurs in the middle of words by hearing words such as "onion" and "canyon," and also by not hearing that sound at the beginning of any word.

It is worth examining an argument that has particular emphasis on how difficult it would be to acquire knowledge of language by inferring or extracting information from the *pld*, as empiricists suggest. Laurence and Margolis (2001, p. 221) provide a rare instance of an outline of the Poverty of Stimulus argument:

> (P1) An indefinite number of alternative sets of principles is consistent with the regularities found in the primary linguistic data.
>
> (P2) The correct set of principles is not in any pre-theoretic sense simpler or more natural than the alternative.
>
> (P3) The linguistic data that would be needed for choosing among these sets of principles are in many cases not the sort of data that are available to

> an empiricist learner in the child's epistemic situation.
>
> (P4) So if children were empiricist learners, they could not reliably arrive at the correct grammar for their language.
>
> (P5) Children do reliably arrive at the correct grammar for their language.
>
> (C) Therefore, children are not empiricist learners.

They define the "empiricist learner" as "one that instantiates an empiricist learning theory" where:

> [S]uch a learner wouldn't have any innate domain-specific knowledge or biases to guide her learning and, in particular, wouldn't have any innate language-specific knowledge or biases. (Laurence & Margolis, 2001, p. 221)

In this passage, Laurence and Margolis present an underdetermination problem that goes beyond the observation that the environment is *missing* crucial linguistic data: They argue that even if there is a sufficient amount of linguistic data in the *pld* from which children can infer principles of grammar, there would be too many alternative sets of principles of grammar, all of which are compatible with the data. Consequently, the child could infer many principles and may never converge on the right set of principles. In other words, children would form too many incorrect hypotheses of grammar. They say, "If a child were an empiricist learner, she'd be limited in the way that she could discern the correct set of principles underlying a target language" (Laurence & Margolis, 2001, p. 221).

To see why an empiricist learner could not infer the right principles, take an example of question formation.[iii] An empiricist learner operates the same way in inferring linguistic principles as when confronted with other nonlinguistic tasks. It seems that the empiricist learner should pick the "simplest or most natural set" of principles by generalizing linguistic rules based off surface structure of language. Suppose an

empiricist learner is confronted with the sentence (a) and (b):

(a) The man is coming.

(b) Is the man coming?

The learner may formulate the following principle: "Move the first 'is' to the front of the sentence to create a question." However, given sentence (c), this principle would yield an ungrammatical sentence (d):

(c) The man who is running is coming toward us.

(d) Is the man who running is coming toward us?

However, children avoid forming ungrammatical sentences such as (d), and instead form the grammatical sentence in (e):

(e) Is the man who is running coming toward us?

The ability to form questions relies on more than the surface order of words in a sentence; it relies on knowledge about deep structure. This knowledge involves concepts such as *main clause* and *auxiliary*, knowledge about movement of phrase (such as noun, verb, and object phrases), and complicated principles about Structural Dependency. However, information about how deep structure works is "not the sort of data that are available to an empiricist learner in the child's epistemic situation" (Laurence & Margolis, 2001, p. 221). Thus, children could not be learning these principles from the *pld*.

All Poverty of Stimulus argument draw a conclusion about how some knowledge of language must be "innate."[iv] The conclusion "knowledge of language is innate" is a claim about how the initial structure of the Language Faculty contains linguistic content. The "Language Faculty" is a cognitive device that inputs linguistic symbols, parses those symbols into linguistic structures such as noun or verb phrases, and produces linguistic

symbols that are then handled by a production system (see Chomsky, 1969, 1980). There is no question that some cognitive structures work to interpret stimuli as meaningful linguistic symbols.

The task for both nativists and empiricists is to determine what the *initial* structure of the Language Faculty must look like. They do this by examining various stages of language acquisition. The mature steady state reflects fully developed linguistic knowledge required for linguistic competence. This is acquired over many years and involves aspects of learning language-specific rules, lexicon, etc. The child starts with an early state (prenatal babies can distinguish voices and linguistic sounds from nonlinguistic sounds), and develops through stages of babbling, forming sentences, etc. The basic elements of grammar are normally acquired around age 6.

The question is how must the human brain be equipped at the initial stage of language acquisition? There are several possibilities: (1) The initial structure of the cognitive system may contain linguistic *content* that represents information such as concepts of noun-phrases, Subject, Verb, Object, etc. (2) It may consist of rules of syntax represented as propositional rules, which I call the *"rule-systems account*," that assumes that the Language Faculty consists in propositional rules such as the principles of Structural Dependency, C-command, or other linguistic concepts.[v] (3) The system may work in systematic ways that can *be described* as following rules, where this description is neutral to any particular theory of representation. This is the view that Chomsky holds (Chomsky, 1986, 1995; Collins, 2008). (4) Empiricists argue that the initial state of the Language Faculty is constituted with general learning mechanisms and contains no linguistic content.

One problem with saying that "knowledge of language is innate" is that the concept of "innateness" is ambiguous: A trait may be innate in the sense that it is present at birth, or that it is insensitive to environmental perturbations, or that it is genetically encoded. I discuss problems with the concept of "innateness" in Chapter 3. In addition, some scholars offer very different definitions of "knowledge of language," while others use it interchangeably with "Universal Grammar" and this ambiguity leads empiricists and nativists to talk past each other.

Another problem is that different POS arguments conclude different things, and sometimes the conclusions are not supported by the right kinds of empirical observations, thus making them easy targets for criticism on the grounds that they are not valid or sound. Empiricists may criticize the group of arguments that make up POS by attacking one version but ignoring other POS arguments.[vi]

Finally, scholars often combine arguments that are not equivalent. The following passage from Seidenberg (1997) illustrates this:

> The input to the child is degenerate, consisting of both grammatical and ungrammatical sentences that are not labeled as such. It is also variable; children are exposed to different samples of utterances but converge on the same grammar. The input does not include reliable negative evidence, that is, evidence about which structures are not allowed by the grammar; logical arguments suggest that in the absence of such evidence there must be strong innate constraints on the possible forms of grammars. Finally, languages exhibit properties for which there is no positive evidence in the input. The claim here is that there cannot be any overt evidence for the kinds of abstract underlying structures characteristic of grammatical theory. That essential aspects of grammar are innate—represented in the brain of the neonate—is said to be the only viable explanation for how languages could be learned so rapidly yet under such impoverished conditions. This hypothesis simultaneously accounts for universal properties of languages. (p. 1601)

The overall conclusion to any Poverty of Stimulus argument is that children are

supplementing their knowledge of language with innate linguistic content, i.e., "the essential elements of grammar are innate," as Seidenberg states (1997, p. 1601). This conclusion has a surface appearance of solving both puzzles, thus explaining the fact that children do acquire grammar in an impoverished environment, and also that they acquire the same grammar despite variation in the environment.

However, it is one thing to conclude that "essential grammar is innate" because there are not enough linguistic stimuli in the environment from which children can learn the grammar. It is, however, another thing to conclude that grammar is innate because children would not be able to interpret whether sentences are grammatical or not from a set of degenerate linguistic stimuli. These two arguments are not equivalent. An empiricist may argue that there is a rich source of *indirect* positive and negative data in the environment from which children can infer rules of grammar, thus solving an impoverishment puzzle. After solving this puzzle, an empiricist may think that all puzzles are solved. When nativists conflate several puzzles, it makes the Poverty of Stimulus weak. If we can separate the various arguments, it will strengthen the overall Poverty of Stimulus strategy.

I argue that there are four separate arguments that have traditionally been called the Poverty of Stimulus Argument. I use the term "Problems of Stimulus" (POS) to refer to this suite of arguments. They are the following:

> (1) *Deficiency of Stimulus* (DOS) arguments indicate the absence of linguistic stimuli in the child's environment, or absence of crucial features from which children could infer linguistic principles.
>
> (2) *Corruption of Stimulus* (COS) arguments indicate that linguistic stimuli are too degraded or poor to be used to acquire consistent linguistic principles.

(3) *Variety of Stimulus* (VOS) arguments highlight the fact that children converge on the same knowledge of universal linguistic principles despite variations in individual's exposure to linguistic stimuli.

(4) *Poverty of Negative Evidence* (PNE) arguments emphasize the child's competency at acquiring correct linguistic principles despite negative evidence in the environment (i.e., some indication that the child's grammar is incorrect).

There are at least two reasons why these arguments are not equivalent and cannot be reduced to one single argument. The first reason is that each argument uses a different set of empirical observations about children's abilities and the environment in which they are raised. The following compiles a list of observations:

(1) *Automaticity*: This observation points to the tacit ability to parse sentences and their deep structures.

(2) *Speed*: Language acquisition is quick and unfolds within a biologically determined timeframe. Babies start babbling and 2.5-years old regardless of environmental stimulation. Children form grammatical sentences by age 3 and master language by age 6 (Klegl, 2004; Radford, 1990).

(3) *Species Specificity*: Language is unique to humans. No other animal exhibits syntactic rules or compositionality found in human languages. Though people have tried to train chimpanzees, they could not learn syntax beyond a rudimentary paring of words (Berwick, et al., 2013).

(4) *Complexity of Syntax*: Syntactic rules are often complex and inextricable. This cannot be explained by saying children invented the simplest or most adaptive approach to communication (Clark & Lappin, 2013; Hauser, et al., 2002; Jackendoff & Pinker, 2005; cf. Christiansen & Kirby 2003).

(5) *Brain Structure*: Certain linguistic abilities correlate to areas of the brain. Double dissociation studies show that the Broca's area correlates with syntactic formation, and the Wernicke's area correlates with semantics or word formation (Grodzinsky, 2000).

(6) *Critical Period*: Language acquisition can only develop in a critical period time. Cases of child abuse make it clear that language cannot be acquired after age 9 or so, if there is no early exposure (Curtiss, 1988).

(7) *Body Structure*: The structure of the speech organs (larynx, etc.) seems

adapted for language production.

> (8) *Acquisition Independent of IQ*: Every normally developed child acquires language despite varying abilities in learning, IQ, motivation, or other aspects intrinsic to a child's biological or psychological make-up (Chomsky 1975, p.144).

Note the diversity of these observations: They deal with brain structure; the speed of acquisition; acquisition of language independent of intelligence, motivation, or ingenuity; the complexity of grammatical, phonological, and morphological structure of sentences; the similarities or differences between natural languages; language change; linguistic disorders; phylogeny; and comparison studies with other animals. It is not clear whether some of these observations can be used to support a Poverty of Stimulus argument. What is clear is that these arguments should be supported with empirical observations about children's abilities despite being raised in certain environments. In the following sections, I separate a list of empirical observations into sets that correlate with the four POS arguments. Considering the wealth of research produced in multiple disciplines, none of these lists is exhaustive, but they are useful to show that each of the arguments addresses a different set of observations and each draws a different conclusion.

The second reason the arguments are not equivalent is that they each draw a unique conclusion about *what* is innate. The conclusion to the general argument outlined above is that "knowledge of language is innate." However, this is highly ambiguous, and as the analysis of the four arguments show, there are actually four separate conclusions about what traits are innate. Some arguments conclude that some kind of *linguistic content* is innate, other arguments conclude that a *cognitive structure* is innate, and some arguments conclude that a *process or procedure* is innate. All three conclusions are found in the various formulations of the Poverty of Stimulus arguments.

The first kind of innate trait is *linguistic content* that refers to language-specific rules, principles, or concepts involved in phonology or morphology or grammar. These may include concepts of phrase structures such as noun phrase, verb phrase, object phrase, main clause, and auxiliary verb. They may also include principles about movement or governing structure, such as subjacency principles, or principles of Structural Dependency (Chomsky, 1980; Khalidi, 2001, 2007; Pinker, 1994).[vii] As the following quotes show, scholars use Poverty of Stimulus argument to argue that there is innate linguistic content:

> (1) Children must be guided by innate principles of grammar…At the very least, [the child] needs to possess the concepts MAIN CLAUSE and AUXILIARY VERB. (Laurence & Margolis, 2001, pp. 221, 223)
>
> (2) Certain aspects of our knowledge and understanding [of language] are innate, part of our biological endowment, genetically determined. (Chomsky, 1968, p. 11)
>
> (3) [A] learner who has no knowledge about auxiliary and lexical verbs (i.e., a simple, unbiased, correlational learner) is almost certain to confuse the two types. (Stromswold, 1999, p. 361)

According to these scholars, children come equipped with innate knowledge, principles, or understanding of language. POS arguments that conclude that the Universal Grammar is innate also fall into this category. Universal Grammar[viii] is a collection of linguistic principles universal to all languages (Chomsky, 1980 1986; Crain & Thornton, 1998; see Evans & Levinson, 2009 for extensive discussion).

Other POS arguments, however, do not warrant a conclusion that linguistic *content* is innate. What is proposed to be innate are cognitive *structures* such as domain-specific modules or the Language Acquisition Device. For example, after describing Poverty of Stimulus arguments, some scholars conclude that what is innate is a "domain-specific

language acquisition device" which is an "abstract linguistic computational system…independent of the other systems with which it interacts and interfaces" (Behme & Deacon, 2008); and that there must be an innate "genetically-determined specialized mechanism that is necessary for the normal development of human language" (Marcus & Rabagliati, 2006). Chomsky (1975), as well, describes an innate cognitive structure: "Individuals of a given community each acquire a cognitive structure that is rich and comprehensive and essentially the same as the systems acquired by others" (p. 144.)

In a third set of conclusions, rather than concluding that "knowledge of language" is innate, scholars use the POS argument to draw conclusions about how the Language Faculty works to interpret, parse, or filter linguistic stimuli. These are given in terms of processes, procedures, or constraints. For example, Khalidi (2001) uses the term "innate cognitive linguistic capacities," and Margolis and Laurence (2013) conclude that there must be an innate "set of structures and processes that isn't restricted to general-purpose learning systems" (p. 698).

Scholars also conclude that ''the capacity of recursion'' is innate, and again this refers to the Language Faculty's ability to parse, combine, repeat, or add certain elements of sentences, such as verb phrases (Hauser, et al., 2002, p. 1571). Crain (1994) describes "constraints" about the kinds of linguistic principles that can be inferred from the *pld*, and thus he is describing how the Language Faculty works to input certain stimuli, and form hypotheses about linguistic principles. These are all examples of how scholars use the Poverty of Stimulus argument to conclude that mechanisms or processes (as opposed to innate content or knowledge) are innate.

Cowie's (1999) description of the Poverty of Stimulus argument concludes both that

there is an innate domain-specific module, and that this module contains linguistic content:

> The [Poverty of Stimulus argument] takes as an example some specific rule of grammar and argues that the data to which the child is exposed during the learning—the 'primary linguistic data' (pld)—are insufficient to enable a learner, endowed only with a general purpose learning ability, to infer that rule. The other rules of the grammar being assumed to be the same in this respect, the argument is generalized to the grammar as a whole, and hence to support the conclusion that language-learning is subserved by a special faculty incorporating the linguistic information specified in Universal Grammar. (p. 178)

A specific set of empirical observations is required to support the claim that there is an innate domain-specific faculty. An entirely different set of observations and arguments should support the conclusion that there is innate content. It is important for scholars to ensure that the conclusions they draw are actually supported by the empirical observations they give about children's abilities and the environment. When we separate the four POS arguments, we get a clear understanding of what each argument concludes about what is innate.

2.2 Deficiency of Stimulus argument (DOS)

The first argument is the Deficiency of Stimulus (DOS). This focuses on the *impoverished quantity* of the linguistic stimuli in the environment. It argues that if the linguistic content is missing from the environment, then a child's knowledge of language could not be learned from the environment. It concludes that a set of knowledge of language is innate. That is, the initial structure of the Language Faculty must be comprised with this set of knowledge of language.

A DOS argument is found in the following:

> The point of the argument from poverty of the stimulus is that if it is the case that during their lifetimes neither the child nor its parents have ever been exposed to or ever uttered the crucial sentences needed to deduce some grammatical property (e.g., properties of structure-dependence, Subjacency, distribution of empty categories, and the like), then no amount of intelligence or ingenuity on the child's part, nor corrections and tutoring on the parents' will yield these properties. In such cases then, it is reasonable to hypothesize that the child must acquire these properties on the basis of internalized principles; e.g., from their genetic program for language. (Jenkins, 2000, p. 80)

Chomsky also argues that children form linguistic principles even when there are no samples in the *pld* that demonstrate the principle. He says:

> A person might go through much or all of his life without ever having been exposed to relevant evidence, but he will nevertheless unerringly employ structure-dependent generalizations on the first relevant occasion. (as cited in Piattelli Palmarini, 1980, p. 40)

Each of these arguments focuses on a deficiency of the right kinds of linguistic stimuli required for children to learn.

> The general structure of the DOS argument is:
>
> (P1) Children have knowledge of language that surpasses linguistic stimuli in the environment in which they are raised.
>
> (P2) This set of knowledge is either acquired by a process of learning, or it is innate.
>
> (P3) If this knowledge is acquired by learning, then the environment must contain enough linguistic stimuli sufficient for the children to infer knowledge of language.
>
> (P4) The linguistic stimuli in the environment are too impoverished to acquire knowledge of language.
>
> (C) Children possess a rich source of innate linguistic content.

Taken together, the puzzle in DOS is about how children could know principles of language if there are no examples of those principles in the *pld*. Since children acquire grammar that is much richer than the linguistic data in the environment would support, it

follows that they cannot acquire grammar via learning. Empirical observations that support Premises (1) and (4) include the following list:

> (1) *Paucity of stimuli*. Linguistic stimuli in the environment are impoverished and missing crucial elements of language from which the child could infer knowledge of language.
>
> (2) *No explicit training*: Children are not explicitly taught language, unlike other knowledge like math or penmanship that must be selectively, explicitly, and painstakingly taught.
>
> (3) *Abundant knowledge*: Children's knowledge of language outweighs the linguistic stimuli to which they are exposed. For example, deaf children who have little exposure to linguistic stimuli acquire language when given the opportunity to socialize with other deaf children (Goldin-Meadow, 2003; Kegl, Senghas, & Coppala, 1999; Pinker, 1994; Singleton & Newport, 2004; Wilbur, Klima & Bellugi, 1983).
>
> (4) *Productivity*: Children acquire an ability to produce or understand any of an essentially unbounded number of sentences. Language is creative and compositional (Chomsky, 1957; Crain, 1994; Hornstein & Lightfoot, 1981).
>
> (5) *Ease of acquisition*: Children acquire language relatively quickly and easily. Babies start babbling and 2-year-olds start speaking regardless of environmental stimulation. The basic elements of grammar are acquired by age 6.

Many arguments fall under the Deficiency of Stimulus heading. It will help to formulate the strongest argument of DOS, namely that the *pld* does not provide enough data for the child to acquire the knowledge. Scholars appeal to extreme cases such as Nicaraguan Sign Language and the case of Simon, to show that children child could not learn from the *pld* since there are no data from which to learn.

Nicaraguan Sign Language provides an example of *paucity of stimuli* (Goldin-Meadow, 2003; Pinker, 1994; Senghas & Coppola, 2001; Singleton & Newport, 2004; Wilbur, et al., 1983). In this population, caregivers who do not know sign language raise deaf children: In effect, it is a case where the environment is devoid of linguistic stimuli,

since they cannot hear spoken language, and they are not exposed to sign language. When the children have an opportunity to socialize with other deaf children, they spontaneously develop a system of communication that qualifies as a language because it consists in fully formed syntactic rules and structure. In other words, this sign language contains elements of linguistic universals syntactic (Kegl, et al., 1999; Pinker, 1994). This case is used to show that children acquire *abundant knowledge* that is sufficient to generate a language which surpasses the linguistic stimuli to which they are exposed.

Kegl describes the case of Nicaraguan Sign Language as a language that emerges *de novo* (Kegl, 2004). The environment is not just missing crucial aspects of language from which to infer principles of language; it is missing *all* linguistic data since no linguistic sounds or signs were present until the children themselves start signing in an attempt to communicate with each other. Children's creation of language offers a striking case of "acquisition of a first language by children in the absence of even fragments of a full language or languages in their input" (Kegl, 2004, p. 1999).

