2001

First words: An investigation of the nature of children's first word productions

Susan Cummings
The University of Montana

First Words: An Investigation of the Nature of Children's First Word Productions

By

Susan Cummings

M.S. Texas Woman's University 1976

B.S. Montana State University 1971

Presented in partial fulfillment of the requirements

For the degree of

Master of Arts

The University of Montana

2001

Dr. Tully Thibeau, Chair

Dr. David Strobel, Dean, Graduate School

5-18-01

Date

Cummings, Susan O., M.A., May 2001 Linguistics

First Words: An Investigation of the Nature of Children's First Word Productions

Director: Dr. Tully Thibeau

Several investigators (Markman 1996; Waxman and Kosowski 1990; Waxman 1994) have proposed that children look for object names first. Their studies reveal that more than 50% of children's first 50 words are object names. I disagree with this thesis. I claim that there is no predictable pattern to very early vocabularies. The patterns seen in vocabularies of 50 words cannot be extrapolated to children's very first words. Children may not initially be looking for object names at all when they begin to produce their first words. They are not initially matching meaning to sound: that is, the sounds they produce are not symbolic words. Rather, children begin talking by repeating frequently heard sounds that are emotionally salient and are socially reinforced. What is salient to one child is not to another. Hence, children's early vocabularies will be highly individualistic. Children may appear to attach a referential meaning to these first words because they say them in the correct context, but these first words are the result of social interaction and reinforcement. In a very short time period, certainly by the time they produce 50 words, the sounds can become symbolic representations – true words. This idea has support in the literature (Aslin et al. 1999; Bates et al. 1984; Gopnik et al. 1995; Gopnik 1981). However, there is other research (Bates et al. 1979; Snow 1981, 1999; Hirsh-Pasek et al. 1999) that suggests a different explanation for early word production than an innate search for object names. A study was undertaken of the early vocabularies of 46 infants, primarily in the 5-word to 25-word vocabulary range. Statistical analysis demonstrates that the percentage of words appearing in word categories (such as "noun") is highly variable and cannot be predicted. Results do not support an innate search for object names in initial word production. Further, early vocabularies can best be described as transitional. A statistically significant increase in the percentage of nouns and a decrease in the percentage of words with social or emotional content were found as vocabulary size increased.

Acknowledgements

This thesis would not have been possible without the encouragement and help of many people. I would like to thank Dr. Tully Thibeau for chairing my committee and being a patient guide, helpful critic, and assistant throughout the process. I thank Dr. Irene Appelbaum and Dr. Wendy Shields for their support, help, and participation. Thanks go to my husband, Steve Cummings, for his patience and occasional helpful suggestions, and to Scott Jones for statistical analysis of the data. I thank the many parents of young children who provided data for my vocabulary analyses and shared with me the wonders of watching infants learning language. Thanks also go to the many researchers who took time from their busy schedules to respond to my e-mail queries, sometimes by sending their bibliographies and answers to my questions by e-mail, sometimes by mailing me drafts and documents, always responding with respect and encouragement.

I would especially like to thank Dr. Nancy Mattina for her encouragement and for convincing me that I could, indeed, begin a new career. Last, but most certainly not least, I owe Dr. Deirdre Black a great debt of gratitude for her encouragement and for the inspiration for this project.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.0 What a Language Acquisition Theory Must Provide

The focus of this thesis is how children begin to use words and what those first utterances might mean to them. Most theories of language acquisition have focused on the acquisition of syntax, but language is not possible without words. In order to understand how children identify part of the noise around them as linguistically significant and come to be language users themselves, one must include a study of the beginnings of word production. Any theory about language acquisition should include an explanation of how children come to produce their first words because the way language begins is related to cognitive development and can affect how it develops into a grammar. An explanation of language learning, then, must describe what words are and how the child understands them. Is the child actively looking for words? Does reinforcement play a role in the storage and production of words? Are first words linked to innate concepts? What is it that children have learned when we say they have learned words? What motivation exists to get the process started? This paper will look for meaningful answers to these questions, focusing on the nature and source of children's early word production.

1.1 Do Children Actively Search for and Attach Meaning to Early Words?

Nativist theories hold that certain fundamental aspects of knowledge are innate. Experience provides triggers to fill in preformed or latent categories. Grammar is treated as an organ in the brain much like other organs in the body. It is said to unfold on a maturational schedule that is governed by genetics. Even in a nativist theory, however, there must be a starting point in the language acquisition process. Some researchers (Markman 1996; Waxman et al. 1990; Waxman 1994) have proposed that children come to language learning with an innate expectation that words will be object names. If this is true, then children's first words must represent a pre-existing linkage between categories of objects and nouns. This supposes that the child matches concepts to words and that this ability, as well as the innate concepts, are both part of the child's innate endowment.

Other researchers argue that children do not start with innate concepts. They argue that not all early words are nouns and that children are not predisposed to look for object labels (Aslin et al. 1999; Bates et al. 1984; Gopnik et al. 1995; Gopnik et al. 1986; Gopnik 1981). Their conclusion has been that linguistic development is grounded in specific types of general cognitive development. Children begin by saying words and this production of words helps concepts develop. This idea has been reinforced by the use of computer simulations showing that certain kinds of these models seem to mimic the gradual acquisition of both word meaning and grammar that has been observed in children (Elman 1999; Plunkett 1995). Termed "emergentist," this view holds that children don't actively search for words but rather that word meanings gradually emerge.

Through a series of experiments with young children, Vygotsky (in Rieber et al. 1987) showed that children appear to talk to themselves only *after* they begin to converse with others. He believed this was evidence that children are first of all social. They develop reflective thinking only after they engage in social interaction. Vygotsky showed that, while adults and children use the same words to refer to the same thing, children sort objects differently than adults and in different ways at different ages. For example, very young children perform a sorting task by putting objects into piles with no ordering principle at all. They progress to sorting by some connection between objects but may use *any* connection at all and sometimes even change what associations they use from one object to the next during the sorting task. He interpreted this to mean that developing or primitive concepts are better described as complexes of features that do not become real concepts until later stages. If Vygotsky is correct, early speech is primarily used for social purposes and the concepts behind early words are not representative of innate concepts since those concepts do not exist until much later in life.

Vygotsky did not think the ideas behind early concepts or complexes, as he called them, were innate because early sorting indicated that *any* resemblance between objects might be used to associate them. If this is true, then perhaps words themselves are stored differently in a child's brain than in an adult's and with different kinds of interconnections. There is recent evidence that this is physically the case. Neville et al. (1992) studied event-related potentials (ERP) in the brain and report that elicitation of words from different semantic classes results in different electrical activity. McDonald (1997) reports that function words are processed in different parts of the brain depending on the age of the child. Young children reportedly process both function words and content words bilaterally, while adults process function words largely in the left-hemisphere anterior temporal regions where processing is more rapid. This process gradually changes as the child grows older, not reaching adult levels until the child is around 15 or 16 years old. Such neurological studies support Vygotsky's conclusion that words are processed differently in the brains of young children than in the brains of older children and adults. While not proof that of changing conceptual development, this evidence is suggestive of the kinds of changes in processing that has been demonstrated in children's behavior.

1.2 What Motivates Children to Produce First Words?

What is it about language that children find so interesting? Why are they so eager to participate in the process? Proof that humans are physically capable of speaking does not explain why they do so. A computer program may be shown to be capable of learning grammar but it seems intuitive that this does not explain the acquisition of language by children because children are not machines that always cooperate with the programmer. Artificial neural networks might be able to mimic word learning, but the programs learn words because the programmer designs the program to do so. If nature has so obviously endowed humans with the ability to acquire language, there should be some reason for it. Is the desire to speak the result of an instinct, a language module in the brain, or a non-linguistic social capacity designed for survival? I suggest that children are not innately predisposed to link certain kinds of words to innately defined categories but rather that the desire to speak is primarily social, which provides the motivation for children to learn to communicate with words.

The problems of how children link meaning to form and what motivates them to begin to speak are difficult to answer. There is a lack of consensus even among people who otherwise generally agree with each other. I believe a multi-disciplinary approach, incorporating background from psychology and anthropology as well as linguistic, may provide answers that have eluded other approaches. This paper examines what theorists currently propose as mechanisms for children's production of earliest words. It proposes a primarily social alternative theory suggesting that first words do not reflect an innate search for object labels but are a social tool to elicit parental response. First words are possible because of both linguistic and non-linguistic cognitive development and an innate desire to interact socially. Infant communication gradually changes from being just social sounds to using sounds in a context, to sounds acquiring meaning as "events." Words begin to acquire real meaning when one word is paired with one context – an indication that a word has acquired indexical meaning. That is, the word is an index to the context in the way the doorbell indicates the presence of a visitor. The child discovers that a word can actually be used in a variety of situations and words begin to acquire, first extended indexical and, finally, symbolic meanings as the child realizes that a word actually represents a *category* of objects, actions, or attributes. The process of acquiring words can be described, then, not as an instinctive search for object labels but as a continuum from a form of social interaction to more exclusively symbolic use of words for communication.

Further, this paper presents a statistical analysis of the production vocabularies of children at various vocabulary levels and suggests that it reveals that the content of early vocabularies is not predictable. Several authors have investigated the vocabularies of babies who know about 50 words. Although babies' personalities can affect the distribution of words in their vocabularies and there are certain universals reflecting children's needs, by the time a child has learned 50 words, the effects of their culture and language appear to be reflected in the proportions of major word categories. Boysson-Bardies (1999) observes that certain characteristics of each language affect what categories of words predominate in children's vocabularies. She says that cultural attitudes influence how parents approach their children linguistically and these same attitudes affect how the children respond. For instance, she says, the types of words in French babies' early vocabularies tend to be words for food, clothing, and "agreeable activities or states." American children name the people around them and use surprisingly few verbs and adjectives. Japanese children use more social terms such as politeness terms and adverbs. My study of children's first few words, however, does not show a stable early pattern reflecting innately defined categories but instead shows that early vocabularies are in a transition period in which words are just beginning to acquire linguistic meaning.

Human language acquisition is immensely complex and most previous studies have focused on small areas but not overall solutions. Providing a complex answer to the question of how children learn words is difficult but I believe there has been sufficient research to make it possible to address. It would be of general interest and considerable benefit to linguistic research if this paper results in a useful consolidation of previous research. If we understand how words are acquired, we will have a better foundation for understanding what language is and where it came from. This paper will provide evidence that children's first words are not the result of an innate search for object labels but, also, it may lead to a better overall understanding of how the process of language acquisition begins and suggest possible directions for research.

1.3 Definitions

The items described here are concepts that are important to the current discussion and an understanding of them as they apply to the current topic is essential. Terminology often contributes significantly to misunderstanding. Many of the more controversial aspects of language acquisition are the result of investigators defining terms in different ways. A discussion of how first words are acquired will depend on standardizing the important terminology. The following discussion of the meanings of these terms may not solve the problem but should provide some resolution as well as an introduction to the difficulties involved in understanding the overall problem.

1.3.1 Innateness

Of all the concepts discussed here, *innateness* is probably the most controversial and difficult to define. A simple definition is that anything that is inborn is innate. In reality, the concept is much more complex. Knowledge you are born with rather than acquired from learning can be called innate (Payne and Wenger 1998). This in-born kind of knowledge is exactly the kind of innateness that early Greek philosophers believed in. Socrates believed that all knowledge was innate. All truth and knowledge, he claimed, was present in the mind and only needed the proper educational environment to emerge fully formed.

Today we try to explain the world in scientific terms. Innateness has come to mean "something present in the genome" – something that nature has endowed us with and that we acquired by the processes of natural selection. The problem is that genetic effects are not usually direct but rather the result of an interaction of many genes as well as interaction with the environment. Essentially everything about us from how we grow toes to how we learn our names is a complex interaction between heredity and environment. Hence, even defining innateness in terms of genetics doesn't provide a clear answer.

Further confusion about innateness stems from the fact that there are different ways to be innate. Elman (1999) outlines three levels of innateness: *representational,* *architectural, and chronotopic*. *Representational* innateness means that the way the synapses between neurons connect with each other is predetermined. This "hard-wired" kind of innateness would result in children being born with the kind of innate knowledge that Socrates talked about. This thesis discusses the use of artificial neural networks as models of the brain. In a neural network, representational innateness is the equivalent to presetting the weights between connections. In other words, the answers to certain questions are predetermined.

Architectural innateness operates on a higher level than representational. It refers to three kinds of limitations on the physical structure of the brain. First, it refers to the types of neurons and how they use chemicals to transfer information. Second, it refers to the density of cells, what kinds of cells are present, and where they are located. Finally, it refers to the way the various pieces of the system are connected together. In architectural innateness, knowledge is not innate but the overall structure of the system determines what kinds of information can be received, what kinds of problems can be solved, and how information will be stored (Elman 1999; Elman et al. 1996).

The third kind of innateness results from developmental timing or *chronotopic* constraints. For all living things, the timing of cell division is critical for the development of a mature entity. In mammals, most of the important developmental phenomena are the result of a complex interaction of internal and external events. Small changes in a developmental schedule may result in very large changes in outcome. In our brains, direct genetic control of timing means that the onset and sequencing of developing neurons and their connections proceed according to a preset schedule. This effect has been used as an explanation for "critical period" events. The assumption is that the functions of regions in the cortex are, early on, plastic and adaptable. Later, these regions become specialized and lose their ability to perform anything other than their assigned task.

There are different ways to be innate and each one is affected by complex interactions with both internal and external environments. When linguists assert that language is innate, they usually are referring to the representational or "hard-wired" kind of innateness. No one denies that something about language acquisition is innate but some researchers believe that children's first words are linked to representationally innate concepts. I will show there is evidence that only the timing and architecture have been provided by our genome.

1.3.2 Sensation, Perception, and Cognition

One definition of *perception* is the processes that create internal representations of objects and events in the environment (Payne and Wenger 1998). In this view, perception is the link between sensation and cognition. There are no clear boundaries between these processes. Items in the environment may be *sensed* but must stand out from background noise in order to be *perceived.* They must be identified and recognized in order to enter *cognitive* processes. Current theory explains the process by postulating iconic (visual) and echoic (auditory) stores for information in the environment to be collected. These storage areas hold information - *all* the information that is sensed.[1] Sensed information is filtered by the central nervous system before it can be perceived.

[1] These areas have been shown to hold much more than a person ever becomes consciously aware of but they contain that information for a very short time – on the order of a quarter of a second for visual and up to four seconds for auditory stores (Payne and Wenger 1998).

What a person consciously hears and cognitively processes depends on surrounding background noise, distractions, expectations, the actual acoustic properties of the sound, and the content of the sound. Once perceived, cognitive processes change, store, recover and use information. Imagery, recall, problem solving, and thinking are all terms that apply to cognition. When we process language, the brain must use cognitive processes to attend to several images or ideas at the same time and be able to associate between them (Deacon 1997).

Put another way, sensation depends on "bottom-up" or "data-driven" processes (Payne and Wenger 1998). The data are in the environment. The task is to collect information about them. Cognition depends on "top-down" or "conceptually-driven" processes in which the brain uses what it already knows to process the input. In between is the process of perception. Top-down and bottom-up processes are both used in perception to recognize what has been seen or heard and match it to what is already known. Language makes demands on the system at all levels of processing and infants must begin without all the information they need initially in place. Linguistic communication requires analysis and production of speech sounds, the eventual recognition of thousands of vocabulary units, and the use of an intricate system of grammar rules (Deacon 1997). Because sensation, perception, and cognition are all affected by genetics and environment and are subject to different kinds of learning, it is important to understand what each one is and how they are different in order to understand how these terms apply to language acquisition. Specifically, sensation is important in language acquisition because we must be able to hear or see in order to enter input into the process. Perception of speech sounds is necessary for the data to be properly accessed. However, it is human cognition and the matching of bottom-up input with top-down internal representations and interconnections of certain sounds or signs that makes language possible for humans. This paper asserts that linguistic sensation and perception are architecturally innate. Word categories that sounds may be mapped to are not representationally innate but the structure of the system may limit what is attended to. Early words are not mapped to innate concepts and human cognition of word meaning develops gradually as the result of an interaction between innate mental structure and learning.

1.3.3 Concept

Concepts, in psychology, are mental representations that group together sets of objects or events (Payne and Wenger 1998). When one sorts a deck of cards with all the hearts in one pile and all the clubs in another pile, we assume that the concept of *suit* has been used to make the categorization. While we can observe the categorization, we cannot observe the concept. This is why Vygotsky was able to claim that young children do not form true concepts. He observed that very young children put things into piles in random order. As their cognitive abilities develop, they begin to sort things but not, according to Vygotsky's interpretation, by using a true concept. He said that any kind of association between objects might be used to put them into piles. He called these associations *complexes* rather than concepts (Vygotsky in Rieber and Carton 1987). He said that for children, an object is only a member of a collection. Each member is related to at least one other member because of its appearance or thematic relationship but may not be related to the original member in the same way and is not based on any abstract idea. While it may *appear* that there is a concept involved, early words represent a concrete relationship, not a symbolic mapping of objects with concepts.

In contrast to Vygotsky, other researchers believe that children begin forming concepts in infancy. Their concepts are different from those of adults and these differences develop gradually. Nelson (1977a,b) describes a *concept-generating process* that begins in infancy. She says that a concept contains a *functional core* that varies from individual to individual, whether child or adult, and which relates an object to the individual, not by its perceptual features but by its place in his or her life experience. Hence, both child and adult may use the same *identificational* features for an item, but the functional cores may differ. Nelson's definition of concept is that it is a collection of features, not an image built up of those features. For example, images of birds might be associated with the concept of 'birdiness' but a single image is not the concept. An adult concept of what a bird is includes the ability to fly and the possession of feathers (Markowitz 1988). The child's problem is then an ongoing one of identifying those features that differentiate one concept from related ones. A concept is neither a concrete and final goal nor as continuous as daily experience but something in-between. Concepts are strongly related to a first occurrence but modified by related episodes. While Vygotsky identified stages of concept development and identified early concepts as complexes, Nelson describes the process as a continuum from early concepts with minimal functional cores to adult concepts with more complex functional cores. For Nelson, this process begins with primitive concepts in early infancy, progresses gradually to adult concepts, and never ends.

