2007

Minimally innate ideas

Michele Merritt
University of South Florida

Minimally Innate Ideas

by

Michele Merritt

A thesis submitted in partial fulfillment
of the requirements for the degree of
Masters of Arts
Department of Philosophy
College of Arts and Sciences
University of South Florida

Major Professor: Daniel Weiskopf, Ph.D.
Stephen Turner, Ph.D.
Alexander Levine, Ph.D.

Date of Approval:
April 10, 2007

Keywords: language acquisition, representational nativism, usage-based linguistics, emergentism, interactivism, connectionism

To Grandma…

The Smartest Woman I'll Ever Know

Acknowledgements:

I am especially thankful for the time, patience, and dedication of my chair, Dr. Weiskopf, and my committee members, Dr. Levine and Dr. Turner. Thank you for your invaluable insights, expertise, praise, and yes, even your criticism.

Thanks to all the professors and staff of the Philosophy Department at USF. I am grateful to have had the opportunity to undertake this project and have learned so much from the seminars, directed studies and teaching assignments.

And to my friends and family: you are perhaps the real cause of my being able to complete this milestone. Whether monetary or emotional, your love and support made it possible to get through even the toughest moments. It will never be forgotten.

Table of Contents

Minimally Innate Ideas

Michele Merritt

ABSTRACT

This project provides a detailed examination and critique of current philosophical, linguistic, and cognitive accounts of first language acquisition. In particular, I focus on the concept of "innate" and how it is embraced, marginally utilized, or abandoned altogether in efforts to describe the way that a child comes to be a competent user of a language. A central question that naturally falls out of this general inquiry is therefore *what* exactly is supposed to be "innate," according to various theories? Philosophically, the theory of innate ideas put forth to explain human learning has existed for centuries and hence, this thesis as it relates to language is discussed. The revival of nativism by linguists like Chomsky is thus a central theme of the first chapter. Universal Grammar and the various arguments for it are closely scrutinized, and I close this chapter with what I take to be the commitments of linguistic nativism, how its proponents conceive of "innate," and several possible objections to the arguments they put forth.

Just as the theory of innate ideas has had its contesters throughout the history of philosophy, so too have linguists and cognitive scientists rejected Universal Grammar and other forms of linguistic nativism. Thus, the second chapter presents several of these alternative explanations of language acquisition. Namely, I divide the chapter into three sections, Usage-Based Linguistics, Emergentism, and Sociolinguistic Acquisition, as it is my suggestion that most of the anti-UG attacks are levied from one of these three fields. In discussing the details of each, two distinctions become of particular concern: first, a large part of the differing conceptions of "innate" seem to hinge on what is meant by "learning" and "acquiring," and therefore second, a fine line between UBL and Emergentism can be drawn, a relationship that is otherwise conflated in the literature.

Because chapter two involves a brief account of the way in which connectionist simulations are often utilized to model or represent language acquisition, particularly from an Emergentist perspective, chapter three begins by examining this feature of Emergentism in more detail. Due to its explanatory power, ability to be effectively modeled, and the evidentiary support found in neuroscience, Emergentism would appear to be the most tenable position to maintain regarding language acquisition. This possibility seems further strengthened when we take into account the neuroscientific data often used to bolster anti-nativist claims. Nevertheless, reflecting on the overarching concern of the project, regarding what is really meant by "innate," it is shown that this attack on nativism might stand on shakier ground than was originally assumed. Finally, based on these considerations, a case is made for an intermediary position, a theory of "Minimally Innate Ideas."

Introduction

The concept of "innate" can undoubtedly be traced as far back in philosophical history as Plato, if not further. In the *Meno,* for example, we find the argument that all human learning is actually nothing more than 'remembering' what was once known in a previous life.[1] Not only is Plato's discussion of innateness one of the first recorded, it is probably, by today's standards, one of the most *extreme,* in that he means to suggest that *all* knowledge, from the ability to walk to the understanding of calculus, is etched onto the minds of infants. Thus, as the child grows and begins acquiring these skills, she is not learning at all, but simply reacquainting herself with knowledge, which was there all along. This radical nativism eventually comes under attack by the empiricist thesis that there can be no innate ideas; rather, as Locke claims, the infant comes into this world, a *tabula rasa*.[2] All learning, therefore, is based on trial and error, observation, or instruction. In another vein of empiricist critique, we find arguments such as those put forth by Hume, that without any experience or sensory impressions, we can form no ideas. Hence, concepts like God, necessary connection, and universality are out of the question for nearly all empiricists.[3] Nevertheless, as the debate between nativists and empiricists has yet to cease, we find, especially in the last few decades, a resurgence of arguments which seek to uphold something not altogether dissimilar to Plato's 'innateness hypothesis' put forth over two thousand years ago. Likewise, these "neo-nativists" have met with their own opponents, those who bolster their anti-nativist claims

[1] cf. 84b-98a
[2] *An Essay Concerning Human Understanding*, Book II
[3] *Enquiry Concerning Human Understanding,* VII, See also, *A Treatise of Human Nature*, I.

with arguments based on psycholinguistic models and neuroscientific data. The position that these new non-nativists are committed to is perhaps not identical to the what the empiricists of the eighteenth century had in mind, however, they share in the tradition of denying any knowledge, specifically of the linguistic sort, to exist prior to experience. And thus, the protracted battle continues.

The cursory history of "innateness" given above was not in any way meant to be exhaustive. Indeed, it does not even skim the surface of the rich historical tradition to which the term owes its fame. I mention some of the major contenders in order to set the stage for the project I am undertaking. Based on this rudimentary history of innate ideas, I have sketched four distinct perspectives on the matter. These are not intended to be the only positions, nor is it always the case that the figures mentioned paid specific attention to language, but insofar as it is possible, at least at this point, to subsume language under each general theory in some way consistent with other phenomena, the list will provide an initial conception. Indeed, it will be edited along the way.

> Plato, Descartes, et al; "The Rationalists": The capacity for and all the representations necessary to language are *a priori*, in the sense that they exist in our minds prior to any experiences, and are known without any recourse to observation, instruction, or experience in general.
>
> Locke and the Empiricists: The mind is a *tabula rasa* and hence, all linguistic knowledge is gleaned through experience. While there might be predispositions or capacities that exist prior to any experience, knowledge of anything, including language, is gained by inductive or *a posteriori* means.
>
> Neo-Nativists: Something is innate if it provides the necessary blueprints for development. All of the principles of language are "hardwired" in the mind of the infant at the onset of development, and are therefore innately given. Experience merely "flips the switches" in the already prespecified language module.
>
> The "New Wave" of Empiricists: Neurobiological evidence suggests that the existence of any "innately specified" information pertaining to language is

> dubious. Based on psycholinguistic models which are able to replicate the language learning experience, it is clear that linguistic representations are not necessary for the process to get off the ground. Perhaps a less robust notion of "innate," which would allow for the architecture of the mind to be pre-specified, will work, but claiming that all the principles of grammar are innate does not jibe with the given data.

This list will be modified as the project unfolds, but for now, it serves as a place from which to begin, insofar as it allows me to make one initial claim and to frame several questions for further inquiry as well. First, to the extent that we ought to choose only one of the abovementioned theses regarding the degree to which language is innate as the correct position is I think, misguided. As we shall find, especially relating to the neo-nativists and the new wave empiricists, there exist data that support both nativist and non-nativist accounts. Thus, one immediate question will be *exactly how are these data being used?* In other words, if evidence from neurobiology can both support and deny nativism (N and ~N), is this really a genuine case of contradictory data, or more a case of utilizing data, as well as the very terms employed in analyzing the data, in a manner convenient to one's argument? Furthermore, assuming that only one of the four positions above captures the truth might be leaving out some fifth option. Even by the very crude nature of my illustration of each, it is clear that no one wants to claim that *absolutely nothing* concerning language is innate, nor does it seem, that anyone, Plato excluded perhaps, wishes to suggest that everything about language, from the ability to speak to the understanding of *wh*-movement, is known prior to experience. Thus, some overarching questions I hope to answer are whether language is truly "innate" in the sense that all acquisition of it proceeds via taking principles already possessed and then mapping them onto discrete samples, which would be a rather deductive account of learning. Or, is it the

case that language truly is *learned*, in the most robust sense of the term, meaning that general principles or rules of the language are only gleaned once the child has experienced enough individual examples? Finally, is it possible for the process of coming to be a competent user of a language to have both deductive and inductive characteristics? Several corollary inquiries naturally fall out from these primary questions, such as whether or not any of the four positions mentioned above adequately expresses the nature of language learning, if there needs to be some other position, whether this position is entirely novel or if it is just an amalgam of several others, and so forth.

In order to satisfy these inquiries, I will look at several theories which purport to answer the question of language and innateness in varying ways. A short glance at any one of these will however, reveal the complex ambiguity and difficulty in ascribing an exact innate/not-innate label to language. It would be nice if I could say that in the following three sections I will look at Universal Grammar, Constructivism, and Connectionism, which argue that 1) Language is Innate, 2) Some of Language is Innate, and 3) Language is not Innate, respectively. While this statement is not entirely accurate, it is not completely off the mark, either. In what follows, I will need to look at not only what each theory has to say about how children 'learn' or 'acquire' language, but more importantly, I will closely examine terms like "innate," "learned," and "acquired," to see whether or not they carry the same meanings across each argument. In addition, I will offer some compelling empirical evidence obtained through experimentation, all of which are intended to prove one of the abovementioned "acquisition theories."

Chapter 1: Nativism and Universal Grammar

1.1 Brief History and Description

By now, it would be rare to meet someone who has not at least heard of the name 'Chomsky.' His theories, both linguistic and political, have been groundbreaking. It is difficult to avoid his looming presence, and this project could conceivably devote several hundred pages to Chomsky's influence on language studies alone. The task at hand however, is to determine precisely what Chomsky's seminal thesis of *Universal Grammar* has to say about "innateness." One might easily pass this question off as already answered, and indeed, the textbook definition of UG calls it a "linguistic theory, which claims that the human mind contains an innate set of principles and parameters which guide acquisition."[4] We can initially say that a Universal Grammarian holds that "principles and parameters" are innate, but determining precisely what that means is somewhat tricky. Moreover, the purported empirical evidence for these "innate" blueprints of linguistic knowledge is subject to serious questioning, as shall soon be discussed.

Understanding the impetus for Chomsky's position might also be helpful, as it brings into focus the status of the innate/not-innate debate on language prior to UG. Behaviorism, the dominant psychological theory in the early part of the twentieth century, was championed by B.F. Skinner, who claimed that all learning proceeds via habit formation. In his book, *Verbal Behavior,*[5] he made a case for language falling under this rule as well. Language, just like any other skill, is best learned by repetition,

[4] cf. O'Grady et al (2005). *Contemporary Linguistics*, Ch. 1 and 11.
[5] 1957

mimicking, and trial and error. This is because according to behaviorism, the human mind does not come equipped with any "blueprints" for language acquisition or any other schemata, for that matter. It is only together with the environment that the infant, in a stimulus-response type of relationship, can come to be a competent user of a given language.[6]

Universal Grammar can be viewed as a direct attack on behaviorism, but that assessment alone would be too narrow. Certainly, the assertion that children do not "learn" languages, but simply "acquire" them presents a devastating blow to the *process* argued by behaviorists to be involved in obtaining language knowledge, but an explicit linguistic critique is made by UG as well. For Chomsky, it became essential to look, not so much at the *differences* among languages, but more at the *similarities* they all seem to share. In his seminal linguistic treatise, *Syntactic Structures*,[7] he demonstrated that all human languages are necessarily generative; that is, the grammar of every language suffices to 'generate' all and only the correct grammatical constructions of that language. Thus, equipped with the principles of a generative grammar, the speaker of that language can discriminate between grammatical and ungrammatical phrases, even if those phrases are entirely novel. While generativity of grammar was a doctrine shared by Chomsky with the Structuralists of that time, he further added that grammars must also be *transformational*, meaning that they have entirely unrestricted rewrite systems. This allows for a comprehensive description of the order of transformations, like the ones governing productive conjunction. While transformational generative grammars (TGGs

[6] cf. Skinner (1938) *The Behavior of Organisms: An Experimental Analysis of Behavior.*
[7] 1957

for short) have elicited much debate,[8] Chomsky has maintained that the descriptive scope of TGGs is not intended to be extensive; it is limited to describing specific principles of human grammar. Therefore, objectors to TGG who claim that it cannot solve the parsing problem are mistaken in their assumption that TGG need concern itself with such issues. Rather, Chomsky claims, problems such as these are for computational linguistics.[9]

The reason that acquainting the reader with Chomsky's early work on TGG is helpful is because it sheds light on the question posed above, concerning exactly what is meant by arguing that the human mind comes equipped with a set of innate principles. One criticism leveled against TGG stems from the very reasoning behind Chomsky's insistence for transformationality. According the 'Chomsky Hierarchy' of grammars, regular grammars, context-free grammars and context-sensitive grammars cannot account for the infinite recursive enumerability of human language, and thus, to explain this infinite production, Chomsky maintained that the grammar of human language must be unrestricted. However, without constraints, the argument that language is rule-governed falls apart, as does the capacity to show that grammar is decidable. Nevertheless, in continuing on after *Syntactic Structures*, Chomsky sought to demonstrate 1.) how child language acquisition can be explained via the TGG model, *and* 2.) that TGG does have constraints, and is therefore, decidable. As was mentioned, Universal Grammar is a linguistic theory which claims that the human mind inherits a universal set of principles and parameters that are responsible for shaping the language eventually acquired. Part of

[8] cf. Shieber, Stuart (1985). "Evidence against the context-freeness of natural language". *Linguistics and Philosophy* 8: 333-343.
[9] *Aspects of the theory of Syntax* (1965)

Chomsky's *Government and Binding* theory[10] is devoted to detailed explication of what sorts of things constitute principles and parameters, but it suffices to say that principles are the unvarying and universal aspects of *all* human languages, while parameters are responsible for the differences we notice among languages. Parameters have a finite number of open values which, once set, will vary from language to language. This parametric variation, or all the possible set values, is not what he argues to be innately present in the mind. This would imply that the child is born with something resembling knowledge of all the languages in the world. Rather, as Chomsky puts it:

> We can think of the initial state of the faculty of language as a fixed network connected to a switch box; the network is constituted of the principles of language, while the switches are the options to be determined by experience. when the switches are set in one way, we have Swahili; when they are set in another, we have Japanese. Each possible human language is identified as a particular setting of the switches-a setting of parameters, in technical terminology. If the research program succeeds, we should be able literally to deduce Swahili from one choice of settings, Japanese from another, and so on through the languages that humans can acquire. The empirical conditions of language acquisition require that the switches be set on the basis of the very limited information that is available to the child. Notice that small changes in switch setting can lead to great apparent variety in output, as the effects proliferate through the system. [2000, 8][11]

Thus, to call Universal Grammar a grammar at all, is really a misnomer. UG refers to an initial state of the brain of any human, according to Chomsky. The grammar a child comes to adopt will depend on her interaction with the linguistic environment, which will in turn, set the "switches" to a definite value. Those definite values, or parameters, are responsible for constraining the grammar of the particular language acquired, and yet,

[10] cf. (1981,1993). *Lectures on Government and Binding: The Pisa Lectures*.
[11] *New Horizons in the Study of Language and Mind*.

regardless of the language, it will be recursive, infinitely generative, and transformational.

With this brief account of the general theory purported by Universal Grammar, we can make a couple preliminary claims regarding the role that "innate" plays. First, as to the meaning of the word itself, it is clear that those in favor of Universal Grammar argue that for something to be innate means that it is genetically encoded information, already present in the brain of an infant, at the moment of birth.[12] To be more specific then, what is meant by "genetically encoded information" is not necessarily a consciously available set of rules or explicit grammatical knowledge, at least not in the sense that an infant can linguistically represent something like phrase structure to herself. Rather, the information an infant is said to come equipped with resembles more a capacity or aptitude, than an already given set of facts. This is precisely why it is crucial that the Universal Grammarians maintain a sharp distinction between principles and parameters, the former being the universally innate linguistic knowledge applicable to all languages, and the latter, the specific information necessary for each individual language. Parameters are often referred to as the 'switches' that exist in the minds of every infant and they are set at some neutral position, much like a zero on an integer line, but once the environment has a chance to make its influence, they get 'flipped' such that they become charged, either on the + or – side of the line. It seems therefore, that if any confusion arises as to what falls under the "innate" umbrella, it stems from not understanding exactly what type of information counts as a "principle" and what counts as a "parameter." Looking at the standard arguments put forth in support of Universal

[12] Or even earlier, depending on at which point in brain development these capacities are argued to exist.

Grammar, it will perhaps be easier to get clear about exactly what these principles and parameters are intended to *be* and *do*.

1.2 Arguments for UG

There have been and remain to be, several prototypical arguments which comprise the core or foundation for all empirical research intended to prove UG's validity. These arguments can of course trace their development back to Chomsky, but several Universal Grammarians, like Pinker,[13] for example, have either added to or modified the original claims as time has progressed. For the sake of economy and organization, I have chosen to follow Geoffery Sampson's account of the most central cases made for UG, as he illustrates both the original arguments, as well as the more contemporary ones. In a highly influential book concerning the language instinct debate, Sampson presents several striking blows to nativists and Universal Grammarians alike. *Educating Eve*[14] is only the first of these attacks from Sampson, and I plan to discuss several of his objections to UG, found mostly in his new book *The Language Instinct Debate*[15] in turn. At present however, I simply want to trace the arguments as he presents them, in order to determine how "innate" gets used among the various formulations.

1.2.1 Speed of Acquisition

Although not the first to point this out, to be sure, Chomsky was impressed by the rate at which children acquire their L1. Language is learned remarkably effortlessly,

[13] 1994, 2002
[14] 1999
[15] 2005

uniformly, and rapidly, he says, which is unlike the way other bodies of knowledge are acquired. Physics, for instance, "is acquired selectively and often painfully, through generations of labor and careful experiment, with the intervention of individual genius and generally through careful instruction…"[16] In another vein of this same type of argument, the "word explosion,"[17] has been used to bolster claims like Chomsky's, suggesting that the speed of acquisition of individual words at a young age is evidence for innately given guiding principles. Between the ages roughly of 18 months and three years, children "explode" with new words, somewhere in the neighborhood of one to two thousand newly produced utterances, many of which they have never heard. Because of the ease and rapidity with which *all* infants achieve this amazing feat, there is certainly cause to believe in something universal about language development. Moreover, the speed with which children acquire language all over the planet provides evidence that social-cultural and environmental factors seem to have very little if anything to do with the process, and hence, there must be an innate ability possessed. Insofar as infants begin erupting with new words without much influence from the environment-indeed, many words they produce were heard only once, and were certainly never encouraged by Mommy and Daddy-there is also evidence suggesting that they have an innately endowed mechanism which allows them to proliferate words with relatively limited exposure.

1.2.2 Age Dependency and Acquisition versus Learning

Another argument that deals with rate of acquisition, but adds to it, an age-dependency, points to a "critical window" of opportunity children have to acquire their

[16] *Reflections* (1975), 144.
[17] cf. Woodward, et al (1994).