The case of Simon is also used as evidence that the linguistic information in the environment to which children are exposed is impoverished (Singleton & Newport, 2004). Simon was born deaf and raised by hearing caregivers whose grasp of sign language was flawed in that it lacked systematic structure. Despite this, Simon developed language normally and went through the same stages and pitfalls as hearing children. His production of ASL surpassed that used by his caregivers, and he was capable of "acquiring a regular and orderly morphological rule system for which his input provides only highly inconsistent and noisy data" (Singleton & Newport, 2004, p. 371).

This strong version of the DOS argument shows that no linguistic data is required for

acquisition of language. There are cases showing that languages develop in the complete absence of *pld*. If this is so, it would follow that all knowledge of language arise innately. The best examples of this are Nicaraguan sign language and the case of Simon, which present challenges to the empiricist. An appeal to standard inductive learning, for example, fails because there are no data from which to infer principles of grammar. That is, if it is true that the environment is missing all linguistic data, and if it is true that children's linguistic abilities relies on knowledge of these sets of linguistic content, then that knowledge cannot be not learned. If children have abundant knowledge of language that outweighs linguistic stimuli, this knowledge must be innate. Not only is there an innate *motivation* to develop language, there also must be innate *linguistic content* that children bring to the task of language formation. The conclusion, then, is that the language faculty must contain a significant amount of linguistic content and concepts.

We can also infer that in cases where there are no linguistic stimuli, innate content must be made available to the child by maturation processes at a particular developmental stage. It cannot be triggered by linguistic stimuli from the environment, since, in extreme cases, there are no linguistic stimuli to trigger knowledge. This makes for a simple theory of acquisition, where linguistic knowledge is present at birth, the environmental stimuli play no role (or only a minor role) in triggering this content, and development is only a matter of revealing content that is already present. A more complicated story about how genes interact with relevant external stimuli will fail to account for cases in which there are no relevant features to be found in the environment. For example, deaf children (who are not raised around sign language) have no exposure to linguistic stimuli, so there can be no interaction between internal and external factors. The conclusion, then, is that *the*

child has linguistic content that emerges independent of environmental stimuli.

I have explained the strongest version of the DOS argument that rests on the observation that the there are too few linguistic data from the *pld* from which children could learn. In principle, it provides the strongest refutation of empiricist theories of learning because it undermines the very possibility of learning grammar from the environment. However, this argument has problems. In its strongest form, it relies on the claim that, "there are *no* linguistic stimuli," but even the extreme cases of Nicaraguan Sign Language or Simon are insufficient to establish this. If we examine these cases further, we find that *some* linguistic stimuli are always present in the environments under study. For example, deaf children do not develop language until they socialize with each other and interact by signing. Thus, children may draw on linguistic stimuli from other children's sign language.

In addition, it seems obvious that no one could acquire language in an environment that is completely devoid of linguistic data. By contrast, Curtiss (1988) argues that severe cases of isolation and absence of linguistic data, such as Genie, show that if a child is deprived of all linguistic input during the "critical period" of development, that child will not acquire language. Genie grew up isolated and abused; she had little contact with language until she was 13. After her recovery, Curtiss describes Genie as a "powerfully effective nonlinguistic communicator," but one who was unable to acquire knowledge of grammar essential for language (Curtiss, 1988, p. 98). Finally, extreme cases like Nicaraguan Sign Language and Simon are rare. Thus, even if these cases do illustrate an absence of linguistic data, appealing to them does not help us explain the common phenomenon of language acquisition. In the normal case, children are exposed to

linguistic stimuli, and even in the extraordinary cases, when linguistic data are as close to completely absent as possible, the child does not acquire language.

A more modest version of the DOS argument is that there are deficient data *of the right kind*. This version also draws on abundant knowledge, but asks how children could know certain principles if there are no instances of the principle in the *pld*. These arguments conclude that linguistic content is innate, but they do so by pointing to the difficulty in acquiring language when the environment is missing important kinds of information.

In the following outline from Pullum and Shultz (2002, p. 18), the Poverty of Stimulus is described as a problem about "crucial evidence" that is missing from the environment:

> (P1) Human infants acquire their first languages either by data-driven learning [i.e., that "relies entirely on generalization from experience by the ordinary methods that are also used for learning other nonlinguistic things"] or by innately-primed acquisition [i.e., that "calls upon inborn domain-specific linguistic information"].
>
> (P2) If human infants acquire their first languages via data-driven learning, then they could not learn anything for which they lack crucial evidence.
>
> (P3) But infants do in fact acquire things for which they lack crucial evidence.
>
> (P4) Thus, human infants do not learn their first languages by means of data-driven learning.
>
> (C) Human infants acquire their first languages by means of innately primed acquisition.

Premise (3) focuses on a "lack of crucial evidence" in the *pld*, a premise that must be supported with observations about how the environment is missing *the right kinds* of linguistic data from which children can infer their grammar. Language acquisition

involves forming hypotheses about principles of grammar or phoneme construction. However, if the environment does not supply evidence of what these principles are, then language acquisition will fail.

An example of the kind of "crucial evidence" required for grammar formation concerns how sentences are structured. Hornstein and Lightfoot (1981) argue that children's abilities to understand and interpret sentences rely on knowledge of principles of grammar. Given that sentences do not come with a set of instructions about how to structure them, it follows 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" (p. 9). This establishes the need for innate understanding of sentence structure, including how to structure sentences into noun, subject, or prepositional phrases.

Other sets of empirical observations in DOS arguments are the *Ease of Acquisition* and *No Explicit Training*. Clark and Lappin (2011) focus on these:

> Language acquisition presents some unusual characteristics…First, languages are very complex and hard for adults to learn. Learning a second language as an adult requires a significant commitment of time, and the end result generally falls well short of native proficiency. Second, children learn their first languages without explicit instruction, and with no apparent effort. Third, the information available to the child is fairly limited. He/she hears a random subset of short sentences. The putative difficulty of this learning task is one of the strongest intuitive arguments for linguistic nativism. It has become known as The Argument from the Poverty of the Stimulus. (p. 33)

This passage combines a number of facts about the children's abilities, noting that children acquire their first language relatively easily, without the aid of instruction, regardless of the quantity or quality of linguistic stimuli in the environment. It does not involve a process of explicit learning and rote memorization, a process required to learn second languages. In addition, children acquire language independent of their motivation

and intelligence. Babies start babbling almost immediately after birth, and by 2 years old they are speaking in sentences independent of environmental stimulation and their own motivation to "learn" it.

Language is unlike anything else acquired via learning. Math or penmanship must be selectively, explicitly, and painstakingly taught. However, language is acquired without any such instruction of this sort. Children are not taught formal rules of grammar and rules of syntax, yet they are competent in forming sentences with complex principles such as c-command, the binding conditions, subjacency, negative polarity items, that-trace deletion, nominal compound formation, control, auxiliary phrase ordering, and the empty category principle (MacWhinney, 2004, p. 888). These general principles of languages are barely understood by linguists, and largely unknown by people who have never studied linguistics; from this, it follows that language is not imparted by explicit training (Laurence & Margolis, 2001).

A final set of empirical observations to show that children's knowledge surpasses available data in the *pld* is *Productivity*. People have an ability to produce or understand any of an essentially unbounded number of sentences (Chomsky, 1957; Clark & Lappin, 2013; Crain, 1994; Hornstein & Lightfoot, 1981; Lightfoot, 2013). Language is creative and compositional. People can generate an infinite number of new sentences that have never been heard before, or to rearrange sentences in new ways. Lightfoot (2013) focuses on productivity in saying:

> Children do not just imitate what they hear but develop a system that is far richer than the fragmentary and limited speech that they encounter in their first few years. For example, they hear a finite number of utterances, and the system they develop must be finite but range over infinity; children develop an internal system that generates an infinite range of expressions. (p. 28)

The output (sentences produced) cannot be learned from the input (sentences heard), given that we have never heard most of the sentences that we produce. Hence, language acquisition cannot be a matter of imitation, mimicking, or even rearranging sentences that have been heard. Theories such as B. F. Skinner's behaviorist theory cannot explain productivity, since it proposes that children's ability to acquire language is proportional to the amount of language to which they were exposed (Chomsky, 1957; Skinner, 1957).

To conclude, DOS arguments cite empirical observations where linguistic stimuli are deficient. It concludes that children have an innate set of linguistic content. If it is true that there are no instances of sentences in the environment that respect principles of subjacency, for example, then the best explanation for why children form sentences that follow the principle is that subjacency principles are innate.

2.3 Corruption of Stimulus argument (COS)

I coin the term "Corruption of Stimulus" (COS) for the second Poverty of Stimulus argument. This argument focuses on the *quality* of the data, not the *quantity*. The distinguishing feature in COS is that there are plenty of linguistic stimuli which could aid in learning a grammar, but those data are not the right sort for inferring the kinds of grammatical rules needed to explain the child's subsequent knowledge of language.

The distinction between COS and DOS is mentioned in Chomsky's 1986 book, *Knowledge of Language* although he does not discuss the distinction elsewhere. He poses two research questions about the nature of knowledge. The first, dubbed "Plato's problem," asks how it is that humans know so *much* given that they could not have learned it from their impoverished environment. The second, called "Orwell's problem,"

asks how it is that humans know so *little* given that their environment has so much information.

We can outline the COS argument as follows:

> (P1) Children have an ability to distinguish relevant stimuli in their acquisition of grammar.
>
> (P2) Either children learn how to distinguish relevant linguistic stimuli, or they have innate means of distinguishing relevant stimuli.
>
> (P3) If children learn how to distinguish relevant stimuli, then the environment must contain indications about whether linguistic stimuli is relevant or not.
>
> (P4) The environment does not contain indications about what linguistic stimuli are relevant or not; instead, the available stimuli are degraded and corrupt.
>
> (C) Children have innate means of distinguishing relevant stimuli.

The set of empirical observations used in COS show that linguistic stimuli in the environment are degenerate, degraded, or corrupt. The following lists several observations used to support Corruption of Stimulus arguments:

> (1) *Degeneracy*: Linguistic stimuli include ungrammatical sentences, stops, pauses, "ums" and a number of nonlinguistic noises.
>
> (2) *Selectivity*: To arrive at the correct grammar, children must select their grammar from among an enormous number of seductive but incorrect alternatives to acquire just the right grammar. Babies only a few days old can distinguish the phonemes of any language and are primed to attend and process linguistic sounds (Seidenberg, 1997).
>
> (3) *Filtering*: Children are able to filter out degenerate linguistic stimuli that would lead to faulty grammar formation. For example, children exposed to pidgins with degraded syntax create Creoles, language with systematic syntax (Bickerton, 1983; Haegeman, 1994; Hornstein & Lightfoot, 1981; Pinker, 2013b).

These observations focus on the degraded nature of linguistic stimuli and the ability of children to reliably acquire correct knowledge of linguistic principles from such

degraded stimuli. *Degeneracy* refers to the degraded quality of stimuli, including nonlinguistic noises such as dog barks, traffic noise, or coughs. The stimuli are *degenerate,* in the sense that they cannot be used as input for grammar formation. Such input includes ungrammatical formulae, some produced by children themselves, and others arising from slips of the tongue, pauses, run on sentences, foreign words and phrases, incomplete sentences, sentences containing "ums," and just bad speaking habits. Laurence and Margolis (2001) point out that,

> A truly unbiased learner couldn't even assume that some of the noises she hears are linguistic…Coughs, whistles, gurgles, and so on must be filtered out. Moreover, all sorts of variations in speech aren't linguistically salient. Each voice has a different timbre…People speak faster or slower, louder or softer, and with different intonations depending on a wide variety of variables. They sing, whisper, mumble, slur their speech, all affecting the sounds they make. None of these variations matters to language per se. But why should an unbiased learner suppose that? (p. 227)

Some linguistic expressions are irrelevant, such as "interrupted fragments, false starts, lapses, slurring," and other distortions of grammar (Chomsky 1962, p. 531). Lightfoot (1982, p. 60) notes that if only 5% of the expressions are ungrammatical, the child would have problems developing grammar on a learning theory in which the child infers her grammar from the available faulty data.

The degeneracy of stimuli shows that children must determine what features of noises in the environment are salient and relevant for language formation. As Laurence and Margolis put it, "speech doesn't come pre-sorted into the categories of reliable data and noise" (2001, p. 230). The problem concerns the ability to pick out the *right* sorts of input from linguistic stimuli in order to form the correct grammar (i.e., the grammar that matches the natural language in the environment). Even a rich set of linguistic stimuli presents difficulty about what stimuli to use when inferring linguistic principles. That is,

the *pld* contains:

> [T]oo many grammars that are compatible with the primary linguistic data and many of these are more tempting to someone in the child's situation than the correct grammar…[I]n addition to the huge number of hypotheses that are consistent with the primary linguistic data, there is also a huge number of hypotheses that are just tempting dead-ends. Since these dead-ends vastly increase what is already an enormous hypothesis space, they make it increasingly likely that the child won't be able to arrive at any plausible hypothesis at all, not even one that generates a significant subset of the primary linguistic data. (Laurence & Margolis, 2001, p. 231)

If the environment contains a rich source of linguistic information, the empiricist learner may infer faulty hypotheses of grammar.

Crain (1994) focuses on cases where the sentence structure or meaning is not often clear, and yet people have no problem parsing or understanding the syntactic structure. He concludes that:

> [E]very child comes to know facts about the language for which there is no decisive evidence from the environment. In some cases, there appears to be no evidence at all. (Crain, 1994, p. 364)

Crain and Nakayama (1987) give an example of coreference. In the sentence, "Jay hurt him," children know that "him" cannot refer to "Jay." This is evidenced by the fact that children react negatively to situations in which Jay is "revealed" as the object of "him." Another example is an ambiguous sentence such as "I rode a black horse in red pajamas." The sentence has two possible sentence structures, depending on whether the subject, "I," or the object, "the horse," governs the prepositional phrase "in red pajamas." Parsing sentences is not a simple matter of understanding the situation or the meanings of words. It involves understanding sentence structure and the rules that govern phrase movement.

Another case that fits COS comes from literature on Creolization. Creoles are

languages formed by children who were exposed to their caregiver's artificial languages, called "pidgins," which their caregivers formed to communicate with others who speak different languages. Pidgins combine words from many different languages and have no consistent syntax: Pinker (1994) notes that there is "no consistent word order, no prefixes or suffixes, no tense or other temporal and logical markers, no structure more complex than a simple clause, and no consistent way to indicate who did what to whom" (p. 34). However, when the children of those caregivers socialize, they create a language, called "Creole," that contains linguistic universals, systematic structure, and syntactic properties such as auxiliaries, prepositions, case markers, and relative pronouns (Bickerton, 1983). Creoles provide an example of corruption of stimulus where there are plenty of degraded linguistic stimuli in the environment. Pinker (1994) uses cases of Creolization to argue that children come equipped with innate linguistic content that they cannot help but use to reinvent the language, noting that "Not content to reproduce the fragmentary word strings, the children injected grammatical complexity where none existed before, resulting in a brand-new, richly expressive language" (p. 33).

DOS and COS arguments conclude that two different kinds of traits are innate. In the DOS argument, there is no information in the *pld* from which children can infer any rules of grammar. If the child acquires knowledge of the rules, knowledge of those rules must have been innate and this knowledge may be triggered by minimal cues, or in response to missing data. The conclusion that children have innate linguistic content explains abundant knowledge of language, for there is no source to infer linguistic principles other than from an innate source. The COS argument deals with another puzzle: Unlike the case of Nicaraguan Sign Language, which illustrates that there is no relevant or reliable

linguistic stimuli in the environment, cases such as Creolization illustrate cases where there is a rich set of stimuli in the environment, but much of the stimuli is degraded or unusable for the purpose of inferring principles of grammar.

The COS argument does not conclude that linguistic principles are innate. It is not clear how a set of innate linguistic principles will help the child filter or sift through degraded stimuli, ungrammatical sentences, or nonlinguistic noises that simply cannot be used to infer linguistic principles. Instead, COS concludes that the Language Faculty works by focusing on linguistic stimuli relevant to acquire the right linguistic principles. This may work by sorting through information in the linguistic stimuli, *ignoring* faulty or ungrammatical data, thereby preventing the child from considering ungrammatical strings as quality samples for her grammar. Conversely, it may work by *priming* children to focus on only relevant linguistic stimuli to use as good input for grammar formation. For example, research indicates that from a very early age, children focus on linguistic sounds, as they look at people who are speaking in their environment (Gousti, 2004). These data show that there may indeed be innate processes that aid children in distinguishing between relevant and nonrelevant linguistic stimuli. Either way, the child must have some way of knowing what counts as relevant linguistic data.

2.4 Variety of Stimulus argument (VOS)

The third argument in POS is what I call "Variety of Stimulus" (VOS). Few scholars mention this particular argument; consequently there is no extensive discussion in the literature. However, there are a number of instances where the argument is alluded to in presentations of other POS arguments. When the argument is raised, it is treated as a

basic Poverty of Stimulus argument; however, as I show, it is not an argument about a poverty of linguistic stimuli. Rather, it concerns the fact that individuals are exposed to a wide variety of linguistic stimuli. I discuss the VOS argument in more detail in Chapter 4.

As Crain puts it, the problem is "to explain how different learners converge on similar mental representations on the basis of dissimilar environments" (1994, p. 365). The argument concludes that children acquire language in specific ways, and which lead all children to end up with the same end-state, namely knowledge of syntax universal to all languages. The basic form of VOS argument is:

> (P1) Children converge on the same set of knowledge of language that includes Universal Grammar (i.e., linguistic principles universal in all natural languages).
>
> (P2) Knowledge of linguistic principles is either learned from linguistic stimuli in the environment, or it is innate.
>
> (P3) Each person is exposed to a different set of linguistic stimuli in the environment.
>
> (P4) If knowledge of linguistic principles is learned, then each individual would acquire different knowledge of principles respective to the linguistic data in their environment.
>
> (C) Knowledge of Universal Grammar innate.

The VOS argument is a puzzle about how children are exposed to different stimuli, such as English, Mandarin, German, or Chinese, and yet they all reliably acquire knowledge of the same set of linguistic principles. The following several observations used to support Variety of Stimulus arguments:

> (1) *Ubiquity*: All normally developed children acquire language. Every society develops language, even in isolation from other societies. Whether the explanation is a nativist or empiricist one, language is an important part of our human nature.

> (2) *Linguistic Universals*: Children acquire knowledge of certain linguistic principles that are universal to every (or most) natural language. The similarities cannot be explained by a spreading of one or two ancestor languages.
>
> (3) *Convergence*: Different children acquire the same grammar, despite variations between each of their environments, *pld*, and upbringing.

The broadest set of data appeals to the fact that the acquisition of language is not confined to a handful of circumstances but is *ubiquitous*. Every society develops a language (or languages) that has complicated syntax, morphology, and phonology, and societies develop languages independent of interaction with other societies. At the individual level, the same is true: All children acquire language (under normal development) and language acquisition is ubiquitous (Pinker, 1994). This observation suggests something about humans that have the capability for acquiring, understanding, and producing language.

The claim in Premise (4) is targeted against empirical theories of learning. Children are exposed to a small sample of linguistic stimuli from their environment. If children acquire knowledge of language from the environment, as empiricists argue, then each child would acquire a different set of knowledge that reflects this sample. In addition, a child might be faced with environments that contain degraded samples, incomplete or nonsystematic languages such as pidgins, or impoverished linguistic stimuli. If children simply learn from the stimuli in their respective environments they should acquire a *different* set of knowledge contingent on the environment. Thus, if empiricism is true, we would predict much more diversity in languages than we see. However, children acquire language with ease, as if these differences do not affect their development. Gleitman and Newport (1995) explain this puzzle as follows:

> Under widely varying environmental circumstances, while learning different languages within different cultures and under different conditions of child rearing, with different motivations and talents, all normal children acquire their native tongue to a high level of proficiency within a narrow developmental time frame. (p. 1)

In addition, children acquire a specific set of knowledge of language that contains *linguistic universals,* aspects of language (syntactic, morphological, and phonological rules); these are universal to every fully developed language (Baker, 1981; Chomsky, 1986; Pinker, 1994).[ix] Linguistic universals are an aspect of language that empiricists have difficulty explaining. One might be tempted to explain some similarities across language by appealing to ancestor languages, the way in which Spanish and French share Latin roots (see Evans & Levinson, 2009, and Putnam, 1967, for discussion). However, there is no single ur-langauge from which all other languages arise. An example of a linguistic universal is the Verb-Object Constraint that states that objects immediately follow the verb whereas subjects do not (Baker, 1981). Even where there are apparent differences such as Turkish or Japanese that have Subject-Object-Verb order, the difference is still consistent with the Verb-Object Constraint.