Further complicating the issue of a concept is that some researchers have suggested different levels of concepts. Savage-Rumbaugh et al. (1998) refer to a Cartesian hierarchy of concept levels. Each concept is dependent on pre-existing subconcepts. For example, the concept of punishment is composed of at least the subconcepts of agency, intention, and responsibility. Partial or minimal concepts such as Nelson (1977a,b) described are not allowed in this theory. If any of the subconcepts are missing, the whole higher concept is missing. The theory also suggests discontinuities or jumps in concept formation. In this view, higher concepts are complex, cannot be mapped to a single object, and are not possible unless all the lower concepts are in place.

Premack (1984) suggests that language itself may allow the development of higher-level concepts. Certain kinds of concepts are not possible prior to the acquisition of language. He showed that chimpanzees normally solve problems on a sensory basis. They understand that one apple is "the same as" another apple. When he used plastic tokens to represent words and taught them some basic sentence structures, they were then able to also match half an apple to half a glass of water illustrating that they understood the concept, "half of something." Chimpanzees that were not trained to use the word tokens failed this kind of conceptual task. Pre-linguistic sensory concepts may be lower level kinds of concepts or may be qualitatively different than linguistic concepts.

Further, in her investigation of category formation, Markman's (1996) results seem to agree with Vygotsky's observation that older children (6 and 7 year olds) sort by taxonomic category but younger children may sort objects by a variety of causal, temporal, or spatial relationships. Waxman and Markman (1990), Waxman (1994), and Smith (1996), however, found that even children as young as 2 years old will sort taxonomically if given a label for one of the items to be sorted. Hence, while a number of studies provide evidence that there that primitive concepts or complexes, as Vygotsky termed them, may form with or without language but that there also may be something special about language that encourages true concept formation, perhaps at a higher level.

There is considerable disagreement between researchers over what concepts children have and when children acquire them. There are probably different kinds or levels of concepts and there may be a continuum between early or primitive concepts and mature or adult concepts. Researchers disagree as to whether some concepts are representationally innate or just acquired very early. Some kinds of linguistic concepts appear to develop after or at the time words are learned so they are probably not innate in the representational sense. This is important because it means it is possible that concepts behind word meanings are not present before the words are learned. This should be testable. If linguistic concepts were acquired only after words begin to be produced, first words would be a reflection of environmental factors rather than innate predispositions.

1.3.4 Categories

Categorization is a cognitive process (Payne and Wenger 1998). All that is meant by general categorization is that a set of objects or events is divided into at least two groups. Categorizing allows us to simplify and order the world. We are able to identify and respond to new objects and events more easily if we can associate them with categories we already know. We can objectively observe and measure categorization when it occurs but we can never know for sure what, if any, concepts are used to make the categorization. Hence, any inference that an innate mapping from concept to object has been made is more intuitive than empirical.

This definition of categories is neat and tidy but categories don't always cooperate in a neat and tidy fashion. For some things in the world, it is easy to identify and define a category - no one would argue about whether or not a particular triangle belongs in the category *triangle*. For most things, however, categories are fuzzier. Almost any category in the real world will contain some items that are more typical members than other items. A common example is the category of birds. Robins are almost always considered typical birds but bats can be included in the category as well because they have wings and fly (Markowitz 1988). In this sense, the members of a category exhibit a "family resemblance." There is usually a continuum between what is definitely included in the set and what is definitely outside it (Givón 1999). Further, categories are defined differently in different cultures (Lakoff 1987). Hence, it is unlikely that word categories could be innate.

It is important to understand what categorization is and what it isn't. Categorization of sounds in the environment is important in the acquisition of words because it allows children to separate out meaningful bits. Categorization is one of a number of cognitive processes that humans need in order to survive because we have to know what sorts of things are good to eat, what are dangerous, and what in the world gets matched up with what other things. The ability to categorize presages the development of the awareness of cause and effect. It allows children to learn that the appearance of one kind of thing predicts the appearance of a related kind of thing (Payne and Wenger 1998).

1.3.5 Learning

Learning means that a change occurs as a consequence of experience. At the neural level, this refers to a change in the weighting of neural excitation (Payne and Wenger 1998). Neurons support learning in three ways: the amount of neurotransmitter may increase or decrease, the synapses between axons may increase or decrease, and new synapses may form. Learning is restricted by the capacity of memory and by time. Learning at the neural level is enhanced and reinforced by repetition but may be lost or inhibited by an increase in time between episodes of exposure. This kind of learning has been simulated in neural networks (Payne and Wenger 1998; Rumelhart and McClelland 1993; Elman et al. 1996). Weights on connections between artificial neurons are strengthened or weakened a small amount each time the network attempts to solve a problem. The result is that the network gradually becomes able to solve certain kinds of problems. When the network is able to arrive at correct answers for the given problems, it is said to have "learned" how to solve these problems.

At a higher level, psychologists recognize three stages in the acquisition of a learned skill (Payne and Wenger 1998). First is the *cognitive stage*, a time when basic facts are consciously memorized and rehearsed. The detailed information or *what* is required to perform the task is important. Second is the *associative stage*, a time when errors in performance are recognized and connections or associations are strengthened. In the associative stage, *procedural knowledge* becomes important as depth of processing increases. That is, *how* the task is to be performed is more important to the learner than *what* is performed. The learner is integrating the stages of the task and connections between the elements of the task become strengthened. Finally, the skill becomes mastered as a routine and automatic in the *autonomous stage*. Performance speeds up and the task can be performed even when actually concentrating on something else. This process of learning in stages applies to many kinds of skills including driving a car, playing a piano, and touch-typing.

Language acquisition is usually considered a different kind of learning than skill learning. Bowerman (1993) differentiates the way that language is learned from other, non-verbal learning by saying that languages use different criteria for classifying referents than non-linguistic tasks. Children must learn a different level of organization in *language learning*. Although many researchers appear to equate the two, there is a difference between *language learning* or *acquisition* and *word learning*. *Word learning* is not equated in this paper with *language learning* or *language acquisition*. *Word learning* refers to the ability of a child to produce and use a word in an appropriate context and does not necessarily reflect linguistic understanding. Early word production is overt evidence of learning in the sense that a change has occurred in a child's brain. While a learning bias may exist, an innate general attentional bias could explain the origins and mechanisms of initial *word learning* (Smith 1996). The fact that the regular pairing of one cue with a second cue will cause the first cue to predict the second is a well-documented property of general learning. Smith maintains that initial *word-learning* biases are not linguistic but made out of the same general associative and attentional processes that are typical of other forms of skill learning.

Because some researchers claim the ability to learn language is an instinct or develops as part of the maturational process (Pinker 1994; Chomsky 1975) it is important to differentiate the learning of words from learning *how* to learn words. The way children learn their first words (that is, learn to *produce* words) is a slow process that beginns between nine to twelve months of age. The process picks up speed a few months later so that by the time a child is three years old, most normal children can say several hundred words (Bates and Goodman 1999; Golinkoff and Hirsh-Pasek 2000; Fenson et al. 1994). This process of learning words seems to follow the pattern of skill acquisition. The *way* that children learn words closely resembles that of classical skill learning as described by psychologists. That is, in the beginning, children learn words very slowly and deliberately as in the *cognitive* stage. Next, while errors may occur, words seem to be analyzed while connections and *associations* between them become strengthened. Finally, word acquisition becomes rapid and automatic or *autonomous*. The process speeds up – children acquire vocabulary rapidly after the age of about two years. Hence, word acquisition may be a skill that is learned rather than innately predetermined, in the sense of representational innateness, even though the architecture of the brain provides the structure necessary for it to occur.

1.3.6 Gesture, Icon, Index, Symbol, Sign

Gestures can be meaningful but are not arbitrary the way words are. There is a wide range of innate gestures that humans produce. These are related to the calls and gestures of other animals. Our "gesture-call" system includes most of our nonverbal communication. It is unlikely that human language grew out of this system because gestures and calls are graded. They have no grammar, cannot be broken down into smaller parts, are usually not learned, and are pretty much the same from culture to culture regardless of language. What this means is that, while a "V" for victory is an arbitrary symbol for something that has nothing to do with fingers, a yawn is universal and usually uncontrollable. A nod and a headshake are both considered meaningful, language-like behavior but they are distinct, with no halfway point, the way there is between a laugh and a giggle (Burling 1993). Although reference may be made to *symbolic* gestures, the default meaning of the word *gesture* is simply a sound or movement that does not have arbitrary meaning.

Icons are essentially pictures of things. *Iconic reference* is defined as a relationship by resemblance. Iconicity is the quality of an image that causes the viewer to recognize, or more exactly, to re-cognize that input image. A picture of a thing and pictograms are iconic. Iconicity generates recognition as in the way a picture or even a caricature of something generates the recognition of the object portrayed (Deacon 1997). A stimulus has iconic reference when nothing more than physical similarity is involved in the reference. Icons have identificational features that indicate similarity but it is not necessary that functional features be noted and a true concept of the thing is not necessarily formed when pre-linguistic iconic reference is recognized. Iconicity forms the foundation of a continuum of meaning because it is the most basic way that something can be re-presented (Deacon 1997). It is the bottom step of the hierarchy of representational process and provides a basis for the development of the next step, indexical relationship.

While iconicity is acquired through perceptual similarity, *index* is defined as relationship by participation (Deacon 1997). It indicates the presence of a thing the way that smoke indicates fire. A child may associate mother putting on her coat with mother leaving and, hence, begin to cry. Indexical reference is a natural function of cognition. Deacon suggests that it is dependent upon iconic reference and that indexical relationships are composed of iconic relationships between sets of icons. Indexical meanings are acquired by a perception of contiguity or correlation (Bates et al. 1979). A word, iconically associated with other occurrences of the sound, becomes able to call to mind an object, iconically associated with other objects in past experience. Golinkoff et al. (1999) call words that have indexical reference "context-bound." These words may seem similar to symbols but are still bound to their referent. Indexes are, however, very important in the development of symbolic meaning, so much so that Deacon (1997) says that symbolic meaning can only be formed after relationships between indices are formed.

A *symbol* is usually defined as something that stands for or represents something else. A word is said to be a symbol because it is an arbitrary sound that stands for a concept. But this is what I just defined as an *index*. Symbolic reference differs because it is relationship by a socially agreed-upon convention and is arbitrary and not causally connected to a referent. Rather than just associating a word with a situation or an object, symbolic meaning allows multiple contexts, multiple situations, and relationships with other words. A symbol is not isolated from other symbols but represents a set of interrelated concepts rather than a relationship with concrete reality.

Deacon (1997) further explains that, while the ability to use a word in a variety of contexts is evidence of symbolic understanding, this can occur with the use of indices as well. There are two things that critically define the difference between an index and a symbol. One is that there is a correlation in time and place between an index and its referent. If a smoke-like smell persists in the absence of anything burning, it will lose its relationship to fire. Symbolic reference persists regardless of any such correlation. The word "smoke," used as a symbol, retains its meaning even when used repeatedly in the absence of smoke. Second, while indices do not affect each other, words are related to other words. Symbols reveal linguistic knowledge because they not only map to the physical world, they can map onto other symbols independently of their physical referents. Deacon uses the example of an animal trained to associate a number of different words with different foods. Each word is associated with a particular food and if one of these word-object associations breaks down or is extinguished, it will have little effect on other associations. Words as symbols, however, can map onto other words. If one of their meanings breaks down, it affects the reference of other words (as when the word *bad* takes on the new meaning, *very good).*

The word *sign* has been used in a variety of contexts. Sign languages are systems of symbolic gestures and, in this sense, a sign is a symbol. There are other meanings for the word, among them the definition that a sign suggests the presence or existence of something not immediately evident. In this sense a sign can be an index.

Because these terms are used in many discussions about word learning, it is important to understand what they mean. As noted above, many people define *symbol* the way we have defined *index*. The two are easy to confuse and, while we speak of children's first words, what we really mean are children's first signs. As noted above and in the following graphic summary, signs may be either symbolic or indexical.

Summary of Referent Definitions	
Icon -	Has one facet of meaning
Index -	Has all facets of meaning in one situation
Symbol -	Has all facets of meaning in any situation
Sign -	A symbol or index

1.3.7 Word

Whether one argues for or against the premise that words are innately linked to categories when children begin to speak, one must have a clear notion of what a *word* is. The term is used indiscriminately by virtually anyone who ever talks or writes about language. Technically, the term *word* can be used in several different senses. Orthographically, a *spelling word* is separated by white space from the rest of a sentence. A *lexical word* is a dictionary item, something you expect to find a separate entry for in a dictionary. A *grammatical word* is the form that appears in a particular syntactic context. When analyzed acoustically, words have no "white space" between them yet, psychologically, we hear them separated (Finegan 1999). The first task of an infant, then, is to separate this continuous language stream into units. These units are usually what we are defining as *words,* but may also be unanalyzed phrases. For example, such phrases as *what's that, want to*, and *have to* are recognized in early speech analysis as single morphemes or meaningful units. No infant needs a definition of *word* to accomplish this and everyone seems to have an intuitive feeling about what is or is not a *word* but no one has a single definition that works in all cases.

Further complicating the issue is the fact that words are constructed differently in different languages. In languages like Chinese and English, words carry relatively few grammatical markers. In polysynthetic languages such as Greenlandic, morphemes may be more important than words (Fortescue and Olsen 1992). In these languages, there may be a single-morpheme stage rather than a single-word stage. It is not surprising, then, that the age of the production of a child's first word is controversial partly because of a disagreement on what constitutes a *word* (Darly and Wintz 1961). Darly and Wintz

(1961) found that the average age of a normal child's first word could be anywhere from 9 to 60 months depending on how the findings were determined and how the term *word* was defined. A study done by having fathers fill out a questionnaire resulted in a median acquisition age of first word as 15.8 months. Experimental observation of infants by another investigator presented the mode as approximately 10 months. Mothers' observations in another study resulted in a median age of 9.8 months. The definition of the term *word* undoubtedly caused a large portion of the variation. Parents so eagerly anticipate the event that any early babbling that happens to occur in the presence of appropriate persons, things, or actions is apt to be considered a "first word." Darly and Wintz (1961) advise caution in interpreting data from parents.

On the other hand, children's early words are not likely to sound much like their parents' words (Darly and Wintz 1961). It is often the case that what parents hear as understandable and meaningful words cannot be understood by anyone else. They developed a working definition of a *word* as *a sound uttered with full consciousness of its meaning and for the purpose of communication.* Recall from the discussion of symbols and indices that this definition of *word* does not require symbolic meaning. It does, however, depend on an observer being able to determine when it happens. Data have actually been collected from parents that seem to be consistent and match research observation. In the course of the development of the MacArthur Communicative Development Inventories, Fenson (2000) found that parental inventories compare well with researchers' observations. Parent reports not only correlate well with observation but also are internally consistent. It would seem that parents and researchers agree on when a child is first using words. These questionnaires define a *produced word* as a sound that a child can use and probably understands. The investigators report, however, that while parents and expert observers generally agree on the definition of the term *word* and reliably report the same early words for children, neither group can always tell when a word is being used meaningfully, when it is merely a repeated sound, or when it is a random sound that sounds like an adult word. What this means is that there is evidence to support the use of parents' reports to determine infants' first words and that this same evidence suggests that the infant's first words may not necessarily have linguistic meaning. In this paper, I will adjust Darly and Wintz's (1961) definition of a *word* to be a sound that appears to have some kind of meaning and is used for communication. In this sense, the term *word* will be considered equivalent to *sign,* which can be either a symbol or index, as earlier defined.

1.3.8 Emergence

Early word production has been described as evidence that a change has occurred in a child's brain. When things change, we look for a cause. Change may occur because something in the environment caused it or it may change because of some internal event or some combination of the two. An outcome is said to *emerge* when it arises for reasons not obvious or predictable from the inputs, either internal or external. Emergence is the result of the interaction between factors. Lacking a theoretical framework, *emergence* had, until recently, been regarded as vague, even mysterious. Something new seemed to arise out of nowhere. Anything without an explanation could be said to "emerge." Recently, however, *emergentism* has benefited from advances in biology, genetics, embryology, brain development studies, and cognitive neuroscience. Science has made great progress toward understanding how genes interact to produce different outcomes and about the plasticity of the developing brain. In addition, advances have been made in the field of computer modeling that have helped to understand mathematically how the interaction of many seemingly simple factors can produce unexpectedly complex results (Elman et al. 1996).

While nativists believe that fundamental aspects of knowledge are inborn, empiricists hold that all knowledge originates in the environment and comes into the brain through the senses. Emergentists do not completely reject either nativism or empiricism. Rather, they conceive of the two older theories as incomplete and consider emergentism a more complete account of language acquisition (MacWhinney 1999; Bates and Goodman 1999). Simply put, knowledge of language is neither completely determined by inborn mental states waiting only to be triggered by external experience nor does it originate entirely in the environment. The theory predicts that early word production is not the result of an innate search for nouns or any other built-in word category. Rather first words are produced as a result of a complex interaction between a species-specific innate architecture and environmentally determined constraints. I suggest in this thesis, that there is something else missing, an innate social predisposition for communication, that all of these theories have overlooked and that is necessary to provide children scaffolding for words to be produced.

1.3.9 Artificial neural network

Much of the foundation for emergentism is based on the development of a type of computer simulation called an *artificial neural network*. This section will define this class of computer programs briefly. A more detailed account of how artificial neural networks function can be found in Appendix I.

Advances in neuroscience and mathematics in the 1940's allowed mathematicians to propose a theoretical model of the way neurons function. The foundations of the artificial neural network models were laid down at that time with theoretical models designed on paper or hard-wired into circuits. Computer programmers were unable to fully implement these ideas until high-speed, large-storage computers became readily available in the 1980's. The goal was to model the brain and use the model to test theories about how the brain worked, including how children learn words.

A mathematical function is a formula that models outcomes by making simple calculations based on the input. Artificial neural networks model changes to the brain's neuronal connections by making a large number of such calculations and adjustments based on comparisons of approximate output to expected output. An assumption is made that a child's brain also makes many comparisons to both internal and external expected output. The models also assume that, like modern computers, our brains can process information extremely rapidly and don't care about how much storage space is required or how many computations are required. The idea is just to get the right answer.