L1. As the familiar case of the language deprived 'Genie'[18] demonstrates, the severely impeded development that occurs as the result of no exposure might lead one to believe that the language acquisition device (LAD) is only fully functional before the onset of puberty. A contrast between L1 and L2 is also helpful here. At any age, an adult can decide to take on a second language, and this "learning" will usually proceed slowly and with difficulty. This is in direct opposition to the way in which the L1 is "acquired" by a normal infant. In a strange scenario like Genie's, however, it is arguably the case that the LAD is no longer operational and instead, the operative mechanism is more of a Language Learning Device (LLD), one tuned more towards experiential learning, much like the way an adult might consciously try to learn a new language. Thus, the infant can be said to tackle the daunting task of learning a language almost entirely unconsciously or without explicit understanding of exactly what she is doing; it just happens. Hence, proponents of UG draw yet another distinction, between acquiring and learning. For Chomsky and his followers, the proper way to characterize an infant's developing ability to use her native language is to call it "acquiring" the L1. Since this acquisition of language is said to be wholly natural and unavoidable given a normal, non-Genie-like environment, it can be juxtaposed to other types of knowledge acquisition, such as computer programming, a skill that is necessarily "learned." The acquire-learn distinction is another way to argue for an innately endowed language faculty, which is consequently only tuned and ready for acquisition during a particular developmental timeframe.

[18] Curtiss, (1977).

1.2.3 Poverty of the Stimulus

Perhaps the most often empirically tested claim for UG is the one stating that the data with which the infant has to work are severely impoverished. The "Poverty of the Stimulus" (POS) argument can be put in this form:

> 1.) The actual linguistic environment for the infant is limited in its correctness and is highly unrepresentative of the complexities of the language itself.
> 2.) Yet, all children acquire the ability to use the language in all its richness, and proceed to do so with relatively little error correction or feedback
> 3.) Thus, there is a "disparity between the generative grammar that expresses the Linguistic competence of the native speaker and the meager and degenerate Data on the basis of which he as constructed his grammar."[19]
>
> Therefore, the infant must be innately endowed with the principles necessary to acquire such a system. (Because with the data alone, it is assumed, there is no way such a system would ever get acquired.)

Two components of POS arguments are worth noting here. First, without explicit instruction and sometimes even with exposure to *incorrect* or *incomplete* instantiations of the grammar, the child comes to produce correct utterances, has equal competence to any other native speaker, and her language is as rich and complex as the grammar itself. In this way, principles such as *anaphora*, although rarely gleaned from input, and certainly never taught, are nevertheless understood at as young as 18-months of age, and produced without thought or question once constructions larger than two words begin occurring.[20] Second, the notion of *error correction* and *feedback* are important as well. You simply do not find parents correcting their children regularly, and when this is observed, it tends to do very little in terms of changing the natural course of development.[21] For instance, a child might utter *Mommy goed to work*, and the Father might recast this statement to the

[19] *Language and Mind*, 102-33.
[20] See the following section for a detailed description of this.
[21] cf. Tomasello (2002).

child: *Yes, Mommy went to work.* However, until the child is ready to accept that her overgeneralization of the –ed rule will not extend to irregular verbs like 'go,' this corrective recast will go unnoticed and unutilized.

1.2.4 Systematicities Among and Across Developmental Sequences

The fact that children acquire correct forms, even when exposed to incorrect ones, is the basis for another argument for UG, stating that intelligence, specific exposure, social situation, and other conditions of learning have very little effect on the grammar that gets acquired. All speakers of a given language, children and elders alike, converge upon the same grammar, with only minor variations among individuals. Regardless of the fact that some children grow up to be nuclear physicists with intelligence quotients of two hundred and others attain only minimal standards of intelligence, their grammars will be nearly identical. And since knowledge of language does not depend on talent, overall intelligence or "learned" skills, we can say that it is a universally given capacity. As Sampson puts it: "With other skills, say, driving a car, we find a few people who keep on taking test after test but who never get a license; we do not find people who never crack the task of first-language acquisition."[22]

1.2.5 Linguistic Universals

One argument that has received a superfluity of positive feedback as well as plenty of negative response has to do with linguistic universals. There are principles

[22] Sampson, 29

which, although they do not necessarily belong to all conceivable languages, do in fact govern all the actual languages of which we know. It is Chomsky's claim that because of the overwhelming evidence this is true, it is an easy leap from this notion to the argument that knowledge of these principles is innate.[23] For example, take the principle of *structure-dependency*, or the idea that rules are applied to strings of words by virtue of the organization of the words into phrases.[24] When formulating a yes/no question out of the declarative sentence, *The job candidates, who will be in town next week, will stay at the Omni,* the phrase structure rewrite rule says that the question will look like: *Will the job candidates, who will be in town next week, stay at the Omni?* To get this rule, "we need to know not just that *will* is a verb, but that the second instance of that word is not part of any clause smaller than the whole sentence."[25] Thus, we never get: **Will the job candidates, who be in town next week, will stay at the Omni?* Since no child ever makes the mistake just illustrated, it is argued that the mind is innately endowed with knowledge of phrase structure, such that the child can not only produce yes/no questions correctly, but can also work out what sort of question is being asked. Chomsky says that there could easily be structure-independent rules that would produce the same forms, but English, and apparently no other human language uses such rules.

The scope of linguistic universals also encompasses errors. In analyzing errors produced by children acquiring their L1, we do not find wild and random mistakes. Instead, errors, like other developmental sequences, are systematic. There are some error sequences through which all children of every language seem to pass, such as

[23] However, it is still an open debate presently in theoretical linguistics, as to whether or not LU are as robust and prevalent as those like Chomsky suggest they are.
[24] Language and Mind, 51.
[25] Sampson, 29

overgeneralizing on past tense –ed constructions, causing things like *Daddy goed to the store* to be uttered. Errors like these might even come *after* the correct form has been produced several times, much to the chagrin of many parents, but then after some time, the child suddenly picks up the correct rule again. While errors like overgeneralization are universal and proceed in a systematic way, another feature or error systematicity prohibits certain errors from occurring. For instance, no child ever makes the mistake of assuming that since the placement of the adjective *red* in *I painted the red barn*, can also be moved behind the noun to say, *I painted the barn red*, that we can do the same thing with the sentence, *I saw the red barn,* such that it would read, **I saw the barn red.* All of these above universals are said to be part of the tacit and innate understanding about language with which children come equipped; therefore, they never get learned, but are instead, acquired.

1.2.6 Argument from Analogy

One can also argue for UG from analogy, as Chomsky has in the past. Basically, the argument runs: The organ that regulates blood sugar, the pancreas, does not need to be "taught" how to do its job; neither do the lungs need to be taught to breathe, the heart to pump blood, and so forth. Similarly, the "language organ" works as any other organ in the human body and does not need to be taught, but only needs to be triggered by the environment to do begin doing its thing. Of course, this type of analogy is only as strong as the actual proof supporting its implicit premise, namely, that there is a "language organ." Another tenet of UG maintains that there is a specific language module in the brain, and although this claim will come under scrutiny in Chapter III when discussing

Connectionism, according to Chomsky and others, the language module most certainly exists.

The language module is also the basis of one final argument that may or may not be put forth in favor of UG. As Sampson points out, Chomsky always places humans in a category all their own, when it comes to language. Even regarding the infamous "chimp" experiments conducted by those like Allen and Beatrice Gardner, ongoing since the 1950s, "Chomsky is one of those who minimizes what was achieved."[26] Even if chimps can learn sign language or bees can give directions, these are not sufficient data to support non-human animal language. It is unclear whether or not Chomsky intends for the exclusivity of language to humans to be explicit evidence in favor of UG. Indeed, he claims that if it were discovered that animals, such as apes, do in fact have a language faculty, then we would have to allow that apes have language.[27] Nevertheless, even if we ascribe a specific language module to every species with a central nervous system, the argument that this module is innate can undoubtedly be maintained. It just so happens, according to Universal Grammarians, that humans are the only ones with this innate genetic endowment.

1.2.7 Language Mutants

In his book, *The Language Instinct,*[28] Steven Pinker devotes a large amount of discussion to many very similar proto-nativist arguments put forward by Chomsky. Since the preceding sections have already dealt in great detail with these claims, it is not

[26] Sampson, 34
[27] *Rules and Representations, 144*
[28] 1994

necessary to trace each of Pinker's reiterations of the arguments, as they simply reemphasize what Chomsky already said. One argument, however, is worth noting, as it presents a more contemporary side of the nativist position by focusing on a recent case study, which has been purported to demonstrate both modularity of language in the mind, and the innate nature by which this faculty is either endowed or left absent.

The study was originally carried out by Gopnik and Crago in 1990,[29] although Pinker's treatment of it is the most extensive to date, amongst all the nativist literature. It centers on a family in Essex, U.K., of whom half the members have a Specific Language Impairment (SLI),[30] namely, the impairment of functionality or inability to control their grammar. For example, when asked to pluralize nonsense nouns, members with the disability would not generalize from analogous 'real' nouns in English. Rather, they would create nonsense plural forms of words like *zat*-which one subject responded that the plural of this word would be *zacko.*[31] A normal speaker of English, however, would know that for this word, you simply add an –s, whereas for words like *sas*, you would need to add an –es, and not simply hold the –s sound at the end, which is what subjects with the SLI did. In this particular family, only half of its members were affected, while the remaining half performed normally on all the language tests.

According to findings like the ones above, the case for genetic specificity of language appears to be quite striking. Insofar as the observed family all lived together, and yet, only some of them had the SLI, it seems highly unlikely that the impairment was a result of environmental or emotional factors. As Pinker and others suggest, the SLI

[29] Published 1991.
[30] Gopnik *The Inheritance and Innateness of Grammar, 1997.*
[31] From Sampson ,1999

must be a genetically encoded deficiency, one which specifically impairs the language module. As well, since no amount of training, 'growing out of,' or error feedback could remedy certain grammatical problems, as most of the above findings were from tests conducted with adult speakers, there is evidence that the impairment is an innately endowed phenomenon. If language deficiencies can be shown to be innate, then, as Pinker argues, so can normal and correct command of the language.

1.3 Empirical Evidence for UG, Empirical Evidence not for UG

The arguments in favor and disfavor of UG are not based solely on qualitative observation, nor are they strictly theoretical. Often, experiments are carried out in an effort to prove or disprove one side of the issue. In particular, testing the validity of the POS argument for UG is highly fashionable. For example, an experiment conducted by Lidz, Waxman, and Freedman (LWF) takes issue with the view that analysis of input is sufficient to explain grammatical acquisition in infants. Instead, they argue that input does not contain enough information to promote unaided learning and furthermore, evidence that 18-month old infants have command of the use of the anaphoric unit 'one' suggests that an innate syntactic structure aids the learner in acquiring a grammar.

In the study, infants were seated in front of a large television screen and were alerted to images that appeared. For the familiarization phase, a single object was shown three times alternating from the left to right sides of the screen. The presentations were accompanied by a recorded voice, for example, *'Look! A Yellow Bottle.'* During the test phase, the same procedure took place except that two images of the same object appeared, one the same color as that of the familiarization trial, the other of a different

color. (The same image of a yellow bottle, now alongside a blue bottle, for instance) There were two conditions within the test phase, a control and an anaphoric group. Both groups were exposed to identical conditions except for the linguistic stimulus. The control condition was accompanied by a phrase such as '*Now look. What do you see now?'* The anaphoric group heard a phrase containing a relative pronoun like, '*Now look. Do you see another one?*' Looking times were recorded for all the phases with the prediction being that the infants would prefer to look at an image that matches the linguistic stimulus.[32]

The results were as follows: For the control condition in which no linguistic cue was offered, infants preferred to look at the novel image (in this case, the blue bottle). In the test phase, however, with the introduction of the anaphoric 'one,' infants preferred to look at the familiar object (the yellow bottle). These findings demonstrate that at 18 months, infants represent 'one' as anaphoric to the category presented in the familiarization phase. In other words, upon hearing the phrase, '*Now Look. Do you see another one?,'* infants are able to make the association between the pronoun 'one' and the category to which it refers, thereby representing a nested concept. Because infants in the very beginning stages of being able to productively combine words directed their attention at the correct image to which the anaphoric 'one' referred, evidence for the "poverty of the stimulus" argument is thereby offered. The input that the infants were exposed to did not unambiguously support the representations they created, but despite the scarcity of information available, the infants demonstrated a rich representational

[32] Lidz, 2003, 6

system, one that is cognizant of both Noun Phrases and anaphoric forms at the very onset of syntactic development. [33]

Although it would be nice if a theory of language development were as simple as an appeal to universal grammar, the objections raised against this position are far too numerous and robust. Tomasello (2004), for example, claims that *another one* is an unanalyzed whole "conventionally associated with certain types of linguistic expressions." He sketches a scenario in which a child is playing with a red block, with no linguistic accompaniment, and is then handed another one, this time accompanied by someone saying, *Here's another one*. In this case, *another one* can only refer to an object similar to the original and thus, the meaning is not determined by syntax, but rather, can be understood entirely among the context of the situation.

Lidz and Waxman's (2004) response to Tomasello's objections are multifaceted. First, they argue that since *another one* is rarely used without a linguistic antecedent,[34] it makes no sense to assume the contexts alone are sufficient for anaphoric understanding. Second, they suggest that on an intuitive level, Tomasello is mistaken in arguing that *Here's another one* only refers to another red block. The expression could just as easily refer to another block of a different color, another toy of the same color, or even a different toy altogether, depending on the situation. For the third part of their reply, Lidz and Waxman cite empirical evidence from Waxman and Markow (1998), who find that when infants are presented with an object (such as a yellow car), alongside an expression like *See this. Look at this*, and then presented with two test objects (a yellow car and a green car), while being asked the question, *Can you find another one?,* the results are that

[33] Ibid, 8
[34] Only about 4% or 33 of the 792 anaphoric uses, Lidz and Waxman, 159.

the infants choose either the yellow or the green car at a rate indistinguishable from chance. According to Tomasello, they ought to have shown a preference for the most similar object based on the contextual information available surrounding *another one*, but this was clearly not the case. In the Lidz, et al study, however, "infants did hear a linguistic antecedent during familiarization and did pick the most similar object at test."[35] This evidence leads Lidz and Waxman to conclude that the infants have an innate appreciation of the internal structure of the noun phrase and their knowledge that *one* is anaphoric to the phrasal category is what led them to pick the most similar object.

Lidz, Gleitman, and Gleitman (2003, henceforth LGG) also performed an experiment in order to show that even in a language which does not use syntax as the most reliable cue to meaning such as the Dravidian, Kannada, for example, an appeal to universal grammar is nevertheless made, namely via Isomorphic Mapping, or what Chomsky earlier called the theta criterion.[36] The idea is that "noun phrase number lines up as simply as possible with argument number."[37] Two sets of experiments were conducted, one on three year olds, the other on adults. Both built on the methodology of Naigles et al (1992) and their work looking at child responses to old verbs in new syntactic environments. While Naigles did well to show that an effect known as "frame compliance," or the extension of already known verb meanings on the basis of syntax, is what guides young language learners, LGG added morphological information to the Kannada studies in order to see if language specific morphology overrides syntax in certain cases. Subjects were presented with familiar verbs in familiar and unfamiliar

[35] Lidz and Waxman, 160
[36] Chomsky, 1981.
[37] Lidz, Gleitman, and Gleitman, 154.

(ungrammatical) contexts and then asked to act out what they heard using toy animals. Since transitivity is not the best predictor of causativity in Kannada, LGG were particularly interested in cases where subjects were asked to act out an intransitive verb in an ungrammatical context. According to a non-nativist account, Kannada speakers should rely almost exclusively on the causative morpheme, rather than the number of NPs. In both experiments however, LGG found that subjects relied on the number of noun phrases much more than they did on the causative morphology, which thereby caused them to conclude that the universalist view of verb learning holds more explanatory power, even cross-linguistically.

In a response to LGG, Adele E. Goldberg (2004) argues that the Isomorphic Mapping Hypothesis is far from universal; one need only look at a languages like Ewe, in which verbs that would appear intransitively in English, always appear transitively with a NP object, or Lao, which allows for special constructions such that three semantic participants are expressed by a verb that only allows for two full NPs.[38] Also, she notes that while a causative morpheme implies a causative interpretation, the converse does not hold. "That is, while the causal morphology may well have perfect *cue validity* as a predictor of causal meaning, it is far from having perfect *category validity* (probability that the causative meaning involves the causative morpheme): many causal utterances do not contain the morpheme."[39] Goldberg then concludes that the mappings between syntax and semantics must be learned, perhaps via pragmatic principles; in any case, such an explanation is at least, if not more plausible than the Universal Grammar account given by LGG.

[38] Goldberg, 78. Cited here was Ameka, 1995, Essegby and Ameka (2005).
[39] Ibid, 82

Naturally, Lidz and Gleitman (2004, LG) responded to Goldberg's objections by arguing that infants' statistics-based discovery procedure must in certain respects be specific to language learning and that while pragmatics and theory of mind might very well aid in language acquisition, the problem of analyzing predicate-argument structures and their multifarious relations to linear strings of words, is solvable only by an appeal to some innate source of computational principles. In short, their reply can be summarized by the following: "While argument structure itself is not directly available to learners as a surface cue leading back to verb meaning, its surface correlate, noun phrase number, is so available and indeed is privileged."[40]

1.4 Sampson's Reply to POS

Geoffrey Sampson presents a series of replies to each of the arguments for UG cited in *1.2*. Although each response is worthy of discussion, I want to focus mainly on his attack of POS type evidence, as it is most often utilized in support of nativist accounts. The chief reason Sampson thinks POS arguments fail is that they are logically self refuting, but he spends some time challenging the details of the position, one such component being Child-Directed Speech (CDS), and its lack of 'grammaticalness,' as Chomsky likes to put it. He references a study by Newport et al (1977) which found that five per cent of the utterances between adults are ungrammatical, which is not significant. Given that, as Hornstein and Lightfoot (1981) and even Pinker suggest,[41] the child discriminates between speech directed to her and speech directed at others, then if it can

[40] Lidz and Gleitman, 89.
[41] As Pinker says: "when given a choice, babies prefer to listen to Motherese than to speech intended for adults." (1994), 279.

be shown that CDS contains even fewer errors, the problem would appear much less significant.

Although there is little relevant data about the grammaticalness of CDS,[42] so Sampson's case thus far rests only on the plausibility that it contains few errors, he goes on to argue that POS arguments can be easily refuted by simply examining their structure. He focuses on Chomsky's explanation of yes/no question formation in English, as it is one of the most highly cited, in both Chomsky's own writing, as well as by others. When forming yes/no questions, Chomsky argues that there are two hypotheses under which a child might operate. (1) Use the first finite verb in the sentence or (2) Use the finite verb in the main clause. A sentence such as *"What I'm doing is in the shareholders' best interest,"* would surely fail under hypothesis (1), as it would read **Am what I doing is in the shareholders' best interest?* But, as Chomsky claims, there is nothing in the data that would enable a child to choose between hypotheses (1) and (2), especially when the sentences contain more than one clause. This is not because CDS is devoid of such evidence, but because language usage in general does not contain the evidence relevant to determining such movements.[43] However, as Sampson suggests, if we were to obtain a corpus of CDS, it seems highly unlikely that it would not be filled with 'evidence' of this sort, questions such as *Would anyone who wants to go see me later,* or *Could a tyrannosaurus that was sick beat a triceratops in a fight?*[44]

Regardless of what the 'real truth' behind CDS is, Sampson still maintains that the nativist position is untenable, as it is based on circular reasoning. Chomsky and other

[42] Save what Tomasello (2004) has compiled, which will be discussed further in Chapter two.
[43] Sampson, 1999, 42.
[44] Ibid, 42

nativists are asserting that language has certain properties, but these properties are not available to or evident in the speakers' corpus, neither spoken nor heard. However, somehow, magically almost, adult speakers of any language do not make mistakes such as the one above very often, if ever. Thus, there must be an innate knowledge of these 'unobservable' rules. But, even the adults' conscious knowledge of the properties of language are based only on finite observations. We might ask then, what about the linguist, who can elicit speakers' beliefs about language and might have more examples and observations than the typical speaker? Ramsey and Stich (1990) as well as Hornstein and Lightfoot (1981) claim that they must have access to some of these unobservable properties. However, Sampson argues,

> they must concede that speakers' beliefs about their language derive purely from experience of examples of the language: in which case, interrogating the speaker does not give the linguist access to any category of data not available to the child.