Convergence refers to the fact that children acquire the same knowledge of language (Pullum & Scholz, 2002). Children acquire knowledge of linguistic universals, despite variations between each of their environments and upbringing. What needs to be explained is how children come to converge on the same knowledge. The puzzle also applies to children raised in similar environments. Given that each child only has access to a unique subset of linguistic stimuli, we can ask how it is that children acquire the same knowledge in a *similar* environment such as children raised around English-speakers. We can compare children from a similar population and still find a VOS puzzle.

Ariew (2006) puts it as follows:

> [D]espite exposure to significantly different samples of data, different children in the same linguistic community end up adopting essentially the same linguistic intuitions, and thus, it is plausible to suppose along with Chomsky that they innately possess essentially the same grammar (more specifically, innate in the context of language cues in the environment of the learner). (p. 8)

The VOS argument also draws on observations that children's internal constitution varies. If children learned language, we would expect variation in speed and ease of acquisition, depending on individual's motivation or intelligence. Laurence and Margolis (2001) argue:

> Plus, children vary as well. If a general learning mechanism is responsible for language acquisition, one would expect to see wildly different outcomes from child to child. Children differ in IQ, responsiveness to environmental cues, eagerness to learn, attention span, memory, and a number of different factors that might affect development of learning mechanisms. The problem that the empiricist faces is not merely to explain an ideal child's learning, but to explain how it is that every child, regardless of the differences in the environment or their own constitution, receives the right kinds of negative data to acquire language. (p. 231)

An ideal learner, they argue, may be able to infer complicated principles, and to spot indications of when they are making errors (i.e., they could use indications such as indirect negative evidence). However, children vary: They may miss important indications, learn at a different pace, or focus on different linguistic stimuli, including nonrelevant features such as cadence, rhythm, volume, and other factors not relevant for grammar formation. It is remarkable, then, that children converge on the same knowledge.

The argument concludes that there must be an innate set of linguistic knowledge that includes linguistic principles, which accounts for the child's ability to acquire the same set of linguistic knowledge. I argue in Chapter 4 that the VOS argument also raises

important inferences about acquisition and development. A more complicated VOS argument involves unique empirical observations about linguistic diversity, insensitivity to variation, and predictable stages of development.

2.5 Poverty of Negative Evidence argument (PNE)

The fourth argument is the Poverty of Negative Evidence (PNE) argument, which is traditionally found in discussions of the "Logical Problem of Language Acquisition." The Logical Problem deals with problems of underdetermination, a problem about how there are an infinite number of possible grammars that are compatible with the linguistic stimuli to which children are exposed. Children may infer possible linguistic rules, but many are incompatible with the target natural language they are trying to learn. The question is how the child converges on the exact grammar needed to fit her natural language. Scholars mostly succeed in keeping the Logical Problem separate from the suite of POS arguments, but they often appeal to empirical observations that are better accounted for by the other POS arguments.[x] Many problems of underdetermination are not strictly about the quantity or quality of the linguistic environment; however, one particular argument, the Poverty of Negative Evidence (PNE), focuses on a poverty of negative evidence, i.e., indications in the environment about when children have made linguistic errors. Negative evidence is crucial if language acquisition is a matter of learning and inferring correct linguistic principles.

Thus, there are two distinct arguments within PNE: The first emphasizes how children do *not* make certain kinds of grammatical errors—its concern is with prevention of missteps. The second focuses on how children are able to *correct* faulty grammar

when they do make such errors—it focuses on recovery from such missteps. Each argument concludes that there must be constraints on the hypotheses that children form about which linguistic principles are correct. I call the first the "Prevention PNE Argument":

> (P1) Children acquire knowledge of correct linguistic principles, among an indefinite range of possible sets of linguistic principles.
>
> (P2) Children either learn which linguistic principles are correct, or they have constraints on the kinds of linguistic principles they can acquire.
>
> (P3) If children learn which linguistic principles are correct, they would make many errors about linguistic principles, errors that are compatible with information in the *pld*.
>
> (P4) Children do not make tempting errors despite lack of negative evidence.
>
> (P5) If children make errors, there must be "negative evidence," i.e., indications in the environment about which principles are incorrect.
>
> (P6) There is a poverty of negative evidence.
>
> (C) Children have innate constraints on the kinds of linguistic principles they can acquire.

The Prevention PNE refers to the puzzle about why children do not make certain kinds of errors that we would expect if they were learning solely from their environment. It views children in the position of tacitly deciding between a set of possible linguistic principles, of which only one will match the target natural language (i.e., English, French, etc.). When children are forming hypotheses of linguistic principles, they may hypothesize that their acquired grammar fits the target language. However, there are many hypotheses that are compatible with linguistic data in the *pld* and hence, grammars are *underdetermined* (Cowie, 1997; Gold, 1967; Garfield, 1994). Garfield (1994) paraphrases the problem: "the examples of the target language to which the learner is

exposed are always jointly compatible with an infinite number of alternative grammars and so vastly underdetermine the grammar of the language" (p. 369).

> The argument uses these sets of empirical observations listed below:
>
> (1) *Poverty of Negative Evidence*: Children acquire language despite lack of "negative evidence" (some indication that the child has made an error). Children are rarely instructed or corrected when they make errors, and are resistant when they are corrected.
>
> (2) *Underdetermination*: Children arrive at theories (grammars) that are highly underdetermined by the data.
>
> (3) *Error Avoidance*: Children's grammatical errors exhibit predictable patterns. There are some errors that are never made, and there seems to be constraints on the kinds errors that can be made.
>
> (4) *Self-Correction*: Children self-correct their faulty grammar despite negative evidence.

Scholars show how there is a *poverty of negative evidence* (Crago, 1992; Crain, 1994; Marcus, 1993; Pinker, 2013a; Ramscar & Yarlett, 2007). "Negative evidence" refers to linguistic features in the environment that would indicate to the child that she has formed faulty grammar, such as caregivers telling the child that her sentences are ungrammatical. Haegeman gives an example involving the that-trace effect, noting that children are not explicitly taught that sentences like, "Who did they think that was available?" are ungrammatical (Haegeman, 1994, p. 10).

These examples show that there must be something preventing children from forming faulty hypotheses of linguistic principles. Many scholars conclude that children have innate "knowledge of language," in their formulations of the argument. However, this conclusion is imprecise, since it is not clear how children would acquire correct linguistic principles, or self-correct their faulty grammar by having an innate linguistic principles such as principles of UG. An innate set of linguistic knowledge does not explain

empirical observations about error-avoidance or self-correction. It is more precise to say that the PNE argument concludes there is some innate process or constraint that works to aid the child in acquiring correct linguistic principles, or that aid the child in self-correcting faulty principles.

In an absence of negative evidence, there needs to be some way to prevent the myriad ways in which children could form faulty hypotheses in the first place. A few scholars talk about these in terms of *constraints* that automatically narrow down the number of hypotheses under consideration and aid in quicker acquisition of the target grammar (Crain, 1991, 1994; Crain & Nakayama, 1987; Cowie, 1997; Hsu & Chater, 2010; Saffran, 2002). For example, Crain (1991) concludes:

> What is innately given is knowledge of certain restrictions on the meanings that can be mapped onto sentences as well as restrictions on the sentences that can be used to express meanings. This knowledge is encoded in constraints. The problem for the learner is that there are no data available in the environment corresponding to the kinds of negative facts that constraints account for. (p. 396)

Having such constraints would prevent the child's forming false hypotheses, and would automatically narrow down the number of hypotheses under consideration and aid in quicker acquisition of the target grammar. Thus:

> Since introducing constraints into a grammar restricts the language, constraints have the desired effect of handling the 'overgeneration' problem that would otherwise arise from highly general rules…Constraints reduce the number and kind of hypotheses children can entertain in response to their linguistic experience. By circumscribing the hypotheses children must contend with, constraints make direct, rapid acquisition less mysterious. (Crain, 1994, pp. 367-368)

The second PNE argument asks how it is possible that correct grammar is acquired given that children *do* form faulty hypotheses of linguistic principles. It concludes that there must be some innate constraints that must be guiding the child to correct faulty

grammar. I call this the "Correction PNE Argument":

> (P1) Children self-correct themselves when they acquire faulty hypotheses of linguistic principles.
>
> (P2) If language is learned from the environment, then given the nature of the available data, learners would overgeneralize and form faulty linguistic principles.
>
> (P3) If children do form faulty linguistic principles, there must be some means of correcting faulty linguistic principles.
>
> (P4) Either acquisition of faulty linguistic principles is corrected by negative evidence in the environment (i.e., indications from the environment about when children produce ungrammatical sentences), or there are innate means of self-correcting faulty linguistic principles.
>
> (P5) There is a poverty of negative evidence in the linguistic stimuli.
>
> (C) Children have innate means of self-correcting faulty linguistic principles.

Children do overgeneralize, but one would expect much more overgeneralization than is actually observed. For example, they may go through stages of over-regularizing verbs (e.g., saying "goed" rather than "went"). Quintero (1992, p. 60) gives a case where children may infer from the well-formed sentences (a) and (b) that the word "with" can be combined with a verb "hit" or "play." Since "read" is also a verb, as indicated by sentence (c), it seems that children might incorrectly infer that (d) is a well-formed sentence:

> (a) What did the little girl hit with the block today?
>
> (b) What did the boy play with behind his mother?
>
> (c) What did the boy read a story about this morning?
>
> (d) What did the boy read with a story this morning?

Sentence (d) is an example of overgeneralizing. However, Quintero notes, children

either do not form such sentences, or when they do they are able to correct themselves without external prompts (Marcus, 1993; Ramscar & Yarlett, 2007). Children's grammatical errors exhibit predictable patterns, and are short-lived. In addition, there are some errors that are never made. These are surprising facts if we assume that learning is a matter of hypothesis testing or induction. MacWhinney (2004) explains:

> Without negative evidence, children are unable to retreat from an overly general grammar to the correct limited grammar. If we believe that negative evidence is not available, and if we believe that children can either escape initial overgeneralization or recover from overgeneralization once it has occurred, we must then assume that they do this by relying not on information that is available in the language they are hearing, but on guidance from other linguistic or cognitive structures. (p. 885)

Many scholars cite cases where caregivers fail to correct children's ungrammatical speech as evidence for Premise (5), thus demonstrating a *poverty of negative evidence* (Brown, 1973; Brown & Hanlon, 1970; Crain, 1994; Crain & Thornton, 1998; Hsu, Chomsky, & Vitanyi, 2013; Pinker, 2013a; Ramscar & Yarlett, 2007). Further, even when feedback is provided, children often ignore it (Brown & Hanlon, 1970, Marcus, 1993; McNeill, 1966). An example from McNeill (1966) illustrates a dialogue where a child ignores feedback:

> Child: Nobody don't like me.
> Mother: No, say 'Nobody likes me.'
> Child: Nobody don't like me.
> (dialogue repeated eight times)
> Mother: Now listen carefully, say 'Nobody likes me.'
> Child: Oh! Nobody don't likes me. (p. 69)

In this case, even when negative feedback is given, the child seems to be "buffered" against the stimuli and determined to hang on to her faulty grammar. However, at some point, she will start saying "Nobody likes me," showing that children acquire adult competence despite not being consistently corrected for linguistic errors. In this case, it

does not seem appropriate to say there are *constraints* that restrict the acquisition of faulty grammar; Instead, it seems that children *correct* themselves, thus converging on formation of grammar appropriate to their natural language (Hsu, et al., 2013; Ramscar & Yarlett, 2007).

Crago's study (1992) of Inuit children demonstrates this. Inuit caregivers do not attend to infants or encourage language development, and young children are expected not to participate in discussions and are often ignored. Despite this lack of attention, the children develop language in the same way and with the same stages of development as children from cultures where they are not ignored or discouraged from participation in discussions.

Crago's case is traditionally cited as a case of Poverty of Stimulus to show how the environment is impoverished. However, on further examination it is clear that linguistic stimuli are not *deficient* in the *pld*. Children are raised around siblings and grandparents who converse in their presence, even if those conversations are not directed at children. In addition, children interact with other children and siblings their age, which gives them the opportunity to practice necessary formation of phoneme, morphology, and grammar formation. The problem arises not with the amount of linguistic input, but with the absence of corrective input the child uses form hypotheses of grammar.

An empiricist theory cannot explain how children would *learn* that their own grammar is faulty, if it is in fact faulty. For this, it appears necessary that child engage in conversations with competent speakers who can recognize ungrammatical sentences and then correct the child. However, Crago's study shows that this impression is mistaken: Children develop language without having such input from competent speakers. The

puzzle addressed by Correction PNE is that of explaining how children learn that certain sentences are *not* grammatical from an environment that contains only sentences that *are* grammatical.

An empiricist may respond by claiming that the environment is rich enough for children to infer what kinds of sentences are ungrammatical by using *indirect* evidence, i.e., the child infers that a sentence is ungrammatical if it is absent from the *pld* (Cowie, 1997; Prinz, 2012). Accordingly, children learn patterns of syntax spoken by competent speakers, and since competent speakers do not utter ungrammatical sentences, children do not learn to produce such ungrammatical sentences. For example, competent speakers do not say, "Is that woman who walking her dog is Tom's neighbor?" Since there are no such ungrammatical sentences in the *pld*, then children can infer that such a sentences would be ungrammatical, and would then refrain from producing it. This appeal to *indirect* evidence, however, does not work. Cowie (1997) explains that,

> If the learner can only learn from the environment, he must extrapolate on an extremely limited string of sentences. The learner has no information about an infinite number of other sentences, grammatical or not. There are infinitely many sentences and words that the learner may never hear, but are nevertheless grammatical sentences. This means that if the learner's hypothesis overgenerates, then, there would be no data that would correct it to the target grammar. (p. 21)

Since there are infinite sentences (both grammatical and ungrammatical) that children do not hear, they cannot infer that a sentence is ungrammatical on the basis of the fact that they do not hear those sentences. Thus, the absent of a sentence is not evidence that it is ungrammatical.

In sum, the PNE argument appeals to the fact that there is impoverished negative evidence in the environment. Empirical observations shows that children either do not

form faulty hypotheses of grammar in the first place (Prevention PNE argument), or, they can self-correct their faulty grammar (Correction PNE argument). The arguments conclude that the Language Faculty is constituted with constraints that prevent children from inferring hypotheses of faulty grammar, either by preventing the child from forming faulty syntactic rules or by self-correct when the child does overgeneralize rules.

2.6 Conclusion

The basic Poverty of Stimulus argument was originally designed by nativists to show that certain parts of grammar were not, or could not, be acquired by learning on the traditional empiricist model, and that any explanation of language learning requires that we posit a set of innate abilities or knowledge. In this chapter, I have shown that there are four arguments that are not equivalent and cannot be reduced to one single argument.

Each argument draws on a different set of empirical observations and each has a different conclusion. All arguments focus on the fact that the environment provides inadequate data for the children to infer the appropriate grammar, thus showing a problem for empiricist theories. Empiricist theories require good input in the environment in order for the child to infer the appropriate grammar.

Empiricists have difficulty explaining DOS cases, as there is a dearth of input from which children can learn. It may appear that DOS gives the strongest support for nativism, since it is striking that children acquire knowledge of language despite missing linguistic stimuli. However, true cases of a deficiency of stimulus are rare, and do not apply to common phenomenon where there is a rich source of linguistic stimuli. Examples of corruption of stimulus are abundant, as there are plenty of linguistic stimuli

on the basis of which to learn a grammar, even though the quality of that data is degraded. The environment provides a potential for faulty hypotheses of grammars. Thus, if the empiricist theory is correct, we would predict an abundance of faulty grammar. The COS argument concludes that there must be some innate means of distinguishing relevant linguistic stimuli, by filtering out nonrelevant linguistic stimuli, or by priming the child to focus on good data that can be used during grammar formation.

Although VOS arguments are rarely found in the literature, they draw on interesting empirical observations about ubiquity, convergence, and the acquisition of knowledge of linguistic universals. VOS concludes that knowledge of linguistic universals is innate. I argue in Chapter 4 that the VOS argument is the most complicated and powerful of the four arguments, as it focuses on unique phenomenon about language acquisition and development.

The PNE argument is a puzzle about how children form correct linguistic principles when there are no indications about whether their hypotheses are correct or not. The Prevention PNE argument concludes that there is some kind of constraint that prevents the child from forming false hypotheses about linguistic principles. The Correction PNE argument concludes that there is some way that aids children in self-correcting their faulty hypotheses when they do form them.

The key conclusion that Problems of Stimulus arguments draw is that there is something more involved in acquisition of language than learning from the environment: Children bring to the task some kind of knowledge of language, whether it is knowledge of rules or principles, or an ability to select good data from a noisy environment, or some kind of constraints that restrict acquisition of faulty grammar. Given that different

arguments conclude that different traits are innate, it is ambiguous to say that "knowledge of language is innate" without further explanation. In the next chapter, I examine the conception of "innateness" and the role it plays in theories of language acquisition.

2.7 Endnotes

[i] The acquisition of "language" is more accurately stated as the acquisition of "I-language" or "I-grammar." The "I" stands for "individual," "internal," and "intensional." The term "knowledge of language" is misleading in that it signals an epistemological view of knowledge. An alternative understanding of "knowledge" is analogous to saying that the immune system "knows" what antibodies to produce, or that the digestive system "knows" how to digest food (See Collins, 2007).

[ii] Empiricist research includes Connectionist Models, where computers are trained to "learn" rules of grammar from stimuli, and these experiments had some successes in advancing the field (Elman, et al., 1999). Advances in Statistical Learning Approach use evidence from these earlier models and empiricists argue that nativists are not warranted in saying that there is a language specific acquisition device (Prinz, 2012).

[iii] Laurence and Margolis (2001) cite this example, which originally comes from Piatelli-Palmarini, 1980, p. 40.

[iv] I follow Chomskian tradition of using the term "acquisition" because it is neutral with respect to nativist and empiricist theories, biological or psychological in nature. This sense of the term is employed when we say that a person acquires knowledge of math via explicitly learning it; or a person possesses competence in forming questions by acquiring knowledge of grammar such as knowledge of the rules for question-formation. Thus, humans acquire knowledge of language, a trait that is not present at time 1, and present at time 2.

[v] Although many empiricists attack a "rule-systems" theory, no one actually defends this theory. For a concise review of how scholars misrepresent nativism, see Jenkins, 2004, p. 317-339.

[vi] See Jenkins, 2004, p. 317-339. He gives examples critiques of the Poverty of Stimulus argument given by empiricists, and then describes how empiricists misunderstand, misapply, or attack straw-man versions of the argument.

[vii] Linguistic content refers to language-specific rules or principles. Note that although empiricists often attribute to nativists the view that rules of syntax are represented in propositional format, no nativist holds this position (see Jenkins, 2000 for review). Chomsky's Principles and Parameters theory, for example, describes the Language Faculty as having principles and constraints that allow or limit the range of knowledge of language that is acquired. However, these principles cannot be said to be propositional rules or content. They describe procedures, not content.

[viii] Examples of principles of Universal Grammar include the following rules, which are

universal to all languages: sentences conform to structural constraints of the X-bar theory; transformations are limited by the principle of subjancency; Case is always assigned to noun phrases; anaphors are bound; and pronominals are free in their Governing Category. Other principles are: that-trace deletion; nominal compound formation, control, auxiliary phrase ordering, and the empty category principle (MacWhinney, 2004; see Evans & Levinson, 2009 for discussion; cf. Christiansen & Chater, 2008).

[ix] See Chomsky 1986, p. 46-7, and Pinker 1994, pp. 237-238 for explicit arguments for Linguistic Universals.

[x] The Logical Problem has been identified as "the projection problem" (Peters, 1972); "the Poverty of Negative Evidence of language acquisition" (Cowie, 1999; Pinker, 2013b); "Plato's problem" (Chomsky, 1989); and the "entailment problem" (Crain, 1994). The Logical Problem is a subset of problems about underdetermination in general: This way of framing language acquisition was introduced by Quine and discussion with Chomsky, Putnam, Goodman, Harman, and others (Piattelli-Palmarini, 1980). It has since spawned its own area of research, as it is a problem for both nativists and empiricists (for a review, see MacWhinney, 2004).