Confusion can arise because of the many different terms associated with and used for these models. Technically a *neural network* is a biological entity – the interconnections of neurons in our brains. Hence, *artificial* neural network (ANN) is the correct name of the model. However, the terms *neural network* and *neural net* are often used to refer to these computer models. In addition, each programmer seems to have a slightly different definition of what an artificial neural network is, depending on what aspect of the process is emphasized. In general, however, an ANN may be understood to be a collection of small processors called *units* or *nodes*, each having a small amount of memory. Each node is connected to other nodes by some kind of communication channel carrying data. The nodes operate only on their local data and on the inputs they receive from the connections. How the nodes are connected, what data they receive, and how they are affected by that data define the kind of artificial neural network a model will be, what kinds of problems it can solve, and what kind of answers it will produce. It is important to remember that an artificial neural network is *artificial*. It is not a copy of our brains, the neurons do not work exactly like biological neurons, and the network cannot self-start – it is not conscious in the sense that a human brain is conscious. We currently have no computer programs that exhibit the qualities of a *mind*. A computerized neural network is capable of simulating only a single task (Jones 1999).

ANN's have been used to model many individual processes related to language learning. Specifically, they appear to be able to imitate children's learning to produce words (Plunkett 1995; Bates et al. 1995). These computer models provide evidence that meaning-form linkages do not have to be hard-wired into a system for word production to begin.

1.3.10 Summary of Definitions

The terms discussed in this section were selected because they are used commonly in linguistic discussions, will be used often in this paper, and are commonly misunderstood. Where different people define their terms differently, arguments have no common foundation. Much of the conflict in the field of language acquisition arises from a misunderstanding of basic terminology. We need to understand the difference between sensation, perception, and concept because, while they represent a continuum, only a *concept* can become divorced from reality. Similarly, icons, indexes, and symbols represent a continuum of abstraction. It is important to understand what a *word* is and that it is not the same as a *symbol* but, when paired with a *concept*, it can become symbolic, separated from reality, able to support and reinforce other concepts and be supported and reinforced by them. Specifically, when we say that a child has learned to produce a first *word*, it may not mean that the child has an adult's kind of concept behind that word or even that the word is being used as a symbol. These are ideas that are not often specified but appear to have been generally taken for granted in discussions about child language acquisition. I have tried to define them here so that the following discussion will be clear and provide a logical and consistent basis for the comparison of theories.

1.4 Historical Background

Theories about how children acquire language have been with us for a very long time. The Greek philosopher Socrates believed that language was present in the mind and only needed the proper educational environment to appear fully formed (Payne and Wenger 1998). Seventeenth century "associationists," thought that knowledge of the world is acquired through experience and stored as associations. Late nineteenth and early twentieth century behavioral psychologists argued that a child is born without any knowledge at all and must be taught words, one at a time. Behaviorists focused on learning acquired by reinforcement principles. In 1957, B.F. Skinner published *Verbal Behavior*, a book that attempted to explain language acquisition by children as a stimuli and response process.

Two years later, Noam Chomsky published a critique of Skinner's theory, arguing that if a child is to construct a grammar on the basis of observation of sentences and nonsentences provided by the verbal community in the way that Skinner proposed, then that child must be capable of constructing an extremely complex mechanism with the properties of an abstract deductive theory. He pointed out the difficulties in Skinner's description of language acquisition and effectively put to rest not just Skinner's theory but all similar theories by showing that behaviorist accounts generally fail to explain language acquisition and use (Payne and Wenger 1998; Chomsky in Allen and Van Baren 1971; Chomsky 1959). While Skinner had portrayed the child as a passive imitator, Chomsky emphasized the learner's active participation (Bowerman 1973). In arguing against Skinner's theories of acquisition, one of Chomsky's main achievements was to point out the difficulties involved in explaining child language acquisition. Chomsky (1959) observed that children seem to acquire complex grammars remarkably quickly and easily and suggested human brains are somehow specially designed to do this.

Those who agreed with Chomsky's proposal that children have a built-in way of approaching language acquisition, as well as those who did not, have continued to try to explain how children acquire their first words. In recent years, dramatic advances in a diverse array of disciplines have made it possible to produce and test theories of language acquisition in ways that have not previously been possible (Elman et al. 1996; Hirsh-Pasek and Golinkoff 1996). Advances in computer technology and programming have produced artificial neural networks that simulate some kinds of word learning and can demonstrate how word production may emerge in a system without being preprogrammed with representationally innate knowledge (Elman 1999; Elman et al. 1996). The history of explanations for how words are first learned has swung back and forth between exclusively internal and exclusively external sources. This paper seeks to describe the process as a complex interaction between innate brain architecture, social precocity, and environmental forces.

1.5 The Developmental Sequence

Children's first words are not spoken in a vacuum. They are spoken only after months of physical and mental development. Therefore, I include a brief summary of language-related and physical events that happen during the time leading up to and including early word production. Certainly, newborn infants do not say any words but two-year-olds normally do. What happens in between is important both to language and physical growth. In addition, such a description is necessary because any theory of child language acquisition must actually match what children do when they begin to talk.

While the information presented in this section represents averages, there is actually great variation among the patterns of how children learn to speak (de Boysson-Bardies 1999; de Boysson-Bardies and Vihman unpublished paper; Fenson et al. 1994; Owens 1984). Most research has worked with averages across children, even treating children as all being the same. Some studies have been done on only a few children or even on just one child. One of the most famous child behaviorists, Jean Piaget, developed his theories about the cognitive development of children based on his diaries of his own children (Payne and Wenger 1998; Golinkoff and Hirsh-Pasek 2000). At the other extreme, results of studies done on large numbers of children present the smooth curves of averaged data (Darley and Winitz 1961). Both methods fail to provide an accurate picture of the great variability between children. It is important to recognize this variability for several reasons. If we ignore variability, the interaction between social and cultural experience and innate abilities cannot be fully appreciated. The variability of language acquisition implies a plasticity of language systems that is not evident in averaged data. Further, the fact that all children do acquire native competence of their own language despite differences in environment and personality must be explained. With these caveats, the following summary of what may normally be expected to occur during the time when children are acquiring their first words is presented.

1.5.1 Newborns

Infants are surprisingly sensitive to speech sounds. Even before birth, fetuses respond differently to the sound of their mother's voice than to other sounds or even other voices (Golinkoff and Hirsh-Pasek 2000). At birth, infants prefer human speech to other kinds of sounds and can distinguish their mother's voice from other female voices (De Boysson-Bardies 1999). Clever researchers (Golinkoff and Hirsh-Pasek 2000; De Boysson-Bardies 1999; Jannedy et al. 1994) have shown that very young infants can tell the difference between different speech sounds by studying how fast they suck on an electronic pacifier. These studies show that infants as young as 3 or 4 days old can distinguish almost all the phonetic contrasts found across natural languages. Even very young infants respond to human speech differently than other kinds of sounds. They respond especially well to their mothers and appear to enjoy and seek human attention.

1.5.2 Two to Five Months

Between two and five months, the infant's vocal tract bends, the tongue changes shape, and the larynx lowers (Deacon 1997; De Boysson-Bardies 1999; Lieberman 1999). These changes allow vocalization to occur in the form of sounds issuing from the larynx and soft palate. Only toward the end of this period is an infant able to modulate the voice voluntarily. Infants laugh and they change the duration, pitch, and intensity of their vocal productions. They appear to delight in their own voices and are responsive to and recognize adults' voices. Infants watch and copy what adults do with their mouths and also try to copy sounds (Golinkoff and Hirsh-Pasek 2000). By four or five months, infants begin to pick out repeated sequences of sounds and can recognize their own names, even in the midst of a speech stream (Golinkoff and Hirsh-Pasek 2000).

1.5.3 Six to Eight Months

Between six and eight months, a baby's cooing and babbling begins to sound more and more like language (Golinkoff and Hirsh-Pasek 2000). Babbling (reduplicated consonant-vowel syllable productions) appears to be practice in saying syllables with the consonants, vowels, and intonation patterns of the language being learned. The early ability to distinguish a wide variety of phonemes disappears by ten to twelve-months. For instance, infants learning English can distinguish between /k'i/ and /q'i/ of Salish at six months but cannot hear the differences between these non-native consonants by twelve months (de Boysson-Bardies 1999).

1.5.4 Nine to Twelve Months

After nine months, babbling becomes non-reduplicated with varied consonants and vowels. Babbling sounds more and more like words (Golinkoff and Hirsh-Pasek 2000; James 1990). At around nine or ten months, many infants begin to use certain sounds in certain contexts such as using /i/ for approval and /u/ for disapproval (James 1990). Infants become increasingly socially aware and learn that their actions can have an effect on people around them (Mervis 1987). They realize they have a power to affect the behaviors of others and will perform acts and wait for a response. They communicate in the form of grunts, whines, points, and body language such as arm waving, vocalization, and eye contact and they expect adults to respond. They will persist in the attempt to communicate even when the adult does not seem to understand. They have learned how to use adults as tools to get what they want (Golinkoff and Hirsh-Pasek 2000). They love playing games that attract attention and one of those games includes imitating sounds.

Toward the end of the first year, games, routines, and other social activities may be accompanied by a high rate of meaningful signaling, pantomimes, and intonational vocalizations. The infant directs acts towards individuals, waits for a response, and is clearly trying to interact with caregivers (Mervis 1987). Some sounds in the infant's babbling are interpreted as first words. These early words can be identified as part of the social routine, not necessarily linked to innate categories or meanings.

1.5.5 Twelve to Fifteen Months

Children normally enter their second year able to produce what their parents have identified as their first word or even several words. It is difficult to identify the moment when the child has productive command of a "first" word. First words are not acquired quickly and easily but are learned and produced one at a time. Fenson et al. (1994) found that vocabulary size increases from an average of fewer than ten words at twelve months to over forty words at sixteen months. This amounts to an acquisition rate of fewer than two words per week and is certainly even slower in the beginning. Enormous variation in vocabulary size characterizes children at this age. In fact, several researchers have reported that variability between children actually *increases* between twelve and fifteen months (de Boysson-Bardies 1999; Fenson et al. 1994; Bates and Goodman 1999).

Normal children may speak anywhere from zero to fifty words at one year of age and anywhere from fewer than ten words to over 150 words at fifteen months. By the middle of their second year many children have already experienced what many call a "word spurt" (Fenson et al. 1994). This acceleration occurs at some point between fifty and one hundred words when, in the following four months, children learn an average of a word a day (de Boysson-Bardies 1999). This change in rate of word learning suggests that there is a qualitative difference between the early stages of the process and later language acquisition, when words are learned rapidly. This paper proposes that a qualitative difference, in fact, does exist and that a child's early words are not linked to innate word categories but that the later rapid word learning reflects that children must, by that point, have discovered language.

Children from twelve to fifteen months of age appear to comprehend many more words than they can produce (Fenson et al. 1994). However, comprehension and production may be dissociated to a high degree in early language acquisition. In fact, the two processes are statistically separable at the time children begin to produce words (Bates et al. 1988). Comprehension of a word is not sufficient for production to occur and does not necessarily mean that an apparently comprehended word has any referential meaning for the child (Bates and Goodman 1999; Fenson et al. 1994). For instance, Grieve and Hoogenraad (1986) noted that a 10-month-old girl was said, by her mother, to understand the word *no* and stopped doing something when her mother said *no*. However, she also stopped doing it when her mother said *yes* in the same tone of voice. Comprehension in this type of circumstance serves to emphasize the importance of social context in the word learning process.

1.5.6 Summary of Chronology

Table 1.1 presents a summary time-line of the acquisition of language skills. For comparison, this summary includes brief descriptions of motor skills that are normally present at about the same time. This table represents averages. All normal infants develop language but there is a great deal of variability between them that is not reflected in this table. The cultures and languages present in the child's environment differ greatly from one society to another. While most children successfully acquire their native language by the age of six, the age at which the first words are spoken, the number of words produced by a given age, and the character of early vocabularies vary from child to child and also from culture to culture (de Boysson-Bardies 1999). This high degree of variability is exactly what would be predicted by a theory that expects children to approach language in a social way rather than search for object labels. In fact, although there seem to be language-specific characteristic proportions of semantic categories in children's vocabularies as they approach 50 words (de Boysson-Bardies 1999), I contend that a study of very early vocabularies will reveal that first words are essentially random and reflect what is socially or emotionally salient to the child. In the next chapter, I will discuss current theories that have a bearing on the question of what first words are, how these theories approach word learning, and what they have to say about a child's first words.

Table 1.1 – Summary of average motor and language development – before birth to 15 months. (Compilation of data from Jannedy et al. 1994; Owens 1984; de Boysson-Bardies 1999; Werker et al. 1996)

Age	Motor Skills	Language Skills
Before birth	Reflexive movement	Reacts to voice. Recognizes changes in sounds. Prefers mother's voice.
1 month	Reflexive movement of limbs, lifts head while on stomach, coordinated eye movement, no reaching.	Responds to human voice, cries for assistance, cooing. Categorical discrimination between sounds. Prefers mother's voice and stories read prenatally.
2 months	Moves arms in circle, swipes at objects, raises head while sitting but head bobs.	Distinguishes different sounds, guttural cooing
3 months	Supports head when in prone position, weight is on elbows, hands open, no grasp reflex.	One-syllable cooing, vocal response to speech of others, vowels predominate.
4 months	Grasps rattle, head self-supported, brings objects to mouth.	Babbles strings of consonants. Pitch, duration, intensity variations.
5 months	Sits supported, rolls from stomach to back, can be pulled up to stand, swaps objects from hand to hand.	Vocalizes to toys, vowel-like sounds interspersed with consonants, responds to name, smiles and vocalizes to mirror.
6 months	Turns head freely, sits straight in chair, balances well, reaches with one hand, creeps.	Varies volume, pitch, and rate, vocalizes pleasure, cooing resembles one-syllable utterances. Decline in discrimination of non-native vowels. Prefers native language prosody.
7 months	Pushes up on hands and knees, rocks.	Vocal play, several sounds produced in one breath, listens to others.
8 months	Stands holding on, thumb-finger apposition, manipulates objects, crawls.	Reduplication of sounds, intonation patterns of native language, may appear to understand some words.
9 months	Stands alone briefly, gets down alone, sits unsupported, explores with finger.	Imitates coughs, hisses, clicks, and social gestures.
10 months	Pulls to sitting position, crawls with bilateral opposition, holds and drinks from cup.	Obeys some commands, imitates adult speech without success. No longer able to discriminate non-native vowels or consonants.
11 months	Stands alone, climbs up stairs, feeds self.	Imitates inflections, rhythms, and facial expressions. Expressive, word-like babbling.
12 months	Pushes to stand from squat, climbs down stairs, uses cup, spoon, pencil, releases object willfully, takes first steps.	Follows simple motor instructions, reacts to 'no', may speak one or more words.
15 months	Unceasing activity, walks a few steps backwards and sideways, carries objects in both hands, throws ball, takes off shoes and socks, scribbles.	Points to named objects, has 4 to 6 word vocabulary.

Chapter 2

Current Theories of Child Language Acquisition

2.0 Introduction

One theory of language acquisition is that because children have a species-specific language "module" or "organ" in the brain, they are predisposed to look for the names of whole objects when they begin language (Waxman 1994; Markman 1993; Waxman and Kosowski 1990). The purpose of this thesis is to examine this theory and compare it to those of other researchers to understand why such an idea has been proposed, to give evidence it is not correct, and to suggest an alternative hypothesis. In this chapter, I will present an overview of several currently held ideas about how children begin to learn language. The first group of theories concerns what are usually called *nativist* or *innatist* theories. Nativists believe that fundamental aspects of linguistic knowledge are inborn, including constraints that limit what a child expects words to be and what a child's first produced words might be. In the sense of innateness as defined in Chapter 1, these theories propose that there is *representationally* innate knowledge in the brain that is unique to the human species and gives children access to language.[2] These theories predict that a child's first words are meaningful and that children look for object labels. I will contrast these ideas with emergentist theories and argue that children's first

[2] That is, this knowledge is hard-wired in or preset – see also Section 1.3.1.

words are better explained by an *emergentist* viewpoint, which accepts that something about language is innate but that innateness is *architectural*[3] rather than representational.

Emergentists do not deny that something about language must be innate. However, emergentists do not believe that there is language-*specific* information hard-wired into the brain. They argue that there are *architecturally* innate brain design features that distinguish humans from even our close primate relatives, making language possible. The human brain has evolved in a way that is structurally extremely plastic, self-organizing, and experience-sensitive in ways that are architecturally unique to the species but there are no special language processing or storage areas, separate from other mechanisms, that store and process *just* linguistic information.

Finally, there are a number of theorists who suggest that both nativists and emergentists have important ideas to contribute and that the similarities between the two approaches outweigh the differences. I have called proponents of this view *synthesists.* The synthesists stress that both nativists and emergentists agree that some part of language learning must be innate and at least some nativists have come to recognize the importance of general learning mechanisms (Hirsh-Pasek and Golinkoff 1996).

2.1 Nativist Theories

The subject of language acquisition has fascinated people for thousands of years and many theories have been proposed (Golinkoff and Hirsh-Pasek: 2000; Payne and Wenger 1998). However, empirically verifiable theories were not developed until the

[3] That is, the structure of the system defines what can be learned. See also Section 1.3.1. This does not mean that children's minds are blank slates with nothing in them. Emergentists contend that many general-purpose abilities are either present at birth or develop before first words are produced. These domain-general capacities include attention biases, size and shape discrimination, object permanence, and imitation.

twentieth century. In the second half of the twentieth century, Noam Chomsky proposed that children are born with a system of principles, conditions, and rules that provides what he saw as a missing link between the linguistic input and the knowledge that children seem to know at an early age (Cook and Newson 1996; Chomsky 1959). His ideas created an atmosphere in which additions and changes to his theory as well as opposing theories have proliferated. What, then are the nativist theories? How do they explain language acquisition as it applies to how children learn their first words?