He goes on to say that if Chomsky were correct about acquisition proceeding via innate knowledge of properties, then a linguist really would be gaining insight into this tacit understanding whenever he or she probed the subconscious of a speakers' metacognitive capacities, but then again:

> An argument for innate knowledge of language which works only if one grants the assumption of innate knowledge of language is less than impressive. [1999, 44]

According to Sampson, nativists assume the very thing they are trying to prove in arguing for innately specified knowledge about language. They then use this assumption as the basis for evidence such as poverty of data found among corpuses of speakers and claim that experience is insufficient for language learning to take place. While Sampson does not mention the experiments conducted by Lidz et al mentioned earlier, it seems clear

that he would have a similar complaint to lodge regarding them. Naturally, every experiment will proceed with the hope of either proving or disproving a hypothesis, but if the hypothesis is the foundation for the design of the experiment itself (ie-assuming that Chomsky is correct and only supplying the infants with 'impoverished data'), then we might wonder if the results obtained are skewed. Under Sampson's understanding, they certainly are, as the very thing we are trying to prove, innate knowledge, is asserted almost as undisputed fact before the experiments begin.

It could be however, that Sampson has it backwards. His claim runs something like 1) nativists presuppose nativism before examining corpuses, 2) They examine the corpuses, look specifically and only for cases of impoverished input, 3) find some, and 4) conclude, fallaciously, that their supposition is correct. First, it should be noted that the accusation of question-begging methodology is applicable only if the nativists are indeed assuming the desired thesis (in this case innate linguistic knowledge) *and also* ignoring or omitting data that would disprove or countermand said thesis. So it is not the fact that there is an assumption being made to which Sampson is taking issue; if it were, then he would be objecting to all scientific investigation.[45] Thus, if his problem is with the manner by which nativists collect data, this is a different story and deserves to be investigated.

While clearly, the data found in these corpuses are too rich and numerous to all be included among discussions of available speech, it does not follow from this fact that what gets left out are samples of speech that would indicate something contrary to POS.

[45] As it might be argued that all science involves hypothesis testing, it can just as easily be argued that science is misguided (See Gould, for example). If this is what Sampson really has in mind however, then his objections would take the discussion too far astray from the present investigation.

Furthermore, POS was not a theory that sprang into existence in the mind of some researcher who then examined speech to try to prove it. Indeed, as Chomsky noticed for example, the language that the child is exposed to does not always follow grammatical rules and is often unrepresentative of the language itself, but it was not until after extensively surveying the speech that POS was considered to be a viable and testable theory about input. In this sense, Sampson's critique seems to go overboard in assuming that the methodology utilized by the nativists is truly circular. Without access to the corpuses, I cannot make the claim that nativists *are not* guilty of pruning the data to fit their purposes, but based on the overwhelming agreement among linguists and anthropologists concerning the impoverishment of speech to which infants are exposed, it is at least plausible to conclude that Sampson's attack is misguided.

However, if the nativist position indeed is in real danger based on Sampson's critique, it will be interesting to see exactly what he has in mind as an alternative way to explain some of the seemingly impossible feats achieved by young language learners. In the rest of both of his books, he explains the process of acquiring language as "Popperian," insofar as it is guided by hypothesis testing, trial and error, and error correction. I will investigate this style of acquisition via alternative theories of language acquisition, but first, a few final remarks should be made about UG and the nativist position in general.

1.5 Conclusion: UG's Commitments and its Construal of "Innate"

Based on the discussion so far, the original characterization of the "Neo-Nativists" can be amended slightly. It seems that Chomsky, Pinker, and others want to

make a distinction between the *state* and the *process* of language acquisition, so as to only concern themselves with the former. Indeed, all they say about the actual process of acquiring a language is that it sets an already wired system into motion. What they are really interested in explaining, therefore, is that pre-wired system. According to Chomsky, the initial state of the human mind has blueprinted onto it, all the principles of grammar, but this does not mean that an infant is born with representation of JAPANESE in its mind. Instead, since all human languages have universal principles, the types of representations that are innately specified are abstract, general, and applicable to every possible language. The parameters, which differentiate among languages, are not set until interaction with the linguistic environment takes place. Thus, it is clear that that the neo-nativist position differs slightly in degree from the traditional rationalist position, insofar as the environment is treated as more of a necessary part of this acquisition. Of course, as Plato claims, we first experience discrete objects and eventually can abstract away to the Form in which they all participate, so experience is an essential component of learning, but we find less of this idea that learning really is just recalling what was known during death and more focus on what the mind-brain of an infant is like at the onset of development.

Nevertheless, the anti-nativist proponents seek to distance themselves from such a position which seeks to claim that linguistic representations are innate. But, according to the 'milder' version of nativism that has been sketched thus far, it would appear that the neo-nativists' claims are not too distinct from arguments put forth by at least some non-nativists insofar as both allow for innate capacities or predispositions. Therefore, one central aim of the following chapter will be to investigate what the disparity is, if nativists

are not as extreme in their views as one might have guessed and hence, their position might be commensurate with some of the non-nativist tenets.

One final point I will consider will be whether or not "innate" is used by opponents of nativism in the same way as the nativists use it. The introduction briefly touched on this and throughout this first chapter it has been apparent that "innate," as it is conceived of by the neo-nativists, means a couple things:

> 1. something present at birth, namely, a specific region of the brain or architecture of it that is innately specified to handle the task of language learning.
> 2. a capacity or predisposition. In the case of language specifically, it is typically argued to be a uniquely human capacity and one that is universal to the species

In the next chapter therefore, I will look at alternative theories of language acquisition, each unique in their specific arguments, but similar in their denial of "innate" linguistic representations or principles, in order to determine if the opposition is well-founded, or if it stems more from different conceptions of "innate" itself.

Chapter 2: Alternative Theories of Language Acquisition

2.1 Introduction

As the debate between Goldberg and Lidz et al discussed in the first chapter suggests, there have been numerous negative reactions to the nativist thesis, and more particularly, to Universal Grammar. These objections vary in type from explicit linguistic critique of Generative Grammar[46] to an attack on UG against its inadequacy, insofar as it fails to account for the plethora of factors that determine how and at what rate a child learns language. For example, Usage Based Linguists like Tomasello (1992) have taken issue with some of the standard arguments for UG and linguistic nativism, such as poverty of the stimulus. Instead, they argue, infants have heard and processed millions of utterances containing thousands of types of constructions and an innumerable number of tokens of these types. It is this repeated exposure, combined with a general learning process extended over a long period of time, which allows the language learner to "construct" increasingly complex forms. Other psycholinguistic theories emphasize the need to more strongly consider the child's social situation,[47] their attentional focus and spatial orientation,[48] or their ability to understand the intentions of others.[49] Thus, in this chapter, I will examine several of these different accounts of language acquisition, with a particular focus on what each has to say about the way in which the child arrives onto the scene of language acquisition. Each theory carries with it a critique of Nativism, UG, or

[46] cf: Ellis (1998)
[47] Cf. Bloom (2000) and for an interesting discussion on childhood bilingual acquisition and social factors, see McCardle and Hoff (eds), (2006). *Childhood Bilingualism*. Surrey, UK: Multilingual Matters.
[48] Cf. Spelke et al. (2001).
[49] Bloom (2000)

both, whether it is an outright denial of any "innate" principles, or an implicit acceptance of *some* "innate" knowledge, but a refusal of UG as a comprehensive account of the *process* of language learning. In turn, I hope to discover 1.) what each has to say in terms of whether language is "acquired" or "learned," 2.) to what extent can knowledge of language be said to be "innate", and 3.) to what degree does the theory in question preclude UG? In regards to this last consideration, I am seeking to determine if any of the abovementioned theories actually rule out UG entirely, or merely claim its limitations for a comprehensive acquisition theory.

2.2 Acquisition as Construction[50]

Perhaps the most pointed attack on UG has been that it accounts only for the "core grammar" available to the child, and fails to provide an adequate explanation of all the other linguistic items, such as irregular or anomalous constructions, idiomatic expressions, and metaphoric extensions. Usage-based approaches, Tomasello argues, in particular, the construction grammar attributed mainly to Fillmore (1989), Goldberg (1997), and Croft (2001), do focus on all kinds of usage patterns.[51] This is because Usage-Based approaches see the child's path to fluency as "gradual, piecemeal, and lexically dependant," as opposed to UG, which views it as structure dependant, "with the acquisition of particular linguistic structures depending heavily on the specific language

[50] In the following section I will group several apparently 'different' theories into a single category. This is because in my research, I have come across no appreciable distinctions among Usage Based Linguistics (UBL), Functional Linguistics, Constructivists, and Emergentistists (This last term is often confused with *Emergentists,* which I think *is* a different theory altogether; therefore, I will treat this entire section as dealing with UBL, and the following with Emergentism)

[51] Tomasello (2003), 98.

to which a particular child is exposed, and with generalizations coming only after a fair amount of concrete linguistic material has been learned."[52]

Central to Usage-Based Linguistics (UBL) is the claim that constructions are not rules. Rather than coming equipped with abstract, semantically empty, syntactic structures, UBL views the infant learner as slowly developing patterns of use, which then can become schematized in their minds as this or that type of construction. While the patterns of usage might become very *rule-like*, in the sense that more predominantly used constructions can be more easily generalized, never do they become entirely abstract "rules," devoid of any communicative function. As Tomasello puts it:

> In usage-based approaches, contentless rules, principles parameters, constraints, features, and so forth are the formal devices of professional linguists; they simply do not exist in the minds of the speakers of a natural language. [2003, 100]

We can also think of UBL as saying something fundamentally different from Generative Linguistics, in that the two view the *process* of language learning oppositely. Although Chomsky himself spent very little time discussing process, the basic assumption of Generative Linguistics, to which UG is intimately tied, sees language acquisition as a *deductive* process, where the child begins with very abstract and general principles or rules, and ends up with concrete ways of applying them. Conversely, UBL argues that the child is only ever exposed to actual instantiations of language and then generalizes over a large amount of data in order to recognize patterns, which in turn allows for "rule-like" representations. The representations of generalizations might be better thought of as exemplars of construction types, rather than resembling rules in any strict sense. In this light therefore, UBL views language acquisition as *inductive.*

[52] Ibid, 98.

Likewise, we could say that while UG argues that the child moves from innately given *types* to an unlimited number of *tokens* available to her in the natural language, in UBL, this process is reversed. The child begins with "tens of thousands of types *presented as* tokens,"[53] and only after repeated exposure, practice, and pattern recognition can she then represent the types in her mind.

The third and final characterization of UBL involves the method by which it studies child language acquisition. Because UBL does not view the directionality of language development as *deductive*, but rather supports the notion that the child must generalize over millions of examples, the process of acquisition is therefore slow and oftentimes, clumsy. Studying language in an ontogenetic manner, however, implies detailed, elaborate, and usually longitudinal experiments. Thus, research is typically guided and benefited by the Child Language Data Exchange System (CHILDES),[54] as it provides a rich database of transcripts, past studies, computer analysis, and tools for linguistic coding. CHILDES also allows for networking among researchers, as they have access to results and findings obtained over long periods of time, such as those of Tomasello's case study of his own daughter, which is a diary filled with one year's worth of detailed analysis. Although I will discuss this and several other similar findings in turn, the underlying outcome a researcher would obtain by looking at Tomasello's data is, according to Tomasello of course, that it demonstrates the "emergence" of new forms on the basis of constructing from simpler ones.[55] In short, the methodology behind testing the validity of UBL is quite different than that of UG. Considering that the former takes

[53] Ellis, 641, my emphasis.
[54] See MacWhinney and Snow, 1990 and MacWhinney, 1995.
[55] Tomasello, 1992.

language acquisition to be "piecemeal" and rather lengthy, it should stand to reason that the way the learner is viewed in experimentation geared towards UBL will be nothing like the manner in which experiments are conducted to prove, that the learner is simply triggering already possessed knowledge of linguistic principles.

2.2.1 UBL: Acquisition or Learning?

Until this point, the terms "acquisition" and "learning" have been used interchangeably in regards to UBL. I have purposely blurred the distinction between the two in order to reflect what goes on in the UBL literature. A few examples:

> "Cognitive Grammar accepts that becoming a fluent speaker involves a prodigious amount of actual learning and tries to minimize the postulation of innate structures specific to language" [Langacker, 2002].[56]
>
> "Virtually all of the formal syntactic properties of grammar are language-specific, and therefore must be learned INDUCTIVELY" [Croft, 2001].
>
> "The child acquires constructions by categorizing utterances into their types, in terms of the grammatical properties of the utterances that the child is able to perceive" [Croft, 2001].[57]
>
> "To understand how children acquire a language we must know something about the language they hear-both in terms of specific utterances and in terms of the constructions they instantiate" [Tomasello, 2003].[58]

Clearly, it is not uncommon to find both terms used to describe one process, and perhaps this is because the terms "acquisition" and "learning" mean the same thing for certain proponents of UBL. For Chomsky however, they are entirely different, and while I am not advocating accepting what he says about the issue as gospel, I do think he has a valid

---

[56] In Barlow, et al. (2002), p3.
[57] Both on page 58, although Croft's "radical" constructive grammar differs from the definition of UBL which I am using here. I will discuss this, in particular, the language-specificity part, in a later section.
[58] 108

concern. Especially when considering the first quote above, from Langacker, who says UBL *minimizes* the number of innate language-specific structures, the natural follow-up question is, *So there are 'minimal' innate structures, right?* If we can posit *any* innate structures whatsoever, then I think we can still speak of acquisition as being at least part of the story of development. Indeed, Chomsky thought that it was the innate principles which provided the conditions for the possibility of tokening infinitely many utterances. Nonetheless, as Ellis claims, regarding UBL, "In these views, language is learned and syntax is an emergent phenomenon, not a condition of development,"[59] we can see the tendency towards a highly anti-acquisitional model.

Since it may not be altogether clear at this point, why it is crucial to pin down what type of process, whether acquisition or learning, UBL maintains language development is, I will briefly explain what I view the significance to be. As I have discussed above, there seems to be an intimate link between innate structures and the process of language development being a deductive one. If Langacker and others want to allow for *any* innate structures, then based on this interrelation, at least *some* of these "rules" must be represented in the infants mind. Of course, this goes directly against Tomasello's version of UBL, as he claims that *no* rules are ever "in the minds of speakers." According to the way I have sketched UBL thus far, I think it is appropriate to argue that it views language development as chiefly a "learning" process, insofar as the "conditions of development" are not language specific, but rely on more general abductive inferences. As we shall see in the following section however, a large number of Usage-Based Linguists, Tomasello included, concede that there must be some innate

[59] Ellis, 642.

principles governing this learning process. For the time being therefore, I will call the process by which UBL characterizes language development as "Learning Lite." Agreeing with Ellis for the most part, who claims that UBL argues that "language is learned," I want to make room for "Radical Constructivists," like Croft, who allow for language-specific structures. If these structures can also be shown to be innate, this version of UBL would simultaneously allow for minimal acquisition.

2.2.2 'Learning Lite' and the Acquisition of Constructions

An important factor governing early childhood language production, according to UBL, is the language to which children are exposed. However, there exists a scarcity of research that seeks to inventory the rich and complex range of expressions making up the child's daily input. Indeed, as Tomasello points out, most of the studies involving Child-Directed-Speech (CDS) focus only on specific aspects of the input, such as tone of voice, with Motherese being a popular topic.[60] There has been recent documentation by Cameron-Faulkner, Lieven, and Tomasello (2003), which examined all of the CDS of twelve mothers and their 2- to 3-year-old children, all of whom were native English speakers. These were catalogued and organized based on constructional categories, but in short, Cameron-Faulkner et al found that:

> Children heard an estimated 5000 to 7000 utterances per day, between one-quarter and one-third of these were questions, more than 20 % of these were not full adult sentences, but rather were some kind of fragment, about one-quarter of these were imperatives and utterances structured by the copula, and only about 15% of these had the canonical English SVO form, supposedly characteristic of the English language, and over 80% of the SVO's had a pronoun subject.[61]

[60] cf. the introduction in Snow and Ferguson (1977), and Galloway and Richards (1994).
[61] From Tomasello, (in press).

Before discussing how this inventory is used to bolster UBL, I want to point out that problem of paucity in cataloguing children's linguistic input has not been entirely satisfied. First of all, CDS is only one type of input; children are exposed to countless other expressions, from passive involvement in adult-to-adult conversations, radio and television, and older sibling or peer speech. None of these can be argued to be child-directed and as the few studies regarding these interactions suggest, a catalogue of them would not identically match that obtained by Cameron-Faulkner, et al. To be sure, this appears to be only a minor point of contention, especially if we assume that CDS is most responsible for directing children's attention to linguistic data, and hence, their ability to produce language. Nevertheless, the non-CDS input still exists; in fact, for some adult language users of other cultures, CDS is an almost non-existent phenomenon.[62] Also, we cannot ignore the stubborn nativist cry, reminding us that *regardless* of input, be it child-directed, incomplete, or adult-directed only, children all over the world learn to speak in their native tongue.

2.3 Language Emerges

Oftentimes, UBL is referred to as "Constructivism," or is listed under the heading of "Constructivist Approaches to Childhood Language Acquisition."[63] In the previous section, I used the blanket name of UBL to include all theories of similar type, regardless of nomenclature. One theory that also sometimes gets folded in with UBL-type theories is *Emergentism*. This, in my opinion, is a mistake.

[62] Infants and toddlers still in the 'babbling' stage in Papua New Guinea, for example, are not addressed at all. It is not until they are able to utter multi-word phrases, that they are entered into the speech community.
[63] c.f. Ellis 1998, Tomasello (1992, 2003)

The majority of emergentists that I have come across all have one thing in common; they are intimately tied to and informed by neurobiology. As a field which is relatively young and constantly changing as advances in science and technology are made, any theory that is beholden to neurobiology must be prepared for radical shifts or upheavals of previous findings. One such shock came in the form of the completion of the human genome sequencing,[64] which showed that there are far fewer genes than was previously thought. Indeed, the revised edition of Sampson's *Language Instinct Debate* duly notes this discovery and utilizes it in a reformulated attack on linguistic nativism. Similarly, Elman, et al (1996) and MacWhinney (1998) have argued that neurobiological research today does not suggest the possibility of any inherited structures which would serve as principles and parameters of UG. Instead, they say, research indicates that interactions at the cellular level, the social level, the environmental level, and every level in between give rise to new forms. As Ellis (1998) notes, "information theory analyses suggest that humans are more than 20 orders of magnitude short of being mosaic organisms, where development is completely prespecified in the genes."[65] Thus, in following along with the current neurobiological paradigm, Emergentists view language development more as regulated by experience and less dependant on innately endowed structures.