CHAPTER 3

INNATENESS AND LANGUAGE ACQUISITION

In Chapter 2, I examined four Poverty of Stimulus arguments. Each argument uses a different set of empirical observations, and draws a different conclusion about what trait is innate. The conclusion that "knowledge of language is innate" is ambiguous because there are heterogeneous properties of the conception of "innateness." This chapter addresses the relationship between the concept of "innateness" and its role in theories of language acquisition. I start by examining criticisms that the conception of innateness is confused and has no place in any theory of language or cognition (Section 3.1).

I suggest three considerations that any theory of language acquisition should address in order to give substantial explanation: Substantial theories should (1) identify proper sets of empirical observations, (2) describe the cognitive mechanisms responsible for the acquisition of traits (Section 3.2), and (3) describe interactions between internal factors of organisms and the environment (Section 3.3).

3.1 Is "innateness" a confused concept?

Several critics argue that the concept of "innateness" is useless in explanation because there are too many differences between the meanings of "innate" that create a level of confusion (Bateson & Mameli, 2007; Griffiths & Machery, 2008; Griffiths, Machery, &

Linquist, 2009; Mameli & Bateson, 2006). Mameli and Bateson (2006) categorized 26 distinct properties of the conception of "innateness" used in theories in the sciences, including the following: A trait is innate if it is genetically determined or encoded; it is highly heritable; it is not learned; it is species-typical; or it is a Darwinian adaptation (p. 177). Some properties of "innateness" appeal to traits that are homogenous, universal, species-specific, pancultural, or evolutionarily adaptive. In addition, traits can be innate because they have fixed ontogeny, are genetically determined, or have universal end-states (Pullum & Schulz, 2002). According to Gottlieb (2003), traits are said to be innate in the sense that they are present at birth; Stich (1973) argues that innate traits have a disposition to appear in the *normal* course of development; and Stromswold (2000) argues that innate traits appear during a *specific* developmental stage.

As an illustration of how the term "innateness" is ambiguous in cognitive linguistics, consider two different accounts of innateness involved in arguments described in Chapter 2: Deficiency of Stimulus (DOS) and Corruption of Stimulus (VOS). The DOS argument concludes that a set of knowledge of language is innate because it cannot be learned from linguistic stimuli. Accordingly, a trait is "innate" if it is "not learned." On this account, "innateness" is defined in relation to the linguistic stimuli (or lack of) in the environment in which the child is raised, emphasizing empirical observations such as *abundant knowledge*, *paucity of stimuli*, *productivity*, and *no explicit training*. The case of Nicaraguan sign language illustrates that when linguistic information is missing from their environment, children supplement content from an internal source. That content is innate because it could not be learned from the environment. Likewise, productivity refers to the ability to understand and create new sentences using knowledge of language

that cannot be inferred from an impoverished environment.

In their different ways, Cowie (1999), Fodor (1974, 1981), Prinz (2002), and Samuels (2002), each provide accounts of "innateness" that involve the property's "not being learned." In broad strokes, Fodor's (1981) account is that innate concepts are those that are not acquired by inferential processes; rather they are triggered "brute causally," i.e., there is no process of inference involved in the acquisition of innate concepts. Prinz (2002) describes "innateness" as "acquisition in the face of poverty of stimulus" (p. 193-194). Samuels' (2002, 2005) proposal is similar: Innate concepts are those that are "primitive," namely they cannot be explained by psychological processes involved in learning such as inference, induction, abduction, conditioning, etc. (see Section 3.3 for further discussion). Each theory takes the property of "not learned" to be crucial to the account of "innateness".

The account used in VOS uses a different conception of "innateness," involving properties such as *ubiquity*, *convergence*, *fixed ontogeny*, and *insensitivity to variation* (see Section 4.1). The conclusion is that a specific set of linguistic knowledge (i.e., universal grammar) is innate, where "innateness" refers to properties of "universality," "appears at a particular stage of development," and "canalized."

Mameli and Bateson (2006, p. 156) argue that when researchers use "innateness," they conflate these different properties, and assume that "the properties must somehow cluster," and this leads to a tendency to equate properties that are not in fact correlated. They urge, "it is wise not to bundle these different scientifically useful notions together under the label 'innateness' unless and until strong evidence is provided for the thesis that the properties…are strongly correlated."

Likewise, Griffiths (2002) argues that the concept of "innateness" is inherently confused. He suggests that the urge to cluster certain properties of "innateness" is a result of folk essentialism, the bias to believe that a species has an inherent nature or underlying essence. He categorizes the list of properties of "innateness" in terms of developmental fixity, species typicality or nature, and intended outcome or teleology, each of which finds basis in folk essentialism. Again the worry is that properties "get bundled together in the innateness concept," which leads to inferring that if one property is present, then others must be as well (p. 60). If researchers rely on their folk intuitions about what counts as "innate," it may lead to "unfortunate" and "pernicious" effects on their theories (p. 76).

Using one term, "innateness," is ambiguous and contributes to confusion. In fact, Griffiths goes so far as to say the concept is so inherently confused that we should eliminate its use as a scientific term. Instead, he suggests, we should substitute "innateness" with whatever properties are being cited. He says,

> Substituting what you actually mean whenever you feel tempted to use the word 'innate' is an excellent way to resist this slippage of meaning. If a trait is found in all healthy individuals or it is pancultural, then say so…If it is developmentally canalized with respect to some set of inputs or is generatively entrenched, then say that it is. (Griffiths, 2002, p. 82)

I agree that appealing to "innateness" does no explanatory work if it is not defined appropriately, and indeed that the term "innateness" should be avoided if it causes confusion. To avoid using "innateness" ambiguously, scholars should define it and indicate in what way traits are innate.

In addition to the problem that "innateness" is ambiguous, theories that conclude traits are innate are insubstantial. Bateson and Mameli (2007) observe that scholars often

"avoid providing detailed developmental explanations" (p. 823). Scholars tend to ignore the developmental story and instead rely on biologists to fill in the details. This view is expressed by Elman, et al. (1999) who see that "calling a behavior innate does very little to explain the mechanisms by which that behavior comes to be inevitable. So there is little explanatory power to the term" (p. 21). What we are missing is a more detailed story about the nature of the developmental processes involved in acquisition.

3.2 Three considerations for more substantial theories

Theories of language acquisition are often limited to acquisition of language in the development of individuals, and thus focus on ontogeny as a proximate explanation. This is only one aspect in the investigation of language involved in ultimate explanations of all aspects of linguistic cognition, including language acquisition, comprehension, and production. Questions about phylogeny, for example, or peripheral aspects of the Language Faculty are part of an ultimate explanation (Tinbergen, 1963; see Mayr, 1982). Ultimate explanations may include the following:

> (1) What the knowledge of language is, i.e., the constitution of the initial state of the language faculty.
>
> (2) How knowledge of language is acquired, i.e., how it develops from the initial state to the mature state.
>
> (3) How knowledge of language is represented.
>
> (4) How the production systems work, i.e. how knowledge of language is used.
>
> (5) Peripheral systems in the language faculty, including auditory or visual systems, etc. that play a role in the comprehension or production of language. Chomsky refers to this as the "Language Faculty broad" (Fitch, 2009; Hauser, et. al., 2002).

> (6) How the language faculty evolved (Berwick, et al., 2013; see discussions in Christiansen & Chater, 2008).
>
> (7) Sociological aspects, including the role of language variation and language change over time.

Current theories of language acquisition can add to an ultimate explanation only if they provide substantial explanations. They should explain empirical observations about children's abilities, rather than merely gesturing that knowledge of language is innate.

If a theory of language acquisition is to be substantial, it should explain specific sets of empirical observations, thus avoiding lumping together traits of organisms that do not actually correlate. In addition, I argue that the study of language should be interdisciplinary, as it will involve theories of psychology and biology. I propose three considerations that theories of innateness can address if they are to offer a substantial explanation. These considerations are hallmarks of a complete and unified theory of language. They suggest ways that scholars can offer a more complete account of language acquisition. These considerations are the following:

> (1) *Empirical Consideration* (EC): An account of language acquisition must give careful attention to different sets of empirical observations. It should not oversimplify language acquisition by underestimating the diversity of observations to be explained.
>
> (2) *Developmental Consideration* (DC): A theory of language acquisition should describe cognitive mechanisms that underlie traits and the developmental processes play a causal role in trait acquisition.
>
> (3) *Interaction Consideration* (IC): A theory of language acquisition should include information about the interaction between factors internal to the organism and environmental conditions.

The *Empirical Consideration* (EC) requires that an account of language acquisition give careful attention to different sets of empirical observations. Often, arguments for innateness refer to a large array of empirical observations without noting the differences

between them, and use "innateness" as a panacea to solve all the explanatory difficulties. This is why Mameli and Bateson (2007) along with Griffiths (2002) criticize the use of "innateness" because it is a heterogeneous category composed of diverse properties. EC addresses this concern by not underestimating the diversity of data that needs to be explained. Different arguments conclude that specific traits are innate, and those conclusions should be supported by the right kinds of empirical observations.

Substantial theories of acquisition can explain why certain I-properties correlate. It seems clear that certain properties of organisms correlate with each other. For example, a trait such as facial recognition is *ubiquitous* in humans and this trait emerges *a very early age,* with *limited exposure to environmental stimuli*. Facial recognition *is not learned*. As another example, the fact that children acquire certain linguistic abilities during a particular stage of development correlates with fixed end-states such as acquiring knowledge of Universal Grammar. These correlations indicate that there is something about the organism that is responsible for the acquisition of those traits.

In order to meet the *Developmental Consideration* (DC), a substantial theory will provide an explanation of developmental processes and offer a theory of acquisition. A conclusion such as "knowledge of language is innate," does not actually explain acquisition. Poverty of Stimulus arguments tell us how linguistic traits are *not* acquired (i.e., it is not acquired via learning); it does not tell us how the traits *are* acquired. There are many kinds of acquisition processes that fall under the heading of "innate acquisition," including triggering, maturation, present at birth, and canalization. A beard is acquired at a particular stage in maturation. The ability to digest lactose is "innate" in the sense that the body produces lactase, though to do so requires exposure to lactose

from the environment. The ability to detect objects is a result of automatic processes of the visual system, and requires relatively minimal exposure to visual stimuli. Other traits require exposure to stimuli. Even "innate" abilities as recognizing the persistence of objects or face recognition take a period of "training" (e.g., Elman, et al., 1996, p. 108). Thus, there may interactions between "innate" traits internal to the organism and environmental stimuli, and these interactions are involved in complicated theories of development. What is required is a fuller description of the kinds of cognitive mechanisms or processes that play a causal role in the emergence of traits.

DC requires a more detailed theory of acquisition that will explain the specific trait concluded to be innate. The four Poverty of Stimulus arguments conclude that different traits are innate. The cases used in DOS arguments, such as Nicaraguan Sign Language and Simon (see Chapter 2), show that when linguistic information is missing from their environment, children supplement content on their own. It concludes that a specific set of linguistic content must be innate because it cannot be learned from the environment. The underlying structure must consist in language-specific content that seems to lie dormant in the child and then comes online at a certain stage of development. Given that children could not have learned the linguistic content from the environment, the kind of mechanism that would explain the acquisition is one where innate content emerges via maturation (Chomsky, 1993).

In COS, the cases show that children are exposed to degraded quality of stimuli such as ungrammatical sentences or nonlinguistic sounds. It concludes that children have the ability to ignore nonlinguistic or irrelevant input, and to attend to linguistic stimuli relevant for grammar formation. In the case of Creoles, children are exposed to pidgins

that exhibit nonsystematic syntactic structure, and hence would fall under the category of degraded linguistic input. An empiricist may offer a theory of learning in which the child infers correct syntactic rules from this input, with a mechanism of statistical learning (Behme & Deacon, 2008; Prinz, 2012).

VOS conclude that linguistic content is innate but can be triggered by linguistic stimuli (Khalidi, 2007; Lightfoot, 2013). In VOS arguments, empirical observations focus on *convergence* or *predictable stages of development*. These data direct our attention to development and they require an explanation in terms of the developmental regularities or abnormalities. In PNE arguments, the empirical observations include lack of negative evidence in the environment, selectivity, and underdetermination, and the child's ability to avoid errors that would lead to faulty hypotheses of grammar. It concludes that there are innate constraints on the kinds of linguistic principles that children can form.

It is interesting to note that COS and PNE arguments conclude that knowledge of language emerges in response to the child's environmental stimuli (or lack of stimuli). The evidence used involves traits that emerge in response to the input (or lack of input) from the environment in which the organism was raised. In contrast, DOS and VOS arguments conclude that knowledge of language develops independent of the quality or quantity of environmental stimuli. Taken together, the four POS arguments point to different traits that are acquired, and scholars must be careful not to use one theory of acquisition to explain them all.

The *Interaction Consideration* (IC) addresses the value of developing theories that give more detailed explanations of the acquisition of traits in terms of the interaction between internal features of an organism and environmental influences in the

development of traits. Language acquisition involves a hybrid of innate traits plus acquisition via learning. We may discover that the best explanation for, say, the creation of Creole languages is that children have innate linguistic content. On the other hand, the best explanation for how we learn the lexicon of particular natural languages may be statistical learning or induction (Behme & Deacon, 2008; Prinz, 2012).

In addition, IC urges an interdisciplinary approach to the study of language that involves explanations from both behavioral sciences and hard sciences. Given that cognitive traits are subject to the same constraints in development and acquisition as other (biological) traits, an ultimate explanation of language will involve knowledge from a number of disciplines, and entail multilevel theories of acquisition, including theories from psychology and biology.

3.3 Biologicizing the mind

It may seem that cognitive linguists would embrace an interdisciplinary approach to the study of language, given that ultimate explanations will involve multilevel theories. However, there was a movement in cognitive science that hindered such interdisciplinary collaboration, called "psychologizing the mind." Supporters of this movement argued that cognitive traits should be explained with higher level theories in behavioral sciences, not with lower level explanations from biology (Cowie, 1999; Devitt & Sterelny, 1999; Fodor, 1974, 1981; Prinz, 2002; Samuels, 2002, 2007). They also argued that it is the job of biologists to explain the nature of innate traits. Consequently, cognitive linguists argued that certain traits are "innate" without further explanation.

The idea behind "psychologizing the mind" is that cognitive traits such as concepts or

knowledge are a special sort of property and should be explained by higher level psychological theories independent of lower level theories in biology, physics, etc. Accordingly, the acquisition of traits such as knowledge or information-bearing mental states are explained by different theories than those used to explain the acquisition of biological traits such as beards or legs. This approach divides psychological explanations from biological explanations. I use Samuels' (2002, 2007) account of "innateness" as an example that most represents "psychologizing the mind." His theory, "Psychological Primitivism" (PRIM), maintains that if there is no psychological theory (of learning, induction, etc.) that can explain a trait's acquisition, then it is innate.

> Psychological Primitivism (PRIM): If a cognitive structure S is innate, then scientific psychology can specify no mechanism or process in virtue of which S is possessed by an individual organism O at a given time t, even though there is no time prior to t at which O possesses S. (Samuels, 2002, p. 251)

A psychological theory may include perception, induction, deduction, conditioning, statistical learning, etc. We may add abduction, inference, or pattern recognition (Behme & Deacon, 2008; Prinz, 2012). A trait is psychologically primitive if "no correct psychological explanation of its acquisition exists" where a "correct psychological explanation" would come from any possible psychological theory (Samuels, 2007, p. 26). A trait is "innate" on Samuels' view, if it is primitive: that is, if an organism possesses a trait at time 2, which it did not have at time 1, and the mechanism for its acquisition cannot be explained by any psychological theory. For example, if we find no psychological theory for the acquisition of face recognition or depth perception, then they are innate.

The worry with Samuels' proposal is that it decides what kinds of tasks would be

appropriate explananda for psychological theory and what is appropriate for biological theory. His approach is one instance of the general claim that cognitive traits require a special sort of explanation, separate from biology. Collins (2008) criticizes this attitude because it is "dualistic in the sense that it presupposes that the methods appropriate for the investigation of the natural world cease to apply when one is dealing with phenomena whose locus is above the human neck" (p. 17).

In contrast to the approach of "psychologicizing the mind," a movement called "biologicizing the mind" seeks to explain the development of traits, whether psychological or biological with the same methodology. "Biologicizing the mind" was introduced by Ariew (1996, 1999, 2006) and supported by many cognitive linguists (Birch, 2006; Berwick, et al., 2013; Chapman, 2000; Chomsky, 2005, 2007, 2012; Collins, 2011; Fitch, 2005, 2009; Hauser, et al., 2002; Hauser & Bever, 2008; Jieqiong, 2014). "Biologicizing the mind" is an approach that seeks to explain any trait with the same empirical methodology, psychological or biological in nature (Ariew, 1996, 1999). Ariew, for example, proposes that we study language acquisition as we would "legs and livers." The reason behind this view is in recognizing that all traits develop from an interaction between genes and the environment: We should not think about the organism's genotype and the development of phenotypes independent of the environment. This applies to both biological and psychological traits, including the growth of the brain, face recognition, pretend play, acquiring beards during puberty, the immune system's reaction to viruses, an accumulation of bacteria in the gut essential for digestion, and the acquisition of bird song.

Biologicizing the mind satisfies the Interaction Consideration. It urges explanations in

terms of higher level psychological categories and concepts as well as lower level biological categories and concepts. Collins (2005) states this sentiment as:

> All else being equal, an account of the innate in cognitive science should also cover non-cognitive structures that are claimed to be innate in scientific theorizing generally…An account of the innateness of cognitive structure that does not depict the structure as falling together with biological structure more generally will have failed to depict adequately what is innate about the cognitive structure. (p. 166-167)

In other words, the approach should be the same in a study of any trait, whether biological or psychological in nature.

The approach of biologicizing the mind takes seriously an interaction between genes and the environment. It is a truism that traits develop in combination with internal factors (genes, for one) and external factors (environmental stimuli). This truism is not taken seriously in the above approach of "psychologizing the mind." A theory of psychological traits should include an explanation of biological processes as well and a theory of acquisition should explain the developmental processes and interaction between the genome and environment. The ability to ride a bike, walk, or develop vision involve both psychological and biological processes, and an explanation of each trait's development should include relevant influences from both environmental stimuli and internal factors (e.g., genome).

Another consequence of considering the interaction between internal and external factors in development is that a dichotomy between nature and nurture breaks down. Jenkins (2004) explains:

> There is no nature-nurture debate in biolinguistics…Every approach to biolinguistics recognizes the critical role of both internal mechanisms and of environment. For similar reasons, it make no sense to talk about a "nativist" approach based on notions of innateness standing apart from other approaches. (p. 324)

The dichotomy between "nature vs. nurture" or "genes vs. environment" leads to simplified theories of the complex processes involved in development.

Biologicizing the mind also emphasizes the need for cooperation between the sciences. Hauser et al. (2002) are hopeful of an integrated approach to the study of language:

> Linguists and biologists, along with researchers in the relevant branches of psychology and anthropology, can move beyond unproductive theoretical debate to a more collaborative, empirically focused and comparative research program aimed at uncovering both shared (homologous or analogous) and unique components of the faculty of language. (p. 1578)

Biolinguistics is an attempt at improving communication between sciences in an effort to explain the complexity of language acquisition. A complete theory of language acquisition may involve unifying theories to integrate higher level and lower level explanations, or it may retain multilevel explanations. Either way, biolinguists aim to:

> Study a real object in the natural world—the brain, its states and functions—and thus to move the study of the mind towards eventual integration with the biological sciences. (Chomsky, 2004, p. 388)

3.4 Conclusion

In this chapter, I offered examples to show how the conception of "innateness" is ambiguous at best, and inherently confused at worst. To avoid confusion, scholars should defining "innateness" in their theories.

I offered three considerations that any theory of language acquisition may address to provide substantial explanation. EC urges scholars to explain specific sets of empirical observations and DC requires a more detailed theory of acquisition that will explain the specific trait concluded to be innate. Ultimately, a good theory of language acquisition

must explain what it is about humans that allow them to understand, parse, and produce sentences with systematic structure. This will involve multilevel theories from biology as well as psychology, and thus an interactionist approach to language acquisition is required (IC).

In the next chapter, I develop an account of language acquisition in terms of Canalization. An analysis of "innateness" given in terms of Canalization meets the criticisms of vacuity and ambiguity, thus establishing a place in our understanding of the complex process of language acquisition.