2.1.1 Noam Chomsky and Universal Grammar

Chomsky observed that all normal people speak some language, that all languages seem to share certain attributes, and that all children learn their native language on about the same time schedule (Cook and Newson 1996). Languages are acquired in a relatively uniform sequence. Practice, reinforcement, and IQ (intelligence as measured by certain standardized tests) appear to have little effect on the outcome. Skinner's earlier behaviorist theories claimed that language is acquired through operant conditioning but Chomsky related language development more to growth than to learning. Chomsky argued that the growth of language is analogous to the development of a bodily organ (Chomsky 1959; Chomsky 1975). Somehow, children are able to pick words out from a continuous stream of sound and, with imperfect and deficient input, match words to their correct reference in the environment. He concluded that there must be something unique and special about the human species that is biological. Children must be born with some kind of innate mechanism that allows them to learn language. Because all children learn language in all societies, input must only provide triggers for this mechanism and does not play a central role (James 1990). In Chomsky's view, there is not enough information in the input that children receive to explain the language knowledge that children acquire at an early age. This apparent discrepancy between input and output has been referred to as the "poverty of stimulus" argument. Chomsky's conclusion was that the source of the knowledge must be the mind itself (Cook and Newson 1996).

As evidence for this theory, nativists (Lightbown and Spada 1999) point out that the acquisition of language is an extremely complicated task, yet children seem to be able to do it easily. Even children with otherwise serious handicaps manage to learn language. Deaf children learn sign language if exposed to it in infancy. Mentally handicapped children learn language more slowly but, nevertheless, seem able to learn it as well. Children who are abused or neglected learn language if they are exposed to it at all.

The nativist claim is that the speech children hear is full of false starts, incomplete sentences, and uncorrected errors, and that no one points out to children which sentences they have heard were correct and which weren't (Cook and Newson 1996). When they do start to speak, they are not consistently corrected and, even when they are corrected, they tend to focus on the meaning of words rather than the corrected form. Children actually seem to ignore the corrections and say things the way they want to (Lightbown and Spada 1999). What children hear underdetermines their observed competence. Children are not provided with examples of all the linguistic rules of the adult language.

When these observations were first made, little serious study had been made of children's early speech. Language acquisition was seen as a logical problem that could be solved without empirically studying children (Cook and Newson 1996). If we observe the input to the child, the *primary linguistic data*, and the output, a *generative grammar*, then what lies in between in the child's mind must contain a Language Acquisition

Device (LAD), now called the *language faculty*, that contains everything necessary for language to be acquired. Chomsky equated the LAD to a species-specific language faculty, Universal Grammar. If we carefully observe what is going in and what is coming out, we should be able to deduce what is going on inside the LAD. For Chomsky, the question to be answered about child language acquisition, then, was what the LAD consists of and how children use it.

2.1.2 Syntactic Bootstrapping

The answer was that syntactic structure plays a central role in mediating word meaning (Cook and Newson 1996). Certain kinds of word meanings are attached to certain positions or functions in sentences and this can provide clues to help children learn word meaning. Gleitman (1993) calls this process *syntactic bootstrapping.* Children need input to be able to access and use the LAD but that they need to hear certain kinds of sentences in order to set the parameters for their own language. They need *bracketed input*, sentences with clear signs of phrase boundaries and subcategorization frames to determine how many phrasal participants are in the clause (Cook and Newson 1996). This kind of input provides a way for children to learn the meanings of words and acquire knowledge of language.

Chomsky assumed that language and thought are autonomous and distinct. Knowledge of a language consists of processing speech in terms of form classes and using combinatorial rules expressed in terms of categories such as noun, noun phrase, and subject of the verb. If thought and language are separate, then the task of the child is to find the relationship between the two (Macnamara 1977). The proposed LAD provides not a full-blown set of grammar rules but a set of constraints on what those rules can be and a set of procedures for learning (Bloom 1993). Chomsky's theory is that the rules of any language are structure-dependent (Cook and Newson 1996). Children know in advance some possible forms that language may take. Hence, they have a head start at linking language and thought, and can begin looking for word meanings at the time they begin to speak.

The theory says that, because they already know phrasal structure, children can exploit sentential information to actively search for meanings of words (Cook and Newson 1996; Bates and Goodman 1999). A child hears a word, notes the clausal context in which the word is used, and formulates a hypothesis about the concept to which the word corresponds. Ideas about what form the concept might take differ. The child's concept may be a single primitive innate concept, a combination of primitive concepts, or a prototype structure. However, a concept in some form must be available to the child in order for the child to match meaning to form (Bloom 1993; Jackendoff 1992; Fodor 1975).

These ideas, which have grown out of Chomsky's Universal Grammar theories, have received support from a number of empirical studies. In the 1990's, several papers provided evidence that children can use syntactic frames to help determine the meanings of novel verbs (Gleitman 1993; Naigles et al. 1995). Gleitman (1993) showed that blind children can learn the subtle differences between *look, watch,* and *see* and that they must be using syntactic frames to do so. Gold (1967) showed, mathematically, that without knowing some critical things about the language being learned, if a child is presented with context-free positive input only, languages would be unlearnable. That is, the learner must be limited to learning only certain kinds of grammars and not others. Such computational arguments lend support to Chomsky's theory because they seem to show that language would be unlearnable unless there were some innate constraints on language learning. That is, awareness of and ability to use the proposed LAD would be impossible unless the child had some inborn knowledge of what language is.

Support for the claim that children do not receive or at least do not attend to negative input comes from research reported by Bloom (1993). Observation of parents and children shows that the grammaticality of children's speech is not correlated with parents' approval or disapproval (Paul Bloom 1993). Parents appear to be more interested in *what* their children say than *how* they say it.

Goldin-Meadow and Mylander (1993) report that deaf children seem to be predisposed to linguistically analyze the words, signs, or gestures they use to communicate. They conclude that several deaf children being raised in hearing families can assign lexical meaning to spontaneous gestures used by their mothers and, in so doing, go beyond the input they experienced. Their work also shows that the proposed LAD is not limited to auditory language but applies to any mode of communication that entails a generative grammar. The study supports the idea that children are following innate patterns of development that appear to be unaffected by what their parents provide for input.

If a Language Acquisition Device helps children begin to learn language, damage to it should result in specific, detectable impairments. Williams Syndrome is a form of mental retardation in which a child is severely mentally handicapped in many respects and may have a measured IQ of only 40 to 60. Surprisingly, language development appears to be unaffected or even enhanced. Children speak fluently and enthusiastically and delight in novel or unusual words (Bates and Goodman 1999; Trask 1997). On the other hand, damage to Broca's area, located just behind the left temple, can cause sufferers to have difficulties producing grammatical structures and understanding grammatically complex sentences. Wernicke's area is located just above and behind the left ear. Damage to this area results in fluent but meaningless speech (Trask 1997). These results provide some neurolinguistic evidence that there is a language module that is separate in the brain from other kinds of mental activity, is different from other mental processes, and that children must have access to it in order to successfully learn to talk.

Noam Chomsky's response to behaviorist explanations of language acquisition reflected his observations that children do not simply memorize and repeat their parents' speech but use language creatively from the time of their earliest speech (Cook and Newson 1996). Chomsky assumed that input is underdetermined, hence, cannot account for a child's relatively easy and rapid acquisition of language. Therefore, there must be a genetically determined program to look for specific information. Chomsky's ideas led to the development of the *syntactic bootstrapping* hypothesis – children's innate knowledge of syntax so tightly limits possible grammars that a small amount of input is sufficient to build word meanings, determine word classes, and develop subcategorization frames for verbs. Syntactic bootstrapping claims that, even at the beginning of word production, children are sensitive to phrase structure and grammatical roles.

2.1.3 Semantic bootstrapping

Macnamara (1977) proposed that the child's cognitive ability to understand something that is happening while listening to adults talk gives the child access to a basis for understanding the meaning of the words used to describe the event (Macnamara 1977;

Snow 1999). The term *semantic bootstrapping* describes the mechanism that allows a child to construct a semantic representation of input with the help of context (Pinker 1987). Pinker (1994) proposed that children are able to determine the meaning of a phrase from a situation. The claim is that children are born with knowledge that allows them to link the labels they hear to their semantic counterpart. The child understands semantic entities such as *thing*, or *causal agent* and expects the input to contain tokens of them in the forms of, for instance, *nouns* and *subjects* (Pinker 1987). This explanation assumes that children are able to parse the input into words and that they possess innate linking rules for joining the words they hear to the appropriate classes (Hirsh-Pasek and Golinkoff 1996). Paul Bloom (1993) describes the process as a one-way mapping from cognition to form. Children learn word meanings by linking their use to a perceptually salient feature of a situation. In this theory, the child is able to learn syntax because of innate cognitive abilities rather than using innate Universal Grammar to help discover word meanings. As a child acquires more word meanings, lexical knowledge can combine with an understanding of events in the world to provide the semantic analysis from which grammar can be bootstrapped (Snow 1999).

While syntactic bootstrapping relies on domain-specific[4] innate knowledge to provide the key to language acquisition, semantic bootstrapping proposes that both

---

[4] As applied to language learning, *domain specificity* or *modularity* is the theory that the brain has a special-purpose learning device that is dedicated to learning language (Fodor 1975; Bates and Goodman 1999). In this theory, language is considered to be a *domain-specific* distinct mental capacity, reflecting knowledge that is not wholly derived from general cognitive capacities (Pettito 1996). On the other hand, if language develops out of *domain-general* capacities, then children's first words are not preprogrammed and they are not searching for object labels. This would mean that early words are the result of a human brain's capacities for information-processing and problem-solving combined with other innately human but non-linguistic processes.

domain-specific *and* domain-general processes are involved (Bloom 1993; Hirsh-Pasek and Golinkoff 1996). That is, general learning devices such as pattern detection are important for acquiring language. However, semantic bootstrapping theory holds that word meanings cannot be acquired unless children are innately predisposed to entertain only certain hypotheses (Markman 1993). Pinker's theories stress learning as a fundamental aspect of language and other cognitive domains, and provide for a seamless continuum from pre-linguistic to linguistic behavior both for an individual child and for the species. Semantic bootstrapping theory places more of the burden of language acquisition on the child's learning capacities than syntactic bootstrapping and accepts the necessity for a large amount of linguistic input to accomplish this (Ulbaek 1998).

The acquisition of word meanings is a central issue for the semantic bootstrapping hypothesis. The theory says that children initially only allow a single meaning for a single word (Markman 1993). Since any single word may have a large number of logically possible meanings even when the child and adult are attending to the same referent (Lederer et al. 1995), there must be a way to place limits on what words can mean. Semantic bootstrapping says that a Language Acquisition Device provides constraints on the child's representational (perceptual and conceptual) biases. These constraints limit the number of hypotheses that the child can entertain and lead a child to favor some hypotheses over others (Bloom 1993). The theory further says that the concepts for word categories and the rules that link words to these concepts are innate and that the beginning language learner can use an emerging understanding of the design of the language being learned as a guide to a new word's mapping (Lederer et al. 1995).

The child's first words are meaningful and are matched to innate prototypes (Pinker 1987).

Golinkoff et al. (1994) say that the process of early word learning has three basic *lexical principles* or constraints. First words can be acquired *because* children know that words 1) refer to objects, actions, and attributes, 2) that they map to whole objects, and 3) that they represent categories of objects, not just the original referent. These constraints *potentiate* learning by limiting the number of hypotheses the learner considers. This is the same as saying that semantic bootstrapping allows children begin word production by mapping word forms with innate concepts. These concepts are linked to words at the time they are learned and children actively look for names of whole objects. In other words, children come to the language-learning task expecting words to be object labels.

There are research studies that support this theory. Studies reported by Golinkoff et al. (1994) show that infants respond very early to pointing gestures and learn to use such gestures to establish joint reference. Establishment of joint reference is important if children use context to learn word meaning. A 1990 study by Waxman and Kosowski (1990) showed that children as young as two years old interpret novel nouns as referring to category relationships. The youngest subjects in these studies also interpreted novel adjectives as taxonomic categories, supporting the theory that children begin by assuming new words are nouns even when used syntactically as adjectives. Waxman (1994) further investigated the process by testing two, three, and four-year old children. In all age groups, children who were given a novel noun and asked to find "another one" extended the meaning to include taxonomic alternatives. When the children were not given a label for the indicated object, they picked alternatives randomly, suggesting that it is the *name* of the object that allows the child to form a category.[5] Markman (1993) proposes that children learn word meanings using two basic principles. First, when presented with an object and a label, children assume that the label they hear is the name of the *whole object*. Even the youngest children in her studies (two years old) followed this rule. Second, her *mutual exclusivity principle* says that children expect only one name for an object and, when presented with a second name for an object, they assume it is either a name for some salient aspect of the object or the name for the substance the object is made of. When presented with a second label for an object, three and four-year olds tend to reject the new label as a name for the object and seek to find another aspect of the object to label. These two principles allow children to first find a name for the object and then provide a way to acquire other kinds of words, such as adjectives, that describe objects for which they already have a name. They provide the mechanism for semantic bootstrapping theory to allow for the acquisition of words beyond object labels. These studies show that children as young as two years old can learn names of things and use real-world situations to match meaning to form. However, all of these studies of two-year old children only show how children may use what they already know about categories to learn new words. They actually show nothing at all about how children begin the process.

---

[5] In February 1999, I ran a small experiment designed to test Sandra Waxman's ideas. Her studies of two-year-olds involved children who were an average of 2 years and 7 months of age with the youngest child being 2 years and 1 month. I tested a boy who was 23 months old at the time of the experiment and used her methods. The child was shown a page with an object in the center of a page and four other objects. Two were related thematically, two taxonomically. He was given a novel name for the center object and asked to find another one. My results exactly repeated her findings. Even though this child was less than two years old, he exhibited a systematic approach to interpreting novel nouns and provided support for the idea that children of this age exhibit bias toward extending object labels to other objects of the same kind rather than to objects that are thematically related.

Semantic bootstrapping models of language acquisition assume that children are able to use special linguistic and general cognitive abilities to understand the meaning of a sentence from its context and link the sounds of the words they hear to innate concepts. In particular, they have concepts for categories, such as agent, and they look for words in their linguistic environment that match these concepts. The theory proposes that children actively search for names of things and that they are constrained to initially assume that the words they hear are names of whole objects.

I will argue that, while there may be substantial evidence that children can and do use context and whole object principles by the time they are two years old, it has not been demonstrated that this is how they *begin* the process. In fact, the evidence shows that children do not have a preference for object labels, that words are not initially mapped to categories, and that an alternative synthesis theory based on emergentism better explains the production of very early words.

2.2 Emergentist Explanations

Bootstrapping theories hold that children need more than verbal input to learn language. Syntactic bootstrapping requires children to have the innate ability to map sentence structure in the form of structural hierarch and possible subcategorization frames to contextual clues and asserts that children need only a minimal number of key phrases to determine which verbs obtain which sets of subcategorizations, as well as other parametric variation. Semantic bootstrapping proposes that children need certain information from context to link meaning to form but that their ability to pick out what they need from the input is enhanced by innate knowledge in the form of semantic cues that correspond to perceptually salient contextual features. Many researchers now believe that context is important in developing meaning but not everyone agrees that the linguistic knowledge children need to process input is *representationally* innate – hard-wired to at least some degree and present at the beginning of language acquisition. *Emergentists* agree that the human language capacity is distinct from other forms of animal communication, but assert that there is no language-specific module in the brain. They propose that language is the result of a large number of complex interactions between general processes. Language is the result of many little things, none of which, alone, would be expected to result in anything very dramatic but, taken together, produce something radically new (Elman 1999). Emergentists reject the claim that there is a special language faculty. They see language as an emergent property of architecturally innate mechanisms rather than the result of representationally innate knowledge about language. That is, it is the structure of the brain, not what it contains, that constrains how learning can occur. Children's first words are the result of physical maturation and domain-general cognitive abilities. Emergentists do not reject the claim that children "know" a great deal when they begin to talk but that the knowledge they have about the world comes from general cognitive processes, not from special language knowledge.

Emergentists rely on computer models of neural networks to show how language learning can take place and how language rules can be learned by a system that does not have the rules already built in. Since the artificial neurons in these networks are heavily interconnected, they are called *connectionist* models. Computer models allow researchers to test their theories of how brains work and have been successful in demonstrating that at least some of the theory is plausible.

2.2.1 Connectionism

Early optimism about the abilities of computers to simulate human language was replaced with disappointment when, in the 1960's, attempts to create language translation programs demonstrated that human language is a complex problem that computers could not easily solve (Crevier 1993). So it is understandable that linguists, knowing that language cannot be processed linearly, are skeptical of claims that computer models can replicate any kind of language learning. However, linear processing models have been replaced by models that researchers think more closely imitate what happens in the brain. One of the earliest such models was McClelland and Rumelhart's *interactive activation model* of word recognition (Elman et al. 1996; Payne and Wenger 1998). In this 1981 model, word recognition is both data-driven at the level of information in the input and conceptually driven from higher-level processes. These processes work simultaneously and in parallel. When a written word is recognized, for instance, information about the features or basic pieces of letters is perceived and passed to the part of the brain that processes letter shapes. Connections to some letters are activated; others are inhibited. Information at the letter level suggests possibilities about what the word might be. Cognitive or "top-down" processes limit and define what word and meaning is activated. Because all the pieces of information are simultaneously connected, this kind of model has been termed *connectionist* and the theory is called *connectionism.*

Rumelhart and McClelland suggested that the mechanisms that process language and make judgments of grammaticality are constructed in such a way that their performance can be *characterized* by rules, but that the mechanism that produces them does not contain any statement of a rule or guiding principle (Rumelhart and McClelland

1993).[6] In other words, the rules are not explicitly written into the system to begin with but they *emerge* because of the way the pieces of the system are connected.

Connectionist theory proposes that the interaction of different modalities, such as vision and hearing, can give rise to developmental phenomena that would not emerge if they acted separately (Plunkett 1995). Computer models using connectionist theory have been used to model the acquisition of grammar and the emergence of phonology, grammaticality, and word meaning (Bates and Goodman 1999; Allen and Seidenberg 1999; Goldberg 1999). Most important for this thesis, computer models show that concepts do not need to be present before words are learned. Concepts may develop slowly after words begin to be produced (Plunkett 1995). Plunkett describes a simulation that was trained to associate images with words. This model demonstrated the early slow acquisition of word meaning as well as a dissociation between comprehension and production and the typical overextension and underextension that have been demonstrated in children's early speech (Plunkett 1995; Bates et al. 1995). Learning curves for both word recognition and production were similar to those of children learning language (Plunkett 1995). The physical images presented to the computer to be linked to words were variations of prototypes but the prototypes were never actually

[6] In the mid 1980's, Rumelhart and McClelland applied their theory, now referred to as *Parallel Distributed Processing* or PDP, to the problem of language acquisition. Selecting a language feature that has traditionally been assumed to be rule-driven, they designed a computer program that used English present tense verbs as input and produced the corresponding past tense as output. Their simple model did a fair job of learning to produce correct past tenses for irregular as well as regular verbs and could produce, with reasonable accuracy, the past tense even for unfamiliar verbs. This model provided early evidence that rule-like behavior could be produced by a mechanism in which there is no explicit representation of rules (Rumelhart and McClelland 1996).

presented to the program while it was learning.[7] Nevertheless, during the testing phase, the model identified prototypes more reliably than the images that the model had been trained with. The model had extracted a central tendency from a group of training patterns. That is, the model recognized the prototype even though it had not seen it before. The implication is that the acquisition of first words need not depend on the presence of innate prototypes. Words can be learned from repeated presentation of examples, and categories can emerge after word production begins.