Before discussing Emergentism's commitments any further, it is worth pointing out that the characterization just mentioned highlights yet another, even finer analysis of the term "innate." Something can be genetic or it can be inherited. These two types of innateness are not mutually exclusive; indeed, it might be determined genetically of me

[64] 2003, see http://www.ornl.gov/sci/techresources/Human_Genome/faq/seqfacts.shtml#sites
[65] 643

that I have blonde hair, but also inherited from my parents, but they are not coextensive either. For example, according to most proponents of developmental systems theory,[66] some inheritance is over and above genetic development, hence it is "epigenetic." Epigenetic inheritance or epigenetic inheritance systems (EIS)[67] are typically used as a demonstration of the contrast between the 'Constructivist Interactionist'[68] understanding of cellular development and a programmatic or preformativist account. On the former view, "the life cycle of an organism is developmentally constructed, not programmed or preformed. It comes into being through interactions between the organism and its surroundings as well as interactions within the organism."[69] Thus, with regards to the possibility of an innately specified language capacity or facility, one can ask if there are genetic predispositions for such an ability, which, based on the Genome project and the various interpretations of it, seems pretty much out of the question. Another question would then be whether or not language learning can be expressed in terms of an EIS. Based on my interpretation of the Emergentist account of language development thus far, this appears to be the case, and as I describe below, I think this particular component is part of what differentiates them from UBL.

Another important component of Emergentism is the correlative relationship between the non-modularity of any "language faculty" and the domain-generality with which language is learned. Neurobiological research has also demonstrated the malleability of the cortex, which, as Elman points out, suggests that specialization of the

[66] Cf. Callebaut & Stotz (2002) and Oyama (2000)
[67] Jablonka (2001)
[68] This is another name Oyama, et al (2001) give to developmental systems theory, but as I will explain in the following section, I think "Constructivism" and "Interactionism" have subtle but important differences with regards to language acquisition and/or learning.
[69] Oyama, et al (2001), 4.

cortical areas is more likely to be the result of learning than the cause of it.[70] Emergentists do not therefore, deny that modularity might in fact characterize the adult brain; they simply insist that this modularization is an "emergent" rather than antecedent property. Furthermore, the protracted debate over a unique "language faculty" in the brain is slowly being won over by researchers like Posner and Raichle (1994) who have shown that the infamous Wernicke's and Broca's areas are not the only language locales. Neural imaging has shown that high frontal and parietal areas, parts of the cerebellum, and some of the subcortices are all valid "language areas."[71] And finally, these areas are not solely designed to handle linguistic information; that is, they are responsible for many other types of processing, from spatial reasoning,[72] to distinguishing between high and low frequency words.[73] In short, these neurobiological findings lead Emergentists to deny the existence of a specific language module at the onset of development. Because of the nonmodular view of the mind they hold, Emergentists are also reluctant to allow for any innately prespecified linguistic principles. Nevertheless, reluctance does not entail outright refusal and as we shall see, Emergentists have quite a different take on what sorts of things could be deemed innate. But first, I want to digress a moment, as I think it is important to differentiate between Emergentism and UBL, two typically conflated theories.

---

[70] While this is a conclusion they draw over and over again in various forms, this particular formulation come early in the book, in the first chapter, as a hypothesis to be proven. It is subsequently followed by neurobiological evidence and connectionists models.

[71] cf. Ackerman, et al. (2000) who offers an account of how the cerebellum is actually *essential* to speech processing, by showing that patients who suffer from cerebellum damage have motor speech disabilities.

[72] While spatial reasoning is usually argued to be processed in the frontal areas, those like Knauff, et al (2002) have shown with fMRI that deductive reasoning not guided by visual input, but that still follows a visual pattern (like in a logical proof) is also processed in the visual association cortex (Broadmann's area 19)

[73] Bates, 2000.

2.3.1 Are Emergentists not just Usage-Based Linguists + Neurobiology?

The reader at this point might reasonably assume that Emergentism is nothing more than a glorified version of UBL. After all, it adds 'super-beefy' neuroscientific data to strengthen the claim that linguistic rules are not innately possessed in the minds of infants. With this point of similarity among the two theories, I concur. It is also the case that both theories view the development of language as an inductive process. However, there are some slight yet significant differences between the two. I will only summarize these distinctions for now, as I hope to explain each in depth, as the following sections unfold.

1.) *Cognition versus Interaction:* While UBL focuses on what the learner *represents* to herself when constructing new forms, Emergentism focuses on the dynamism between learner and world. In other words, UBL sees the child as a somewhat passive receptor of the millions of language instantiations and then seeks to understand how, with only very primitive forms available, the child is able to construct more complex ones. Instead, Emergentism views the child as actively involved in social, environmental, physical, and biological interactions, all of which are integral components of development.

2.) *Product versus Process:* As the first distinction illustrates, UBL seems to be more concerned with 'what gets learned' rather than *how* it gets learned. This is not to say that the theory of "Constructing a Language" is not about the process involved in learning. UBL is however, more invested in determining how the

child arrives at a final state, or to put it another way, how the child comes to possess and be able to utilize complex constructions. Emergentists do not view language as a 'thing' to be learned; it is rather an emergent phenomenon, resulting from interactions at various levels. One analogy drawn by Ellis, suggests that "complexity of a solution emerges from the interaction of problem and solver. Apparent complexity may come more from the problem than from the system that learns to solve it."[74] Similarly, Bickhard and Terveen (1995) focus on the way in which Interactivism (Emergentism) provides a functional model of representation. "That is, it presents a functional explication of representat***ion*** (or representing), rather than a characterization of representat***ions***."[75]

3.) *Linguistic Universals:* While it is difficult to pin down a unified theory of linguistic universality in UBL, there is at least a substantial amount of disdain for the "convergence of grammars," prevalent in the UBL literature.[76] Indeed, the linguistic anomalies found in languages like Lao and Ewe are often cited as evidence that there can be no universal principles of language, especially none that are innately specified in the minds of infants. It is also questionable whether or not Emergentists buy into the notion of linguistic universals; however, if such universality is found to exist, Emergentism is not in danger, as its basic tenets are not based on the denial of such universals. Instead, Emergentists explain away universals as they do any other phenomenon; linguistic universals have emerged,

[74] Ellis, 642.
[75] 57, their emphasis. See also Van Gulick (1982), who claims that only insofar as a representation functions appropriately for an epistemic system can we say that it is a representation at all.
[76] c.f. Sampson (2005), Goldberg (1995), etc

just as the universality of human transport has emerged.[77] This universality comes not from a prespecified code, but rather, from the interactions of users, their needs, the environment, and the constraints placed thereupon.

To summarize, Emergentism is often found in the literature as being synonymous with Interactivism, and based on the discussion above, it ought to be perfectly clear why. UBL, on the other hand, focuses less on the learner's dynamic relationship with the incoming data, and more on the complexity of the data itself and the means by which the child is able to 'construct' her own language, thereby emulating her mature-speaking counterparts. While UBL and Emergentism might share many similarities in terms of describing the initial state, the process of learning is viewed in slightly different ways, as I have attempted to show. These difference have all too often been disregarded or ignored, thereby allowing the two theories to be subsumed under the general heading of Interactivism, Constructivism, or some similar alias.

2.3.2 Analogies: Emergence in Nature and Beyond

A great source from which Emergentists draw their conclusions about language learning is the natural world. It is chock full of examples analogous to the language-as-emergent model they advocate. For instance, the geometry of snowflakes reflects a highly complex organization of water molecules, but this complexity was not the condition for development of the snowflake. Rather, the intricate patterns and shapes emerge from what were once unorganized water particles. Other examples include the hexagonal shapes of cells in a honeycomb (MacWhinney, 1998, Elman et al, 1996), the formation of

[77] Simon (1983)

the jaws of the soldier ant (Campbell, 1990), and fingerprints (Ellis, 1998). For all of these biological phenomena, it cannot be said that they *explain* anything regarding how development takes place. The hexagons comprising the honeycomb, the piercing design of the teeth of soldier ants, and the concentricity of skin ridges on the fingertips are all emergent properties. They do not explain development, but are the result of it.

Biological analogies are not the only ones out there. Gould and Lewontin (1979) have noted the prevalence of such emergent phenomena in human artifacts. Take for example, the "spandrels of St. Mark's Cathedral" in Venice. Spandrels are formed by the intersection of two rounded arches, and as they add such beauty when placed on the façade of cathedrals or when worked into triptychs of stained glass, it is often thought by onlookers that spandrels were part of the architectural design, a response to some need to create triangular shapes as special effects. However, as Elman, et al point out:

> In fact, this explanation is the exact reverse of the correct one. The spandrels are simply the emergent product of mounting a dome on rounded arches: a triangular shape necessarily emerges. What must have happened is that the artist recognized post factum the design potential of the resulting spandrels. [1996, 112]

One other case of human-created analogies to emergentism is given by Ellis (1998), who claims that the "principles" and "rules" by which we operate can also be shown to be emergent. Consider the systematicity with which cars line up at a red traffic light on a multilane highway. As time passes and the light remains red, cars pile up, but not just randomly and unpredictably. The cars always seem to fall into a lane in such a way that the lines are more or less of equal length, but this is not something prefashioned by the Highway gods. "Instead, the 'rule' that equalizes the number of cars in the carriageways

emerges from satisfying the constraints of the more basic goals and behaviors of drivers, traffic planners, cars, and time."[78]

Although these analogies are not the sole driving force behind Emergentists' arguments, they do provide powerful illustrations of similar processes found all over the world, and therefore, further strengthen the bond between Emergentism and the sciences. A keystone of Emergentist support comes from connectionist simulations, which will be discussed at length in the third chapter, but there are a couple other important tenets of Emergentism worth noting. One has to do with reformulating "innateness," and the other involves imposing these constraints, as mentioned by the traffic queue above, on learning systems.

2.3.3 Learning and Emerging: Is Nativism Still Lurking?

Since Emergentists are so interested in the transactional feature of language learning, there would seem to be little to no room for any discussion of "innate" knowledge which exists prior to such interactions. It would also appear, at least according to my characterization of learning and acquisition, that Emergentists view language development as unequivocally inductive, thus as a learning process. Of course, so do Usage-Based Linguists, so it is important to consider again, some of the differences between the two theories, in order to determine whether or not any innate knowledge is posited, and if so, what kind of knowledge it might be.

As was mentioned earlier, Emergentists closely follow neurobiological research so as to be able to talk about language development in a way that mirrors human development in general. In their book, *Rethinking Innateness*, Elman et al discuss at

[78] Ellis, 643.

length, how cellular development occurs, and citing Edelman's (1988) work on embryology, they divide development into two camps; *mosaic* and *regulatory.* In mosaic development, the cells' development is highly determined by the genetic code and is thus, affected very little by environment or intercellular interactions.[79] Regulatory development, conversely, is highly dependant on cellular-level interactions. "The orchestration of cell differentiation and the final outcome are under broad genetic control, but the precise pathway to adulthood reflects numerous interactions at the cellular level that occur *during development.*"[80] An example of a mosaic system is found in the nematode, *C. Elegans*, while more complex organisms are usually shown to be under regulatory control. The point of delineating all of this is to set the stage for a more specific argument that there is no "language gene." As Ellis (1998) pointed out, the human species is twenty degrees of magnitude away from being a mosaic organism, so too do Elman et al (1996), capitalize on this fact in support of an interactivist approach. To further draw a parallel, we can add that as the human genome project has demonstrated, there are far fewer genes than we had previously thought. Elman and his colleagues certainly exploit this datum as well, by citing numerous examples of how genes can be responsible for many forms of development, are not rigidly fixed or determined, and can even change their main function over time and through development.[81]

All of this discussion of the human genome as a regulatory system versus mosaic systems like that found in *C. Elegans* leads emergentists to conjecture that it would be

[79] Elman, et al (1996), 13
[80] Ibid, 15
[81] Elman et al, 10-12 and 16

irrational to assume that any linguistic principles are prespecified in particular genes of the human organism. It is not however, necessarily the case that Emergentism rules out nativism altogether. Indeed, at the end of their book, Elman et al state, "There can be no question about the major role played by our biological inheritance in determining our physical form and our behaviors."[82] Thus, the disparity between linguistic nativists and Emergentists depends not upon the total acceptance or denial of genetic predisposition; rather, the dispute concerns what precisely, those dispositions are supposed *to be like* on the one hand, and what they are supposed *to do*, on the other. Emergentists tend to view the nativist position as misguided when it takes the fact that there are certain behaviors which seem inevitable, language use being one of them, and equates this with a "language instinct," suggesting that all those outcomes are contained "innately" in the genes of the neonate. As Elman et al point out, the problem with this conception is two-fold, having both to do with the *mechanism* by which behaviors come to be, and the *content* of what is presumed to be innate.[83] Just because we call a behavior innate does not entail an understanding of the mechanisms responsible for making that behavior inevitable. Furthermore, positing universality as demonstrative of an innate mechanism does not have sufficient evidence behind it. The fact that systematic and universal principles of language exist is not in itself proof that those principles were simultaneously the mechanism responsible for making themselves inevitable. The absurdity of accepting such an argument has given rise to the common Emergentist critique of UG; that it has raised the systematicities and universality of language from explanandum to explanans.[84]

[82] Ibid, 357.
[83] Elman, 21.
[84] Ellis, 633

Elman et al therefore posit a broader definition of "innate," one adapted from Johnson and Morton (1991), which removes the language-specificity aspect of the term given to it by Universal Grammarians. Basically, "given the normal environment for a species, outcomes which are more or less invariant between individuals are innate."[85] This alone does not suffice however, as it leaves one wondering how this characterization of innate differs at all from UG. Determining what counts as a "normal environment" and an "invariant outcome" are surely problematic. An invariant outcome of human development surely seems to be the comprehension and production of language, and according to this definition, such abilities must be innate. Likewise, the problems surrounding UG also apply to Johnson and Morton's definition. There are many outcomes of human development that could be argued to be invariant; for example, we all pretty much believe that snow is cold, that sticking your finger in an electrical socket will probably shock you, and that eating too much ice cream will induce stomach woes. Claiming that these invariant behaviors are somehow innate seems to be too much of a leap. Thus, Elman et al clarify the above characterization slightly. They argue that development is the result of interactions within and across many levels internal to the organism and that for an outcome to be innate it must be constrained at one or more of these levels. Constraints are specified to operate on one, two, or all three of these levels: *Representations, Architectures,* and *Timing*. While a large part of the third chapter will be devoted to explaining the way connectionist models attempt to simulate these constraints for emergent interactions, I will offer a brief explanation of each, as conceived by Elman, et al.

[85] Elman, 22

Representational: if a network had prespecified weights, then it is constrained representationally[86]

Architectural: when individual nodes or neurons are prespecified, the network is architecturally constrained

Chronotopic: if timing of inputs is manipulated in developing a network, this would be an exogenous chronotopic constraint, while endogenous timing constraints would occur internal to the network (or the brain), usually as a result of the learning itself, or because of genetic control.

2.4 Social Factors in First Language Acquisition

Although Emergentism seeks to account for interactions at a multitude of levels, all of which are said to be integral to the language learning process, one interaction rarely discussed by the theory is that among learner and other language users. This is not to say that interpersonal dynamics are entirely omitted from Emergentist models; indeed, as Elman, et al and Ellis suggest, the social aspect of language learning cannot be ignored and that models that attempt to integrate this crucial element are important.[87] Despite this recognition, sociolinguists and social constructionists demand more focus on the language-as-social (LAS) component. The case could be made that UBL, in its rich interpretation of the child as generalizing over millions of instances to construct systematicities is doing so in a social context. To be sure, all these tokens of language do not arise from thin air, but are from other language users. In this light, it is reasonable to characterize the child's learning as parasitic on hearing others speak and that all the constructions made by the child are built upon the foundation of other speakers. Nevertheless, UBL does not emphasize the social *dynamic* of language so much as it does

[86] Which of course, is a strange distinction to make, as all networks have at least *some* prespecified weights.

[87] Elman et al, 394 and Ellis, 622.

the social *output* of the speakers, and even that research is lacking, as the limited corpus of CDS data reflects. Instead, UBL really centers on how the *output* provided by speakers becomes comprehensible *input* for the child and then gets internalized, generalized over, and "acquired." Social theories of language learning take all these considerations to heart and suggest we reevaluate the situation in which the child truly is when learning a language. I have summarized their concerns into three basic categories of inquiry, which are as follows: 1) How and to what extent is CDS used by the child, and what other types of speech interactions are useful to the child?, 2) How are errors in children's speech responded to by adults or older children? and 3) what other factors, such as nonverbal cues, play a shaping role in learning?

2.4.1 What do Children Actually Hear?

Besides involving a change in pitch and intonation, CDS researchers have suggested that interactional modification often takes place between caretaker and infant, which results in positive effects regarding production. Pine (1994) for example, criticizes Gleitman et al's (1984) conclusion that CDS must be understood in terms of the language learner's capacity to extract and exploit linguistic data from the input.[88] While this much is true, according to Pine, it is not very informative. Undoubtedly, children make use of input; no UG theorist, UBL advocate, nor Emergentist would contest such a claim. Therefore, "the crucial question is not how facilitative CDS might turn out to be, but rather how it is used by the child at particular points in the acquisition process."[89] This concentration on the microgenesis of language is central to sociocultural acquisition

[88] Gleitman et al, (1984), 76. Cited in Pine (1994), 25.
[89] Pine, 25.

theorists. Instead of a sweeping claim about CDS and its effect on language learning in general, the focus is on studying language development at various snapshots and then asking to what degree CDS facilitates control of production.

Pine cites Newport et al's (1977) original study, in which they sought to explain relationships between CDS and children's analysis of input. In summary, it was quite clear that frequency in general (ie-how often auxiliary verbs are used) played little role in determining production control, while frequency of specific factors, such as saliency, reduced processing load, and referent matching did. For example, while frequency of auxiliary verbs in CDS reflected no correlation with children's production, the caretakers' use of inverted yes-no questions (*Do you like chocolate?* vs. *You like chocolate?*) was positively correlated with children's developing control of verbal auxiliaries. "This was taken to reflect the increased perceptual salience of sentence-initial auxiliaries to the child, who was seen as having a processing bias in favor of attending to the beginnings of utterances."[90]

As the short discussion above illustrates, the social perspective on language development takes CDS very seriously, unlike the theories mentioned thus far. While the proponent of UG views input as necessary for the "switches" to begin flipping which then set a predetermined course of outcomes in motion, LAS theorist argue that language development is dependant on the *type* of input received, is not universal, and is highly variable at any given microgenetic stage. Perhaps the most readily cited objection to such overt reliance on CDS in determining development comes from studies of cultures or societies wherein CDS is limited or absent. A study by Heath (1983) with the people of

[90] Ibid, 25.

Trackton, a rural community in the south-eastern U.S. has revealed that children are not even seen as "conversation partners" until they can produce multiword utterances. Similarly, as Schieffelin (1990) points out, the Kaluli of Papua New Guinea view infant babbling as "bird talk." They not only refrain from curtailing speech so as to be more "child friendly," but speaking to children in any form is discouraged.[91] In societies such as these, CDS appears to play no facilitative role. Nonetheless, the Kaluli children and the children of Trackton learn to speak just as flawlessly as middle class Anglophones who are the typical subjects of CDS.

While UG proponents are quick to cite findings like that of the Kaluli, to bolster the claim that there must be an innate endowment for linguistic development, it is not altogether clear that they have a sound case. Lieven (1994) points out that in these very same speech environments, where CDS is rare or non-existent, children are socialized into language learning in alternative ways. In the Kaluli culture specifically, children imitate conversational routines, which are explicitly taught by adults.[92] It is also common to find children participating in contextualized talk routines or in group settings, where peer-to-peer language learning is encouraged. Lieven therefore concludes:

> The study of child language development cross culturally supports the idea that children will only learn to talk in an environment of which they can make some sense and which has a structure of which the child is a part; on the other hand, children can clearly learn to talk in a much wider variety of environments than those largely studied to date. This is…only partly because of the repertoire of skills that the child brings to the task of learning to talk. It is also because there are systematic ways in which the structure of within which the child is growing up gives her/him access to ways of working out the language. [1994, 73]

[91] 710-11
[92] Lieven, (1994) In *Input Interaction and Language Acquisition,* Galloway, et al, Eds.