CHAPTER 4

A CANALIZATION ACCOUNT OF LANGUAGE ACQUISITION

In Chapter 2, I argued that there are four Poverty of Stimulus arguments. Each argument addresses a specific set of empirical observations, and each reaches a different conclusion about what is innate. A theory of language acquisition that attempts to explain *all* of the empirical evidence by a generalized appeal to "innateness" ends up being insubstantial. At the same time, a theory of language acquisition should explain observations about children's linguistic competence and how they come to acquire their knowledge of language. From this, it follows that there needs to be a more nuanced and complex treatment of any innate component of language acquisition. The suite of POS arguments may point in the right direction for this; they constrain the possible explanations of empirical observations about children's competence within the realm of the innate, and they may help define the different senses of "innate" that are used. In this chapter, I develop an account of language acquisition that explains the empirical observations in VOS arguments.

The VOS argument is important as it addresses several sets of empirical observations about children's linguistic competence. Scholars have underestimated the value of the VOS argument; however, as I show, it is important because it involves unique empirical observations about acquisition and development. It is the only argument that draws on

phenomenon about linguistic diversity, fixed ontogeny, and insensitivity to variation. It also challenges empiricists and nativists in different ways: Empiricists must explain how children converge on the same knowledge; nativists must explain why there is so much diversity in natural languages. Finally, the VOS argument highlights the importance of development, and interaction between environmental stimuli and internal factors. In Section 4.1, I discuss these features of VOS arguments.

In Section 4.2, I give an account of language acquisition in terms of canalization. I define canalization and combine it with Chomsky's Principles and Parameters account of language acquisition (Chomsky, 1986; Chomsky & Lasnik, 1995). In Section 4.3, I argue that the resultant account offers the best explanation of empirical observations used in VOS arguments. I also argue that while canalized traits may development in two ways, fixed ontogeny and plastic ontogeny, language acquisition demonstrates plastic ontogeny. In Chapter 5, I will discuss a debate concerning Ariew's (1999, 2006) suggestion that "innateness" is best defined in terms of canalization.

4.1 Six features of the VOS argument

The VOS argument highlights some unique features of language acquisition. I will review the three sets of empirical observations about language acquisition discussed in Section 2.5 and then add three more sets empirical observations. Only by taking in account of all six sets can we hope to give a complete explanation of language acquisition.

In Section 2.5, I described three sets of empirical observations:

> (1) *Ubiquity*: All children acquire language and every society develops language.

> (2) *Linguistic Universals*: Children acquire knowledge of certain aspects of language (phonological, morphological, or syntactic principles) that are universal to every natural language.
>
> (3) *Convergence*. Children acquire the same knowledge of grammar (i.e., linguistic universals), despite variations between each of their environments and upbringing.

To these, I add three more sets of observations:

> (4) *Linguistic Diversity*: Principles of phonology, morphology, and syntax of natural languages vary between each other, and change over time.
>
> (5) *Fixed Ontogeny*: Language acquisition develops in ordered stages, and in relatively the same pace and pattern, with predictable errors.
>
> (6) *Insensitivity to Variation*: Children are insensitive to some range of stimuli that would impede acquisition of knowledge of linguistic universals.

Although other POS arguments may refer to these sets of empirical observations, the VOS argument uses them in a unique way. The feature of *ubiquity* invokes a sense of "innateness" that is species-specific, universal, monomorphic, and pan-cultural. These involve a sense of "innateness" that apply species-wide, and indeed, VOS is the only argument that has us explain how *populations* of children acquire knowledge of language. In DOS, COS, and PNE, the puzzle is to explain how a child can acquire knowledge of language given that the linguistic stimuli do not provide sufficient information. These arguments ask how an individual acquires knowledge, and there is no need to compare different individuals.[i] However, the appeal in VOS arguments comes from noting that the environment varies *between* children. It is necessary to compare populations of children and take account of their similar (or dissimilar) traits and patterns of development. The theory must explain how *any* child can acquire language raised in *any* environment (e.g., whether the child is raised around pidgin-speakers, French-

speakers, multilanguage environments, ASL, etc.), and it must also explain how, despite this, all children converge on the same sets of knowledge of language.

4.1.1 Linguistic diversity

The empirical observations of *ubiquity*, *linguistic universals*, and *convergence* are used to support claims that certain sets of knowledge of language are innate. The idea behind this is that *all* children must come equipped with the same knowledge of linguistic universals. Nativists argue humans have innate knowledge; in light of the evidence, this seems like a reasonable conclusion. It is tempting to think that the more impoverished the environment, the greater the amount of content that must be posited as innate. However, if children have an abundant source of innate content, then we would expect to see more uniformity in languages across the world. Positing too much innate content presents problems in our account of the variation in languages, and indeed, it is the variety of language that creates the puzzle of VOS in the first place. Thus, we have another source of empirical observations that needs to be addressed:

> (4) *Linguistic diversity*: The phonology, morphology, and syntax of natural languages vary between each other, and change over time.

Linguistic diversity is an issue that is often neglected in Poverty of Stimulus debates (Christiansen 2003; Christiansen & Chater, 2008; Evans & Levinson, 2009). There are over 6,000 natural languages that differ in lexicon, phonology, morphology, and semantics. Indeed, it is the variation in natural languages that is at the heart of the challenge which VOS arguments present, both for nativist and empiricist accounts. The diversity at issue is not that found in dialect, speech patterns, intonation, rhythms, and so on; these are superficial and have no impact on the universal grammar that is present in

every language.

If, as nativists suppose, humans come equipped with a rich source of linguistic knowledge, then we would expect to see *more* uniformity in natural languages. After all, if language acquisition were a matter of triggering innate rules, then the products (i.e., natural languages) would reflect what is innately known, and if innate knowledge were uniform, it would follow that natural languages would be uniform. However, this is not so: Language does not develop like other innate traits such as the growth of limbs, whose phenotypic expression varies little between individuals. The diversity in language indicates that it is not a matter of triggering innate rules. The theory of language acquisition must explain the fact that people have the capacity to produce a wide variety of natural languages. Chomsky (1986) notes this problem when he states that an adequate theory of the language faculty should be:

> [R]ich enough to account for the attested variety of languages and, indeed, for their possible variety. A second task is to show that these devices are meager enough so that very few I-languages are made available to the language-learner, given data that, in fact, suffice for language acquisition. (p. 51)

In other words, the theory of language acquisition should describe the ability to produce a variety of linguistic expressions, thus explaining linguistic diversity. At the same time, it also must constrain the possible number of grammar hypotheses that a child may form, so that they do not acquire faulty principles of grammar. *Linguistic diversity* then, poses a challenge for nativists to explain why there is so much diversity if language acquisition is simply a matter of triggering innate knowledge.

In addition, *linguistic diversity* poses a problem for empiricists. Unlike the other POS arguments, the VOS argument does not appeal to a *poverty of stimulus* in the

environment. Rather, the variety in natural languages actually shows that there is a *rich* source of possible linguistic stimuli. This may seem ideal for empiricist theories, but it presents a problem in view of the fact that children converge on the same knowledge, rather than producing diverse sets of knowledge. Recall that empiricists propose a similar initial state, namely, that children come equipped with a general learning mechanism that responds to specific stimuli to develop different sets of knowledge. The motivating idea is that the initial mechanism can detect patterns in the environment and gain knowledge from these patterns. A general learning mechanism is a good account only if learners contend with a small number of possible hypotheses of grammar. However, it is difficult for empiricists to explain the *convergence* of knowledge in the face of diversity of linguistic stimuli.

The puzzle about why children converge on the same knowledge of language traditionally arises at the level of the individual. The VOS argument, however, compares populations of children. If each child only has access to a limited set of data, then each child should acquire a different grammar depending on the stimuli; these grammars would be as diverse as the stimuli. The empiricist account predicts that this should be so, and that there should be more errors in grammar, incompatible communication, or just plain "dead-end" grammars than there are.

An empiricist might respond by saying that the reason we find linguistic universals in all languages is because those universals are present in the linguistic stimuli in which empiricist learners can input. In other words, if linguistic universals are found in all natural languages, then there is a common source of input in each child's environment (Putnam, 1967). The problem with this response is that there are no clues in the

environment indicating which sets of linguistic stimuli happen to be universal in all languages (Baker, 1981; Chomsky, 1975; Cowie, 1997). Children do not have access to all languages, and they are not explicitly taught universal features or general principles of language. In addition, these universals are highly abstract and do not fit any sort of "easy solution," so learning by induction would still lead to an underdetermination problem (Laurence & Margolis, 2001, p. 203).

What we have seen is that the empirical observations indicating diversity of linguistic input creates a challenge for both nativists and empiricists. Nativists must explain why there is such diversity: If language acquisition were a matter of triggering innate (and universal) knowledge, we would not expect so much diversity in natural languages. Empiricists, on the other hand, must explain why, despite the diversity of input and the absence of clues about what features of linguistic stimuli happen to be universal to all language, children *converge* on the same knowledge of linguistic universals. I suggest in 4.2 that Chomsky's Principles and Parameters account works well to explain these two sets of empirical observations.

4.1.2 Fixed ontogeny

The next feature in the VOS argument highlights the importance of development. In the following passage, Chomsky gives a VOS argument:

> That transition from the initial state to the steady state takes place in a determinate fashion, with no conscious attention or choice. The transition is essentially uniform for individuals in a given speech community despite diverse experience. (1986, p. 51)

Chomsky invokes a VOS argument when he appeals to uniform development of child language acquisition despite their diverse experiences. The argument is as follows:

> (P1) Children acquire the same set of knowledge of language (i.e., knowledge of linguistic universals).
>
> (P2) Knowledge of language is either learned from the environment or there are innate developmental regularities that guarantee acquisition of knowledge of language.
>
> (P3) Children are exposed to different sets of linguistic stimuli from their respective environments.
>
> (P4) If knowledge of language is learned, then children exposed to different set of linguistic stimuli would acquire different sets of knowledge of language.
>
> (P5) Children exhibit similar stages of development in language acquisition.
>
> (C) Innate developmental regularities lead children to converge on the same knowledge of linguistic universals.

The argument is that children acquire the same set of knowledge of language despite their exposure to disparate stimuli. If children were empiricist learners, then it seems that each child would acquire a different set of knowledge, and exhibit different patterns in their stages of language acquisition. However, evidence that supports Premise (1) and Premise (5) show that children end up acquiring the same set of knowledge of language of linguistic universals in roughly the same developmental stages. The fact that children exhibit predictable stages of development is a surprising fact if children were empiricist learners.

One empirical observation that supports Premise (5) is that children exhibit predictable stages of development:

> (5) *Fixed Ontogeny*: Language acquisition develops in ordered stages, and in relatively the same pace and order regardless of chronological age, with predictable errors.

Radford (1990) categorizes several stages of acquisition: Prelinguistic stage (0-12

months) where children "babble" and develop important phonological attributes; single-word stage (12-18) months; and multiword stage (18-30 months). Children achieve milestones in the formation of questions, negative constructions, passives, datives, case markings, embedded sentences, causative constructions, verb-particle constructions, and relative clauses, in regular order (Brown, 1973; Petitto & Marentette, 1991; Stromswold, 2000). The same stages also occur in deaf both children who learn sign language (Goldin-Meadow, 2003; Kegl, et al., 1999; Klima & Bellugi, 1979; Singleton & Newport, 2004; Wilbur, et al., 1983), and blind children (Gleitman & Newport, 1995). Given the variety of linguistic stimuli in each child's environment, an empiricist would predict a more haphazard set of developmental stages.

In addition to displaying similar stages of acquisition, all children make predictable errors during each of those stages, and these errors are universal to children learning any natural language. There are errors in syntax and sentence formation, phonology, and morphology, including over-regularizing verbs (e.g., saying "goed" rather than "went," or "holded" rather than "held"). These errors are universal: Bickerton (1983) observed several similar errors from English-speaking children and Creoles. For example, an English child may say, "Look it a boy play ball" or "Nobody don't like me," whereas in Jamaica Creole, the child says, "Luku one boy a play ball," and in Guyana Creole the child says, "Nobody no like me." The phrase "I no like do that," is found in English, Hawaii Creole, and Guyana Creole. Nativists explain similar patterns of errors by saying there are constraints that limit the kinds of hypotheses in grammar formation. An empiricist would, it seems, predict a wide set of errors that children make given the indefinite space of possible ways children could go wrong.

4.1.3 Insensitivity to variation

The last set of empirical observations used in VOS arguments concerns the acquisition of knowledge despite a range of environmental stimuli:

> (6) *Insensitivity to Variation:* Children acquire linguistic universals despite wide variation in linguistic stimuli, and regardless of their education or upbringing.

The VOS argument addresses empirical observations that children are sensitive to some range of stimuli, and insensitive to others. On the one hand, the argument depends on the fact that there is a wide variation in stimuli, i.e., the *linguistic diversity* in natural languages. In order to have linguistic diversity in the first place, children learn features particular to natural languages such as lexicon, syntax, etc. This is what creates linguistic diversity, and indeed, variation in the linguistic stimuli in the first place. In other words, children must be *sensitive* to some range of stimuli in order to learn a natural language. Indeed, any linguistic theory must explain the enormous amount of learning, whether or not there is a rich source of innate structure to aid this learning process.

On the other hand, children must be *insensitive* to some stimuli that would impede their acquisition of knowledge of linguistic universals. The puzzle is to explain how children acquire the same knowledge given that each child has access to a different possible set of stimuli. An empiricist solution might be that every child "picks up" linguistic universals from their environment. This possibility would explain how all children acquire linguistic universals despite each child being raised in a different environment. According to this account, children could be particularly sensitive to linguistic universals in order to acquire them, and insensitive to other linguistic stimuli that impede this acquisition.

The problem with this solution is that the child's environment would have to be guaranteed to contain linguistic universals in order for children to learn them. However, the strength of the VOS argument is found in the fact that there is no guarantee that the environment will contain rich stimuli or linguistic universals. Children may be exposed to English, Japanese, multilingual homes, pidgins, Nicaraguan sign language, or any other number of languages, regardless of whether the environment contains the right linguistic stimuli or not. An account of language acquisition must explain how children are sensitive to some range of environmental stimuli, but insensitive to other stimuli that would impede their acquisition of knowledge of linguistic universals.

In this section, I have offered six sets of empirical observations that an account of language acquisition must explain, and showed how the POS arguments use these observations to support a legitimate use of "innateness." In addition, I showed how VOS is unique in that it involves puzzles about the complexity of development and the interactions between an organism and its environment. The VOS argument tells us about what the cognitive mechanism of language acquisition must look like: viz., it should be able to generate a range of possible languages, thus explaining the *diversity in language,* but it should also constrain grammar formation to generate a set of core principles universal to all languages, thus explaining acquisition of *linguistic universals*. In Section 4.2, I describe an account that explains these features well.

4.2 Language acquisition is canalized

In this section, I give an account of language acquisition based on canalization. I first define canalization. In Section 4.2.2, I use the Principles and Parameters theory to

develop an account of language acquisition. Language acquisition, I argue, is a matter of acquiring features of language during certain stages, and the developmental pathway is canalized. In Section 4.3, I show how this account explains empirical observations cited in VOS arguments.

4.2.1 Canalization

"Canalization" is a term introduced by Waddington (1957) and revised by Ariew (1996, 1999, 2006). It describes a process by which phenotypic variation (among a population of organisms) is reduced by regulation in development. Canalization measures the degree to which a phenotype is expressed regardless of variations in the environment during the course of the individual's development. We can define canalization as:

> (CAN) A developmental pathway is canalized to the degree to which development of a particular end-state is insensitive to a range of environmental conditions under which the end-state emerges. (Ariew, 1999, p. 128)

A trait may be canalized to various degrees depending on whether or not the trait is sensitive or not to environmental perturbations. A trait is *highly canalized* if the phenotype is expressed regardless of perturbations from the environment. A highly canalized trait is one where the phenotypic expression is more or less guaranteed, barring radical perturbations. Canalization allows for:

> The capacity to produce a particular definite end-result in spite of a certain variability both in the initial situation from which development starts and in the conditions met with during its course. (Waddington, 1975, p. 99)

Canalization is usually discussed in terms of *genetic* canalization that tracks regularities in phenotypes that develop despite genotypic variations (Gottlieb, 2003). In contrast, Griffiths and Machery (2008) explain *environmental* canalization as regularities

in phenotypic expression that develop despite variations in environment:

> A phenotypic outcome is environmentally canalized if those features of the surface which direct development to that endpoint are relatively insensitive to the manipulation of the environmental parameters. (p. 397)

For example, a strain of corn seed may grow to a certain range of height despite variations in soil nutrients, sun, or water. For our purposes, we are concerned with environmental canalization only, since the question asked in connection with language concerns how knowledge of language arises via interaction between genetic expression and a wide range of environmental features.

A theory of canalization is a type of invariance theory, like heritability (see Sober, 1998), or flat-norm reaction (Lewontin, 1974). Heritability measures the proportion of phenotypic variance that is due to *genetic* variance. For example, we can measure the degree to which height of a strain of corn is heritable, by comparing different strains of corn (different genotypes) in a *similar* environment. Norms of reaction can measure the converse, namely, it can hold the genotype fixed and vary the environment, thus noting the degree to which the *environment* affects the trait. By definition, norms of reaction measure the proportion of phenotypic variance produced by a given genotype under a range of environmental conditions. For example, we can measure body size in environments raised in high or low temperatures.

Both of these invariance theories suffer a limitation that is important to our discussion. Both accounts measure adult traits while assuming that the environmental conditions are fixed throughout the course of development (Ariew, 1993, 1996). However, it does not indicate crucial information about how the environment influences development:

> A flat norm of reaction…depicts the pattern of adult phenotypes produced by a given genotype under a range of environmental conditions that are fixed throughout the course of development. What's missing is the possible effect a fluctuating environment during the course of development has on the phenotype in question. That is, what is needed is to plot individual life histories across a variety of environmental conditions, not norms of reactions. (Ariew, 1996, p. 24)

For example, we can measure the height of corn when it is grown in either low or high temperatures. However, this does not capture what we want to know, namely, how temperature fluctuations affect the height during development. This problem leads Ariew to object that, "a flat norm of reaction can mislead one to conclude wrongly that a trait is stable for a genotype when in fact it is not" (Ariew, 1996, p. 24).

What we want is to know *how* an individual acquires a trait when there are fluctuating conditions in the environment that might affect the phenotype during development. For this, we need an account of the course of development of individuals that takes into account the various environmental conditions to which the individuals are exposed. This may involve important stages of development, important internal factors (including, but not limited to genes), and relevant external factors in the environment that affect the trait. These measurements are missing in heritability. For example, if we find that height is 60% heritable (i.e., 60% of the differences in height is attributed to genes), we only know *that* genes affect height among a population; we have no information about *how* genes are affecting the phenotype, or how genes and environment interact. As Ariew (1996) says:

> Heritability is a measure of the variation of traits in a population; it does not explain why individual members of a population have the traits they do…The heritability estimate provides no information about how genes and the environment interact to express height in an individual. (p. 23)

The problem is not that the analysis is at a population level. Rather, the problem is that it gives us no information about development and interaction between genes and

environment.

Ariew enhances the theory of canalization in order to meet these objections by taking the interaction between the organism and its environment seriously. Canalization measures the range of possible environmental features that act to influence the trajectory of development; hence, it is important to specify what these environmental features are. In order to identify factors relevant to the expression of a phenotype, a theory of canalization requires a description of both internal and external factors relevant during development.

We can see the similarity between traits that are canalized and language acquisition. Highly canalized traits will be displayed in any organism raised in any environment, thus exhibiting *ubiquity* among members of a species. In addition, developmental pathways among individuals will *converge* on a particular end-state, i.e., the phenotype that is expressed will be similar among individuals. Convergence can only occur if, at some stage of development, the pathways are similar among individuals.

Language acquisition seems to be canalized, and some scholars have seen a potential to develop a theory of language acquisition in terms of canalization. Chomsky (2005), for example, describes canalization as:

> [P]rocesses adjusted so as to bring about one definite end result regardless of minor variations in conditions during the course of the reaction, thus ensuring the production of the normal, that is optimal type in the face of the unavoidable hazards of existence. That seems to be a fair description of the growth of language in the individual. (p. 5)

He describes language acquisition as a process of "growth" where certain universal knowledge is attained (a 'definite end-result") despite "minor variations in conditions" such as exposure to particular languages such as English or French. Both canalization and

language acquisition seem to be processes that are determined, "much as it determines that we will grow arms not wings" or "undergo sexual maturation at a certain stage of growth" (Chomsky, 1993, p. 519).