Plunkett's (1995) model exhibited a lag of production behind comprehension that has also been recognized in young children.[8] Plunkett suggests that this lag might be the result of the development of categories. He proposes a mechanism for how the two might be related and how words might be produced without categories being present at the beginning of word production. During the early stages of training, comprehension is exhibited because labels may be better predictors of category membership than image patterns. Labels are discrete and images tend to be what Plunkett calls "fuzzy" predictors of a category. Labels can trigger several possible correct images so the network exhibits understanding of a word meaning when it is processed while image categories are still diffuse and may even overlap. Hence, it will be more difficult to select and produce a correct label in a given situation for production. Until the network establishes an accurate prototype of the clustering of images, labels provide more reliable cues to category membership than images. We know from many studies that this asymmetry is,

---

[7] In psychology, the *prototype theory* of categorization assumes that a summary representation is preserved in memory. A new memory trace is not stored for each item. Rather, a running average or central tendency of presented examples is stored. This "average" item is called the *prototype*. (Payne and Wenger 1998)
[8] See Section 1.5.5 for a discussion of this phenomenon.

in fact, very large in the early stages of word learning but gradually disappears (Plunkett 1995; Hirsh-Pasek and Golinkoff 1996; Plaut and Kello 1999; Gupta and Dell 1999). The dissociation between comprehension and production in computer models demonstrates that concepts need not be present before words are comprehended but that they may develop gradually as production begins.

Since emergentist theory holds that categorization and meaning emerge from the way children process input rather than innate prototypes, one implication of this theory is that there should be as much variation between children's initial word production as there is variation between children's environments. In fact, several researchers have noted that there are indeed very great differences in learning style (Fenson et al. 1994; Darley and Winitz 1961; Bates et al. 1995). Emergentist theory says that language is the product of interactions between many modalities, some of which can have enormous differences in developmental rates. These differences could cause significant variation in the timing of word comprehension as well as production (Bates et al. 1995).[9]

Studies with small numbers of children show what kinds of variation in linguistic development are possible (Darley and Winitz 1961). But the extent and nature of such variation is only evident when a large population is investigated (Bates et al. 1995). Evidence from a sample of more than 1,800 children relates the observed variations to many factors including gender, birth order, and social class. In addition, the variables

[9] For example, there is evidence that children with Downs Syndrome and at least some children with Specific Learning Impairment may have one or more impairments in the auditory processing system as well as general cognitive delays. Hence, it is not surprising that these individuals have the most difficulty detecting, storing, and retrieving linguistic input that is lowest in acoustically phonological salience (Bates and Goodman 1999).

tested included cognitive abilities such as tool use and categorization, attention to spoken words during the babbling stage, use of intonation during babbling, and interest in mimicking parental actions and sounds. Their research confirmed that there is enormous variability in the rate of children's linguistic development and that only a combination of factors could explain the differences.[10] These factors are both maturational and environmental and cannot be explained by a single, language-specific module in the child's brain.

The emergentist position proposes that no special language faculty is necessary to explain early word production. Concepts do not need to be pre-programmed but can be acquired gradually along with word comprehension and production. Word meanings and the categories they represent may be the result of word production rather a prerequisite for it. The prediction that such emergent processes would result in a great deal of variation in the way children begin to learn their language has been verified by studies of learning styles. I will provide evidence to support this claim and, further, suggest that children's very early vocabularies are an extension of social interaction. However, any theory of early word production assumes that children can segment some sounds out of the speech stream. The next section will address how emergentism might explain how

[10] Two of the prototypical strands of learning styles that could be identified were: Strand 1 which tended to be first-born females, word-oriented, object oriented, and less interested in imitation. These learners had a high proportion of nouns in their first 50 words. A typical Strand 2 learner was a later born male, intonation-oriented, person-oriented, more likely to imitate spoken words, and had a lower proportion of nouns in the first 50 words. These strands are similar to what other researchers' have variously termed *analytic* v. *holistic, referential* v. *expressive*, and *nominal* v. *pronominal* that have been used to classify children's personalities. However, they emphasized that these categories are *examples* not *definitions* of learning style. The important result was that there is demonstrably huge variation among children's learning styles.

children separate some sounds from the speech stream and why the sounds that become words are identified.

2.2.2 Word Segmentation

Before they can begin to speak, infants have to be able to pick out meaningful units from the input. This process of dividing up the speech stream is called *segmentation*. What seems to an adult like a series of words is actually an acoustically continuous stream of sound. How are infants able to segment this stream? Regardless of the meaning content that may or may not be attached to separate segments, children must be able to separate out the meaningful bits contained in spoken language. In fact, this unbroken speech stream contains language-specific clues. For instance, languages may contain differences in pitch, in vowel lengths, or in stress that even pre-linguistic infants can detect and that may be necessary for them to discover how their language is constructed. Infants as young as eight months already know what kinds of sounds can occur at the beginnings and endings of words (Golinkoff and Hirsh-Pasek 2000).

The cognitive processes that help infants segment words begin to develop very early. Fernald (1993) suggests that paying attention to speech is a survival strategy. Prelinguistic infants prefer speech to other sounds but show no preference for words pronounced in their own language over any other. By nine months, however, infants begin to show preferences for their own language by turning their heads toward a loudspeaker more often when they hear words in the language they are learning than some other language (Golinkoff and Hirsh-Pasek 2000).[11] Although not yet able to pick

[11] For example, Golinkoff and Hirsh-Pasek (2000) report that Dutch children prefer to listen to Dutch words and English children prefer English words.

out individual words, as a first step, they become tuned to the *sounds* of their own language and they are sensitive to acoustic cues of phrase boundaries (Golinkoff et al. 1999). Before they are able to process words linguistically, infants do some kind of acoustic analysis of their own language.

The next step is to be able to segment the meaningful chunks in speech. Emergentists believe that the acoustic properties of speech are statistical in nature and help the child pick out words (Aslin et al. 1999). They have been able to show that the transitional probabilities that certain sounds follow other sounds can predict word boundaries. These statistical properties may be what guide the infant to segment the speech stream. Aslin et al. (1999) showed that adults, children aged seven to eight years old, and even eight-month-old babies can pick out nonsense words from an unbroken stream of nonsense syllables when the targeted nonsense words reappear predictably in the speech stream.[12] As little as 45 seconds of listening to a series of "sentences," within which a target word is embedded, is sufficient to induce a difference in the amount of time eight-month-old infants spend listening to that word versus an unknown word. In other words, infants are able to pick out repeated linguistic units from a stream of sound simply by the probability of the occurrence of sound combinations and, moreover, they can do it with very little exposure. When pure tones were substituted for syllables, adults and infants alike extracted repeated sequences the same way they had done with nonsense

---

[12] Aslin et al. (1999) presented adults with "word" pairs – a nonsense word that had been presented in the sound stream and one that had not. After only 21 minutes of exposure, subjects were able correctly choose the previously heard word 76% of the time. Eight-month-old infants were tested with the preferential listening technique. Sounds were played through a loudspeaker and repeated until the infant looked away or until a predefined maximum was met. The test phase recorded how long infants listened to a word before getting bored and looking away. Listening times were significantly different for word and non-word test items, an indication that infants discriminated between them.

words. The authors conclude that the ability to segment words out of the speech stream is at least partly due to a general statistical learning mechanism that can operate on a wide range of acoustic input rather than from innate knowledge.[13]

Redington and Chater (1997) reviewed a number of computer simulations and distribution studies, concluding that there is enough statistical information in normal speech to account for word segmentation and the development of word classes and meanings. Gupta and Dell (1999) have shown that a variety of sequential tasks, including both repetition of nonsense syllables and a skill task of keying numbers, benefit by having sequences repeat *even when no feedback is given* to indicate whether they are correct or incorrect.[14] Analysis of error-production shows that occurrence constraints arise at different levels - those within words or numbers, those within lists, and those across all lists. They equate these constraints with syllable level, experiment level, and language-wide level.

The implication of these experiments is that there is a general human sensitivity to sequential constraints and that these constraints emerge at multiple levels in procedural memory (Aslin et al. 1999). This is the same kind of sensitivity that arises in the processing of sound sequences. Other studies have shown that infants are sensitive to the

---

[13] Aslin et al. (1999) acknowledge that many clues are necessary to help infants learn where meaning unit boundaries occur. Their studies were designed to eliminate other clues such as context, prosodic regularity, and stress patterns and test what role statistical data might play in predicting word boundaries and whether that ability was strictly linguistic or whether it might be a general learning mechanism. They designed their data to contain transitional regularities. This statistic involves calculating the probabilities that certain sounds follow other sounds. For example, if the transitional probability between two syllables were close to 1.0, those syllables would be within a word. If the transitional probability between two syllables were close to 0.0, this indicated a likely word boundary.

[14] Participants were given lists of 5-digit numbers to key into a computer. No feedback was provided for the response. Numbers were keyed and the enter key pressed. One second later, the next number was

prosodic regularity of their own language at an early age (Aslin et al. 1999; Hirsh-Pasek and Golinkoff 1999), and Cutler (1994) demonstrated how stress might be used as a cue. The information needed to segment words from the speech stream is almost certainly contained within that stream itself. While this does not explain the development of meaning, infants must be able to find words and we do not need to rely on innate knowledge or bootstrapping to explain how they do it.

Advances in cognitive skills at around ten months allow analysis of speech information in new ways. At six months, infants could discriminate changes in stress pattern or changes in syllable pattern but are not able to coordinate both. By nine months, they began to be able to integrate prosodic grouping with distributional information, detecting changes in sequences of syllables based on predictable stress patterns. At ten months, children lose the ability to discriminate between *non*-native speech sounds. Researchers suggest that this is because children are now able to predict what acoustic variability is functional in their own language-learning environment (Werker et al. 1996).

Certainly the ability to segment repeated sequences from a continuous stream of sound is not all that is necessary for children to learn to speak their language. However, the studies reviewed here show that it is, in principle, possible to explain how the acoustic stream might be segmented by processes that may mature slowly and that are domain-general in nature. Speech segmentation alone does not result in the acquisition of language knowledge and does not require domain-specific processes or innate linguistic

---

presented for entry. Some numbers appeared multiple times during the experiment while others appeared only once. Errors declined gradually and progressively for items that repeated (Gupta and Dell 1999).

knowledge. However, the ability to divide up the input into meaningful pieces is necessary for children to begin to learn words and the research shows this can happen from statistical segmentation processes that occur without conceptual mapping.

2.3 Syntheses

This paper has thus far reviewed two main ideas about language acquisition. Nativist theories hold that there is some kind of representationally innate, or preprogrammed, language knowledge present at the time that children begin to learn language. Emergentists point to the success of artificial neural networks in modeling features of language acquisition and claim that structural or architectural innateness can fully account for language. Although it seems that these two views are inherently incompatible, there are researchers who accept many of the principles from both views. These researchers believe that the way children learn language is best explained by combining the two views and they have tried to provide synthesis accounts of language acquisition, including how children learn their first words.

For instance, while syntactic bootstrapping theory holds that there is a language-specific organ in the brain responsible for language acquisition, Pinker (1994) and Gleitman (1993) suggest that children need to use *both* domain-specific and domain-general processes to analyze input. Pinker (1987) likens the language acquisition task to a puzzle with multiple pieces. Solving one piece of the puzzle would lead to solutions for all the other pieces but none of them can be solved in isolation. The problem for the child is finding a place to start. For Pinker, the solution of this puzzle is that innate constraints, which he calls *rule prototypes,* give the child a head start in mapping form to meaning. But he believes the rest of the solution lies in the input.

Gleitman (1993) agrees with Pinker's claims that children can induce noun meanings by linking their use to a situation. However, as discussed in Section 2.1.1, she has shown that this "word to world" mapping cannot be the *only* way that children learn words. She proposes that children often understand the meaning of a sentence first and deduce word meanings by analyzing the content and structure of that sentence. This kind of mapping provides a way for children to learn verb meanings when observation fails them, as she says it eventually must. The process is categorical rather than probabilistic in that the children need only a small database to conclude that a word has a particular meaning in a particular syntactic environment. She concludes that children are able to play two imperfect and insufficient processes (semantic bootstrapping and syntactic bootstrapping) against each other to derive the best result.

At the other end of the spectrum, emergentists must acknowledge that language is universal among normal human beings. They take this to mean that some kind of prerequisites must be present for language to develop. Among the emergentists, there are two groups – those who believe socialization is a special innate prerequisite and those who believe that general human cognitive development makes language possible. If a synthesis is possible, it will come from the understanding that language acquisition must be the result of the interaction of many factors. I have suggested that general cognitive development is responsible for children's early word production rather than innate knowledge and that children are especially sensitive to social factors, so it will be important to understand what roles socialization and cognition might play in that synthesis.

2.3.1 Socialization and Cognition

Researchers who emphasize socialization as a way for language to begin emphasize that children are active communicators virtually from birth. First words, say Grieve and Hoogenraad (1986), do not spring from nowhere but represent a point on the continuum of the child's developing ability to communicate. The child's first words are continuous with what has been happening before words first appear. Early utterances are frequently employed to initiate and maintain the sort of social interaction the child has become used to. Social roles must play an important part in language acquisition (Ochs and Shieffelin 1995). Cross-culturally, people rely on similar linguistic means to accomplish similar social ends. However, these ends are culturally organized according to their situational scope and significance, that is, who, when, where, and how these ends may appropriately be accomplished and what they mean to the participants. Cultures differ in how children are addressed and how children are expected to speak.[15] Ochs and Shieffelin (1995) suggest that the primary goal of language is to socialize infants into culturally appropriate persons. Repeated interactions with their environment provide children with "scripts" that serve as the substrate for initial language learning. Social processes, such as joint attention, script construction, and correction, form a foundation for language acquisition (Hirsh-Pasek and Golinkoff 1996).

---

[15] For instance, children may be addressed differently based on their age and/or sex. Children may only rarely hear certain words that they are, nevertheless, expected to use when they address adults. In some cultures, parents expect their children's first words to be certain culturally determined and highly conventionalized forms. For example, in Kaluli society, everyone's first words are "mother" and "breast." In Samoan communities, the child's first word is part of a curse (Ochs and Shieffelin 1995). In European and American societies, *mama* and *papa* or *dada* are expected to be and very often are a child's first words (de Boysson-Bardies and Vihman unpublished).

Some social interactionists believe socialization provides a complete explanation for language acquisition but others see it as only a way of facilitating word learning.[16] Catherine Snow (1981; 1986; 1999) agrees that syntactic and semantic bootstrapping are important for children to acquire language knowledge but sees social precocity as a more likely innate feature that allows children to learn their first words. By the end of their first year, children have developed the capacity of *intersubjectivity*, the understanding that other people have minds like theirs. The understanding that other's minds are like their own leads to an acceptance of conventionalized communication, something Snow considers a trigger for the children to begin word learning. Children's expectation that other people will interpret the signals they produce as intentional attempts to communicate provides them with a motive to produce words.

Several researchers believe that cognitive development plays a primary role in the acquisition of language (Bates et al. 1979; Slobin 1992; Peters 1986). The processes a child uses to learn words parallel those of other cognitive developments. These include paying attention to perceptually salient stimuli, remembering the stimuli, and classifying the stimuli. Children use these abilities to make sense of words the same way they make sense out of other cognitive domains. Bates et al. (1979) suggest, for instance, that the ability to use tools is accompanied by cognitive advances such as planning, sequencing of steps, and part-whole analysis. Tool use is related to language – both begin the same way cognitively but diverge in their requirements of substitutes (Bates et al. 1979). Other cognitive abilities that have been correlated with word learning include imitation and

---

[16] That is, learning to produce words, not necessarily the same as *language* learning. See also Section 1.3.5.

means-ends abilities (Bates et al. 1979; Uzgiris 1981). These cognitive prerequisites for language develop continuously, and handicaps in one cognitive area may be overcome by strengths in other areas (Bates et al. 1979; Ingram 1981). Bates speculates that these cognitive skills may even develop independently of each other until each reaches a critical threshold, allowing the child to begin to use words.

2.3.2 Does the Synthesis Approach Work?

We can now list a number of things about language acquisition that many researchers agree on. For instance, language acquisition is generally believed to be the result of a complex interaction of factors (Slobin 1992; Elman et al. 1996; Hirsh-Pasek and Golinkoff 1996). Learning a language depends on culling information from a number of different inputs (Hirsh-Pasek and Golinkoff 1996). Solving the language puzzle requires input from both innate constraints and environmental input (Pinker 1987). Clearly, the solution to any human problem, linguistic or otherwise, lies in the interaction of three factors – genes, environment, and the structure of the problem (Bates et al. 1979). Genes and environment provide causal input while the task structure provides the rest of the solution. Children are sensitive to properties of their own language such as prosody, stress, and phonemes at a very early age, earlier than has previously been recognized. Slobin's (1992) cross-linguistic research on child language acquisition provides evidence that some sort of maturational process is also involved.[17] While not agreeing on exactly what they are, there is at least agreement among researchers that

[17] For example, in a polysynthetic language such as Greenlandic Eskimo, input consists of complex verbs, and children go through a single *morpheme* stage, progressing to a *two-morpheme* stage.

many factors, some innate, some derived from the input, are involved in language acquisition.