As is obvious, the attention paid to CDS in sociocultural theory reflects not a broad determination of the effectiveness of "baby talk" on language learning in general, but rather, places CDS in a class among other input types to which a child might be subjected. That particular type of input in turn, reflects the society of which the child is a part, which is said to shape language learning, as opposed to innate and universal principles.

2.4.2 Error Corrections are Uncommon, but Recasts are not

Another typical argument put forth by nativists has to do with the limited amount of feedback children receive, especially when it comes to grammatical errors produced by the learner. Moms and dads simply do not provide explicit corrective feedback when their children make mistakes. This fact, although researched in-depth, is easily recognized intuitively by anyone who has spent time around children. If a child utters the phrase *Daddy goed to the store*, we would all gasp in shock to hear the child's mother reply, *No, son. You have overgeneralized the add –ed rule in this case. 'Go' is an irregular verb, which means that its past tense form does not follow the standard rule. The correct way to say that is "Daddy went to the store."* Again, because of this lack of error correction, it is argued by nativists that developmental sequences are innately prescribed in the genes. In fact, even if corrections like the one above were offered, claims the nativist, children would still proceed by first using the correct form of the past tense of 'go,' followed by the overgeneralization of the add –ed rule (goed), and then even perhaps the double inflected "wented," until finally arriving at the correct "went." Input, particularly error correction, according to UG, really plays no acquisitional role for the child.

Simply ruling out the importance of input based on the scarcity of explicit error correction is misguided however. Not all error correction comes in the form of discrete grammatical commentary. Perhaps the most common type of corrective feedback found in CDS are recasts. Broadly defined as any repetition of a speaker's utterance with the intent to correct, expand upon, or draw attention to a particular aspect, either morphosyntactic, semantic, or phonological, it has typically been more common to find research pertaining to recasts in second language acquisition studies. Some examples of recasts utilized between teachers and students or native speakers (NS) and non-native speakers (NNS) include:

Focus on Form (FonF): *Student*: Why does the aliens attacked earth?
Teacher: Right. Why did the aliens attack earth?[93]

Corrective Recasting: *Jose*: I think that the worm will go under the soil.
Teacher: I *think* that the worm *will* go under the soil?
Jose: (No response)
Teacher: I *thought* that the worm *would* go under the soil.
Jose: I thought that the worm would go under the soil.[94]

Expansive Recasting: *Teacher:* What did you do in the garden?
NNS child: Mm, cut the tree
Teacher: You cut the trees. Were they big trees or were they little bushes?
NNS child: Big trees[95]

Recasts in second language learning are not only frequent, they are operationalized by many researchers and instructors alike, in attempts to increase rate of learning and the

[93] Mackey, et al. (2000), 11.
[94] Doughty and Varela (1998). The task was administered in a science classroom, where students were asked to report on their initial hypotheses. Thus, the recasts are an example of FonF, insofar as they emphasize only grammatical errors, but they are presented by repeating exactly what was said by the students and then placing emphasis on the incorrect form by raising the pitch of the voice when repeating the incorrect phrase. (124)
[95] Oliver (2000), 140.

students' affective filters by avoiding explicit error correction, which is often argued to raise frustration levels and unnecessarily embarrass. In first language acquisition however, there has been significantly less work aimed at documenting recasts in CDS. Tomasello (in press) does point out nevertheless, that Sokolov and Snow (1994) have extensively argued that recasts are quite common among CDS and that these recasts offer the child potentially useful negative evidence with which they can reformulate their hypotheses about the way the language actually gets used by its members.[96] This evidence flies in the faces of those like Janet Fodor and Carrie Crowther (2002), who claim that the 'poverty of the data' argument has been misunderstood by critics of UG. What is impoverished, they suggest, it not just *positive* data; in fact the child often hears what he or she *can* say in a given language. It is never told to the child explicitly what *cannot* be said.[97] Thus, since there can be no other explanation for how the child learns that *she may have been singing* is okay, but **she may being have sung* is not, we must accept that there are innate guiding principles which disallow certain constructions. But, as Sokolov and Snow argue, negative evidence *is* available to the child. Although the question remains, to what extent, if any, do recasts become part of the uptake of the child, thereby actually improving rate of development? While this positive correlation is speculative at best, the notion that children are exposed to an array of both positive and negative examples of the language is central to the LAS model. When children learn a language, they are not just learning a set of rules in a strictly linguistic manner; they are learning about social norms, interactions, conceptual schemes, and the thoughts of others.

[96] See Sokolov and Snow (1994), page 47.
[97] Fodor and Crowther (2002), in Sampson (2005), 89.

2.4.3 Non-Linguistic Cues and TOM

In his book *How Children Learn the Meanings of Words,* Bloom (2000) rightfully claims that:

> Learning a word is a social act. When children learn that rabbits eat carrots, they are learning something about the external world, but when they learn that *rabbit* refers to rabbits, they are learning an arbitrary convention shared by a community of speakers, an implicitly agreed-upon way of communicating. When children learn the meaning of a word, they are-whether they know it or not-learning something about the thoughts of other people. [2000, 55]

Although the scope of this paper does not permit an in-depth discussion of word-to-object mappings or concepts, I think Bloom's point can be extended to language learning more broadly. Indeed, as socio-cultural theorists such as Ohta (2001) have argued, both about first and second language acquisition, learning is first social and then individual. The child begins by *intermentally* co-constructing new language with others. Then, and only then the child can *intramentally* hypothesize about forms and rules. The idea that children read intentional cues provided by other speakers and thereby learn something about the structure of the mental representations of those speakers, which then allows for language learning on an individual level, is typically referred to as Theory of Mind (TOM). TOM is not seen as a comprehensive picture of language development by social theorists however, but instead is one of the necessary conditions for the possibility of language learning getting off the ground.

Besides intention reading and TOM, other non-linguistic cues exist which are argued to facilitate development. Direction of gaze, for instance, is important because at only nine months, infants will naturally follow their caretakers' gaze as well as finger

pointing.[98] From a sociocultural perspective, gaze following or gesture recognition provides a scaffolding mechanism for the infant learner. The child's learning is mediated through social acts such as pointing, which serve to capture attention, highlight features of the environment, or supplement the linguistic data that may or may not accompany such acts. Arguing that children follow gazes, and children learn language, therefore gaze following aids in language development would be an oversimplification. It is not entirely clear what kind of work non-linguistic cues actually accomplish in the grand scheme of ontogenetic development, let alone language development. TOM advocates would suggest that during attentive gaze following, infants have an implicit understanding that whoever is doing the gazing must be interested in some object or person to which their gaze is directed. If true, then it could be argued that following a caretaker's gaze is a lesson in mind reading for the growing infant.[99] Conversely, it could be said that gaze following is only the product of an "automatic orienting procedure, either innate or learned, that is initiated by exposure to certain stimuli,"[100] which would mean that reading non-linguistic cues has little or nothing to do with intentional attribution. These considerations are interesting, but reconciling them is well beyond the aims of this project. Suffice it to say that for LAS models things like reading facial expressions, following gaze, and responding to gestures are all shared social acts which mediate learning. Language is not the condition for this shared meaning; it is rather, the product of it.

[98] Cf. Moll & Tomasello (2004).
[99] Bloom (2000), 62, and see also Johnson, Slaughter and Carey (1998), for an experiment argued to demonstrate that gaze following has to do with intentional attribution.
[100] In Bloom (2000), 62, who cites Butterworth, (1991), Corkum and Moore (1995), and Perner (1991).

2.4.4 Socialization of the Infant: Why Worry?

The point of discussing the LAS model was two-fold. First, it provides us with a serious critique of strictly linguistic theories of acquisition. Thus, in some ways, it is a direct attack on Generative Grammar and UG. It does not however, provide a pointed criticism of UBL or Emergentism in a strict sense. After all, UBL proponents and Emergentists alike want to account for social interactions. Whether or not they succeed in doing so will be explored in Chapter Three, but the LAS model cannot reasonably be concluded to stand as an explicit objection to the aforementioned cognitive approaches. It merely *expands* upon certain themes running through said theories. As the well-known social theorist, Vygotsky, originally argued that the task of psychology is to understand the manner in which social and mental activity is organized through culturally constructed artifacts and social relationships, the same is true of "neo-Vygotskyan" acquisition theories. Just as humans use hammers, forks, and automobiles to mediate between self and physical world, they use symbolic tools to mediate and regulate relationships with others. Language is just one tool among many-art, arithmetic, music- and thus should not be studied *detached* from these phenomenon, but rather, in a more holistic manner.[101]

The second reason for mentioning the LAS model is perhaps more relevant to the overall point of this project. Rarely, if ever, do the terms "innate" and "predisposed" arise in the sociocultural literature. Despite this fact, it would be logical to assume that principles and parameters are not close friends of the LAS model. There is definitely a sense in which language is "learned" under the social framework of development, but

[101] cf Lantolf (2000), p78-84.

even more interesting is that LAS theorists insist that shared social acts provide the very conditions for language itself. This would seem to rule out any possibility of innateness, when it comes to language. The only thing that could be argued to be innate according to the sociocultural perspective, I suppose, would be some type of predisposition to behave socially. Therefore, part of the following chapter will address this issue. Specifically, I will examine Emergentist theories, simulated by connectionist models, in order to determine whether or not social interactions are in fact included in the simulations. One other question to explore will be under which of the three constraints, if any, should social interactions be placed?

Chapter 3: Emergentism, Connectionism, and Minimal Nativism

3.1 Introduction

At this point, it would seem as though I have sketched a rudimentary overview of the major acquisition theories available, pointed out some strengths and weaknesses of each, and then left them all hanging midair. Nevertheless, as the close of the preceding chapter suggests, the direction of this project has been ever so slightly moving towards the Emergentist position, for three reasons. First, Emergentism takes the bold step of trying to account for all the interactions that provide the basis for language learning, unlike UG, which simply posits that some innate knowledge somehow interacts with the environment and amazingly, kids start talking. While innateness is not altogether ruled out by Emergentism, if there is any innate language faculty or representational knowledge of grammar, then clearly, more needs to be said about the mechanism behind putting such foreknowledge to use. Second, there is an abundance of evidence which suggests that Emergentism can be effectively modeled via connectionist simulations, and can therefore, stand as a testable theory. As Elman (1996), Rumelhart and McClelland (1987), and others point out, when a theory is supported by models which actually run and provide substantial proof of that theory, there is much less room for accusations of hand waving. Last and most importantly, I have devoted a great deal to exploring what happens to the term "innate" in each acquisition theory and I realize that to assume Emergentism is the correct story is also to beg the question when it comes to the correct conceptualization of "innate." However, if I am successful in demonstrating the validity

of connectionist research providing accurate analogs of learning, as well as lending support to language as an "emergent" property of these simple learning mechanisms, then it seems reasonable to conclude from there, that whatever we want to call "innate" should be restricted to architecture, both in terms of networks and the mind. In other words, while specific representations, such as the principles of grammar for example, are not innate, the way the brain is organized or 'designed' could still be innately specified. More specifically, the architecture of the mind, much like that of the connectionist networks I will discuss, seems to be intrinsic to the system; that is, the human mind comes equipped, not with the *representations* needed for language acquisition, but with the learning system which is predisposed to take on this task. Consequently, this characterization which allows for an LAD that is not representationally specified, but nonetheless is innate, will not run contrary to what many nativists have in mind when asserting linguistic nativism. However, I will spend some time at the end of the chapter discussing the problematic nature of such a definition of innate, particularly when it comes to denying Representational Nativism (RN), or the argument that representations are indeed innate. Based on these compelling reasons to maintain a nativism regarding representations, contrasted with equally compelling reasons to reject such a position, it might make sense to choose an alternative explanation of what we mean by "innate."

3.2 Nodes, Neurons, Networks

The allure of connectionism has a lot to do with the analogue to the brain the models are argued to provide. Nevertheless, it remains contested exactly how the models

are supposed to line up with neurons, synapses, and regions in the brain.[102] It is also dubious to try and pin down one clear definition or account of Connectionism Proper, as it is not a specific theory on its own, nor does it emphatically support one particular theory over another. Even more challenging is the fact that connectionism does not even relegate itself to one specific field of inquiry; indeed, it provides models which explain learning in general, not just language learning, and is thus employed in cognitive science, linguistics, neuroscience, and every discipline in which the study of learning is at stake. This chameleon-like utility of connectionism does not, however, preclude some universal basics. In order to approach the question of whether or not connectionist models support Nativism, Emergentism, or some other language acquisition story, it will be helpful to clarify these basics and then discuss how they can be used in the specific arena of language learning.

In all connectionist models, whether recurrent, feed-forward, or some other complex architecture, their most basic components are nodes and weights. Typically, a node is analogous to a single neuron and like neurons, nodes send and receive inputs. Also like neurons, nodes might be responsible only for sending and receiving inputs to and from other nodes, or they might receive input from the outside world, and last, they might send activation outward. Depending on the network, nodes might do one, two, or all three of these tasks. There are arrows between the nodes which are meant to represent the direction of either *excitation* or *inhibition* and these connections between nodes are measured in real numbers, which are called weights. Weights are either unidirectional or

[102] Cf Elman et al, who suggest that typically, nodes are supposed to share a one-to-one mapping with neurons and the weights are said to be like synapses between neurons. However, in the case of hidden nodes, it would be incorrect to map them onto "interneurons." In fact, Elman et al simply skirt the issue of to what hidden nodes are supposed to be analogous.

bidirectional, and serve to multiply the *output* of sending nodes. For example, if node P has an output of 0.4 and the connection between P and Q is a weight of 1.0, then node Q will receive a signal strength of 0.4, or the product of P's output and the connection strength between it and Q. When signal strengths are positive, they are considered excitatory and when they are negative, the connection is inhibitory.

Input is calculated by the taking the product of a node's activation and the weight on the connection it shares with another node. Since input might come from a multitude of sources, we want to know the *net* input to the node in question, which is just the summation or input from all node sources. Output, or the nodes activation value, on the other hand, cannot be determined strictly from input, except for in the cases of nodes which are entirely linear. In these cases, the activation value just is the net input. For other networks, what the node "does" is captured by a response function that involves a threshold. For example, node P might output a 1.0 if and only if the net input exceeds threshold value X and will emit a 0.0 for all other threshold values. Outputs which are at or close to 1.0 and 0.0 are more rigid and are examples of more rule-like behavior, while outputs closer to 0.5 represent more sensitive nodes which behave in a more graded fashion.[103]

Because the connections between nodes are at least partially constitutive of knowledge,[104] architecture plays a key role in determining what kind and how much knowledge a network can have. A simple two-layer network consisting of one output

[103] Elman, 52. According to Elman, et al, these nonlinear responses are what allow the networks to make more subtle distinctions among categories, rather than the all-or-nothing categorical responses of linear outputs.
[104] Although we might go so far as to say that weights just are knowledge, but such a strict identity thesis relating to connectionists nets might create more problems than it would be intended to solve.

node and two input nodes can easily implement the logical AND, for example, if both input nodes are 1.0, or "on." In a network like this, a pseudo-node, or what is sometimes called a *bias node* might present a negative weight, such that by default (in the absence of external input), the output will be "off." It is only when the net input of both nodes is large enough to overcome the negative bias that the output can be turned on. This network cannot, however, implement the Exclusive OR (XOR) function. According to Elman et al, the only way a network can implement such a function is to allow for a third layer of nodes called "hidden units," which are equivalent to internal representations.[105] Thus, the number of layers a network has predetermines the complexity of representations it can have.

There are several types of networks employed in connectionist research for various purposes. In the case of simulating language learning, it is not uncommon to find a all of them used for one element or another. We might, for instance, find a feed-forward model that simulates phomemic acquisition or a complex model consisting of several recurrent models interconnected that simulates syntactic constructions. The point of the following section is not therefore, to decide upon a superior type of model for language learning in general, but to examine several studies that simulate specific instances of language learning, such as morphosyntactic acquisition, in order to determine what, if any, type of acquisition theory these studies support.

[105] Ibid, 64, although Ramsey (1994) disagrees with this notion of hidden nodes as representations.

3.3 From Simple to Complex: Emergent Linguistic Phenomena

3.3.1 Phonological Sequences

It was mentioned above that a two layer network cannot implement the XOR function, as a set of hidden units are needed in order to make the task of prediction possible. This is because for the network to be able to predict the correct output for a string of inputs, such as 11000111001, it must rely on both the input itself and the network's previous state. Thus, a network learning the XOR function will have both a hidden layer which represents the previous state of the input immediately preceding the current string, along with its surface layer that processes the current string. According to Elman (1990), what is most interesting about XOR is how temporal structure affects learning. When fed input all at once, output is no better than randomly correct; but, when given each bit one after the other, the network is actually able to develop memory, which enables it to call upon previous strings in order to predict the next output. The network adjusts the weights at every time step based on the difference between the predicted state and the target output. In other words, the network gets better with experience. Learning in this sense is not, however, very complex, since it requires just that the network's memory extend back one temporal bit and the network is only ever exposed to four different input patterns.

In order to expand upon the XOR simulation, Elman (1990) reproduced the model, this time with 6 input nodes, 20 hidden nodes, 6 output nodes, and 20 context nodes. In this case, the inputs were binary vectors, each corresponding to speech

sounds.[106] While the input was random in the sense that consonants were highly unpredictable, it was constrained insofar as the vowels that followed consonants were regular. As the network was fed one phoneme at a time, it could easily predict both the identity and number of vowels which would follow. It was trained with 200 sentences containing no word or sentence boundary information, but the sequences wrapped around so that the first pattern immediately came after the last. An example of this type of input is:

Manyyearsagoaboyandgirllivedbytheseamanyyearsagoaboyandgirl…

The task of the network was to predict the next letter in the sequence and, much like in the XOR simulation but to a more complex degree, it was expected that to do so, it would process the input recursively, such that at Time 2, the hidden nodes would rely on the input of Time 2 as well as the results of processing at Time 1 from the context nodes. Errors made suggested that while the network might not be able to predict the exact output that should follow in the sequence, it could, with great regularity, predict the correct *type* of output. On a global level this means that that following consonantal input, error is very low, as the network is consistently able to predict precisely what and how many vowels should follow. At the end of a vowel sequence, however, error is high, as the network cannot predict with accuracy the correct consonant to follow. Thus, on this level, the network has learned to abstract orthographic sequential probabilities up until the boundary of a word is reached at which point, error shoots back up.[107] Nevertheless, since error at the global level is the sum squared error over all six bits of input, an

[106] Although, they were not derived from real speech, as there are far more than six phonemic dimensions in articulatory patterns.
[107] Elman (1990), 189.

examination of error at the bit level shows us something more interesting. While the network cannot know with certainty which consonant follows, it can know *that* a consonant follows. It therefore moved from processing surface regularities to processing more abstract categories, such as "consonant," "vowel," "bound morpheme," and so on. However, these linguistic constraints were not in any way programmed into the original architecture of the network. They simply emerged as learning took place. Even more fascinating is that in contrast with the XOR function, the phonological sequences are longer, depend more on temporal context, and are 6-bit rather than 1-bit vectors. In short, this task should have been harder for the network,[108] as the complexity of the sequential dependencies might have confounded or even overloaded the processing capacity. But, as Elman points out, "The fact that there are subregularities (at the level of individual bit patterns) enables the network to make partial predictions, even in cases where complete prediction is not possible."[109]

3.3.2 Word Order

Elman also explores the constraints placed on word order in sentences. It has long been argued that word order cannot be accounted for linearly, or, as Chomsky puts it, there must be a "deep structure" underlying surface structure which allows for such

[108] And in a sense, one could argue, that because of the training time being longer and the sequences being more complex, that the task was in fact harder. However, because something takes longer does not entail that it is more difficult. As Elman finds, the error was still high for the precise vowel or consonant that followed, but was quite accurate in determining what *type* of phoneme (vowel or consonant) followed, which was of more interest to his study.
[109] Ibid, 191.

constraints on sentential organization.[110] Nevertheless, the language learner is not directly exposed to the abstract order, but rather has visible and audible access to surface strings. The deep structure is cued by the surface forms and is therefore implicit in them.[111] Thus, in order to reconstruct the abstract structure from the surface form, it would seem that a certain amount of innate knowledge is required, perhaps a symbolic knowledge that extends beyond the capacity of nonsymbolic connectionist models. However, if a network can learn the abstract structure, even if no elements of that structure were available in the input (ie-the network is given strictly surface cues), then this latter assumption might appear questionable.