Though it is promising to think of language acquisition as being canalized, scholars have not pursued this account in great detail. The notable exceptions are Ariew (2006) and Collins (2011). Ariew's main argument is to define "innateness" as canalization, and Collins (2011) expands Ariew's account to defend it against a number of criticisms. Other scholars have briefly noted that empirical observations about language acquisition can be explained by canalization (Birch, 2009; Dor & Jablonka, 2010; Khalidi, 2009), but they do not develop the theory in any detail. At present, there is a need for a more detailed account of how canalization can be used in explanations of language acquisition. In what follows, I offer such an account.

4.2.2 A canalization account of language acquisition

A canalization account of language acquisition uses Chomsky's theory of Principles and Parameters (Boeckx, 2006; Chomsky 1986, 1988; Chomsky & Lasnik 1995; Hornstein, Nunes, & Grohmann, 2005). Accordingly to this theory, the initial state is composed of a network that "turns on" certain principles of language when they are triggered by linguistic stimuli in the environment. Chomsky (1986) describes this in terms of switches:

> We may think of the language faculty as a complex and intricate network of some sort associated with a switch box consisting of an array of switches that can be in one of two positions…The fixed network is the system of principles of universal grammar; the switches are the parameters to be fixed by experience. (pp. 62–63)

When children are exposed to a particular language (such as English or Japanese or ASL), the syntactic, phonological, or morphological elements in their respective environments trigger certain switches.[ii] For example, someone who hears English might be triggered to form a Subject-Verb-Object typology, whereas someone exposed to Japanese will form Subject-Object-Verb sentences.

On this account, the initial state is common to everyone who can acquire language to the extent that each of them is equipped with this network of switches (Chomsky, 1993, p. 528). This initial state constitutes the Language Acquisition Device (LAD), which includes all of the cognitive mechanism responsible for acquiring, understanding, interpreting, and producing language. On this account, each individual comes equipped with this device that is constituted with a network of switches set at a certain value. The LAD consists in subsystems or modules, which display systematic behavior: It accepts certain input (linguistic stimuli from the environment), and interprets that input in systematic ways that map onto phonological, morphological, or syntactic categories. This reflects the ability to interpret linguistic stimuli in certain constrained ways.

The mature state is achieved via an interaction between internal factors and external features of linguistic stimuli. Linguistic stimuli act to turn on (or off) principles, which, in turn, constrain or limit further options for activation of other principles. Dove (2012) uses Wimsatt's (1986) analogy of a combination lock: When you initially buy the lock, it is possible to set it for any of a number of combinations; however, as you have set each number in the combination, each decision limits the number of possible combinations. If a child is exposed to a Subject-Verb-Object language, then principles compatible with this language are available, and others (for Subject-Object-Verb or Verb-Object-Subject)

are not. These constraints on possible combinations of principles are known as "parameters." In other words, as the child develops, the grammar is increasingly constrained to follow certain pathways.

The next section describes these features, in terms of explaining empirical observations invoked in VOS.

4.3 Canalization explains empirical observations in VOS arguments

So far, I have developed a theory of language acquisition based on Chomsky's Principles and Parameters account and argued that certain sets of knowledge of language are canalized. I will now show how this explains the empirical observations that the VOS argument addresses.

Any theory of language acquisition should explain empirical observations concerning children's linguistic competence. For example, to explain linguistic diversity, an adequate theory must include an explanation of how children interact with linguistic stimuli to learn principles of phonology, morphology, and syntax of their particular natural language from their natural language. However, the conclusion drawn from DOS, COS, and PNE arguments paint a simplistic account of acquisition where interaction with the environment is not necessary.

The DOS argument, for example, concludes that children come equipped with an innate set of linguistic knowledge, and this knowledge is acquired independently of exposure to any linguistic stimuli. This conclusion follows from premises that use extreme cases of deficient stimuli, such as the case of Simon or Nicaraguan Sign Language. If true, these cases show that children acquire knowledge that far outstrips the

limited linguistic stimuli to which they were exposed. The problem with such a theory, i.e., a theory that claims that children have innate knowledge, is that it presents a static view of development and acquisition and is inadequate to explain other sets of empirical observations such as linguistic diversity. According to the innate knowledge theory, linguistic knowledge "comes online" at some point in maturation: This knowledge does not change during development, nor is it informed or triggered by environmental stimuli. The contribution of the environment only adds minor and relatively unimportant details to a pre-existent structure, and the interaction between the organism and environment is simply not important. Although this theory may explain those extreme (and rare) cases where children are raised in dramatically impoverished environments, it does not explain the more common cases where children are exposed to linguistic stimuli that allow them to acquire knowledge of vocabulary and lexicon. If a theory of language acquisition does not respect the fact that children learn certain elements of language from their environment, that theory will fail to explain linguistic diversity.

We find a similar problem with the conclusion in COS where the child has an innate mechanism that sorts linguistic stimuli and filters out certain information, including nonlinguistic noises, that would lead the child to faulty grammar. A theory that posits that there is an innate sorting mechanism may explain the ability to select the right grammar when exposed to degenerate stimuli. However, this theory presents development in a shallow way: There is no sense of development in which children's acquisition of knowledge emerges *because* of linguistic stimuli in the environment.

The PNE argument shows that the cognitive mechanism operative in the child acquiring language either prevents or self-corrects faulty hypotheses of grammar. It is

self-contained in that it does not need to rely on negative evidence to correct faulty grammar. Therefore, no interaction is needed between internal factors of the child, and environmental stimuli making development a matter of maturation, not interaction. It, too, fails to explain linguistic diversity.

The canalization account respects the fact that language is acquired as a result of an interaction with environmental stimuli. It is not appropriate under the Principles and Parameters account to claim that children have innate knowledge, if by "knowledge" we mean some set of principles or rules that simply emerge during some stage of maturation. What is innate is the network of principles and parameters, not the content. A Chomsky notes, "There are no rules at all, in the conventional sense, in the central areas of syntax" (1986, p. 102). Thus, describing linguistic knowledge in terms of a static set of "innate rules" is a gross simplification.[iii]

The following sections explain how the canalization account can explain empirical observations that used in the VOS argument.

4.3.1. Linguistic universals and diversity of language

The canalization account explains why we see a wide diversity across languages, all of which are constrained by the same *linguistic universals*. Language acquisition is not a matter of an innate set of rules that emerges as the child acquires language; rather, it is a matter of setting the values of parameters based on exposure to linguistic stimuli. This involves a process of interaction with linguistic stimuli. Development is initially plastic, and it is sensitive to certain aspects in the environment where much knowledge of language is achieved via learning. At the same time, we see a steady maturation towards

a fixed end point (acquisition of linguistic universals), and the development is insensitive to environmental perturbations.

The same mechanisms that explain ubiquity, universal initial state (UG), and production of linguistic universals can also account for linguistic diversity. As Chomsky (1980) states:

> If the system of universal grammar is sufficiently rich, then limited evidence will suffice for the development of rich and complex systems in the mind, and a small change in parameters may lead to what appears to be a radical change in the resulting system. What we should be seeking, then, is a system of unifying principles that is fairly rich in deductive structure but with parameters to be fixed by experience. (p. 66)

Small changes in parameters lead to large effects concerning how the system interprets linguistic stimuli and what is produced. When one value is set, it alters the options for setting other parameters. For example, setting a Subject-Object-Verb parameter constrains the possible values covering principles about movement order, and thus affects how phrases are interpreted. Linguistic diversity is, in part, reflected by these variations. Thus, "linguistic diversity is not an obstacle to the uniformity of the process of acquiring a language because every language is just a superficial variation of an innate linguistic scheme" (Lorenzo & Longa, 2009, p. 1302). This is not to discount the other aspects of linguistic diversity due to peripheral systems, production, and a capacity to generate an infinite range of sentences. Rather, the idea is to show how the same system that ensures that *linguistic universals* can also create *linguistic diversity*.

4.3.2 Fixed ontogeny

In order to explain how children can acquire the same knowledge of linguistic universals despite disparate upbringing, the Language Faculty must *initially* adapt to

various possible environments in which the child may be raised. It must be *sensitive* to some range of linguistic stimuli in order to learn certain aspects from the language in the environment. This is true at early stages; however, later stages of developmental must be *insensitive* to possible perturbations from the environment because this would impede acquisition of knowledge linguistic universals. Therefore, although the initial stages of development are plastic and adaptable to a range of stimuli, later stages are determined and are insensitive to possible perturbations.

The canalization account is compatible with the view that initial stages of development are plastic. Initially, there is a potential for many possible developmental pathways, namely, different possible settings of parameters. Those parameters are set via exposure to natural languages; hence, a significant component of "learning," understood as setting parameters, depends on environmental input. Different combinations of these settings result in natural languages which express diversity in phonology, morphology, and syntax.

Dor and Jablonka (2010) explain the importance of plasticity in initial stages of acquisition:

> The entire language out there, which is spoken by the adults around the child, is the 'attractor,' and in order for the child to be able to reach the attractor, the child must explore at all levels: the child must try different ways of communication, different ways of usage of language, different interpretations for the utterances heard around him/her. Moreover, the child's brain goes through a whole series of explorational and selective stabilization processes, in which neural pathways, allowing for successful comprehension and production, are stabilized. (p. 138)

The initial stage is one of probing various aspects of language, perhaps by trying out different phonological, morphological, or syntactic possibilities. This allows the language mechanisms to interpret input and thereby set principles.

At some point in development, each child will express similar stages of development, i.e., *fixed ontogeny*. In the normal course of development, there may be many possible courses of development that branch out, but then return to a determined end-state. Subsequent development becomes stabilized by setting parameters and once these parameters are set, it limits the choices for possible grammar formation. The child may use both positive and negative evidence in the environment as clues.

> Exploration and selective stabilization mechanisms are the most prominent mechanisms that lead to open-ended plasticity. They may occur at the cellular, physiological, behavioral, and social levels. All are based on a similar principle – the generation of a large set of local variations and interactions, with only a small subset eventually being stabilized and manifested. Which output is realized depends on the initial conditions, the ease with which developmental trajectories can be deflected, and the number of possible points around which development can be stably organized…Selective stabilization thus involves both the constraining of certain aspects of the response and extensive plasticity (output variability) within this range. (Dor & Jablonka, 2010, p. 137)

A large range of possible features in the linguistic stimuli can influence the developmental pathway, especially during initial stages of the development. However, during the course of development, the trait is stabilized. That is, it displays predicable stages of development, such as the prelinguistic stage (Petitto & Marentette, 1991; Radford, 1990).

Ariew describes the process in terms of "triggering," where some linguistic stimuli act to set parameters. Once the parameters are set, development becomes rigid, in that it is insensitive to further perturbations. For example,

> The adoption of 'head-first' languages are 'triggered' or phenotypically switched by a few linguistic cues. Once the triggering environmental cue is encountered, development of one or the other pathway is relatively unaffected by the presence or absence (or poor quality) of further linguistic cues. Perhaps post-trigger development proceeds independently of linguistic cues. If so, we would say that post-trigger development of

> 'head-first' or 'head-last' grammar is innate across linguistic cues. Otherwise we would say that it is simply to some degree canalized. Either way, compared to learning models of grammar acquisition Chomsky's switchbox model predicts that the development of specific grammar rules is relatively robust. (Ariew, 2006, p. 16)

The robust aspect of language acquisition occurs because development is in some sense insensitive to environmental perturbations, thus ensuring stabilization towards a mature state.

So far, I have sketched a picture of language development that starts out flexible and plastic, and then is stabilized into a fixed mature state. This sort of plasticity may seem incompatible with canalization. After all, a theory of canalization only requires that the expression of a phenotype be guaranteed by being insensitive to a range of environments. This says nothing about how those traits develop. However, this is precisely what we want to know: How does the organism interact with the environment during development?

It seems, then, that there are two possible ways in which a trait can be canalized. I will call the first way, "fixed ontogeny," i.e., development that follows rigid and predictable stages, where each stage of development is insensitive to perturbations. In the second way, "plastic ontogeny," there is a range of possible pathways that development may take to reach a mature state. A trait with plastic ontogeny may display some degree of plasticity during initial stages of development by being more sensitive to environmental stimuli. For each option, we can say that a trait is canalized because it is measured by the extent to which the environment influences expression of phenotypic end-states. In either case, the expression of the phenotype is achieved despite environmental perturbations.

Fixed ontogeny displays developmental rigidity. Some traits such as the acquisition of eyes or hands are highly canalized because at each stage of development, they are largely insensitive to a wide range of environmental perturbations. These traits develop in a predictable fashion from the initial stages to maturity. At each stage of development, the trait is insensitive to stimuli, and because it is not influenced by variations in the environment, they display a small range of variations in phenotypic expression (such traits are monomorphic).

Plastic ontogeny, on the other hand, displays sensitivity to the environment, at least in initial stages of development. Traits of this kind may not necessarily display fixed or predictable ontogeny since their development changes in response to environmental perturbations. Nevertheless, the phenotype is reliably expressed at mature stages of development despite variations in sensitivity to stimuli during development. Development reaches a certain outcome to express a phenotype, regardless of variations in ontogeny. Plastic ontogeny, then, displays more sensitivity to stimuli at earlier stages.

Dor and Jablonka (2010) highlight the importance of plasticity and make a strong claim that, "[A]lmost every case of canalized development (in the face of genetic and environmental noise) requires plasticity at underlying or overlying levels of organization" (p. 139). They give an analogy of acclimating to high altitudes that involves increasing the number of red blood cells in order to maintain the level of oxygen in the blood. The body's response is plastic in that it can handle and adapt to environmental stimuli by adjusting the number of red blood cells. It is also an illustration of canalization, if we consider that the concentration of oxygen remains stable despite perturbations in the environment. It is because the body is able to adjust the amount of red blood cells that

makes it possible to maintain oxygen stability.

Language acquisition exhibits features of plastic ontogeny at initial stages of development (Maratsos, 1989). Knowledge of language is acquired, in part, by being sensitive to variations in the different environments (variations to natural languages such as English or Japanese), and this knowledge is acquired via different pathways. In addition, each stage of development is sensitive to linguistic stimuli: Initial stages require stimuli that will set parameters such as head-first or head-last grammars, and these settings influence the options available at later stages. Initial stages of acquisition, then, are sensitive to the variations in natural languages.

Because initial stages of acquisition are sensitive to stimuli, we see variations in the stages of development between children, such as differences in chronological age at which the multiword stage is reached (Radford, 1990). These variations indicate that the system is sensitive to stimuli, unlike traits that are triggered by internal maturation processes, such as the growth of optic pits at approximately 22 days of fetal life.

So, too, we can see that language also requires plasticity. Knowledge of linguistic universals is acquired because children are able to interpret a large variety of linguistic stimuli.

> Although all normal children acquire the languages of their communities there is plasticity in that the particular routes of linguistic development differ, as does the specific output—the individual idiosyncrasies of one's language production. Different children come to the world with different genetic makeups, different learning capacities and different embryological histories, and they are exposed to different sets of linguistic inputs. The very fact that they eventually manage to zoom in on the target language and produce a relatively invariant behavior means that they must manifest great plasticity at the neural level. (Dor & Jablonka, 2010, p. 136)

To put it another way, Dor and Jablonka are giving a VOS argument by pointing out that

a trait is acquired despite variable external stimuli and internal features among children. They note the variability in genetic makeup, learning capacities, and differences in brain physiology level, even among identical twins. Each child arrives at this competence via different ontogenetic pathways, i.e., language acquisition is modality-independent. The fact that each has acquired similar abilities in linguistic competence means that those developmental pathways must have converged to a similar end-state (in Chomskian terms, features of their I-language are similar or universal to all children). Thus, "It is the ability to generate neurological variability…that allows for the construction of different developmental trajectories that lead to something that everyone recognizes as language" (Dor & Jablonka, 2010, p. 136). This entails that the child develops in a slightly different way in response to variable environments, i.e., language acquisition is canalized via plastic ontogeny.

I have described how language acquisition displays plastic ontogeny at initial stages of development and fixed ontogeny at later stages. Language acquisition is canalized since the trait (in the mature stage) is acquired despite variations in internal constitution or external stimuli.

4.3.3 Insensitivity to environmental stimuli

A trait that develops despite a wide range of relevant environmental features is "buffered" against those features. Indeed, in describing canalization, the focus is on how phenotypic outcomes are "buffered" against, or insensitive to, a possible range of environmental conditions.

To say that language acquisition is canalized means that children will exhibit

linguistic competence regardless of the environment in which they are reared. This is not to say that the environment does not matter. The term "*insensitivity to environmental stimuli*" is misleading. It is not that language acquisition is *insensitive* to any stimuli, nor does language develop *independent* of the environment. Some linguistic stimuli are necessary and required to trigger or inform linguistic knowledge. In this sense, external factors are critical, but that does not mean that there are *specific* factors in the environment required for all children to acquire language. Rather, these aspects are acquired despite exposure to a *range of possible stimuli*.

Although some scholars talk about canalization in terms of traits arising *independent* of the environment, this cannot be the case. No trait develops independent of environmental influence. Instead of saying that canalized traits are *insensitive* to the environment, it is more precise to say that they are "modality-independent," meaning the trait is acquired via multiple avenues of sensation (sight, sound, touch, etc.). Collins (2011) describes canalization as "modality-independent," since there are many possible ways in which knowledge of language (and more particularly, knowledge of linguistic universals) can be acquired. Linguistic competence is achieved "independent of the possible modalities of the consumption and production of language" (p. 198). In other words, language acquisition can develop by the child either seeing sign language, or hearing acoustic sounds. In either modality, the language faculty has the ability to adapt to the stimuli, and to interpret or "map" those stimuli as phonetic properties, morphemes, etc.

> In short, there appears to be no specific external stability that maps onto the shared linguistic competence that is invariant over the species, even though each organism presumably exploits some local external stability. (Collins, 2011, p. 198)

An analogy can help illustrate modality-independent traits. Tuna live in salt water and cannot tolerate fresh water; goldfish have the reverse tolerance. Both fish can only tolerate narrow ranges of salinity and are highly sensitive to any changes in their habitat. Conversely, Euryhaline fish can live in both. Salmon, for example, are born in fresh water and spawn in fresh water but migrate to the ocean to live in salt water. They need exposure to both habitats early in development (i.e., there is a critical period of acquisition), and they must undergo an acclimation period during each migration. The mechanism that is responsible for the Euryhaline fish to acquire high salinity tolerance (mechanisms of osmosis and acclimation to salt) must be plastic in the sense that fish can adapt to various possible environments. In addition, although the acquisition of high salinity tolerance depends on certain features of the environment (exposure to fresh and salt water), the acquisition of high salinity is buffered against variations in salinity.

Theories of acquisition should be characterized in terms of the conditions relevant to the development. Scholars often talk about picking out relevant features in the environment in terms of pragmatically selected contrast class of alternatives (Birch 2009; O'Neill, 2014). For example, Birch (2009) says, "only factors that make the difference between T and T* are relevant to the judgment" (p. 298). However, this is not to say that the issue is a pragmatic one: There is a real effect that certain environmental stimuli have on the acquisition of traits, and the task is to identify these effects. We should always view the trait in terms of environmental factors that are relevant to the trajectory of the development. This is part of what it means to say that a trait is canalized relative to some range of environments (see Birch, 2009; Collins, 2005; O'Neill, 2014). Some features of the environment do not play a causal role in development (that range is irrelevant), and

some do. For example, the exposure of oxygen is not particularly relevant to a bird acquiring a particular song, since oxygen exposure leads to acquisition of *either* trait T or T*. However, exposure to English rather than Japanese does make a difference in acquiring grammar G or G*. More research is needed to discover what those particular factors are that play a causal role in the acquisition of grammar. Cognitive linguists should embrace the challenge of identifying relevant features in the environment.

4.4 Conclusion

I developed a canalization theory of language acquisition, based on the Principles and Parameters account. I argued that it explains empirical observations addressed by VOS arguments. The canalization account accounts for the fact that children acquire knowledge of linguistic universals despite being raised in a variety of environments. The canalization account also respects the complexity of development, and it takes seriously the role that the environment plays. I argued that language acquisition displays plastic ontogeny, where initial stages of development are more sensitive to linguistic stimuli, which in turn constrains options in further stages.