Hirsh-Pasek and Golinkoff (1996) suggest that a compromise is possible between nativistic and emergentist theories by collapsing the dichotomies that separate them. They say that ideas about child language acquisition all fall on a continuum. Theories of initial language *structure* range from the purely linguistic to social/cognitive; theories of language acquisition *mechanism* range from the domain-specific to the domain-general; theories about the *source* of the initial structure of child language range from innate predisposition to construction from input. Theories tend to group themselves into families - the nativists who argue that critical parts of early linguistic structure is domain-specific and innate, and the emergentists who argue that the cognitive/social features of early language come from domain-general, constructed sources. The development of a synthesis theory of initial word production depends on showing that both sides share common ideas.

For instance, emergentists agree that a minimum requirement for word learning is the linkage of "acoustic packages," containing prosodic and phonemic properties, with events (Hirsh-Pasek and Golinkoff 1996). This sound-event mapping is not necessarily linguistic but may even be the sound of music associated with the movement of a mobile. Cognitive and perceptual limitations may provide the source of some of the constraints on what a child can learn and use (Griffiths 1986). At least some nativists (Pinker 1987; Pinker and Prince 1988) recognize the importance of domain-general processes, and emergentists (Karmiloff-Smith 1993) agree that some parts of the language acquisition process are innately specified, although they don't *necessarily* have to be domain-specific in the beginning.[18] Karmiloff-Smith (1993) suggests that domain-specificity may be a developmental process – initially domain-general functions become domain-specific as the child matures and acquires language. Many researchers have suggested that social interaction (Grieve and Hoogenraad 1986; Ochs and Shieffelin 1995; Hirsh-Pasek and Golinkoff 1996; Snow 1999) and cognitive development (Bates et al.1979; Slobin 1992; Peters 1986; Ingram 1981) are important factors in the ability of infants to learn words.

Hirsh-Pasek and Golinkoff (1996) argue that the disagreements about whether or not language is innate would collapse if emergentists were to grant that children are predisposed to search for certain kinds of information in their environment at the time of first word production and can derive a limited class of linguistic generalizations at that time. Whether innate or the product of domain-general processes, certain language-relevant information must be available to children by the time grammar learning begins. For Hirsh-Pasek and Golinkoff, the question is not whether language is innate and domain-specific, but rather how much language-specific knowledge and what kinds of domain-specific processes are necessary to ensure that language learning takes place? In the next chapter, I will point out specific problems with each of the presented theories and argue that a synthesis is only possible if we assume that what is innate is the structure of the brain and the desire to communicate. Children's first words represent learning but are not language in the generative sense; they are not mapped to innate categories, and are not indicative of real language knowledge when they are first produced.

[18] For example, Karmiloff-Smith (1996) says that the early plasticity of the infant's mind suggests that a radical nativist position must be wrong. Special *attention biases,* such as those for shape and size discrimination, can channel the way in which even a newborn processes linguistically relevant inputs.

Chapter 3

Critique of Reviewed Theories

3.0 Introduction

In Chapter 2, I discussed several theories which purport to explain how children begin to learn language. Each theory addresses the issue of what children's first words represent and how children acquire them. Each theory makes a reasoned and rational contribution to the problem. In this chapter, I will investigate some of the difficulties with each theory. Early word production is not satisfactorily explained by genetic programs or as an instinct. While emergentists have argued that the input is richer than Chomsky earlier claimed, computer models are unable to explain some features of the process. Some of the proposed theories explain the acquisition of first words better than others, but there are parts of each that seem to be able to answer different parts of the problem. The solution to the problem of child language acquisition may lie in a synthetic approach.

3.1 Nativist Weaknesses

Whether syntactic bootstrapping or semantic bootstrapping is claimed as the trigger that allows children to gain access to a Language Acquisition Device (LAD), nativist theories hinge on some kind of specific linguistic knowledge existing in the genetic code. This is a major claim of the nativist viewpoint. The apparently most compelling nativist argument, however, for the innateness of linguistic knowledge is the *poverty of stimulus argument.* According to this argument, there is not enough information in the linguistic input that children are exposed to, to account for the language competence that children achieve. Hence, it is logical to conclude that what is missing from the input was already present (Cook and Newson 1996). In the following sections, I will discuss nativist suggestions that language knowledge could be coded for genetically and the necessity for a poverty of stimulus argument. I will evaluate syntactic and semantic bootstrapping arguments, and look at the need to explain language learning as an instinct. Finally, I will evaluate claims that first words are linked to innate prototypes.

3.1.1 Why Genetics Can't Account For Grammar

Chomsky's conclusion that Universal Grammar must exist because language is too complex to be accounted for by the data that children receive seems, at first glance, like an incontrovertible argument. Surely, given the circumstances, there is no other way to account for the rapid acquisition of language. There are problems with this argument, however. One problem is that of genetics. If language knowledge is built into the child's brain, even if it develops slowly as part of the maturational process, it must somehow be coded for in our genes. And if it is coded in the genes, then one must be able to show that it is plausible to account genetically for a complex behavior to develop in the human being.

Elman et al. (1996) ask us to consider several important features of genetics that were not known until recently. First, the amount of DNA that must account for preprogrammed language knowledge is surprisingly small. Scientists have always assumed that there were thousands, perhaps hundreds of thousands of genes. Completion of the Human Genome Project produced the surprising result that there are only about

30,000 human genes (Pääbo 2001a). We already knew that, genetically, we are almost identical to our primate cousins (Pääbo 2001a; Pääbo 2001b; Deacon 1997; Dunbar 1997). The difference between human and chimpanzee DNA is only about 1.6%, now recognized to be only a few hundred genes. These few genes must account for all the many differences between the species, including language. The problem, however, is more complex than just the amount of DNA. It is not just a question of how *many* genes it would take to program grammar but a matter of what *kind* of genetic development is necessary for language in the human species.

Scientists studying DNA have discovered two kinds of genetic development. In *mosaic development*, cells develop relatively independently of each other.[19] However, while mosaic development works for small, simple creatures like worms, higher organisms have opted for the second kind of development – *regulatory development*. In regulatory development, the outcome of cellular division is under the broad control of DNA but the final outcome depends largely on interactions that occur while the young organism is growing. There are considerable advantages to regulatory development. Neighboring cells can often compensate for damage to another group of cells. Greater complexity in genetic expression is possible when genes interact with each other. The

[19] For instance, a species of worm, C. elegans, has been studied extensively. It develops rapidly and it appears that most of its cells seem preprogrammed to grow into particular organs with a particular organization. Each cell grows up to be what it is supposed to be, regardless of what the rest of the body does. While this might seem like the safest way for an organism to develop, there is a price: the genome must be close to a complete blueprint for the entire body. Single-celled animals have 10^6 to 10^7 base pairs making up the structure of the DNA code and there is a predictable increase in the number of base pairs up to the complexity level of the mollusks. After that, more complexity is not associated with more DNA. Higher animals level off at around 10^9 to 10^{11} base pairs. Flowering plants actually have more base pairs than mammals. It seems there is a limit to how much DNA can be stored in a cell and safely replicated across generations (Elman et al. 1996).

trade-off is that the process of development is slowed down greatly to allow these interactions the time they need to complete their processes. If humans developed the mosaic way, the entire organism would be created in just 47 binary divisions (Elman et al. 1996). Instead, our genes are algorithmic – they develop by way of algorithms rather than by descriptions. That is, they depend on regularities in the input rather than carrying encoded information. The interaction of genetics with the environment provides the solution to the developmental problems of higher organisms with only a small amount of pre-wired biases (Elman et al. 1996). Studies of DNA activity support the conclusion that it is the way the structure of the human brain develops that is responsible for the difference between humans and chimpanzees (Pääbo 2001b). Human and chimpanzee gene activity is very similar in the blood and liver but the overall rate of activity in the human brain is three times that of the chimpanzees. Because the differences in human and chimp DNA are so small, it cannot be the difference in the information that they carry but rather the way they are expressed that creates the very large differences in structure.

Another key point in the nativist argument is that there is an "organ" in the brain that is dedicated to language. Scientists used to think they had found such an organ when they discovered that patients with damage to Broca's and Wernicke's areas often have language difficulties (Payne and Wenger 1998; Pinker 1994). However, language problems are not inevitable. Studies have shown that damage to these areas does not always result in aphasias (Bates and Goodman 1999). Most important for early language acquisition, studies of infants with frontal lobe damage showed no delay in language if the damage occurred prior to 19 months of age. Further, MRI, PET, and CAT scan studies show that many parts of the brain are activated when speech is processed, not just these areas. Also, Broca's area is used for other tasks – adult damage to it results in loss of general cognitive abilities and all parts of Broca's area show activation during *non*verbal motor planning tasks. Finally, different parts of the brain are activated in different people even when they are processing the same linguistic data (Deacon 1997). Broca's and Wernicke's areas appear to be important to language and may be specialized in some way but are not entirely dedicated to speech processing and have no special status at all until some time after language begins to be acquired (Bates and Goodman 1999). In summary, recent discoveries about DNA and how genes are expressed cast doubt on the theory that language knowledge is encoded in our genes. Further, neurological evidence suggests that there is no physically separate language organ dedicated to linguistic processing.

3.1.2 The Stimulus is Not Impoverished

The poverty of stimulus argument is that the language input available to children is not sufficient to account for what a native speaker knows about grammar (Cook and Newson 1996). The strongest counter to this argument is the kind of computer modeling described by Elman et al. (1996) that was discussed in section 2.2.1. It is possible to design a computer program that can learn both grammar and semantics when presented with a large number of sentences (Elman 1999). Elman used 10,000 sentences as input data to his computer model. Although this may seem like an excessive amount of data, infants actually hear far more than that in the months before they say their first words (Kita and Dickey 1998). Infants between the ages of six and nine months hear over 65,000 utterances, most of which are either fully or partially grammatically correct (Allen and Seidenberg 1999; Meisel 1995. There is also evidence that children who hear more adult speech learn earlier and faster (Snow 1999; Fowler 1981). The more they hear, the larger their vocabularies and the sooner they acquire grammar. Hence, there is enough data available to children, most of it is well formed, and there is evidence that they must be using that data when they are beginning to learn words.

One argument thought to support the poverty of stimulus argument comes from Goldin-Meadow and Mylander (1993). They claim that deaf children raised in hearing households can develop a grammatical gestural communication with minimal input. They believe their study supports the theory that children come equipped with some sort of innate predicate calculus. Their interpretation of the data is that the spontaneous gestures hearing parents use are enough to trigger the development of a grammar system. Bates and Volterra (1984), however, point out that the input these children receive may be richer than claimed or even realized by the researchers. The children that Goldin-Meadow studied were not profoundly deaf and the most advanced children also tended to be children with the most residual hearing. Mothers were sending "complex multimodal messages" with children getting signals from lip-reading, facial expression, sound, gesture, and context. Psychologists have found that a pairing of two cues that are both below the threshold of perception can boost each other when they occur together (Payne and Wenger 1998). The question is not "How are children inventing language?" but rather "How are children taking a 'mixed media' message and converting it to gestural output?" Bates and Volterra also point out that Goldin-Meadow and Mylander assume that points are nouns and pantomimes are verbs and that this is not the same way as ASL categorizes where points are treated as pronouns and pantomimes are part of nouns.

Further, there were essentially no gains in grammar complexity across the duration of the study. The grammar that was observed was very primitive, probably the result of input that was not as impoverished as claimed, and it did not progress to a truly generative grammar. Hence, it cannot be considered good support for the poverty of stimulus argument.

There is also evidence that the input that children receive is specially structured to get their attention and provides them with linguistic clues at a time when they are beginning to notice vocal communication. Infant-directed speech (IDS) is different in several ways from adult directed speech (ADS). It uses higher pitch, greater pitch variability, shorter vocalizations, and longer pauses (Fernald 1993; Hirsh-Pasek and Golinkoff 1996). These differences hold cross-linguistically and may be biologically relevant signals that have evolved to ensure survival (Fernald 1993). IDS appears to be sculpted to the child's developmental level because its character changes as the child develops. Werker et al. (1996) note that four-month old infants prefer infant-directed speech to adult-directed speech and they prefer native to non-native speech. These infants not only prefer IDS, they seem to require it to make sense out of the speech stream since their ability to discriminate between different phonemes depends on the presence of IDS: By the time they are nineteen months old, children do equally well on phoneme discrimination tests whether they hear IDS or ADS. IDS may help infants pick out words as well as assisting in the development of syntax since they can detect phrase boundaries only in IDS, not ADS. Hirsh-Pasek and Golinkoff (1996) report that their preliminary attempts to study the effect of ADS on speech acquisition in infants indicate that this kind of research may inevitably fail because very young children simply don't pay attention to it. Prelinguistic infants and children who are just beginning to produce words require input that attracts their attention and primes them for their entry to language.

3.1.3 Problems With Syntactic Bootstrapping

Syntactic bootstrapping (see Section 2.1.1) proposes that children come to language knowing something about the structure of sentences. They learn their own grammar by setting certain parameters for grammatical principles. These parameters can be set with a minimum of input. The major problems with this theory concern how those parameters could be correctly set with the tools that children have at the time they begin to produce words and the timing of the development of prototype concepts.

Syntactic bootstrapping claims that children can locate triggers in the input they receive and set parameters for their grammar (see Section 2.1.1). Chomsky did not specify how children identify the triggers in the input. If children can identify these triggers, however, why can't we just present them with the ones they need and have them learn language immediately? Since there must be default parameters, some languages should be easier to learn than others. In fact, children never learn language instantly or even very quickly and all languages seem to be equally easy for children to acquire (Meisel 1995; Slobin 1992; Pye 1990).[20] Meisel (1995) questions how the theory might

[20] That is not to say that all features of all languages are learned equally easily. For instance, a number of investigators have noted that children learning English have some difficulty learning case distinctions of English pronouns (James 1990; Goodluck 1991). Japanese children experience difficulty learning to use accusative case markers on Japanese noun phrases. German marks case and gender on articles, adjectives, and pronouns. German children learn to make the appropriate distinctions by age 2;2 on adjectives but make errors with articles as late at 5;6. German children over-generalize the nominative form of the articles to objects. Polish children begin to acquire case marking on nouns as early as 1;7 and quickly develop the markers for nominative, accusative, and genitive. Turkish nouns are acquired early and children use the accusative case marker productively by age 2;0 (Pye 1990).

allow for parameters to be reset if they are initially wrong. Meisel surveyed data from American children and their parents that showed four percent of parental replies were ungrammatical and sixteen percent were acceptable but not fully grammatical. The notion of triggering implies that, in principle, an extremely small number of examples should suffice to set parameters. In principle, twenty percent or even four percent ungrammatical sentences are likely to make the task impossible because the child has no way of knowing whether a given sentence is ungrammatical or whether the parser needs to be adjusted.

Several researchers have found that children can learn new verbs by using syntactic frames (Gleitman 1993; Goldberg 2000; Goldberg 1999). Goldberg (1999; 2000) has shown that speech to children contains certain basic templates that children can use to infer the meanings of novel verbs. Syntactic frames can be shown to drive vocabulary acquisition as early as 20 months (MacWhinney 1998). However, this research has been unable to demonstrate that very young children use an innate knowledge of structure to learn their first words. The problem that this thesis addresses is the question of how children acquire their very first words, not how they acquire additional vocabulary once they know grammar. Children who are learning their very first words are neither semantically nor syntactically precocious. They do not parse, analyze, or understand reality, and they are notoriously conservative in their use of novel structures. Once they have progressed beyond the single-word stage, syntactic bootstrapping can help children quickly acquire new words and grammar but this is probably not how children learn their first words (Snow 1999; Bloom 1993).[21]

3.1.4 Problems With Semantic Bootstrapping

Semantic bootstrapping theory differs from syntactic bootstrapping in that children are seen as innately constrained to make certain kinds of guesses about word meanings and they can get these meanings from context. They listen to the words used to describe things in the world and try to figure out their meanings by observing the associated situation (Pinker 1994). If a child figures out the meanings of the words, figuring out syntactic structure becomes easier – the child can use what is known about meaning to induce syntactic frames (Pinker 1994; Snow 1999). Semantic bootstrapping theory predicts that children produce words when they understand what they mean. Hence, for these theorists, word production is equated with language acquisition.

The kind of reasoning that allows a person to guess word meanings from a few specific examples is called *inductive* reasoning. Children, says Pinker (1994), are able to induce word meanings because they are innately predisposed to make certain kinds of guesses. One of these kinds of guesses is that a word said in the presence of an object will be the name of that object rather than one of an infinite number of alternatives.[22]

The main problem with this theory is that infants are not particularly good at inductive reasoning (Snow 1999). For children to use semantic bootstrapping, they would need to understand what is going on in the world. But they usually learn about the

---

[21] *Word learning* is not equated in this paper with *language learning* or *language acquisition*. See also Section 1.3.5.

[22] For example, if a child hears the word *rabbit* while watching a rabbit running, alternatives might include such ideas as "scurrying rabbit," "scurrying thing," "undetached rabbit parts," etc. (Pinker 1994).

world through language rather than learning about language through their knowledge of the world. Further, speech to children is often simplified but not limited to canonical descriptions of prototypical scenes (Bowerman 1993). There is, in fact, evidence that children can't use semantic bootstrapping to learn the meanings of all words.

For example, Gleitman (1993) argues that children cannot learn the meanings of even simple verbs by attending to the context of use because there is both too much information and not enough information to learn meanings. She focuses on verb learning because it is in this area that it becomes clear that the observer who notices *everything* about a situation can learn *nothing*. There is no end to the number of ways a situation could be described. However, even blind children, who cannot use visual observation, learn subtle differences in the meanings of *look, watch,* and *see*. Further, many verbs are only remotely related to the observable world. Gleitman suggests that words such as *think, guess, wonder*, and *know* cannot be learned by observation and, hence, must be acquired from clausal syntax. Gleitman's argument leaves us with the conclusion that if children were to attach meaning to words at the time they first produce them the way semantic bootstrapping claims, they would not be able to do so for all word meanings without some additional cues.