Elman trained a recurrent network on strings of words from a sentence generator, which produced random two- and three-word sentences from a template, consisting of 29 lexical items from 13 classes of Nouns and Verbs. One entire sequence was composed of 27,354 31-bit vectors, and each word was input one at a time. Training involved cluster analyzing the network's ability to predict the 31-bit vector equivalent to the next word in the string after six complete passes through the sequence. This analysis showed that the representations the network had formed across its hidden unit activations corresponded to major categories of words. The category structure was hierarchical, insofar as "mouse" belonged to the class of small animals, but also to the non-human, animate, and finally, noun classes. This hierarchy was also implicit and graded, such that while some categories are very distinct, and thus far apart from one another in the hidden unit activation space, other categories might share properties, leaving the boundaries between

[110] Cf. *Syntactic Structures,* 1957.
[111] Elman (1990), 195.

them blurred. Membership in a category can therefore range from clearly demarcated to marginal.

Perhaps the most noteworthy point to take away from Elman's network is that it was able to move from processing surface regularities to representing a more abstract structure, but without any syntactic or linguistic rules built into the original architecture. The network was also not provided with any real world information or semantic cues, which means in a sense, according to Elman, "the network has much less information to work with than is available to real language learners."[112] Because the architectural constraints were general and yet, the network grew to represent linguistic-specific constraints, it would make sense to say that these latter constraints "emerged" during learning. Ellis (1998) is quick to presume as much, arguing that models such as Elman's provide the psychological model to support the theory of Emergentism. Linguistic regularities emerged out of a simple learning algorithm, which is analogous to the way language emerges as a result of the interactions in the brain of the language learning child.

One must take caution however, as Finch and Chater (1994)[113] suggest, in simply inferring from a simple model to the overall reality of language acquisition. A simulation similar to Elman's conducted by Chater and Conkey (1994) showed that recurrent networks which rely solely on prediction, become "extremely inefficient and slow" as the language grows in complexity. Furthermore, statistical information seems to be limited as far as providing decidability among grammars. As we recall, Elman's network, along

[112] Ibid, 201.
[113] Finch, S., and Chater, N., (1994). In Oaksford, M., and Brown, G.D.A., (1994) *Neurodynamics and Psychology.*

with that of Chater and Conkey, were not exposed to real-world stimuli (the networks obviously could not hear or see, etc), which means that negative evidence was entirely left out of the input. I have pointed out in the previous chapters that exposing children to 'what not to say' is a crucial factor in acquisition and Finch and Chater want to suggest that it is the *only* way a learner can disconfirm the plausibility of a grammar. Positive evidence alone will never entirely rule out one string of words over another, they argue, and therefore:

> while the use of statistical information in place of negative evidence is attractive in principle, its practical application is very difficult because the space of possible grammars is so vast. Thus strong innate constraints on the nature of language may still be necessary to explain how language is acquired. (1994, 319)

3.3.3 Lexical Semantics

In a somewhat different simulation, Landauer and Dumais (1997) trained a network to recognize word patterns from reading texts. The Latent Semantic Analysis (LSA) model, although a purely mathematical simulation, was employed to solve the problem of "excessive learning," or the fact that people seem to know a lot more than what is contained in the information to which they are exposed.[114] In particular, the amount of vocabulary children acquire by high school is astounding, considering that only a small fraction of that number is found in everyday speech. Quite clearly, children learn the meanings of many words through reading, but it is not exactly clear how. Most experiments conducted to date, in which children read and reread paragraphs and are tested on the vocabulary therein, find that new words are acquired at a rate of about 2.5 per day. However, as Landauer and Dumais point out, simple math will show us that if

[114] 1997, 214, 211.

the average total vocabulary of 20-year olds is somewhere between 40,000 and 100,000 words, then on average, 7-15 words must be added daily. Between fifth and eighth grades, they suggest, the rate is closer to 10-15 new words a day.[115]

The LSA is described conceptually by Landauer and Dumais as "a simple but rather large three-layered neural net. It has a Layer 1 node for every word type (event type), a Layer 3 node for every text window (context or episode) ever encountered, several hundred Layer 2 nodes, and complete connectivity between Layers 1 and 2 and between Layers 2 and 3."[116] Since the model was intended to simulate young children's learning of vocabulary from text, sample paragraphs were taken from Grolier's *Academic American Encyclopedia,*[117] the words were analyzed by Singular Value Decomposition, (SVD)[118] and the text data were cast into a complex matrix, the cells of which marked the frequency word types appearing in the text sample. Simply put, the model was trained to treat words as similar if they co-occurred with the same words over and over again in various contexts. Training consisted of exposure to 30,000 text samples from the Encyclopedia.

In order to test the model's word knowledge after training, Landauer and Dumais administered a portion of the TOEFL (*Test of English as a Foreign Language*), which focused on synonyms. Typically, the TOEFL section on reading will contain 30-40 questions, about 30% of those being synonym questions, and the listening section will be

---

[115] 221
[116] Ibid, 217
[117] 1994
[118] SVD is the mathematical realization of the process behind LSA. In short, it uses factor analysis to find the most concise representation among all the relationships between a set of variables. In this case, SVD is used to parsimoniously represent the correlations between the columns and rows of a matrix, the former standing for words, the later standing for the contexts in which those words appear. (See page 218 in Landauer and Dumais)

about the same length, with only 20% of the questions relating to vocabulary. Thus, roughly 15-17 questions on the TOEFL will deal with synonyms. The model scored 64.4% of these questions correct. By comparison, the average score of a sample of non-native English speaking applicants to U.S. colleges is 64.5%, which is a score suitable for admission. Based on these results, the model can be said to closely mimic the proficiency and behavior of intermediate-level English readers. What is perhaps even more noteworthy is that LSA achieved this status by "using text samples whose initial representation was simply a 'bag of words'; that is, all the information from word order was ignored, and there was, therefore, no explicit use of grammar or syntax."[119] The model was also unable to make use of phonology, morphology, orthography, or any other real-world information, as it was devoid of sensory-motor capacity. To Ellis and other Emergentists, this simulation arguably demonstrates that lexical semantic acquisition can "emerge" simply from the analysis of word co-occurrence.[120]

However, it must also be kept in mind that the network was only trained on synonyms. It did not have to learn all the other skills relevant to the TOEFL-recognizing coherence patterns, inferences, writing essays, etc-and thus, its score reflects its performance on an isolated task. While it is true that the average score for admittance into college for an ESL student is about 80 points out of 120 (66.67%) and the network achieved close to this performance rate, it would be a far cry to assume then that the network is as good at English as an international student who has just successfully passed the entire TOEFL. Sixty-six percent is considered a strong score when all four components-reading, speaking, listening, and writing-are tested over a 3.5 hour period.

[119] Ibid, 220
[120] Ellis, 1998, 651.

Sixty-six percent on a single set of synonyms after much training seems far less impressive.

3.3.4 Morphosyntactic Acquisition

Rumelhart and McClelland (1986) were the first to simulate the acquisition of inflectional morphology which subsequently has paved the way for several other compelling, if not controversial simulations of a similar type. Morphosyntactic acquisition is unique in that it builds upon processes of acquisition discussed above, such as phonological sequencing, which itself generates letter and word sequences (-ed, -ing, -s, etc) that provide cues to tense, case, number of referents, and so forth. Perhaps one of the most fascinating aspects of early childhood morphological acquisition is the overwhelming tendency to overgeneralize, such as the infamous case of –ed marking the past tense for *every* verb, including irregulars, resulting in constructions such as *Mommy goed to the store.*

The original simulation conducted by Rumelhart and McClelland which used a simple recurrent "pattern associator"[121] was designed to mimic this particular behavior common to young children, what is known as the U-shaped learning curve.[122] It was also intended to capture the aspects of varying performance on different types of regular and irregular verbs. Bybee and Slobin (1982) catalogued nine different types of irregular verbs and three different regulars; for example, Type I irregulars, where forming the past involves no change whatsoever (beat/beat, hit/hit), versus Type VIII irregulars such as *grow*, whose past tense form *grew* involves a vowel change and a dipthongal ending

[121] 1986, 226-30 for a full description of the model.
[122] 240-241

sequence. Thus, for Rumelhart and McClelland, their task was to simulate findings similar to Kucazj's (1978), which show that children perform with much more accuracy when forming the past tense of Type I verbs, but tend to overgeneralize and use the *past+ed* construction for Type VIII verbs. After a training sequence, the pattern associator model had successfully simulated these two aspects of morphosyntactic acquisition; it began by accurately forming the past tense of the "easier" types of irregular verbs, but overgeneralized the rule of *+ed* for regular verbs, so that its initial responses involved outputs such as "*wented*." As training proceeded, the model produced more and more correct forms, progressing through the three-stage learning curve through which Kucazj and others have shown children progress.

The third and final aspect the pattern associator was intended to capture was the ability to respond correctly to unfamiliar verbs. Rumelhart and McClelland found that:

> On no occasion does the model assign a strength greater than, 2 an incorrect variant of the past tense. Thus, the model has clearly learned the substructure of the regular correspondence and adds the correct variant to all different types of base forms. These results clearly demonstrate that the model acts in accordance with the regular pattern for English verbs and that it can apply this pattern with a high level of success to novel as well as familiar verbs. [1986, 263]

Of course, there were rather surprising responses given, such as changing *squat* to *squawked*, or implementing the *past+ed* form of seven of the irregular verbs. They speculate that children and adults make similar errors, although the literature to date does not support such a claim.[123] Nonetheless, the model performed with greater-than-expected accuracy, thereby capturing the aspect of acquisition which involves dealing with novel verbs.

[123] 265

Rumelhart and McClelland's study sparked much criticism,[124] but it also paved the way for some significant improvements in connectionist modeling of the past tense, simulations which either build upon or add to the original. For example, Ellis and Schmidt (1997), et al "have closely simulated the error patterns, profiles of acquisition, differential difficulties, false-friends effects, reaction times for production, and interactions of regularity and frequency that are found in human learners, as well as acquiring default case allowing generalization of "wug" tests."[125] Like the original simulation by Rumelhart and McClelland, these current connectionist explorations support a view of acquisition of past tense morphology as the result of associative learning principles that use optimal inference to abstract regularities. Sequence analysis, for example, between phonological form of lemma and phonological form of inflected form, leads to correctly formed surface strings, both morphologically and syntactically. However, these patterns of association are never programmed into the models; they emerge through training and repeated exposure to regularities in the input.

3.3.5 Syntactic Constructions

Nativists draw a large amount of support for their claims from data which show children's ability to use syntactic frames as reliable cues to word meanings. For example, Landau, Smith, and Jones (1992) found that when children are presented with a novel word like "xerillium" within a phrase such as *Give me the xerillium one, not the red one,* they treated "xerillium" as a color word. Or, as MacWhinney points out, if we ask a child to "repulsate" the tub with water, they will inevitably take "repulsate" to be similar in

[124] Cf Pinker and Prince, 1988.
[125] Ellis (1998), 652

meaning to the verb "fill" that has as its goal, a direct object, and a transferred object in an indirect instrumental phrase.[126] Findings such as these, argues the nativist, suggest an innate and tacit understanding of the role syntax plays, as the input is far too impoverished for such knowledge to be imparted.

Connectionist models nevertheless suggest something quite different about how syntactic constructions are learned. MacWhinney (1997) developed a network intended to simulate the abstraction of regularities based on surface cues. When presented with a sentence such as *The girl eats the apples,* there are several cues, some stronger than others, which direct the learners' attention to form and meaning. For example, the morphological marking –s on *eats,* cues one to the agreement between the singular *girl* and *eats*, rather than *apples* and *eats.* MacWhinney used cues such as these for the input layer and for the output layer used nodes representing functional interpretations such as topicality, definiteness, and perspective. Hidden unit activations allowed for associations to be made among all the various input and output nodes.

3.4 Connectionism and Emergentism: Two ways of saying the same thing?

So far, I have painted Emergentism in much more flattering light than any Nativist or even Constructivist accounts. As the second chapter showed us, there are analogous examples of emergent phenomena found in the natural world as well as among human artifacts. Besides the abundant examples available in nature and elsewhere that endorse an emergent theory, there are also, as we have seen, numerous reasons to reject a UG-type approach. In particular, UG is limited in its explanatory power; it is not that the theory offers nothing in terms of linguistic explanation, but more that it *only* describes the

[126] MacWhinney, 1998, 14.

linguistic element to language and really provides very little in terms of the process of acquisition itself. As Goldberg (1995) and others have attested, we ought to focus more on the Language Acquisition Process (LAP) than the device or mechanism (LAD). While Chomsky and his followers would have us focus on a specific module or "language organ" that is said to come equipped with all the necessary blueprints for language acquisition, Emergentists suggest that explaining and understanding this 'initial state,' although an interesting and important endeavor, is dubious at best. We can reap more, says the Emergentist, from attempting to explain what actually goes on once language learning begins.

The negative evidence against nativism coupled with positive examples of emergent phenomena as analogies to language acquisition are further strengthened by the fact that it is possible to create working models of Emergentism. As the previous sections have just illustrated, Connectionism provides us with simulations for many aspects of language learning, including but not limited to phonological acquisition, lexical-syntactic acquisition, and morphosyntactic acquisition. This is not to say that simply because we build a model of a theory it gives the theory viability over and above others. Quite clearly, a model of UG might be constructed and for all intents and purposes, it might work very well. There are however, two elements that I think separate the above connectionist models from other types of models. First, their analogous nature to the human brain, albeit not an exact one-to-one relationship, offers a strong parallel, one that is useful not only because the models actually work, but also because they purport to

mimic behavior exhibited by the human brain as it faces the task of learning language.[127] Second, the simulations are not restricted to language learning; rather, connectionist models offer simulations of many other behaviors such as mathematical computation.[128] Just as in the language acquisition simulations, these other connectionist models suggest many other forms of learning "emerge" out of simple systems with little to no prespecified representational knowledge.

Therefore, a preliminary conclusion I will draw at this point is to argue that Emergentism and Connectionism are two sides of one coin. As described above, while the former provides the theoretical description of the process involved in language acquisition, the latter offers a model which is analogous to the processing unit involved in language learning, thereby simulating what actually goes on during particular aspects of learning. Of course, there are plenty of connectionist models that arguably do not simulate any kind of emergent features of learning. Likewise, there are opponents to the argument I have just marshaled, like Fodor and Pylyshlyn (1988) who argue that connectionist models most closely resemble Constructivism, or Quartz (1993), who asserts that the models are "nativist in a robust sense." These objections not only threaten the viability of the preliminary claim just made that Emergentism and Connectionism are inseparable aspects of a larger picture of language acquisition; they also raise the more fundamental question of whether or not talking about language as an "emergent" phenomenon is really accurate. If there are interpretations of connectionist simulations that go so far as to call them 'robustly nativist,' then it seems pertinent to re-examine

[127] Although models such as the LSA are so vastly different from the brain, both in construction and behavior, that 'purporting' to model the brain might even be too strong a claim.
[128] Cf. Horgan & Tienson (1994, 1991) and Waskan & Bechtel (1997).

Connectionism per se, in order to determine the cause of such wildly varying construals. I think it is necessary to revisit constraints-*representational, architectural,* and *chronotopic*-in order to solve this problem. As Elman et al suggest that "innate" applies to these constraints, then it seems quite clear that nativist renderings of the simulations are going to argue that more constraints are present, while Emergentist interpretations will want to say that the constraints are minimal. In particular, *representational* constraints tend to be the most contested, as they are widely denied by Elman and other Emergentists. As we shall see in what follows however, there are reasons not to wholeheartedly embrace a denial of Representational Nativism. Thus, I will explore this opposing account in the final section, seeking once more to pin down an exact meaning of "innate" and discussing exactly what we ought to assign to this category when looking at Connectionist models.

3.5 What have we done with Innateness?

In this last section, I am hoping to tie together what appear to be loose and unrelated aspects of the nativist versus non-nativist accounts of language acquisition. While to be sure, the three chapters above have provided an overview of the major contenders in the race for a comprehensive story of language acquisition, I have also, along the way, maintained a focus not simply on the arguments put forth by each, but to the very terms utilized in defending each particular claim. Most importantly, the term "innate" has been examined and its definition, depending on the theory in question, has proven to be capricious at best. To recap the rudimentary history of "innateness" sketched

thus far, we have added several elements to the neo-rationalists and neo-empiricists, such that we now have:

> Contemporary Nativists (Chomsky, Pinker, et al): Something is innate if it provides the necessary blueprints for development. All of the principles of language are "hardwired" in the mind of the infant at the onset of development, and are therefore innately given. Experience merely "flips the switches" in the already prespecified language module.
>
> The "New Wave" of Empiricists (Elman, et al): To be innate simply means to be constrained in some way. There can be representational constraints, architectural constraints, and timing constraints. Based on neurobiological research and connectionist modeling, there is reason to believe that the mind is only constrained in its basic architecture (ie-the way in which neural networks are ordered, connected to one another, and numbered) However, to claim that specific nodes (or neurons) have prespecified weights (representations) is not supported at all by neurobiological evidence in favor of the plasticity of the brain. Furthermore, connectionist simulations of language acquisition suggest that with minimal architectural constraints and no representational constraints, behavior that is directly analogous to human behavior nevertheless emerges.

Based on the last two conceptualizations, coupled with the neurobiological evidence, there is one component of the term "innate" which seems to be agreed upon, and would probably even convince the staunch Empiricists such as Locke. The architectural constraints placed on development are innately specified. In the same way that for any connectionist model to get off the ground, there must be *some* preprogrammed "set-up" (ie-how many output nodes there are, whether the model is recurrent or feed-forward, etc), so too is the brain "designed" in an architecturally constrained manner. Certain architectural constraints are often assumed by Emergentists to be the only innately specified 'information' available to the learner. But this characterization has very little cash value when it comes to settling a two thousand year

long dispute. The problem remains: For the nativist, not only is the architecture prespecified, so too are the *representations*, but for the non-nativist, in particular the Emergentist who utilizes Connectionist simulations as evidentiary support, representations need not be innate for something like language learning to take place. In fact, claims the Emergentist/Connectionist (henceforth EC), neurobiological data are sufficient in number, kind, and relevancy to entirely rule out Representational Nativism (RN). Thus, I now turn to one final perspective on the meaning of "innate" in order to determine whether or not the proponent of EC has a valid case against RN, or if it is nothing more than a straw man attack, based upon a careless equivocation.