This is not to say that canalization is the best account for all arguments for innateness in cognitive linguistics, and I do not argue for this position. The POS, COS, and PNE arguments require appropriate accounts of cognitive mechanisms to explain a diverse set of empirical observations about properties of organisms. I showed how DOS concludes that innate linguistic content comes online via internal maturation processes. This may explain why children have abundant knowledge in environments deficient of linguistic stimuli. However, this solution is insufficient to explain other observations about

children's linguistic competence. A canalization account can explain numerous sets of empirical observations: Despite being raised in a wide range of environments, children reach linguistic competence and converge on the same knowledge of linguistic universals. Although certain aspects of development are insensitive to environmental perturbations, linguistic competence cannot happen independent of environmental stimuli. Indeed, the force of the VOS argument is in highlighting that, for better or worse, there are plenty of linguistic stimuli to which children are sensitive. Language acquisition occurs by fixing parameters of language, thus constraining developmental pathways, and leading towards linguistic competence.

4.5 Endnotes

[i] The task is not to describe individuals as such, but rather the *idealized* individual at an appropriate level of explanation. This is because individuals vary given that "the actual state of one's language faculty is the result of interaction of a great many factors." An account of the cognitive mechanism, on the other hand, will describe the initial state "idealizing from actual states of the language faculty" (Chomsky 2000, 123). This task is similar to developing an account of anatomy that describes ideal human physiology even though there are differences in the physiology of individuals.

[ii] Principles of linguistic universals include (Chomsky, 1986 Chapter 3) X-bar theory, C-command and Government, binding theory, Case theory, theta theory, bounding theory, etc., each subsystem with its own principles. Overriding principles include projection principle, FI, principles of licensing. Principles such as c-command, principles about domain, and government play a central role throughout.

[iii] See Jenkins, 2000, for discussion about how nativism has been construed as claiming that children represent propositional rules of language. This particular construal of nativism continues to be attacked by both nativists and empiricists alike, though no one actually holds this straw-man view.

CHAPTER 5

A CANALIZATION CONCEPTION OF INNATENESS

In the last chapter, I developed a theory of language acquisition and showed how some sets of knowledge of language are canalized. We are now in a position to examine whether these canalized traits may also be said to be "innate." In Chapter 3, I agreed with Griffiths that if the theoretical term "innateness" leads to confusion, then it should be eliminated. I showed how several leading theories of acquisition that employ "innateness" are insubstantial, ambiguous, or misleading because they attempt to explain too many diverse empirical observations by appealing to "innateness"; the result is that nothing is explained. If we could define "innateness" in a substantive way, then we might be able to make progress in explaining at least a portion of the complex processes of language acquisition.

In this chapter, I consider Ariew's proposal that "innateness" should be defined in terms of canalization (1996, 1999, 2006). In Section 5.1, I explain Ariew's definition of "innateness" as canalization. In Section 5.2 I, discuss Ariew's suggestion that we define "innateness" as the "canalization conception of "innateness"" or "c-innateness." I use an example from Griffiths and Machery (2008) to show that "c-innateness" is not a complete definition of "innateness." I argue that similar counterexamples show that some traits, which are often thought to be innate, cannot properly be called "c-innate." At the same

time, I argue that the definition "c-innateness" is an appropriate definition to describe some linguistic abilities, viz., the set of empirical data invoked by VOS. This shows that "c-innateness" is a useful and explanatorily valuable concept in this context. It is not, however, applicable to the other arguments in POS. If defenders of "innateness" wish to continue to use these arguments to prove a version of the innateness hypothesis, they will have to provide a substantive analysis of "innateness" for their purposes.

5.1 C-Innateness

Ariew (1996, 2006) argues that traditional accounts of "innateness" fail to take into consideration important details about the *development* of traits. These accounts of "innateness" (he cites Samuels' account of Primitivism as an example, 2006, p. 4), focus on what the environment does not provide: They infer that a trait is innate only if it is acquired *independently* of environmental stimuli. However, they do not provide an account of what the environment *does* provide, especially in cases where the expression of phenotypes vary depending on fluctuations from environmental stimuli. The problem, he argues, lies in explaining acquisition without regard to the complex interactions between organisms and the environment. Without considering the nature of these interactions, a theory of acquisition will be insubstantial. Ariew argues that we should not ignore environmental influences; instead, we should ask, "What difference does the presence or absence of certain environmental factors have on development of the trait?" (2006, p. 4).

In order to illustrate different conceptions of "innateness," Ariew explains three types of birdsong acquisition, using data from Gould and Marler (1991; see discussion in

Sober, 1998). In Type 1 birdsong acquisition, birds sing their species-specific song even if they are not exposed to the song during development. This type of birdsong acquisition provides a powerful case in support of the Deficiency of Stimulus argument, which invokes properties of "innateness" such as *paucity of stimuli* and *no explicit training*. It shows that a trait can be acquired independently of appropriate stimuli such as exposure to other birdsong. This is not to say, of course, that birds acquire their song independently of all stimuli, as some stimuli are required for the bird to survive (nutrients, water, etc.). However, it is a case where relevant stimuli (exposure to other's birdsong) are deficient.

Unlike Type 1, which seems to indicate some kind of innateness, Type 2 birdsong acquisition occurs when birds acquire any song to which they are exposed. In these cases, birds mimic any song they hear, whether or not it is native to their species. This case seems to be a good candidate for an empirical theory of learning since the bird acquires the song to which he is exposed to during training.

The last type of birdsong acquisition, Type 3, demonstrates an interesting case of song acquisition, and it is one that Ariew uses to argue that birdsong is innate because it is canalized. These birds acquire their species-specific song given exposure to *any* song, including another species' song.

> All that is required to produce their song is contact with some song or other. They do not require a tutoring period; they require only exposure to some song. They will not respond to silence. Songs from other species or even other bird-like songs suffice to "trigger" their song capabilities. (Ariew, 2006, p. 3).

Thus, this case indicates that exposure to *some* auditory cue is required as silence is not sufficient, but tutoring is not required for acquisition of their birdsong.

Type 3 birdsong acquisition is a prime example of a canalized trait. The song is acquired reliably despite a range of environmental perturbations. Since the development

of the trait depends on stimuli from *any* relevant environmental feature (instead of a *particular* environmental feature), we can infer that their birdsong-acquisition mechanism must be able to handle a large range of possible inputs (exposure to any possible birdsong) and adapt to their environment. Their mechanism is, to use Collins' description, modality-independent (see Section 4.3.3). Thus, Ariew argues, Type 3 birdsong acquisition is canalized.

Type 3 birdsong acquisition is an illustration of a trait that is acquired when exposed to a variety of stimuli. It is not a case of deficiency of stimuli: The birds' environments are not impoverished in the sense that they are receiving too little stimuli, so it is not a case of deficient stimuli. Rather, birds are exposed to a lot of stimuli, albeit from different sources than what their normal habitats would provide. That is, the trait is acquired despite a variety of stimuli. There is a *ubiquitous* acquisition of birdsong, thus indicating a trait that is *species-specific*. There is a fixed end-state, in the sense that birdsong is reliably acquired. Each bird acquires knowledge of a specific birdsong (analogous to I-language), and this process converges on a mature state of the species-specific birdsong (analogous to linguistic competence). The mature state displays features common to all species-specific birdsong (analogous to linguistic universals). In addition, we find *fixed ontogeny* in the sense that there are predictable stages of development. The white-crowned sparrow, for example, begins to produce and experiment with song notes at 1 month old, and crystallizes its song by 4 months old. There is a critical period for acquiring the trait, and once the bird song is acquired, it is fixed (Gould & Marler, 1991, p. 13).

I conclude that Type 3 birdsong acquisition is innate in the sense that it is canalized.

The difference between Type 1 and Type 3 is found in the extent to which the song is acquired independent of environmental perturbations, which is measured by the degree to which each kind of trait is canalized. The more canalized a trait is, the more it strikes us as being innate. As Prinz puts it, "Being robust across environments may be behind intuitions supporting the proposal that innate traits are independent of the environment" (2002, p. 190).

We can give a canalization definition of "c-innateness" as follows:

> The degree to which a biological trait is innate for a genotype is the degree to which a developmental pathway for individuals possessing an instance of that genotype is canalized. (Ariew, 1996, p. 25)

A trait is c-innate when it is canalized, where "a pathway is canalized to the 'degree to which development of a particular end-state is insensitive to a range of environmental conditions under which the end-state emerges" (Ariew, 1999, p. 128).

Ariew argues that other accounts of "innateness" do not capture the differences in acquisition among the three types. For example, Chomsky (1993) uses a growth analogy:

> Language learning is not really something that the child does; it is something that happens to the child placed in an appropriate environment, much as the child's body grows and matures in a predetermined way when provided with appropriate nutrition and environmental stimulation. (p. 520)

The problem with a growth conception of "innateness" is that it does not distinguish between the three types of birdsong acquisition. Ariew (2006) explains,

> [A]ll three types of birdsong involve growth. If innateness means growth rather than non-growth, then innateness ascriptions will fail to pick out interesting differences between the three types of birdsong development. (p. 3)

Another attempt at defining "innateness" is given by Samuels' account of Primitivism (see Section 3.3). This, however, is incomplete: A Primitivism account can distinguish

between Type 1 birdsong (which is not-learned) and Type 2 birdsong (which is learned, and hence, not innate). But the Primitivism account does not distinguish between Type 1 and Type 3 birdsong acquisition. Neither type of birdsong is learned: Both types are "…explained fully within biology rather than cutting across the psychology/biological domains" (Ariew, 2006, p. 4). Both accounts are innate in the sense that they are primitive, and so conception of Primitivism is unhelpful in distinguishing the two types of acquisition.

An invariance account is also inadequate to distinguish between Type 1 and Type 3 birdsong acquisition. For example, Sober's (1998) account defines "innateness" as follows:

> A phenotypic trait is innate for a given genotype if and only if that phenotype will emerge in all of a range of developmental environments. (p. 796)

According to Sober's account, both Type 1 and Type 3 birdsong acquisition would be innate, since both species of birds acquire their birdsong in a range of environments. However, his account does not distinguish between Type 1 birdsong acquisition that emerges independent of stimuli from their species' song, and Type 3 birdsong acquisition that depends on some stimuli that act to "trigger" phenotypic expression.

Type 3 birdsong acquisition is interesting precisely because it *depends* on some cue in the environment, unlike in Type 1. What distinguishes between Type 1 and Type 3 birdsong acquisition is the degree to which environmental factors influence development and canalization can give us a measure of such degree. Type 1 is *highly* canalized, given that external stimuli have little effect on the expression of the trait. It develops independently from any auditory cue, and hence, it is "innate with respect to auditory

cues" (Ariew, 2006 p. 16). Type 2 birdsong acquisition is not canalized: Its development is plastic, and the trait depends on environmental stimuli, and differences in phenotype mirror their environment. It is a prime example of empiricist or experiential learning. Type 3 falls somewhere in the middle. It depends on some environmental cue though the range of cues has little effect on its outcome. It is "contingent upon the presence on the presence of auditory cues at some stage of development…If an auditory cue is encountered, then its subsequent song development is canalized, otherwise, not" (Ariew, 2006, p. 16).

Ariew compares Type 3 birdsong acquisition to language acquisition and concludes:

> As evidenced by the [Poverty of Stimulus argument] the development of specific grammar rules appears relatively unaffected by fluctuations of quality and quantity of linguistic cues, suggesting that the development of grammar rules is to some degree canalized. (2006, pp. 16-17)

In other words, acquisition of rules of grammar is c-innate. We are now in a position to ask whether "c-innateness" is a good conception of "innateness."

5.2. Canalized traits are c-innate

In Chapter 3, I showed how "innateness" is a heterogeneous category: A trait can be "innate" in the sense that it is genetically determined, highly heritable, not learned, present at birth, species-typical, or it is a Darwinian adaptation. The problem with using the term "innateness" is that scholars may assume, without evidence, that "properties must somehow cluster" (Mameli & Bateson, 2005, p. 156; see Section 3.1 of this dissertation for further discussion).

One way to avoid this confusion is to eliminate the use of the theoretical term "innateness" and to replace it with appropriate substantive definitions. Griffiths and

Machery (2002) argue that if what we mean by "innateness" is "canalized," then we should just say "canalized" and avoid confusion: "If [a trait] is developmentally canalized with respect to some set of inputs or is generatively entrenched, then say that it is" (2002, p. 82).

I agree that we need appropriate definitions of "innateness" to avoid confusion. Ariew proposes that "innateness" should be defined as "c-innateness." This initially seems like a good definition, given that canalization can be empirically tested. However, the problem is that the definition of "c-innateness" cannot be generalized to explain the acquisition of any trait that may be innate. Ariew's analysis only applies to traits that are canalized and thus it ignores other forms of acquisition, such as maturation, triggering, or simply being present at birth.

I will use Griffiths' and Machery's (2008) counterexample to illustrate that not all traits that we consider innate are canalized, and thus, "innateness" should not be defined solely in terms of canalization. They use research from psychobiologist Moore (1984), who shows that the ability of male rats to copulate depends on a surprising environmental factor, namely maternal licking of the genital area during development. Mother rats respond to chemicals released by male pups, prompting maternal licking. Copulation is not c-innate because it depends on an environmental condition, one that happens to be in all environments. However, "intuitively the rat's ability to copulate *is* innate" (Griffiths & Machery, 2008, p. 398). This case shows that the definition of "c-innateness" is not a complete analysis of "innateness." The rat's ability to copulate is innate even though it is not canalized. Therefore, there are more cases of innateness, but these cases cannot be defined as "c-innate."

Although "innateness" should not always be defined in terms of "c-innateness," it is an appropriate definition when referring to traits that are, in fact, canalized. "C-innateness" is an appropriate conception for the VOS argument: The argument focuses on empirical observations about certain linguistic abilities. Insofar as the acquisition of those abilities is canalized, then they can be said to be c-innate. However, the other POS arguments focus on other linguistic abilities that may not be innate, and hence, may not be c-innate.

Recall that the conclusion of the VOS argument is that a set of linguistic knowledge (i.e., the set that includes knowledge of linguistic universals) is innate. In Chapter 4, I showed how this set of knowledge is canalized, i.e., it is acquired independent of a range of environmental stimuli. The empirical observations about *ubiquity, convergence, linguistic universals, developmental stages, and fixed ontogeny* are all properties of organisms that correlate precisely because they are acquired via canalization. These traits are canalized, and hence are c-innate.

However, even if we restrict ourselves to the use of "innateness" in connection with language acquisition, the canalization analysis of "innateness" cannot be used to explain all instances of alleged innateness. As I have shown, the term is used in at least four contexts within the POS arguments; these four contexts (COS, DOS, VOS, and PNE) each offer "innateness" as the solution to problems that arise from separate sets of data. While "c-innateness" explains the data in VOS arguments, it does not work for the other POS arguments.

The DOS argument concludes that linguistic knowledge is innate, but this sense of "innateness" is that of being "present at birth." Children acquire language without the aid

of linguistic stimuli. Children have *abundant knowledge* that exceeds information found in the environment. Like Type 1 birdsong acquisition, the cases of DOS show that no linguistic stimuli are needed since there is an impoverishment of any sort of "triggering" cues from the environment. This implies, then, that linguistic content is innate, in the sense that it is already present (at birth, say), and emerges via maturation processes.

The COS argument states that knowledge of language is reliably acquired even though the environment is noisy and degraded. When the child is forming her grammar, she has the ability to filter through stimuli, pick good samples (those that correlate with linguistic universals), and ignore bad stimuli (ungrammatical sentences, nonlinguistic noises, etc.). The conclusion of the COS argument is that children have an innate means of distinguishing linguistic stimuli that is relevant for formation of linguistic principles. The proposed innate trait is a process or mechanism that filters or sorts through linguistic and nonlinguistic stimuli. This innate trait is built into the Language Faculty. The sense of "innateness" in this argument is unlike that of "c-innateness:" There is no sense that linguistic competence is acquired by being insensitive to a range of environmental conditions during development, as "c-innateness" states.

The PNE argument states that children reliably acquire knowledge of language despite negative evidence, i.e., indications from the environment that would indicate whether the child's grammar is correct or not. The conclusion of the PNE argument is that children have innate constraints on the kinds of linguistic principles they can acquire. These constraints that prevent the child from forming faulty linguistic principles, or that aid children in self-correcting their faulty grammar. These constraints are not acquired or canalized; rather, they are part of the initial structure of the Language Faculty.

The VOS argument concludes that knowledge of linguistic universals is innate. I argued that the conception of "c-innateness" is appropriate for this argument because knowledge of linguistic universals is acquired via canalization. While my arguments show that the definition of "c-innateness" works within the context of VOS, it is premature to conclude that it is the only analysis of innateness. At this stage, we can only say that the definition of "c-innateness" cannot and should not be substituted for "innateness" across the board; future research may produce other analyses of "innateness" that account for DOS, COS, and PNE arguments.

CHAPTER 6

CONCLUSION

This dissertation aimed to strengthen theories of language acquisition advanced by cognitive linguists by identifying ways of making them more substantial. We have seen some problems in traditional innateness theories of language acquisition that characterize the approach of cognitive linguistics. Scholars frequently gloss over the complexity of language acquisition by employing the blanket term "innateness" to explain diverse set of observations. One problem lies in the ambiguity of the concept; it is used to refer to properties of organisms with any of the following features: the trait is present at birth, not learned, reliably appears during a stage of maturation, genetically determined, highly heritable, or canalized. The concern is that scholars cluster these properties together and assume, with insufficient evidence, that if one property is present, then others will be as well.

To avoid using "innateness" in an insubstantial way, theories of language acquisition must address three considerations:

> (1) *Empirical Consideration* (EC): An account of language acquisition must give careful attention to different sets of empirical observations. It should not oversimplify language acquisition by underestimating the diversity of observations to be explained.
>
> (2) *Developmental Consideration* (DC): A theory of language acquisition should describe cognitive mechanisms that underlie traits and the

> developmental processes that play a causal role in trait acquisition.
>
> (3) *Interaction Consideration* (IC): A theory of language acquisition should include information about the interaction between factors internal to the organism and environmental conditions.

I use the Poverty of Stimulus argument as one example of an ambiguous and insubstantial use of "innateness." The reason for the ambiguity is that advocates of the argument appeal to diverse sets of empirical observations about properties of organisms and the environment in which they are raised to motivate the conclusion that there must be some innate trait that accounts for language acquisition. Different arguments concluded that different traits are innate, and some conclusions do not follow given the diversity of empirical observations. This argument can be made more effective by identifying sets of empirical observations, and creating appropriate conclusions. To this effect, I distinguished between four strands of arguments that constitute a class called "Problems of Stimulus" (POS). These strands are Deficiency of Stimulus (DOS), Corruption of Stimulus (COS), Variety of Stimulus (VOS), and Poverty of Negative Evidence (PNE).

Distinguishing the four strands of poverty of stimulus arguments is beneficial in four ways. First, it helps identify a diverse set of empirical observations about traits and the environment. Each strand of argument raises puzzles about language acquisition and we cannot assume that if we sufficiently explain one puzzle about language acquisition, then we have explained them all. Whereas the DOS argument focuses on the impoverished quantity of linguistic stimuli, the COS, VOS, and PNE argument is about the quality of linguistic stimuli. In addition, each strand focuses on different traits of the organism. The DOS argument focuses on children's abundant knowledge on which they draw to create

language despite limited exposure to linguistic stimuli. The COS argument focuses on the ability to sort and distinguish linguistic stimuli from nonlinguistic noises or ungrammatical samples, and to use good samples in grammar formation. The VOS argument focuses on children's acquisition of the same knowledge despite exposure to diverse linguistic stimuli. The PNE argument focuses on the ability to form correct grammar despite an indication about whether the grammar is correct or incorrect.

I categorized the list of empirical observations that are used to support each POS argument. However, more empirical work is needed to identify proper sets of observations, and to ensure that the traits in each set are actually correlated. Insofar as some traits are correlated, work is needed to discover whether the correlation is a result of common causal developmental pathways. Separating the strands avoids the danger of lumping together empirical observations and offering one solution—"innateness"—to explain them all.