3.1.5 Is Language an Instinct?

Pinker (1994) has said that language acquisition acts like an instinct. It is no more related to culture than standing upright. Further, he says that it bears all the hallmarks of an instinct because it is a complex skill, acquired without instruction, and is the same in every individual. He says it is acquired much like birds learn their songs and that there is a critical or *sensitive* period when they must be exposed to input for them to acquire it. I suggest that Pinker is wrong. While it is certainly a complex skill, language is also intricately related to culture and social expectations (Bate 1979, 1995; Hurford 2000; Snow 1999; Vygotsky in Rieber and Carton 1987; Deacon 1997; Grieve and Hoogenrad 1986; Boysson-Bardies 1999; Ochs and Shieffelin 1995). Language also does not compare well to birdsong. There are great differences among birds' needs for learning. Some birds can sing their songs without any input at all while others must hear their species' songs during a critical period in their life or they can sing only a very primitive kind of song (Marler 1993). Assigning language the status of an instinct means that children must learn language this way during a narrow window of time.

While they do learn a lot in the first 2 or 3 years of life, children still have much to learn (Bloom 1993). James (1990) reports that significant language learning normally occurs even after the age of 5. Moeser (1977) specifically challenged the notion that children have special language-processing abilities that are lost by puberty. Her studies with mini-languages show adults can learn to use languages as well or better than children.[23] It appears that, whether children break into language through some kind of bootstrapping or not, whether they have pre-programmed language knowledge or not, it is unlikely that language acquisition is the result of an instinct. In the discussion of infant development (see Section 1.5), we saw that children vary widely in their cognitive development, word learning, and even language acquisition and, while language is a

[23] Miniature artificial languages allow the experimenter to simplify the task of studying language acquisition by restricting the number of variables in the experiment. These languages consist of a symbolic system, a rule system, and referential system corresponding to the phonological, syntactic, and semantic components of natural language. Each component, however, is less complex than in natural language and, hence, easier to analyze. Investigators assume their findings are relevant to language acquisition and use because the psychological processes they study, such as word order, rule systems, and the acquisition of syntactic code, are equivalent to the processes involved in natural language (Moeser 1977).

complex skill, it is not subject to the same kind of critical period restrictions as the acquisition of birdsong.

3.1.6 Innate Prototypes?

Nativists claim that children have innate prototypes and concepts that they link to new words (Pinker 1987). In an earlier discussion on prototype development (see Section 2.2.1), computer modeling showed that prototypes are not necessarily innate but may emerge *after* a word can be produced rather than before. Gopnik and Meltzoff (1993) focus on assessing cognitive development and correlating it with linguistic development. They find that, although children as young as nine months appear to have some early understanding of categories and can identify named objects, only children who have experienced what they call the *naming spurt* are able to sort objects into groups. Further, they found cognitive differences between children learning English and children learning Korean. Korean mothers use more verbs when they speak to their children which, Gopnik and Meltzof conclude, accounts for the fact that Korean children have more verbs in their early vocabularies and may experience a *verb spurt* before a naming spurt if, in fact, they have a naming spurt at all. They conclude that since English-speaking children learn more names of things at an earlier age, they develop the ability to sort by categories earlier than Korean children. What children hear apparently influences what they say but, more importantly for concept development, what they say influences how their cognitive abilities develop.

Semantic bootstrapping theory suggests that children are able to untangle syntax from the meanings of words. But this would require that 1) children learn words in sentential context and 2) when they do learn words, these words fill grammatical categories for which children have innate concepts. There is evidence that neither of these requirements is satisfied when children first begin to produce words. Ninio (1992) collected data from 48 mothers and 24 children who averaged 18 months of age. He recorded the utterances of both mothers and infants and found a direct correlation between the single word utterances of the mothers and the vocabularies of the children. In other words, the single words the children spoke were apt to reflect the single-word utterances of the mothers. While both mothers and infants spoke these words in situations where there was shared attention to named objects, there was no sentential context. Barrett (1995) also studied infants' early single-word production. He found that the context of their use indicates that, at the beginning of lexical development, children do not group words into grammatical categories in the way the words later function in the children's grammar. Griffiths (1986; 1979) says that first words may be *holophrases* - not words, as we understand them, but *word-sentences*. Words cannot initially fill grammatical categories because, when first acquired, they are *proto-sentences* that are just one *proto-word* long. Children produce early words in the absence of sentential context and these words do not necessarily define grammatical categories. Hence, innate prototypes must not be necessary for children to begin learning words.

3.1.7 Summary of Nativist Theories

Bootstrapping theories place a heavy requirement on children's early abilities in order to account for how they first begin to learn words. Snow (1999) concludes that infants are neither semantically nor syntactically precocious and can't use semantic bootstrapping or syntactic bootstrapping to begin producing words. There is not enough information storage in DNA to account for what nativists credit children with knowing.

Language learning does not compare well to instincts because humans are not limited to learning language during a critical period of time and children do not seem to learn language as quickly and easily and with as little input as once assumed. In fact, children work hard at learning language and they probably require a very large amount of input over a long period of time.

3.2 Emergentism and the Problems with the Computer Models

In recent years, a number of researchers have developed the theory that language emerges from the interaction of small differences in non-specific abilities. This idea has been supported by the development of a number of computer models designed to mimic the acquisition of certain language features in children. Although promising, these models and the ideas behind them suffer from problems of their own. Among them is the continuing argument that without innate constraints, children would have too much information to be able to decide how to attach meaning to a referent. There are also specific problems with the way the models work. One of these problems has to do with causality and the other is the lack of sufficiency of explanation by the models. Computer models ignore social and emotional factors that will be seen to be of critical importance in the acquisition of first words.

3.2.1 The Causality Problem

Throughout the discussion of how children learn their language, the different explanations have omitted one critical and important issue. *Why* do children talk? What *causes* word production to begin? The most cleverly designed computer program cannot answer this important question. Bates et al. (1979) describe the four kinds of Aristotelian causality and their importance to language acquisition. *Efficient* causality is the "push-from-behind" kind that can explain why a book fell off the table. Natural selection is considered an efficient cause of evolution. Behaviorist psychology seeks efficient explanations of behavior. Bates does not believe that any efficient cause is sufficient to "push" symbolic communication into existence.

Material causation – the presence of material conditions necessary for something to occur - is also not sufficient to explain how language develops. Having flour, eggs, and water available does not make a cake happen. Similarly, having the physical equipment to produce language does not make language happen.

Final causation is the "pull-from-the-front" kind of cause. Desire to reach goals can be a form of causation as in the operant conditioning experiment where a rat expects to get food when it pushes a lever. Computer programs are subject to a kind of final causation when they are programmed to run until a final goal or answer is achieved.

The last kind of causality is *formal* – the laws governing the range of possible outcomes in a given situation provides the structure or form required. For example, the laws of physics combine to cause a rock to fall in a given place when it has been thrown. The amount of stirring required and effects of heat limit how a cake may be baked.

Language learning may require all four kinds of causality and we know that computer models do not provide them in the ways that a child's brain and the world do. The programmer has too much control over the program to be certain that the answers he gets represent human language and the way that it is acquired. We do not know enough about brains, children, and computers to be able to assume that what causes word production to begin in a model is the same in a model as it is for a child.

3.2.2 Are Neural Networks a Satisfactory Solution?

The connectionist models presented in Section 2.2 are very good at categorization but have difficulty predicting other forms of behavior, including social and emotional behavior. Since the system does not record what items have been presented, one cannot predict performance. No matter how good computer models might be at categorization, there are empty categories that children seem to know about. Computer models have not yet been able to explain how things that are absent in the surface structures of sentences could be acquired at an abstract level (Hirsh-Pasek and Golinkoff 1996). Simple connectionist models can only learn categories that are linearly separable. Models with hidden nodes can solve more complex categorization problems but they can only learn word meanings when the words are presented at the same time as the referent. Hence, while they can model learning, it may not be the way that children learn since children attend to the object the adult is talking about only 30 percent to 50 percent of the time (Bloom 2000). Also, the models are so flexible that minor alterations in their design can cause major differences in behavior (Payne and Wenger 1998). If the addition of one node to the model makes it possible for the model to categorize in ways that the previous model couldn't, was the previous model a failure? Is the new model really a new model? If two different models behave the same way, how can the programmer know which one is better? These questions point out that a serious problem with computer models is that they may be able to learn words but may not represent how children learn words.

Pinker and Prince (1988) and Pinker (1993) criticize the Rumelhart and McClelland connectionist model of the acquisition of English past tense[24] by pointing out serious general problems with the neural network computer model and its abilities to explain language acquisition. The model was unable to represent certain words, it had a tendency to learn rules that no human language contains, and it was unable to explain morphological and phonological regularities. Pinker concludes that the model actually failed in its task because it gave an incorrect explanation for some developmental phenomena and gave accounts of others that were indistinguishable from rules. Pinker argues that the model mimicked the classic U-shaped learning curve only because of unrealistic assumptions about the input. He says that children do not overregularize based on any properties of the input – rather they overregularize when they learn the regular rule. Further, the model did not actually produce *words* or even *phonemes*, but only the *features* that the answers would contain. There was also no attempt to relate these answers with lexical features. Pinker and Prince (1988) and Pinker (1994) argue that a past-tense rule must apply to a verb stem stored in the lexicon, not just a set of phonemes. Hence, the model could never differentiate between such words as *break* and *brake* or between *wring* and *ring*. Further, the model treated both regular and irregular verbs qualitatively the same way, which is not how they occur in the language. Pinker notes that when verbs are formed from words of other parts of speech, they are subject to

[24] Rumelhart and McClelland (1993) designed a model that translated words into phonetic features and then predicted what phonetic features the past tense would contain. Presented with a small number of irregular verbs, the model quickly learned to produce the phonetic features of a correct past tense. When they added a large number of regular verbs to the input, the model first made more mistakes and then gradually learned to produce mostly correct answers for both regular and irregular verbs. The model also produced answers such as *membled* for the past of *mail* and *squawked* for the past tense of *squat*.

the rules for regular stems regardless of their sound. Computer models are powerful learners but this is more a liability than an advantage. If computers are to correctly model human learning, they must account for the constraints and rules that human learners appear to have.

Neural networks have been used with success in the modeling of many specific individual linguistic processes and many researchers now agree that they represent part of the answer to language acquisition questions. For instance, in the case of past tense learning, there may be two processes involved, the acquisition of a regular past tense rule and the learning of irregulars by associative memory (Pinker and Prince 1988; Pinker 1994). However, computer models suffer from a number of technical and theoretical problems. All of them involve the fact that, while they can mimic small parts and properties of human learning, they are not human and are either too powerful or not powerful enough to account for language acquisition. Redington and Chater (1997) discuss this problem and conclude that there must be many sources of information involved. They have suggested that a combination of different kinds of models hold promise for future research but it is currently impossible to achieve a human level of performance with any single computer model.[25] Language may emerge from domain-general processes as the models suggest, but it has not yet proven that they can explain everything about early word production.

[25] Other models include probabilistic models that use statistical properties of language and distributional models based on relationships between linguistic units (Redington and Chater 1997).

3.3 Is There an Alternative Solution?

In the discussion on synthesis solutions in Section 2.3.3, I showed that some researchers believe the differences between these theories may be collapsed (Hirsh-Pasek and Golinkoff 1996). These theories can be arranged on a continuum and none of them are completely opposed to the ideas of the others. There is more agreement between the sides than is first apparent.

First, everyone agrees that there is something special about humans that must be present when word production begins (Hirsh-Pasek and Golinkoff 1996). The child must have some kind of sensitivity to input information and be able to arrange that information. Nativists focus on innate knowledge, and emergentists suggest that general processes allow children to pick out words, associate them with a context, and begin to produce them.

Second, there must be something that drives the process forward. Hirsh-Pasek and Golinkoff (1996) propose that there is no single impetus that causes language to be acquired but rather a large number of things that encourage children to learn words. These include environmental events, social interaction, prosody, and syntactic patterns. They suggest that children "live in a benevolent world in which these input sources covary reliably with one another." Word production is not only affected by many things in the child's world but actually depends on the predictable presence of all of them in varying degrees at specific times.

3.3.1 Problems With the Current Synthesis Proposal

There are several problems with the synthesis as currently proposed. First, the suggestion that there is something special about humans because they learn language is not an answer. The question is not whether there is something special about humans, but what it is that is special. Here, the nativists and emergentists disagree substantially. Nativists claim that evolution resulted in special language-processing mental structures that encode language knowledge, including the appropriate mechanisms to find meaningful units in the speech stream. Emergentists claim that small differences in general human mental processes work together to create something that did not previously exist. Hirsh-Pasek and Golinkoff have not reconciled this difference in theoretical view and, in fact, they reveal that they side with the nativists when they say, "Our opinion is aligned with a kind of process-oriented, inside-out view in which children come to the learning task with some sensitivities to properties in the input that are informed by internal grammatical knowledge."[26] They claim that both the "something" that is present when word production begins and the mechanism that allows children to bootstrap into language must be answered by the nativist stance. For them, a synthesis solution needs a domain-specific module with built-in language knowledge.

Finally, they argue that a well-timed coalition of environmental factors drives the process forward. They propose that children are sensitive to a range of environmental

[26] "Inside-out" theories are the nativist viewpoint - those theories that contend language is the result of a domain-specific language module in the brain and that there is hard-wired, innate knowledge about language in a child's brain. This is contrasted with "outside-in" theories – those that claim language structure is in the environment. These theories claim children construct language from general cognitive and social processes.

inputs that, presented in the proper sequence,[27] drive the process forward once it gets started, but they don't explain why children learn their very first words.

3.3.2 Other Possible Syntheses

While there are other synthesis theories that attempt to provide answers to the question of how children acquire language, they start from the point at which word meanings have already been acquired and then attempt to answer the question of how grammar arises. Elman et al. (1996) argue that human brains evolved to be able to solve problems and language is one of them. They believe that children are born with the general processing tools and a drive to solve the language problem. The grammar they develop is the result of a natural solution. Pinker's (1987) Constraint Satisfaction model is a similar explanation but with a nativist view. He agrees that language is a problem that children find a natural solution for, but believes that children need innate constraints in order to solve it. Pinker's argument assumes children have innate prototypes for word categories that they match to input representations as words are learned. Emergentist theory holds that prototypes develop *after* words begin to be learned. Pinker and Prince (1988) allow that a connectionist model might be able to answer some kinds of language acquisition questions but maintain that symbolic rule-based learning will always be necessary to explain grammar. While these alternative ideas suggest that some kind of synthetic theory might answer language acquisition questions, they do not appear to

[27] For example, they claim that the infant is especially sensitive to the basic units of language (words, phrases, clauses) and to basic relations such as agency and location and that the environment provides these things at the right time and in the right order.

address the primary problem of this thesis, which is how and why first words are produced and what they represent to the child.

3.4 Summary of Current Views and Their Weaknesses

In summary, nativist theories say that children learn too much too quickly to be accounted for by the input they receive. What is missing is attributed to innate knowledge and this has been called Universal Grammar. What UG actually consists of has been the subject of much debate with claims ranging from built-in rules for grammar to simple skeletal limits on what can be learned. On the other hand, emergentists believe that the input children receive is sufficient to account for early word production. Computer models have shown that at least some of the constraints necessary may be contained in the input. However, computer models do not explain what motivates children to learn to say words in the first place. Synthesis arguments that propose to take the best of both sides invariably either fall back on innate knowledge or assume the child can somehow get the process started without explaining how.

3.5 An Alternative

Do we know enough about genetics and children to be able to develop a theory of language acquisition that does not collapse into either a nativist theory or a connectionist model? I propose to look at how a child develops cognitively during the time when first words appear and relate that understanding to what we know about the evolution of our species and of language, borrowing some ideas from the fields of psychology and anthropology to develop an argument that will include innate social precocity and emergentist ideas. I will look at how cognitive development interacts with general attention-getting mechanisms leading to early word production and the importance of social interaction in this process. We will find that certain kinds of cognitive development are prerequisites to language and may be innate but others emerge only after first words are produced. Researchers generally agree that older children can use both syntactic and semantic bootstrapping processes to acquire additional vocabulary and grammar, but *first* words are learned through domain-general processes and are neither tied to a specific referent nor are they the result of an innate search for object labels. Certain innate but general cognitive skills and innate social skills provide the ability and motivation to produce first words.

3.5.1 Cognitive and Social Development in the Pre-Linguistic Infant

What occurs before the infant can begin to talk? In addition to generally acknowledged physical maturation, it appears that certain kinds of cognitive development must occur. I am claiming, however, that the cognitive skills necessary for children to learn to produce words emerge from general processes that work together. An infant must first be able to pick out significant chunks from the speech stream in the environment and must be motivated to do so. There is evidence that the ability to segment words arises from general mechanisms, and that innate social skills of humans motivate parents and infants to communicate with each other.

What we know about the segmentation of the speech stream indicates that an infant needs and uses several different kinds of clues to pick out meaningful units. As discussed in Section 2.2, it is not just one feature of language, rather a multitude of clues combine to give children the tools to pick out salient parts of the speech stream by the time they are ready to begin producing their own words. What drives this process, however, is the infant's developing cognitive abilities and need for social interaction.

Lock (1997) speculated on the relationship between the cognitive development of infants and the acquisition of language and suggested that the newborn's world is primarily child-centered. A very young infant cries because of discomfort of some sort and wants that discomfort to go away. Very quickly, the infant begins to focus on events, objects, and people in the outside world and learns that an action achieves a goal, it can be used as a tool to achieve that goal again. In addition, both the infant and the caregiver are willing participants in a social game that has evolved because it results in better survival (Fernald 1993). The mother is genetically programmed to ensure species survival and is willing to expend energy to get an infant's attention with sounds as well as physical interaction. The infant pays attention, not only to get fed, but also to get the kind of social interaction needed to thrive. This includes certain kinds of game playing that has verbal content. In Sections 2.2.2 and 3.1.2, I discussed evidence that shows the speech of adults to infants is generally different from speech to adults in acoustic content and meaning, and is effective in establishing emotional communication (Fernald 1993). An infant is predisposed to interact with others from birth (Grieve and Hoogenraad 1986). The child's world is a social world and mother is usually a constant social companion. Only in institutions is the child likely to be isolated from social constancy and stability, and it is a common observation that linguistic development is slower when otherwise normal children are institutionalized. Bates et al. (1979) find that the way mothers interact with their children has an impact on cognitive development and, ultimately, language development. The beginnings of communicative behavior lie in early infancy.