3.5.1 Innate Specification versus Intrinsicality

The argument against RN that Elman et al put forth is based on evidence gleaned from neurobiological studies in which the brain is shown to be quite malleable and adaptable to change. For example, as was referenced in the second chapter, if auditory cortical tissue is transplanted into the visual cortical region early enough in development, then the transplanted tissue will actually take on the new role of surrounding visual cortices. These sorts of findings suggest that the representational properties of the tissue in question change, and are therefore, not innate.

However, as Richard Samuels (1998) points out, the case against RN might be confounding *innately specified* properties of cortical tissue with *intrinsic* properties.[129] Most simply put, "intrinsic" refers to any property or characteristic that inheres in something independently of the relationships it shares among other things. Prima facie, this definition seems more in line with what Chomsky, Pinker, and others have in mind

[129] Samuels (1998), 555.

when they characterize the principles of grammar as "innate." It would be a misconception nevertheless, to say that Nativists purport that these interactions play no role in the development of these interactions. Elman et al characterize a property as being "innate" if it has no recourse to information outside the organism,[130] but Nativists insist that it is a necessary condition for acquisition of representations that information from "outside the organism," in the form of environmental interactions, "triggers" the representations. They say very little about the learning process itself, but instead, seek to explain the initial state, which is comprised of innate representations, but which also require environmentally derived information in order for the "switches to be flipped." Thus, "innate," as it is defined by Elman et al, does not provide an accurate account of what Nativists have in mind when using the term.

Samuels goes on to suggest that in addition to this initial misrepresentation of RN, "there is no reason to suppose that properties cannot be both innately specified and *extrinsic*."[131] Properties are extrinsic to an object if the relationships that the object bears to other things are at least partially constitutive of those properties. For instance, it is a property of my skin to be on the outside of my musculoskeletal system. This property of "being on the outside of" is relational-the relation that my skin bears to my musculoskeletal system *constitutes* this very property-thus, it is extrinsic. This is not to say, however, that we can rule out that it is innately specified of my skin that it be on the outside of my musculoskeletal system. In fact, it seems quite plausible, at least according to this example, that this is both an innately specified *and* extrinsic property of my skin. Assuming that RN implies that representations are *intrinsic*, therefore, creates a straw

[130] Elman (1996), 22.
[131] Samuels (1998), 555, my emphasis.

man version of Nativism and masks the fact that RN is perfectly consistent with the neurobiological evidence cited by Elman et al. As Samuels claims:

> If we adopt Elman et al's assumption that some neural structures are representations (or, at any rate, implement representations), then it should once more come as no surprise that altering the inputs to a neural structure or transplanting it from one cortical region to another should affect the content of the representation. After all, such experimental manipulations may radically alter the causal relations that the neural structure bears to other representations that are implemented in the brain and, hence, alter the content of the representation. [1998, 557][132]

Another point to consider is the difference between *whole organism nativism* (ON) and *tissue nativism*. (TN) Samuels thinks that the inference made from neural plasticity to the absence of innate representations hinges on a denial of tissue nativism, or the thesis that there are representational properties that specific pieces of cortical tissue are innately specified to possess.[133] Organism nativism, on the other hand, claims that whole organisms possess innately specified mental representations. More specifically, we can say that mental representations are individuated by their contents insofar as there are types of representations, BELIEF or COLOR for example, such that all human organisms are said to possess tokens of those types.[134] While it seems clear that the neurobiological evidence for neural plasticity rules out the possibility of TN, it is not as apparent that we must also reject ON. The assumption made by Elman et al is ~TN and therefore ~RN, but again, this appears to be a misinterpretation of RN. Rejecting TN should hardly worry the nativist. After all, the claim behind ON is that *people,* and not individual pieces of tissue,

[132] Samuels rightly notes that Elman et al are not clear about whether we ought to think about representations as *identical to* or *implemented by* neural structures.
[133] Ibid, 559
[134] Adapted from a footnote in Samuels, p 559. Note that where he mentions RED as an innately specified *type*, I have changed this to COLOR, so as not to suggest that RED is a universal innately specified representational type.

possess innately specified concepts.[135] This is a psychological claim, not a neural one and it is therefore easy to imagine that ON true, even if TN is not.

We cannot accept this attack against Elman et al wholesale just yet. A point one might already have considered is that ON can be true while TN is false *only if* mental entities and mental states are *multiply realizable* (MR), or, if "tokens of the same type of psychological entity can be realized by different kinds of neural entity[s]."[136] Orignally a thesis put forth by Putnam in 1967, philosophers of mind now broadly use MR to argue against a strict type-type identity between psychological states and neural states. While it might not be innately specified that a *specific* region of the cortex represent, say, BELIEF, it might be innately specified that *some* region represent this mental state. Thus, according to multiple realizability, BELIEF might be implemented by neuron N1 for Dan, but by N2 for Tom. Samuels suggests that multiple realizability is "a common assumption in contemporary cognitive science," and that the alternative, a strict *type-type* identity is "widely agreed" to be "implausible."[137] It is to this assumption that I now turn.

3.5.2 Considering Multiple Realizability

In their paper, *Multiple Realizability Revisited*,[138] Bechtel and Mundale suggest that the irresistible urge to claim that psychological states are multiply realizable stems from the typical philosophical usage of psychological measures to map the brain. More specifically, the inconsistency with which these correlative mappings are made might hint at a case against MR. Typically, researchers have used a coarse grain to classify mental

[135] cf. Fodor, 1981.
[136] Samuels, 1998, 560. But clearly, the original thesis of MR comes from Hilary Putnam, 1967.
[137] Ibid, 560, 562
[138] 1999.

states as "the same" across individuals and species, while a finer grain is used to differentiate among brain states. However, as Bechtel and Mundale note, if we keep the grain consistent for both the psychological and the neural classifications, the case for MR does not fair so well. A more fine grained mental classification would allow for differentiating mental states among individuals or even within the same individual over time. Likewise, utilizing this same fine grain to analyze the brain, it is easy to see how one could map mental differences onto brain differences and vice versa. It is usually the case that in apparent instances of different brain activity failing to produce apparent differences in mental states, one has usually just not used a fine enough grain to analyze mental states.[139] Determining the correct grain size however, depends on the context in which the comparison is being made.

Lack of context, Bechtel and Mundale claim, is another reason that MR is so widely assumed true. It might be that Lisa and her pet octopus have the same mental states, realized by different neural structures, for example, if they are both hungry and this hunger is associated with some general food-seeking behavior. However, regarding other factors, such as how the food is sought, what food is sought, under what conditions, etc, then the food-seeking behavior and its corresponding mental state are different. As Bechtel and Mundale rightly point out, asking whether two things are the same or different only makes sense with respect to some third thing. If we use different 'third things,' it is reasonable to assume that Lisa and her octopus have the same mental states, but if we consider the context and keep it fixed while doing comparative analysis, both

[139] Bechtel and Mundale, 1999, 177-8.

psychologically and neurally, then the case for multiple realizability appears dubious at best.[140]

Regardless of whether we want to accept this case against MR,[141] the point remains that MR is not as widely accepted as Samuels assumes. The example above is just one among many. Lawrence Shapiro, for instance, makes a similar argument, focusing his attack on the Block and Fodor's argument that a psychological entity is often associated with several distinct neurological states.[142] He claims that in an attempt to forge an existence proof for multiple realizability by neural plasticity, the reality of the neurobiological evidence is overlooked. The rewiring experiments actually show that the rewired rodents do *not* see exactly the same as normal rodents-i.e. their visual acuity for recognizing light gradations is largely impoverished. Furthermore, the auditory cortex of a rewired rodent comes to look identical to the visual cortex possessed by its unwired friend. Thus, he argues, we have neither a case in which the *same* function is realized in two different ways, nor a case in which the same function is realized in two *different* ways.[143]

Nevertheless, while Samuels might be overzealous in claiming that multiple realizability is the norm, he is correct that at least it is quite a popular account of mentality. Furthermore, hemming and hawing over whether or not MR is the case misses the point, claims Samuels, as he maintains that it makes no difference if we adopt MR or a strict type-type identity thesis about mental states. I don't think we ought to just accept

[140] Bechtel and Mundale, 1999, 202-3.
[141] And my sketch of Bechtel and Mundale's argument is cursory at best, thus the reader not feel as though she is in the position to make such a decision.
[142] Block and Fodor, 1972, 238.
[143] Shapiro, 2000

this argument wholesale however, as I think the difference between the two positions does make a difference, namely, a difference between one version of nativism and another. Thus, in the following section I will briefly explore Samuel's attempt and then add to it what I take to be an essential point.

3.5.3 Why Ontology Matters

The idea that representations can be innately specified even if mental states are identical to brain states is illustrated by the following example. Suppose that a piece of cortical tissue, call it C1, encodes the representation RED at time t1. C1 is then transferred to another region of the brain, such that it no longer encodes RED. Meanwhile, another piece of cortex, C2 alters its organization so as to become type-identical to C1 at time t1 (like the auditory cortices in the rewiring experiments aforementioned). Clearly, it is not innately specified of either specific piece of cortex that it encodes RED, and therefore, a tissue nativism must be false. However, it may still be true that *some* piece of cortex represents RED, or that the *whole organism* is innately specified to represent RED. In other words, if representations *just are* brain states, it can still be innately specified of Dan that he possess the representation RED.

To be sure, this argument is convincing, as I see no problem with assuming that mental states can be identical to brain states and yet still be innately specified. In fact, I think it can be argued that *we must* accept that representations can be innately specified if we adopt a strict type physicalism. My explanation is this. Imagine that a simple representation such as the feeling of hunger is identical to some brain state, call it B1. Surely, the brain state which would be identical to "hunger" must be innately specified of

a human infant. It seems innately specified of all human infants that they possess "hunger." This particular representation runs us into trouble however, because as soon as we say that more than one infant represents the same type, "hunger," we are flirting with multiple realizability once again. It is especially a problem, as Bechtel and Mundale pointed out, when we begin talking about octopi as tokening the same type "hunger." This can be easily solved nonetheless, by rephrasing the claim slightly. For instance, we can say of each individual person that insofar as they are born with some brain states, they must then also be born with some mental states.

But the innate representation of hunger is not the best example. It might be argued that "hunger" is not a representation at all, but is really just a physical state, in a sort of eliminative materialist vein. Thus, if hunger is innately specified, there is still no argument for Representational Nativism. Furthermore, "hunger" is not interesting. Even the most radical non-nativists are content to say that "hunger" is innate, regardless of whether or not it is a mental representation. The real controversy arises naturally, when people start claiming that representations such as language are innate.

So let's consider another representation, say, a linguistic principle, such as ditransitivity. The principle (call it the Rule of Ditransitivity) says that in any clause, if the verb valency pattern is ditransitive, then a direct and an indirect object are obligatory. For the sake of argument, assume that this principle, or something like it, is represented by any speaker of a language whose verbs are both transitive and intransitive. But wait, do we really want to say that the representation of the Rule of Ditransitivity is identical to some brain state or sequence of brain states? That seems a bit odd, if not absurd. Even if we do reduce the mental state of representing the rule to some neural sequence, then is it

possible for two people to represent the same rule? Suppose Dan and Tom are both discussing verb valency patterns (apparently out of lack of anything better to do), and they are effectively communicating to one another facts about the Rule of Ditransitivity. Of course, in Dan's brain, the representation of the rule will be identical to brain states in Dan's brain, and the same for Tom, thus making the representations different. However, to suggest that they are not thinking about the same rule is an unfavorable explanation, and hence, this is where the proponents of multiple realizability step in and point out the flaws in type-physicalism.

Of course, multiple realizability is not without its problems, as Shapiro and others have rightfully pointed out. Moreover, what sorts of things representations are according to MR is neither clear at this point, nor is it obvious that they are the same kinds of things as type-physicalism would have them be. The typical MR thesis claims that mental states are 'realized' by physical states, but *not* identical to them. Functionalists feed on this negative definition and conclude that mental states must simply be functions of the brain, and hence, the same types of functions can be realized or represented by many different organisms or systems. Difficulties arise however, when the question is posed, *just what kinds of systems are sufficient to possess certain functional states?* Without delving too deeply into the abyss of the Artificial Intelligence debate, the claim that only a select number of systems can realize complex representations such as language seems relatively uncontroversial. The Rule of Ditransitivity, for instance, can be represented by people and maybe some connectionist networks, but not by much else, at least not by things like my Australian Sheppard or a thermostat. Nevertheless, the typical MR story does not clearly define the boundaries whereby we can draw such a distinction. It simply suggests

that when things are organized in the right way, they can be conscious. Even more puzzling is how Samuels is able to make the assumption that multiple realizability allows for representational nativism. The fact that many people can represent the same type of belief or linguistic rule entails the *possibility* that those representations are innately specified of the organism. If I were trying to maintain nativism, I would surely opt for the ontology that necessitates it, and not the one that only permits its possibility.

Based on my discussion so far, it would appear as if we have a dilemma: either accept type-physicalism so as to entail nativism, and then face the trouble of explaining how representations of things like the Rule of Ditransitivity just are brain states, or take the multiple realizability route and all the obscurities that come with it, along with only a possibility of nativism. On the other hand, perhaps this is only the appearance of a dilemma. The implicit assumption being made here is that type-physicalism and multiple realizability are mutually exclusive, and to be sure, this is a commonly held position. But what if type-physicalism were true *and* mental states are multiply realizable? I think such an option is not only possible, but is a way to resolve these difficulties, as well as provide a powerful argument for representational nativism.

In his article, *Putnam's Intuition,* Tom Polger discusses four varieties of multiple realizability "worth wanting," in attempts to satisfy this very possibility.[144] They are as follows:

> 1.) Weak-MR: At least some creatures that are not exactly like us in their physical composition can be conscious
> 2.) SETI-MR[145]: Some creatures that are significantly different from us in their physical composition can be conscious
> 3.) Standard-MR: Systems of indefinitely (perhaps infinitely) many physical

[144] 2002, 146-7.
[145] The acronym which stands for Search for Extra-Terrestrial Intelligence project.

compositions can be conscious
4.) Radical-MR: Every organized system, regardless of its physical composition, can be conscious

According to Polger, the Standard-MR story is hiding behind the mask of Radical-MR, or that in nearly all cases, Standard-MR will support Radical-MR.[146] And just as he is not willing to endorse such a wildly unconstrained thesis, neither should we, if we are to avoid the supposed dilemma above.

If we adopt Weak-MR however, not only does the constraining worry disappear, but there seems to be no necessity to drop type-physicalism either. As Polger puts it:

> "The relationship between mental states of a kind is merely that they are the same with respect to those properties characteristic of the kind…Weak-MR is therefore no threat to Identity Theory, for it is no problem if some creatures not physically identical to us have conscious states, as long as they share some physical properties in common with us" [2002, 150].

In other words, if we go back to Dan and Tom, they are two creatures not exactly alike in their physical composition, although they are pretty darn close, but nonetheless, they can have the same types of mental states and these mental states can still be identical to brain states. Think of it this way: Dan and Tom both represent the Rule of Ditransitivity. Dan's brain realizes this rule by the neural sequence D1…D2, and Tom's brain realizes it by T1…T2. Dan and Tom's brain states are thus different tokens of the same type, according to Weak-MR; that is, they are not identical physically, but they share similar properties. Therefore, the mental representations had by Dan and Tom, while identical to the brain sequences they are having, are also multiply realized because the mental states themselves are also tokens of the same mental type.

[146] Ibid, 148.

Forging this bond between type-identity and Weak-MR thus dissolves the problematic nature of positing their mutual exclusivity. We don't have to think about complex representations such as linguistic principles as being unshared, nor does the worry about allowing thermostats to represent the Rule of Ditransitivity creep into the picture. Furthermore, based on my picture of the type-identity thesis, we must allow for *at least some* innate representations. The moral of the story: ontology does matter.

With the consideration of ontology in mind, one further illustration might be helpful. The above discussion seems to suggest a need to revise what we mean by "representation," at least as it applies to nativism. Andy Clark suggests for example, that we need not think of representations as knowledge of the propositional sort. In a paper titled *Minimal Rationalism*,[147] he seeks to defend a weak version of nativism and illustrates this possibility by way of connectionist models which allow for minimally innately specified weights. When weights are pre-set in a network, they represent knowledge, but not in the usual "propositionally specifiable kind." Nevertheless, the pre-set weights on such a net are "in a real sense *informational*. It "knows" things which stop it from inducing certain conclusions (local minima) from the training data."[148] Assuming for the sake of argument that these nets are in some way analogous to human learning, we have a unique position which straddles rationalism and empiricism. The inductive process places the learning closer to empiricism, while the fact that the initial weight manipulation needed to avoid the minima is problem specific then places the learning back into the space of rationalism. In this "minimal rationalism," we can speak of representations as being innately specified, under the Samuels model of innate

[147] 1993
[148] Clark, 600.

specification of an entire organism. The problem has thus shifted from quibbling about "innate" to changing what we mean by "representation." If we accept Clark's position regarding the interpretation of "representation," it seems that we can take both a nativist and non-nativist stance on the acquisition of representations.

Conclusion: Minimal Rationalism + Language Acquisition = Minimal Nativism

Regarding language acquisition, under a minimally rationalist account, I suggest that connectionism can provide not only a working model of the way language *emerges*, but also of the way in which it is *innate*. As Clark suggests, if we have a model whose initial weights are set only strongly enough so that the system avoids local minima (error), but has no propositionally specified knowledge, it is clear that anything the network learns 'emerges' through interaction between it and outside stimuli. It is also the case however, that it is innately specified of the entire network that its weights be initially set, so we know ahead of time some things it will *not* do (choose a contradiction to the XOR problem, for example). Drawing an analogy to human learning, this seems to be at least part of what Chomsky and others have in mind when they insist that there are certain errors children will not make. Of course, the more robust sense of UG which claims that all the principles of grammar exist in the mind of the infant probably will not jibe with a Clarkian learning model. After all, the kind of information that is innately specified, he says, is non-propositional. Nevertheless, an outright rejection of nativism would appear to be hyperbolic.

In light of making commensurate what might otherwise be considered two contradictory positions on language acquisition, I conclude with two considerations.

First, the discussion by Elman, et al and Samuels specifically, and by non-nativists and nativists generally, presents a false dilemma with regards to "innate." and how it figures into the story about representations in the mind. The dilemma: either side with Elman et al and face the difficulties of responding to the compelling counterarguments, or side with the nativists and be forced to prove that mental states are multiply realizable, or even worse, as Samuels suggests, if they are not multiply realizable, still insist that representations are innate. It turns out that there is a third possibility, insofar as we can alter our conception of representation and thus, the dilemma disappears.

Second, in developing a position about language acquisition that represents both an amalgam of several theories and models, as well as a denial of several of their major tenets, I have shown in a sense, that *none* of these acquisition theories are correct or comprehensive on their own accord. While it is true that I have effectively taken what I assume to be the strongest points from each, this alternative explanation is indeed not identical to UG, UBL, nor Emergentism. This is not to say that I have therefore developed an entirely new and systematic theory for language acquisition, but simply, that I have forced the reader to reconsider the very terms in use among the dominant approaches. By doing so, it is clear that accepting or rejecting one over the other wholesale would be a mistake.

References

Ackerman H. & Hertrich, I. (2000). The contribution of the cerebellum to speech processing. *Journal of Neurolinguistics 13*, 95-116.

Ameka, F.K. (1995). The linguistic construction of space in Ewe. *Cognitive Linguistics 6*, 139-181.

Barlow, M. & Kemmer, S., eds. (2002). *Usage Based Models of Language*. New York: Cambridge University Press.