The second benefit in separating POS arguments is that it helps identify the cognitive structures and developmental processes that play a causal role in acquisition of linguistic traits. We cannot assume that, even if they are innate, different traits are acquired in the same way. Separating these strands will aid in creating appropriate theories of the cognitive structures and developmental processes.

The third benefit in separating the strands of POS is that it allows us to see that some strands are more effective or better supported by empirical observations. Scholars overlook the unique aspects of the VOS argument that raises a puzzle about how children can converge on the same knowledge despite any possible environments in which a child could be raised. In order for children to be able to adapt to a wide range of possible

environments, the language faculty must be plastic, at least in initial stages. I made a distinction between "plastic ontogeny" and "fixed ontogeny" and argued that language acquisition appears to have plastic ontogeny. This would explain how the same knowledge of language emerges in different environments.

In addition, the VOS argument is unique in that it requires a theory of acquisition that takes seriously the interaction between internal factors of the organism and environmental stimuli. The VOS argument challenges the empiricist and nativist account in different ways: Empiricists need to explain how children converge on the same knowledge that constitutes linguistic universals (i.e., syntactic, phonological, or morphological features common in all natural languages). This fact cannot be explained by a theory of learning that proposes that children infer this knowledge since linguistic input varies between children. Conversely, nativists must explain why there is so much diversity in natural languages. This cannot be explained with a theory that proposes that a rich set of innate content is triggered by maturation, independent of environmental stimuli. Instead, the theory must propose a cognitive structure that ensures acquisition of knowledge of linguistic universals, but is also has the capacity to generate linguistic diversity. I described Chomsky's Principles and Parameters account as a good candidate.

In Chapter 4, I develop a theory of language acquisition in terms of canalization as an example of a substantial theory that meets the criteria. It explains a certain set of empirical observations used in VOS arguments, thus satisfying EC. I showed how some properties such as ubiquity, convergence, and fixed ontogeny might be correlated. Insofar as they are correlated, it is evidence that those properties may share a common developmental pathway, e.g., canalization. This theory described a cognitive structure

based on the Principles and Parameters account that is constituted by a network of switches that correspond to principles of syntax, phonology, or morphology. I described how knowledge of language is acquired at various stages of development where the values of switches are set with exposure to linguistic stimuli.

A fourth benefit in separating the strands of POS arguments is that it helps us recognize that there are different types of acquisition, and that we must avoid lumping them together under the term "innate." In Chapter 5, I argue that the conception of "innateness" is a broad category. This means that different concepts of "innateness" will be appropriate for different strands of POS arguments. A definition of "innateness" in terms of canalization ("c-innateness") is a useful and explanatorily valuable concept only insofar as it applies to processes of acquisition that are, in fact, canalized. It works well to solve puzzles about VOS and properties such as ubiquity, convergence, fixed ontogeny, knowledge of linguistic universals, and insensitivity to variation. This is not to say that the other strands of POS arguments are less effective, but further work is needed to elucidate the arguments and implications of the other strands.

Finally, I hope to contribute to the effort in understanding the complexity of language acquisition. I endorse "biologicizing the mind" as the right framework for this task. Traditional innateness theories make a distinction between internal factors of the organism and environmental conditions. This underestimates the complexity of language by creating a false dichotomy between nature vs. nurture, or genes vs. environment. The problem is not with this distinction itself, but in prioritizing the role of one over the other. Traditional innateness theories prioritize the importance of cognitive structures, content or concepts and underestimate the value of environmental stimuli. Biolinguists, however,

aim to develop theories that explain the complex interaction between internal mechanisms and environmental influences in trait acquisition.

The efforts made by biolinguists should also help unite nativists and empiricists. Their theories are constrained by empirical observations about language acquisition and discoveries about the structure of cognitive system. Consequently, empiricists and nativists will eventually converge on the same theory. This is not surprising, as each camp is involved in the joint task of discovering what it is about the mind/brain that accounts for how humans acquire language. Separating the four POS arguments will help nativists and empiricists identify empirical observations in need of explanation. I hope that the criteria for more substantial explanations will aid in efforts to create appropriate theories of cognition that explain the complex process of language acquisition.

REFERENCES

Ariew, A. (1996). Innateness and canalization. *Philosophy of Science, 63,* 19–27.

Ariew, A. (1999). Innateness is canalization: A defense of a developmental account of innateness. In V. Hardcastle (Ed.), *Where biology meets psychology: Philosophical essays* (pp. 117–139). Cambridge, MA: MIT Press.

Ariew, A. (2006). Innateness. In M. Mathen & C. Stevens (Eds.), *Handbook of the philosophy of science* (Vol. 3, pp. 567–584). Oxford: Elsevier.

Baker, C. L. (1981). *The logical problem of language acquisition.* Cambridge, UK: MIT Press.

Bateson, P., & Mameli, M. (2007). The innate and the acquired: Useful clusters or a residual distinction from folk biology? *Developmental Psychobiology, 49,* 818–831.

Behme, C., & Deacon, S. H. (2008). Language learning in infancy: Does the empirical evidence support a domain specific language acquisition device? *Philosophical Psychology, 21*(5), 641–671.

Berwick, R. C., Chomsky, N., & Piattelli-Palmarini, M. (2012). Poverty of the stimulus stands: Why recent challenges fail. In M. Piattelli-Palmarini & R. C. Berwick (Eds.), *Rich languages from poor inputs* (pp. 19-42). Oxford, UK: Oxford University Press.

Berwick, R. C., Friederici, A. D., Chomsky, N., & Bolhuis, J. J. (2013). Evolution, brain, and the nature of language. *Trends in Cognitive Sciences, 17*(2), 91–100.

Bickerton, D. (1983). Creole languages. *Scientific American, July 1983*, 116–122.

Birch, J. (2009). Irretrievably confused? Innateness in explanatory context. *Studies in History and Philosophy of Biology & Biomedical Science*, *40*(4), 296–301.

Boeckx, C. (2006). Linguistic minimalism: Origins, concepts, methods and aims. New York: Oxford University Press.

Brown, R. (1973). *A first language: The early stages.* Cambridge, MA: Harvard University Press.

Brown, R., & Hanlon, C. (1970). Derivational complexity and order of acquisition in child speech. In J. R. Hays (Ed), *Cognition and the development of language* (pp. 11-53) New York: Wiley.

Chapman, R. S. (2000). Children's language learning: An interactionist perspective. *Journal of Child Psychology and Psychiatry, 41*(1), 33–54.

Chomsky, N. (1975). Linguistics and philosophy. In S. P. Stich (Ed.), *Innate ideas* (pp. 188–198). Berkeley, CA: University of California Press.

Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.

Chomsky, N. (1986). *Knowledge of language*. Westport, CT: Praeger Publishers.

Chomsky, N. (1993). On the nature, use, and acquisition of language. In A. Goldman (Ed.), *Readings in philosophy and cognitive science* (pp. 511–534). Cambridge, MA: MIT Press.

Chomsky, N. (2000). *New horizons in the study of language.* Cambridge, UK: Cambridge University Press.

Chomsky, N. (2004). Language and mind: Current thoughts on ancient problems. In L. Jenkins (Eds.), *Variations and universals in biolinguistics* (pp. 379–406). Cambridge, MA: Elsevier Books.

Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry, 36*(1), 1-22.

Chomsky, N. (2006). *Language and mind* (3rd ed.). Cambridge, UK: Cambridge University Press.

Chomsky, N. (2007). Biolinguistic explorations: Design, development, evolution. *International Journal of Philosophical Studies, 15*(1), 1–21.

Chomsky, N. (2012). *The science of language.* Cambridge, MA: Cambridge University Press.

Chomsky, N., & Lasnik, H. (1995). The theory of principles and parameters. In J. Jacobs, A. von Stechow, W. Sternefeld, & T. Vennemann (Eds.), *Syntax: An international handbook of contemporary research* (pp. 506–569). Berlin: Walter de Gruyter.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences, 31*(5), 489–509.

Christiansen, M. H., & Kirby, S. (2003). Language evolution: Consensus and controversies. *Trends in Cognitive Sciences, 7*(7), 300–307.

Clark, A., & Lappin, S. (2011). *Linguistic nativism and the poverty of the stimulus.* New York: Wiley-Blackwell

Clark, A., & Lappin, S. (2013). Complexity in language acquisition. *Topics in Cognitive Science, 5*(1), 89–110.

Collins, J. (2005). Nativism: In defense of a biological understanding. *Philosophical Psychology*, *18*(2), 157–177.

Collins, J. (2008). *Chomsky: A guide for the perplexed*. Cambridge, UK: Bloomsbury Academic.

Collins, J. (2011). Innateness, canalization, and the modality-independence of language: A reply to Griffiths and Machery. *Philosophical Psychology, 24*(2), 195–206.

Cowie, F. (1997). The logical problem of language acquisition. *Synthese, 111*(1), 17–51.

Cowie, F. (1999). *What's within: Nativism reconsidered*. Oxford, UK: Oxford University Press.

Crago, M. (1992). Communicative interaction and second language acquisition: An Inuit example. *Tesol Quarterly, 26*(3), 487–505.

Crain, S. (1994). Language acquisition in the absence of experience. In P. Bloom (Ed.), *Language acquisition: Core readings* (pp. 364–409). Cambridge, MA: MIT Press.

Crain, S., & Nakayama, M. (1987). Structure dependence in grammar formation. *Language, 63*(3), 522–543.

Crain, S., & Thornton, R. (1998). Investigations in universal grammar: A guide to experiments on the acquisition of syntax and semantics. Cambridge, MA: MIT Press.

Curtiss, S. (1977). Genie: A psycholinguistic study of a modern-day "wild child." Boston, MA: Academic Press.

Devitt, M., & Sterelny, K. (1999). Language and reality: An introduction to the philosophy of language. Cambridge, MA: MIT Press.

Dor, D., & Jablonka, E. (2010). Plasticity and canalization in the evolution of linguistic communication: An evolutionary developmental approach. In R. Larson, V. Deprez, & H. Yamakido (Eds.), *The evolution of human language: Biolinguistic perspectives* (pp. 135-147). Cambridge, UK: Cambridge University Press.

Dove, G. (2012). Grammar as a developmental phenomenon. *Biology & Philosophy*, *27*(5), 615–637.

Elman, J., Bates, E. A., Johnson, M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1999). *Rethinking innateness* (4 ed.). Cambridge, MA: MIT Press.

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences, 32*(5), 429.

Fitch, W. T. (2009). Prolegomena to a future science of biolinguistics. *Biolinguistics, 3*(4), 283–320.

Fodor, J. A. (1974). Special sciences (or: the disunity of science as a working hypothesis). *Synthese, 28*(2), 97–115.

Fodor, J. (1981). The present status of the innateness controversy. In Fodor (Ed.), *RePresentations: Philosophical essays on the foundations of cognitive science* (pp. 225–253). Array Montgomery, VT: Bradford Books.

Gleitman, L. R., & Newport, E. L. (1995). The invention of language by children: Environmental and biological influences on the acquisition of language. *An Invitation to Cognitive Science, 1*, 1–24.

Goldin-Meadow, S. (2003). *The resilience of language: What gesture creation in deaf children can tell us about how all children learn language*. New York: Psychology Press.

Gottlieb, G. (2003). On making behavioral genetics truly developmental. *Human Development, 46*(6), 337–355.

Gould, J. L., & Marler, P. (1991). Learning by instinct. In D. W. Mock (Ed.), *Behavior and evolution of birds* (pp. 4–19). New York: W. H. Freeman and Co.

Gousti, M. T. (2004). *Language acquisition: The growth of grammar.* Cambridge, MA: MIT Press.

Griffiths, P. E. (2002). What is innateness? *The Monist, 85*(1), 70–85.

Griffiths, P. E., & Machery, E. (2008). Innateness, canalization, and 'biologicizing the mind.' *Philosophical Psychology, 21*(3), 397–414.

Griffiths, P., Machery, E., & Linquist, S. (2009). The vernacular concept of innateness. *Mind & Language, 24*(5), 605–630.

Grodzinsky, Y. (2000). Neurology and syntax. *Behavioral and Brain Sciences, 23*, 1-21.

Hauser, M. D., & Bever, T. (2008). A biolinguistic agenda. *Science, 322*(5904), 1057–1059.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science, 298*(5598), 1569–1579.

Hornstein, N., & Lightfoot, D. (1981). *Explanations in linguistics.* London, UK: Longman.

Hornstein, N., Nunes, J., & Grohmann, K. (2005). *Understanding minimilism.* New York: Cambridge University Press.

Hsu, A. S., & Chater, N. (2010). The logical problem of language acquisition: A probabilistic perspective. *Cognitive Science, 34*(6), 972–1016.

Hsu, A. S., Chater, N., & Vitányi, P. (2013). Language learning from positive evidence, reconsidered: A simplicity-based approach. *Topics in Cognitive Science, 5*(1), 35–55.

Jackendoff, R. (2003). *Foundations of language: Brain, meaning, grammar, evolution.* Oxford, UK: Oxford University Press.

Jackendoff, R., & Pinker, S. (2005). The nature of the language faculty and its implications for evolution of language (Reply to Fitch, Hauser, and Chomsky). *Cognition, 97*(2), 211–225.

Jenkins, L. (2000). *Biolinguistics: Exploring the biology of language.* New York: Elsevier Science and Technology.

Jenkins, L. (2004). *Variations and universals in biolinguistics.* San Diego, CA: Elsevier, B. V.

Jieqiong, W. U. (2014). An overview of researches on biolinguistics. *Canadian Social Science, 10*(1), 171–176.

Karmiloff-Smith, A. (2009). Nativism versus neuroconstructivism: Rethinking the study of developmental disorders. *Developmental Psychology, 45*(1), 56–63.

Kegl, J. (2004). Language emergence in a language-ready brain: Acquisition. In L. Jenkins (Ed.), *Variation and universals in biolinguistics* (pp. 195–236). San Diego, CA: Elsevier, B. V.

Kegl, J., Senghas, A., & Coppola, M. (1999). Creation through contact: Sign language emergence and sign language change in Nicaragua. In M. DeGraff (Ed.), *Language contact and language change: The intersection of language acquisition, Creole genesis, and diachronic syntax* (pp. 179–237). Cambridge, MA: MIT Press.

Khalidi, M. A. (2001). Innateness and domain specificity. *Philosophical Studies, 105*(2), 191–210.

Khalidi, M. A. (2007). Innate cognitive capacities. *Mind & Language, 22*(1), 92–115.

Khalidi, M. A. (2009). Should we eliminate the innate? Reply to Griffiths and Machery. *Philosophical Psychology, 22*(4), 505–519.

Laurence, S., & Margolis, E. (2001). The poverty of the stimulus argument. *The British Journal for the Philosophy of Science, 52*(2), 217–276.

Lewontin, R. (1974). The analysis of variance and the analysis of causes. *American Journal of Human Genetics, 26*, 400–411.

Lightfoot, D. (1982). The language lottery: Toward a biology of grammars. Cambridge, MA: MIT Press.

Lightfoot, D. (2013). Types of explanation in history. *Language, 89*(4), 18–38.

Lorenzo, G., & Longa, V. M. (2009). Beyond generative geneticism: Rethinking language acquisition from a developmentalist point of view. *Lingua, 119*(9), 1300–1315.

MacWhinney, B. (2004). A multiple process solution to the logical problem of language acquisition. *Journal of Child Language, 31*(4), 883–914.

Mameli, M., & Bateson, P. (2006). Innateness and the sciences. *Biology & Philosophy, 21*(2), 155–188.

Maratsos, M. P. (1989). Innateness and plasticity in language acquisition. In M. L. Rice & R. L. Schiefelbushch (Eds.), *The teachability of language* (pp. 105-125). Baltimore, MD: Paul H. Brooks.

Marcus, G. F., (1993). Negative evidence in language acquisition. *Cognition, 46*, 53-85.

Marcus, G. F., & Rabagliati, H. (2006). What developmental disorders can tell us about the nature and origins of language. *Nature Neuroscience 9*, 1226-1229.

Margolis, E., & Laurence, S. (2013). In defense of nativism. *Philosophical Studies, 165*(2), 693–718.

Mayr, E. (1982). *The growth of biological thought: Diversity, evolution, and inheritance*. Cambridge, MA: Belknap, Harvard University Press.

McNeill, D. (1996). Developmental psycholinguistics. In F. Smith & G. A. Miller (Eds.), *The genesis of language: A psycholinguistic approach* (pp. 15-84). Cambridge, MA: MIT Press.

Moore, C. (1984). Maternal contributions to the development of masculine sexual behavior in laboratory rats. *Developmental Psychobiology, 17*, 346–356.

O'Neill, E. (2015). Relativizing innateness: Innateness as the insensitivity of the appearance of a trait with respect to specified environmental variation. *Biology & Philosophy, 30*(2), 211-225.

Petitto, L. A., & Marentette, P. F. (1991). Babbling in the manual mode: Evidence for the ontogeny of language. *Science, 251*(5000), 1493–1496.

Piattelli-Palmarini M. (1980). *Language and learning: The debate between Jean Piaget and Noam Chomsky* (Ed.). Cambridge, MA: Harvard University Press.

Pinker, S. (1994). *The language instinct*. Cambridge, MA: MIT Press.

Pinker, S. (2013a). *Learnability and cognition: The acquisition of argument structure* (2nd ed.). Cambridge, MA: MIT Press.

Pinker, S. (2013b). *Rich languages from poor inputs.* Oxford, UK: Oxford University Press.

Prinz, J. (2002). *Furnishing the mind: Concepts and their perceptual basis*. Cambridge, MA: MIT Press.

Prinz, J. (2012). *Beyond human nature: How culture and experience shape the human mind*. New York: W. W. Norton & Co.

Pullum, G. K., & Scholz, B. C. (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review, 18*(1-2), 9–50.

Putnam, H. (1967). The "innateness hypothesis" and explanatory models in linguistics. *Synthese, 17*(1), 12–22.

Quintero, K. W. (1992). Learnability and the acquisition of extraction in relative clauses and wh-questions. *Studies in Second Language Acquisition, 14*(1), 39-70.

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax: The nature of early child grammars of English*. Cambridge, MA: Basil Blackwell.

Ramscar, M., & Yarlett, D. (2007). Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition. *Cognitive Science, 31*(6), 927–960.

Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language, 47*(1), 172–196.

Samuels, R. (2002). Nativism in cognitive science. *Mind & Language, 17*(3), 233–265.

Samuels, R. (2004). Innateness in cognitive science. *Trends in Cognitive Sciences, 8*(3), 136–141.

Samuels, R. (2007). Is innateness a confused concept. In P. Carruthers, L. Laurence, & S. Stich (Eds.), *The innate mind: Foundations and the future* (Vol. 3, pp. 17-36). Oxford, UK: Oxford University Press.

Seidenberg, M. S. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science, 275*(5306), 1599–1603.

Senghas, A., & Coppola, M. (2001). Children creating language: How Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science, 12*(4), 323–328.

Singleton, J., & Newport, E. (2004). When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognitive Psychology, 49*, 370–407.

Skinner, B. F. (1957). *Verbal behavior*. Acton, MA: Copey Publishing.

Sober, E. (1998). Innate knowledge. In E. Craig & L. Floridi (Eds.), *Routledge encyclopedia of philosophy* (pp. 794–797). Cambridge, MA: Routledge.

Stich, S. P. (1978). Empiricism, innateness, and linguistic universals. *Philosophical Studies, 33*(3), 273–286.

Stromswold, K. (1999). Cognitive and neural aspects of language acquisition. In E. Lepore & Z. Pylyshyn (Eds.), *What is cognitive science?* (pp. 356–400). Cambridge, MA: MIT Press.

Stromswold, K. (2000). The cognitive neuroscience of language acquisition. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 909–932). Cambridge, MA: MIT Press.

Waddington, C. H. 1957. *The strategy of the genes: A discussion of some aspects of theoretical biology*. London, UK: Ruskin House.

Wilbur, R. B., Klima, E., & Bellugi, U. (1983). Roots: On the search for the origins of signs in ASL. *Chicago Linguistic Society* 19, 314–336.

Wimsatt, W. C. (1986). Developmental constraints, generative entrenchment, and the innate acquired distinction. In W. Bechtel (Ed.), *Integrating scientific disciplines* (pp. 158–208). Dordrecht: Martinus Nijhoff.