Bates, E., & Elman, J.L. (2000). The ontogeny and phylogeny of language: A neural network perspective. In S. Parker, J. Langer, & M. McKinney, (Eds.), *Biology, brains, and behavior: The evolution of human development* (pp. 89-130). Santa Fe: School of American Research Press.

Bates, E., Elman, J., Johnson, M., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1998). Innateness and Emergentism. In W. Bechtel & G. Graham (Eds.), *A Companion to Cognitive Science* (pp. 590-601). Oxford: Basil Blackwell.

Bechtel, W. & Graham, G., Eds. (1998). *A Companion to Cognitive Science.* Oxford: Basil Blackwell.

Bechtel, W. and Mundale, J. (1999). Multiple realizability revisited: linking cognitive and neural states. *Philosophy of Science*: *66*, 175-207.

Bickhard, M. & Terveen, L. (1995). *Foundational Issues in Artificial Intelligence and Cognitive Science: Impasse and Solution*. New York: Elsevier.

Block, N. and Fodor, J.A. (1972). What psychological states are not. *Philosophical Review*: *81*, 159-181.

Bloom, Paul. (2000) *How Children Learn the Meanings of Words*. Cambridge: MIT Press.

Butterworth, G. (1991). The ontogeny and phylogeny of joint visual attention. In A. Whitten (Ed.), *Cognition and the Development of Language.* New York: Wiley.

Callebaut, W. & Stotz, K. (2002). *Bioepistemology and the Challenge of Development and Sociality.* Cambridge: Cambridge University Press.

Cameron-Faulkner, T., Lieven, E., & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science* 27*,* 843-873.

Campbell, D. T. (1990). Levels of organization, downward causation, and the selectiontheory approach to evolutionary epistemology. In Greenberg, G., and E. Tobach, (Eds.),Theories of the Evolution of Knowing. Hillsdale, NJ: Erlbaum.

Carey, S. and Spelke, E. (1994). Domain-specific knowledge and conceptual change. In L. Hirschfield and S. Gelman (eds), *Mapping the Mind*. Cambridge University Press.

Chater, N. & Conkey, P. (1994). Sequence processing with recurrent neural networks. In G. D. A. Brown & M. Oaksford (Eds.), *Neurodynamics and Psychology* (pp.269-294). London Academic Press.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge: The MIT Press.

Chomsky, N. (1965). *Cartesian Linguistics*. New York: Harper and Row.

Chomsky, N. with Place, U.T. (2000). The Chomsky-Place Correspondence 1993-1994. *The Analysis of Verbal Behavior* 17 (2000): 7-38.

Chomsky, N. (1993). Language and Thought. Wakefield, RI: Moyer Bell.

Chomsky, N. (1987). *Language and Problems of Knowledge*. The Managua Lectures. Cambridge, MA and London: The MIT Press.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: The MIT Press.

Chomsky, N. (2000). *New Horizons on the Study of Language and Mind.* Cambridge, MA: Cambridge University Press.

Chomsky, N. (1959) A Review of Skinner's Verbal Behavior. *Language* 35 (1); 26-58.

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

Chomsky, N. (1982). Some concepts and consequences of the theory of government and binding. *Linguistic Inquiry Monograph Six*. Cambridge: The MIT Press.

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

Clark, A. (1993). *Associative Engines: Connectionism, Concepts and Representational Change.* Cambridge: MIT Press

Clark, A. The cognizer's innards: A psychological and philosophical perspective on the development of thought. *Mind and Language* 8 (4); 487-519.

Clark, A. (1993). Minimal rationalism. *Mind 102*, 587-610.

Corkum, V., & Moore, C. (1995). Development of joint visual attention in infants. In C. Moore & P. Dunham (Eds.), *Joint Attention: Its Origin and Role in Development.* Hillsdale, NJ: Erlbaum.

Croft, W. (2001). *Radical Construction Grammar: Syntactic Theory in Typological Perspective.* Oxford: Oxford University Press.

Curtiss, S. (1977). *Genie: A Psycholinguistic study of a modern day 'wild child.'* New York: Academic Press.

Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114-138). Cambridge: Cambridge University Press.

Edelman, G.M. (1988). *Topobiology: An Introduction to Molecular Embryology.* New York: Basic Books.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996) *Rethinking Innateness: A Connectionist Perspective on Development.* Cambridge: MIT Press.

Ellis, N. (1998). Emergentism, connectionism, and language learning. *Language Learning*: 48, 631-664.

Ellis, N. C., & Schmidt, R. (1997). Morphology and longer-distance dependencies: Laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition*, *19,* 145–171.

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

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA: MIT Press.

Essegbey, J. & Ameka, F. (2005). Serializing languages: Satellite-framed, Verb-framed or neither? *Proceedings of the 32nd Annual Conference on African Linguistics*, University of California, Berkeley.

Fillmore, C. J. (1989). Grammatical construction theory and the familiar dichotomies. *In* R. Dietrich & C.F. Graumann, eds. *Language processing in social context*. Amsterdam: North-Holland/Elsevier. (17-38).

Finch, S., & Chater, N. (1994). Learning syntactic categories: A statistical approach. In M. Oaksford & G.D.A. Brown (Eds.), *Neurodynamics and Psychology* (pp. 295–321). London: Academic Press.

Fodor, J. (1975). *The Language of Thought*. New York: Crowell.

Fodor, J.A. (1981). The present status of the innateness controversy. In *Representations.* Cambridge, MA: MIT Press.

Fodor, J. (1979) *Representations: Essays on the Foundations of Cognitive Science.* Cambridge, MA: MIT Press.

Fodor, J. & Pylyshyn, Z. (1997). Connectionism and cognitive architecture: A critical analysis. [Chapter] Haugeland, J. (Ed). (1997). *Mind design 2: Philosophy, Psychology, Artificial Intelligence.* (pp. 309-350). Cambridge, MA: MIT Press.

Fodor, J. D. & Crowther, C. (2002). Understanding stimulus poverty arguments. *Linguistic Review* 19 (1-2); 105-145.

Galloway, C. & Richards, B., eds. (1994). *Input and Interaction in Language Acquisition*. Cambridge, MA: Cambridge University Press.

Gardner, R. A. & Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science 165*, 664-672.

Gleitman, L. R., Newport, E. L., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language*, *11*, 43–79.

Goldberg, A. E. (2004). But do we need universal grammar? Comment on Lidz et al. *Cognition* 94, 77-8.

Goldberg, A. E. (1997) Construction Grammar. *In* E.K. Brown and J.E. Miller, eds. *Concise Encyclopedia of Syntactic Theories*. New York: Elsevier Science Limited.

Gopnik, M. (1997). *The Inheritance and Innateness of Grammars.* Oxford: Oxford University Press.

Gopnik, M. & Crago, M. (1991). Familial aggregation of a developmental language disorder. *Cognition* 39 (1); 1-50.

Gould, S.J. & Lewontin, R.C. (1979). The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptionist programme. *Proceedings of the Royal Society of London, 205,* 281-88.

Greenberg, G., & Tobach, E., Eds. (1990). *Theories of the Evolution of Knowing.* Hillsdale, NJ: Erlbaum

Grube, G.M., trans. (1980). Plato's *Meno*. Indianapolis: Hackett.

Heath, S.B. (1983). *Ways with Words: Language, Life, and Work in Communities and Classrooms.* Cambridge: Cambridge University Press.

Horgan, T. & Tienson, J. (1991). *Connectionism and the philosophy of mind*. Dordrecht: Kluwer.

Horgan, T. & Tienson, J. (1994). A nonclassical framework for cognitive science. *Synthese*, *101*, 305-45.

Hornstein, L. & Lightfoot, D., eds. (1981). *Explanation in Linguistics.* London: Longman.

Jablonka, E. (2001). The systems of inheritance. In S. Oyama, P.E. Griffiths, & R.D. Gray (Eds) *Cycles of Contingency: Developmental Systems and Evolution.* Cambridge, MA: MIT press.

Johnson, M.H. & Morton, J. (1991). *Biology and Cognitive Development: The Case of Face Recognition.* Oxford: Blackwell.

Johnson, S., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? The elicitation of gaze following in twelve-month olds. *Developmental Science 1,* (2), 233-38.

Karmiloff-Smith, A. (2002). *Beyond Modularity: A Developmental Perspective on Cognitive Science.* Cambridge: MIT Press.

Karmiloff-Smith, K., & Karmiloff-Smith, A. (2001). *Pathways to Language: From Fetus to Adolescence.* Cambridge: Harvard University Press.

Knauff, M., Mulack, T., Kassubek, J., Salih, H.R., & Greenlee, M.W. (2002). Spatial imagery in deductive reasoning: a functional MRI study. *Cognitive Brain Research 13*, 203-212.

Kuczaj, S.A. (1978). Why do children fail to overgeneralize the progressive inflection? *Journal of Child Language 5*, 167-71.

Landau, B., Smith, L., & Jones, S. (1992). Syntactic context and the shape bias in children's and adults' lexical learning. *Journal of Memory and Language*, *31*.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211–240.

Langacker, R. (2002). A dynamic usage based model. *In* M. Barlow & S. Kemmer., Eds. (2002). *Usage Based Models of Language*. New York: Cambridge University Press.

Lantolf, J.P. (2000). Second language learning as a mediated process. *Language Teaching 33,* 79-96.

Laurence, S. & Margolis, E. (2002). Radical concept nativism. *Cognition*, 86 (1); 25-55.

Lidz, J., Gleitman, H. & Gleitman, L. (2003). Understanding how input matters: Verb learning and the footprint of universal grammar. *Cognition* 87, 151-178.

Lidz, J., Gleitman, H., & Gleitman, L. (2004). Kidz in the 'hood: Syntactic bootstrapping and the mental lexicon. *In* Hall, D. Geoffrey and Waxman, S. R., eds. (2004). *Weaving a Lexicon.* Cambridge: MIT Press, 603-637.

Lidz, J. & Gleitman, L. (2004). Yes, we still need Universal Grammar. *Cognition* 94*,* 85-93.

Lidz, J. & Waxman, S. R. Reaffirming the poverty of the stimulus argument: A reply to the replies. *Cognition* 93, 157-165.

Lidz, J., Waxman, S., & Freedman, J. (2003). What infants know about syntax but couldn't have learned: experimental evidence for syntactic structure at 18 months. *Cognition* 89 (3); 65-73.

Lieven, E. V. M. (1994) Crosslinguistic and cross-cultural aspects of language addressed to children. In C. Galloway and B. J. Richards (eds.) *Input and Interaction in Language Acquisition.* Cambridge University Press.

Lieven, E., Behrens, H., Speares, J., & Tomasello, M. (2003). Early syntactic creativity: A usage-based approach. *Journal of Child Language* 30 (2); 333-367.

Mackey, A., Gass, S. & McDonough, K. (2000). How do learners perceive implicit negative feedback? *Studies in Second Language Acquisition 19,* 37-66.

MacWhinney, B., & Snow, C. (1990). The Child Language Data Exchange System. *ICAME Journal* 14, 3-25.

MacWhinney, B. (1995). *The CHILDES Project: Tools for analyzing talk*. (Second ed.). Hillsdale, NJ: Lawrence Erlbaum.

MacWhinney, B. (1998). Models of the emergence of language. *Annual Review of Psychology, 49,* 199-227.

MacWhinney, B. (1997). Second language acquisition and the competition model. In A.M.B. de Groot and J.F. Kroll (Eds.), *Tutorials in Bilingualism: Psycholinguistic Perspectives* (pp. 113–144). Hillsdale, NJ: Lawrence Erlbaum.

McCardle, P. & Hoff, E., eds. (2006). *Childhood Bilingualism*. Surrey, UK: Multilingual Matters.

Moll, H., & Tomasello, M. (2004). Twelve- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science, 7*(1), F1-F9.

Naigles, L. R., Fowler, A., & Helm, A. (1992). Developmental shifts in the construction of verb meanings. *Cognitive Development, 7,* 403–427.

Newport, E. L. (1977). Motherese: The speech of mothers to young children. In N. J. Castellan, D. B. Pisoni, & G.R. Potts, eds. *Cognitive Theory* (vol. 2). Hillsdale, NJ: LEA.

Nidditch, P. H., ed. (1975). Hume, David: *Enquiry Concerning Human Understanding.* Cambridge: Oxford University Press.

Norton, D. & Norton, M., eds. (2007) Hume, David: *A Treatise of Human Nature,* Vol. 1. Oxford: Clarendon Press.

O'Grady, W., Archibald, J., Aronoff, M., & Rees-Miller, J. (2005). *Contemporary Linguistics.* Bedford: St. Martin's.

Ohta, A.S. (2001). *Second Language Acquisition Processes in the Classroom: Learning Japanese.* Mahwah, NJ: Lawrence Erlbaum.

O'Leary, D. and Stanfield, B. (1989). Do cortical areas emerge from a protocortex? In M. Johnson (ed) *Brain Development and Cognition: A Reader.* Oxford: Basil Blackwell.

Oliver, R. (2000). Age differences in negotiation and feedback in classrooms and pairwork. *Studies in Second Language Acquisition 17,* 459-81.

Oyama, S., Griffiths, P.E., & Gray, R.D., Eds. (2001). *Cycles of Contingency: Developmental Systems and Evolution.* Cambridge, MA: MIT press.

Oyama, S. (2000). *Evolution's Eye: A Systems View of the Biology-Culture Divide.* Durham, NC: Duke University Press.

Parker, S., Langer, J., &. McKinney, M., Eds. (2000). *Biology, Brains, and Behavior: The Evolution of Human Development*. Santa Fe: School of American Research Press.

Perner, J. (1991). *Understanding the Representational Mind*. Cambridge, MA: Cambridge University Press.

Pine, J. M. (1994). The language of primary caregivers. In C. Gallaway & B. J. Richards (Eds.), *Input and interaction in language acquisition*. Cambridge: Cambridge University Press.

Pinker, S. (1994). *The Language Instinct.* New York: Harper Collins.

Pinker, S. (1999). *How the Mind works*. Cambridge: W. W. Norton & Company.

Pinker, S.,& Prince, A. (1988).On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition 29*, 195–247.

Polger, T. (2004). *Natural Minds.* Cambridge, MA: MIT Press.

Polger, T. (2002) Putnam's Intuition. *Philosophical Studies 109:* 143-170.

Posner, M.I & Raichle, M.E. (1994). *Images of Mind*. Scientific American Library.

Prinz, J. J. (2002) *Furnishing the Mind: Concepts and their Perceptual Basis*. Cambridge: MIT Press.

Pullum, G. K., & Scholz, B C. (2002). Empirical assessment of stimulus poverty arguments. *Linguistic Review* 19 (1-2); 19-50.

Putnam, H. (1967). Psychological predicates. In W.H. Capitan and D.D. Merrill (eds.), *Art, Mind, and Religion*. Pittsburgh: University of Pittsburgh Press, 37-48.

Putnam, H. (1988). *Representation and Reality*. Cambridge, MA: MIT Press.

Quartz, S.R. (1993) Neural networks, nativism, and the plausibility of constructivism. *Cognition* 48, 223-242.

Quartz, S. and Sejnowski, T. (1994). Beyond modularity: neural evidence for constructivist principles in development. *Behavioral and Brain Sciences:* 17.

Quine, W.V.O. (1960). *Word and Object*. Cambridge: MIT Press.

Ramsey, W. & Stich, S.P. (1990). Connectionism and three levels of nativism. *Synthese*, *82* (2), 177-205.

Rumelhart, D. & McClelland, J. (1987). *Parallel Distributed Processing* (Vol. 1). Cambridge: MIT press.

Rumelhart, D. & McClelland, J. (1986). On learning the past tense of English verbs. In J. McClelland & D. Rumelhart, (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition Vol 2: Psychological and Biological Models.* Cambridge, MA: MIT Press, 216-271.

Samet, J., Flanagan, O. (1989). Innate representations. In S. Silvers, ed. (1989). *Representation.* New York: Kluwer Academic.

Sampson, G. (1999). *Educating Eve: The Language Instinct Debate.* London: Continuum International.

Sampson, G. (2005.) *The Language Instinct Debate.* London: Continuum International.

Samuels, R. (1998). What brains won't tell us about the mind: a critique of the neurobiological argument against representational nativism. *Mind and Language* 13(4); 548-570.

Schieffelin, B. (1990). *The Give and Take of Everyday Life: Language Socialization of Kaluli Children*. Cambridge University Press, 1990.

Scholz, B. C., & Pullum, G. K. (2002). Searching for arguments to support linguistic nativism. *Linguistic Review* 19 (1-2); 185-223.

Shapiro, L. (2000). Multiple realizations. *Journal of Philosophy*: 97, 635-654.

Shieber, S. (1985). Evidence against the context-freeness of natural language. *Linguistics and Philosophy* 8: 333-343.

Simon, P.(1983).Cars are Cars. On *Hearts and Bones* [cassette].WEARecords Ltd.

Skinner, B.F. (1938). *The Behavior of Organisms: An Experimental Analysis of Behavior.* New York: Appleton-Century-Crofts.

Skinner, B.F. (1957). *Verbal Behavior*. New York: Appleton-Century-Crofts.

Snow, C.E. & Ferguson, C.A., eds. (1977). *Talking to Children: Language Input and Interaction.* Cambridge, U.K.: Cambridge University Press.

Sokolov, J.L & Snow, C.E. (1994). The changing role of negative evidence in theories of language development. In C. Galloway & B.R. Richards, (Eds.), *Input and Interaction in Language Acquisition.* Cambridge: Cambridge University Press.

Spelke, E. S. (2000). Nativism, empiricism, and the origins of knowledge. *In* D. Muir & A. Slater, eds. (2000). *Infant development: The essential readings.* Malden, MA: Blackwell Publishers.

Spelke, E. S., von Hofsten, C. & Feng, Q. (2001). Predictive reaching for occluded objects by 6-month-old infants. *Journal of Cognition and Development,* 2(3). 261-281.

Sternley, K. (1989). Fodor's nativism. *Philosophical Studies,* 55, 119-141.

Sur, M., Garraghty, P.E. and Roe, A.W. (1988). Experimentally induced visual projections into auditory thalamus and cortex. *Science*: 242, 1437-1441.

Tomasello, M. (2006). Acquiring linguistic constructions**.** *In* D. Kuhn & R. Siegler, eds., *Handbook of Child Psychology.* New York: Wiley.

Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition.* Cambridge: Harvard University Press.

Tomasello, M. (1992). *First Verbs: A Case Study of Grammatical Development.* Cambridge, MA: Cambridge University Press.

Tomasello, M. (2004). Language and Representation. PsycCRITIQUES. [np]. American Psychological Assn, US.

Tomasello, M. (2004). Syntax or Semantics? Response to Lidz et al. *Cognition* 93, 139-140.

Van Gulick, R. (1982). Mental Representation: A Functionalist View. *Pacific Philosophical Quarterly*, 3-20.

Vygotsky, L.S. (Cole, et al Eds.) (1978). *Mind and society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.

Waskan, J. & Bechtel, W. (1997). Directions in connectionist research: Tractable computations without syntactically structured representations. *Metaphilosophy* 28 (1-2), 31–62.

Waxman, S. & Markow, D. (1995). Words as invitations to form categories: Evidence from 12-to 13-month old infants. *Cognitive Psychology* 29 (3); 257-302.

Winkler, K., ed. (1996). Locke, John: *An Essay Concerning Human Understanding*. Indianapolis: Hackett.

Woodward A.L., Markman, E.M., & Fitzsimmons, C.M. (1994). Rapid word learning in 13- and 18-month olds. *Developmental psychology 30* (4) 553-66.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*: 358, 749-50.