Researchers in other fields have also suggested that it is social factors that prepare the infant for entry into the world of language. Anthropologist Robin Dunbar (1996) links social communication to the evolution of the human species. He points out that humans have evolved the largest and most complex social groups of any species. His research shows that social group size is correlated with brain complexity. Our ancestors, like most primates, undoubtedly used grooming to maintain social bonds. Grooming works well in small groups but the amount of time required makes it impractical as a means of social communication in groups of more than around fifty individuals. Dunbar claims that larger groups require both an efficient means of communication between members and a brain that can keep track of many complex layers of social interaction. He suggests that language provided the kind of communication necessary to maintain the large social groups that humans participate in. Hence, says Dunbar, language capacity evolved initially as a form of *social* communication.

Studying the evolution of language, Hurford (2000) designed a computer model that simulated the sharing of language in social settings. Hurford claims that social transmission is required for passing language from one generation to the next whether or not grammar is assumed to be innate. His simulations go through a "one-word" stage of communication with no grammar rules, only lexical items. One version of his model assumed that categories such as subject and predicate are understood (pre-programmed) at the time single words were learned and one version assumed that first words are whole propositions. That is, word classes develop after word acquisition from analysis of propositions. Both models succeed in developing language rules. These models point out that language rules can emerge either if they are innate or from analysis of input. The models also spotlight one of the difficulties with computer modeling. If both versions of the model work, is one better than the other? Or whether, as Hurford suggests, both processes are somehow involved, along with the social interaction that, he stresses, is required in either case.

Very young infants are weak and defenseless when compared to infants of other species. The main strength and survival tool that human infants have is their talent for social interaction. Even newborns prefer to look at adult faces and listen speech-like stimuli (Snow 1999). Snow reports that they can make eye contact and use smiling, quieting, and their own facial expressions to make and keep their parents' attention. Infants maintain joint attention with caretakers through gestures and context-embedded games and routines (Snow 1981). Cross-linguistic studies show that there are may be an instinctive way of speaking that gets their attention (Fernald 1993). They respond to the affective features of prosody when their caregivers talk to them using infant directed speech (Fernald 1993; Snow 1986; Thoman 1981). Non-verbal communication is important in the development of affective communication. For instance, *smiles* can communicate affect to the mother. Facial expressions are not random but occur in predictable patterns that mothers look for and *assign meaning to*. Mother-infant non-verbal communication, while necessary to survival, is not directly analogous to language but Thoman's research shows that successful affective communication correlates with cognitive development and later linguistic success.

Humans have evolved complex social groups and a need for social interaction. Infants are genetically programmed to respond to their mothers' vocalizations. They pay attention to and respond to these vocalizations. They enjoy and participate in social-oriented games and routines. This kind of social interaction reinforces learning and cognitive development and prepares them for later acquisition of language by allowing them to practice vocal and social skills that precede word production. Most importantly for the beginnings of language, infants are predisposed to attend to and attempt to participate in various kinds of social communication with their caregivers.

3.5.2 Cognitive Prerequisites for Word Production

No child is born with the ability to speak. Language is neither physically nor cognitively possible for the very young infant (Lieberman 1999; Lock 1997). Spoken language is physically impossible for the newborn because the vocal tract length must mature. Language is not cognitively possible until the child recognizes that people and objects can act on each other. Snow (1999) and Bates et al. (1979) propose that it is this cognitive and physical development that provides the already socially precocious infant with the tools for language. The cognitive capacities that have been identified as essential for children to begin speaking are tool use, imitation, and means-ends abilities (Bates et al. 1979; Uzgiris 1981). These general cognitive abilities appear to be necessary for children to be able to use words.

For example, Bates' et al. (1979) studies of cognitive development and language skills in a group of children between the ages of nine and twelve months showed that tool use, imitation, and means-ends abilities at nine months are correlated with word use at twelve months. Bates discovered that not all cognitive abilities are correlated with language. Object permanence and spatial measures do not predict later word production. However, means-ends abilities, imitation and symbolic play activities do presage later word learning.

There is also a correlation between means-ends task solution and attachment style, tying cognitive development to social interaction (Bates et al. 1979). Infants who are either not well attached to their mothers or overly dependent on them fare poorly in cognitive tests of means-ends abilities. Infants who have a secure relationship with their mothers are more successful on such cognitive tests. Since Bates et al. also show correlations between attachment style and gestural communication, and between gestural communication and language development, there must be an important relationship between mother-infant interaction and language development. Infants who have secure social relationships develop cognitive skills better and, subsequently, begin to speak earlier than infants who have poor interactions with their mothers.

This research shows that there are cognitive links between socialization and certain kinds of cognitive development (Bates et al. 1979). These cognitive skills – imitation, means-end abilities, and tool use abilities – can predict later word production. Before infants can speak, they must notice how they relate to their world learn that they can affect it. They care about their relationship to their caregivers because they are basically social beings and they actively work to find their role in this social arena. Their world-view is changing and it includes the people and things around them.

3.5.3 First Word Productions

At the time that infants learn they can control peoples' actions, they have already been practicing and doing it unconsciously for a long time. Bates et al. (1975) suggest that eye contact, gestures, and prelinguistic vocalizations have a systematic effect on caregivers without the conscious intent of the child. The next step is the use of these means to purposely control that effect. Finally, the child utters sounds that have the same purpose and effect as the earlier nonverbal actions. The questions that have not yet been answered include how the infant decides what sounds to produce and what meanings they may have.

We know that infants have learned to use a range of cues to pick out sound units from the speech stream and they have been babbling and practicing the sounds of their language. Eleven-month old infants prefer to listen to speech that is interrupted at word boundaries rather than at random spots in a stream of speech. This does not imply they understand the meanings of the words. Anything in the environment that is repeated, however, can acquire significance. Children are exquisitely sensitive to repeated sound/event sequences and can pick out repeated sounds and associate them with repeated events (McDonald 1997).

Furthermore, children have been learning how they relate to their social environment and they have learned that what they do affects the people around them. It is natural, then, to expect that the child's first use of words will be to elaborate further the process of regulation and structuring of social interaction with mother and others (Grieve and Hoogenraad 1986). First word productions are highly idiosyncratic and variable across children (Fenson et al. 1994; de Boysson-Bardies 1999; Boysson-Bardies and Vihman unpublished paper; Bates et al. 1995; see also Section 1.5). Early word productions have significance to the infant that may not be readily apparent to an observer. They may have no *referential* meaning to the child. Whether children eventually develop generative grammar from a domain-specific language module that is already present in the brain or language-specificity emerges from domain-general abilities, the kind of word meaning that is associated with language must arise later from use, environment, social interaction, and short-term memory development, and it is gradual. Computer simulations in artificial neural networks have suggested that this kind of meaning may emerge (Redington and Chater 1997; Gupta and Dell 1999) while nativists have made strong arguments that there are innate features that suggest certain kinds of linguistic knowledge are present in the mind of the child (Cook and Newson 1996; Pinker 1994; 1993; 1987). When the sounds that children make during babbling begin to acquire the phonology and prosody of the language spoken around them, mothers begin searching for meaningful words in the sounds their infants make. At some point, infants say things that sound like words. Mothers accept these sounds as first words and respond to them as if they were meaningful, especially when the child repeats them in appropriate settings. There is a great deal of evidence indicating that, before infants begin to speak, they are accustomed to interacting with their parents in contingent response patterns (Snow 1995; 1981) so that when they begin to use word-like sounds, they are already able to use the responses that they receive to strengthen their use of the sounds in appropriate contexts. First word productions, then, reflect an extension of social interaction routines with caregivers, not the linguistic import of the words themselves. Mothers' interpretations of these productions and subsequent response to them may help reinforce the associations that infants make between the form and the context.

While first words may seem appropriate in a given situation, the child is still experimenting and has learned to pair sounds with a situation but not with a concept (Capirci et al. 1996). Early word production happens in a setting that is both social and reinforcing. Actual concept formation develops gradually and is not dependent on an innate prototype. Even if words appear to have a referent, the first uses of those words are invariably are *context-bound* – associated with some highly specific behavioral event in the child's social world (Barrett 1986; Harris et al. 1988).[28] There are a few documented cases where infants seem to be using words in a referential sense even when they are used for the very first time (Harris et al. 1988; Bates et al. 1979). However, even in these cases, the non-verbal context at the time the mother used a word could be linked to the child's use of that word (Harris et al. 1988). In other words, whether early words are clearly context-bound or appear to have a referent, they are linked to the mother's use of them in a particular situation.[29]

While they expected to find a complex of factors affecting early word use, Harris et al. (1998) find that their data points to a single relatively simple factor – frequency of use by the mother. Some early word productions might seem to be referential but are usually initially used in only one context, that being the one in which the mother uses the word most frequently. Further, at least in some cases, the child uses a particular word because it is used in a context to which the child is particularly responsive. Both the child and the mother, through frequent repetition and familiarity with certain routines, jointly influence which words became salient. Hence, Harris et al. conclude that first words are produced because they are used often by a caregiver, have emotional salience

---

[28] For example, initial use of *duck* was only while hitting toys ducks of the bathtub (Barrett 1986); initial use of *car* was only while looking at a car moving on the street (Bloom 1973); initial use of *catch* was only while throwing a ball to another person (Barrett 1986).
[29] Harris defines *context-bound* words as those used only in one highly specific behavioral context. *Referential words* are those not tied to a context but used in at least two different contexts (Harris et al. 1988). See also Section 1.3.6.

for the child, and that these words are probably context-dependent whether or not they appear to have a referent.

This paper proposes that early word productions are acquired as part of a means-ends game and this works because children are socially precocious. First words are not the result of an innate search for object labels but the result of a desire to participate in a social game that has evolved to meet the needs of the species. The earliest vocabulary can be anything at all. The composition of children's very early vocabularies should not be predictable because children do not recognize categories of words, only that sound units can associate with certain situations in the world that they wish to participate in.

3.6 Summary of the Alternative Account of First Words

Nativist claims that input is not important and is usually ill formed are wrong (Snow 1981; 1999). Observed speech to children is usually simpler and *more* grammatical than adult-to-adult speech and it is abundant. Children ignore complex speech and only attend to what they can understand. They readily participate in social situations. By the time they produce their first words, infants are adept at social communication. They already understand feedback, how to pay attention, and how to get someone's attention. They first say words for the purpose of participating in social interaction. Social games lay the foundations necessary for later language learning. Humans are highly motivated to be social, to learn culture, and to "fit in." There is a strong pressure to gain group membership and approval. The establishment and maintenance of social identity is an essential human need (Dunbar 1997). Words are children's entry tickets to an exclusive club.

None of the current theories addresses all these issues – not even 'synthesis' theories. Here I have tried to develop an alternative theory – that the unique innate features in human brains that help children begin the process by learning words are the result of the specialized architecture of human brains and the social nature of the species. There are certain features of language learning that can be successfully modeled by computer programs but children need social interaction to develop necessary cognitive skills. It is in the area of social interaction that children are especially precocious and innately predisposed to respond and say words. Children's first words are probably not referential and are not initially linked to categories or concepts but are learned because they are a successful way to interact and communicate with their caregivers.

My theory claims that children's first words do not represent innately programmed categories as has been suggested by nativists. I believe that the variability noted in early vocabularies (Fenson et al. 1994; de Boysson-Bardies 1999; Boysson-Bardies and Vihman unpublished paper; Bates et al. 1995; see also Section 1.5) is significant because it provides evidence that first word productions are not predictable and that children are not innately predisposed to look for object categories nor are there innate linkages between words and categories as claimed by Waxman (1994). My theory predicts that an examination of children's very early vocabularies will reveal that their content is so variable as to be considered random. Further, the evidence suggests that first words are the result of social interaction and that they become meaningful language gradually as children acquire larger vocabularies.

In the following chapter, I report the results of a study investigating the statistical nature of children's very early words. The vocabularies of 29 children knowing ten

words or fewer and the vocabularies of 17 children knowing between 10 and 44 words were collected. A procedure for sorting the vocabularies into word-types was developed and the data statistically analyzed for deviations from averages to determine whether there were predictable patterns in these data sets. The results provide evidence that initial word production is not the result of a search for object labels. Rather first words are sounds that children hear often in particular contexts and learn to use for social interaction.

Chapter 4

A Statistical Analysis of Children's First Words

4.0 What the Study Tests

The main claims of this thesis are that children are not seeking to link innate categories to object labels when they first begin to learn words but that, instead, they learn their first words because they are seeking social interaction and these first words are those sounds that occur frequently in situations of social or emotional interest to them. This thesis makes two separable claims. The first is that, contrary to the claims of some nativist researchers (Pinker 1987; Markman 1993; Waxman 1994; Waxman and Kosowski 1990), children are not innately predisposed to link innate prototypes to word when they start to say words – specifically they are not innately looking for whole object labels. The second is that children begin to talk for social and emotional reasons. Although some researchers have shown that children as young as two years old appear to be able to use their knowledge of categories to help learn the meanings of novel nouns (Waxman 1994), these studies start late in the word learning process, at a time when words already have linguistic meaning. Hence, they do not accurately portray what happens when children produce their very first words. The present study is designed to test the first part of the thesis – to determine whether the very early lexicons of infants will reveal categorical patterns that indicate that they are searching for object labels or whether there is so much individual variation that no prediction can be made regarding what children's first words represent.

4.1 Method

4.1.1 Subjects and Data Collection

Twenty-five subjects were normally developing American infants living in monolingual English speaking households in Kalispell, Montana. Kalispell is a small community in Northwest Montana. Since the population is not large, any normal infants with vocabularies in the appropriate range whose parents were willing to participate were accepted. Four of the mothers were teenagers attending high school. These mothers were interviewed in person. Data samples were collected by having the mothers fill out a questionnaire in the presence of the interviewer. The remaining twenty-one subjects were children attending a local church day-care. These mothers were interviewed by telephone. The ages of the infants in Kalispell ranged from 8 to 23 months. Unless the parents had actually recorded and dated their children's earlier productions, data was only accepted if the children were producing the words at the time of the study. That is, unrecorded memories of earlier productions were not acceptable. Although it can be argued that mothers are not reliable sources for this information (Darly and Wintz 1961), Fenson (2000) showed that parent inventories are at least as reliable as observations made by trained researchers. Parent reports correlate well with observation and are internally consistent.[30]

Additional data were taken from several published studies. Twelve data sets were from a study conducted by Vihman (1996). These subjects were normally developing infants between the ages of 10 and 16 months from three different cultural backgrounds. The four American sets (infants aged 10-16 months) were collected at Stanford

University, four French sets (infants aged 10-14 months) were collected in Paris, and four English sets (infants aged 10-16 months) were collected at Rutgers University in England. The data were drawn from videotaped half-hour unstructured play sessions with the mother, collected bi-weekly or monthly in the children's homes. Data published by de Boysson-Bardies (1999) contained five normally developing French infants, aged 10 to 17 months. These data were also drawn from videotaped play sessions. The final data were from Barrett et al. (1991). Subjects were four monolingual British children who were followed from 6 months to two years of age. Data in this study were collected from videotaping play sessions and from mothers' inventories.

4.2 Analysis

The data were initially classified according to sex, nationality, language, and the number and types of words that the child knew at the time of data collection. These variables were included in preliminary data analysis because these factors have been linked to language development by some authors (Bates et al. 1995; Fenson et al. 1994). Because language development has been shown to correlate much better with vocabulary size than with age, age was not included in this study (Bates and Goodman 1999). For each word type, statistical models were developed to predict membership in that word category and to determine which, if any, variables were significant to explain the data. Additionally, the resulting models were tested to determine how well they explained the observations. For each model, the population of interest was children knowing between three and 44 words who were one of the three included nationalities – American, British, or French.

---

[30] See also Section 1.3.7.

4.2.1 Word Categories

The process of analyzing collections of first words is complicated by the fact that every researcher has a slightly different way of "coding" first words. In fact, there is no evidence that children themselves group words into categories (Barrett 1995). Nevertheless, some form of categorization was necessary to analyze the data. Researchers categorize early vocabularies differently depending on the focus of their research. For instance, James (1990) and Barrett (1995) report that researchers who emphasize the importance of pragmatic function categorize by communicative intent.[31] Barrett (1995) focuses on patterns of developmental change and categorizes words as context-bound, decontextualized, and over- or under-extended. Most researchers studying lexical development use standard word classes corresponding to parts of speech in traditional grammar (Smiley and Huttenlocher 1993; Fenson et al. 1994; De Boysson-Bardies and Vihman unpublished study).

My purpose was to compare the use of object labels with other kinds of words, particularly words used for social communication, such as interjections and greetings, so I categorized words as: nouns used as object labels, proper nouns, events (verbs), adjectives, pronouns, imitative sounds, and words with social or emotional content. I believe it is important to separate proper nouns from other nouns because they are cognitively different, at least for adults: proper nouns are singular referring expressions and do not represent a category of things. If there are innate categories and innate linking

---

[31] Dore's list *of primitive speech acts* includes *labeling, repeating, answering, requesting, calling, greeting, protesting,* and *practicing.* Halliday's list includes *instrumental, regulatory, interactional, personal, heuristic,* and *imaginative* (James 1990).

rules, these words could not be included in the same class as object names. I believe there are more dimensions or layers of meaning to proper nouns because they can be part of means-ends routines and may involve social feedback and reinforcement. They carry emotional content not associated with object labels and they may index situations, emotions, rewards, social and emotional gratification, and satisfaction of physical needs. Object labels represent only the object, its qualities, and the context in which it occurs. Like proper nouns, words categorized as having emotional or social content index situations, social and emotional gratification, and even satisfaction of physical needs. Words in this category include the interjections and greetings that often entail social routines. These words included *hi, hello, hiya, bye bye, nite-nite, thank you, scuse me, pat-a-cake*, and *boo*. Other words considered to have emotional content included *no, wow, ouch, owie, ow, whee, yes, yeah.*

Object labels include animate and inanimate objects such as *kitty, doggy, duck*, and *cracker, water, shoe.* Included in events were verbs such as *go, ride, walk* and also any word associated with action or request for action, a change, or a motion. For example, *go potty, go bye-bye,* and *down* were included in this group because they are either actions or requests for action. Although some of these "words" are written as two words, they are included because they function as single words rather than as combinations. The word *more* was included in the verb category because early use of this word has been documented as a request for action (Barrett et al. 1991).

The published data that I used were not previously categorized. Vihman's (1996) study focused on phonological development so word categorization was not important.

De Boysson-Bardies (1999) also studied phonological development and, although her work commented on the relationship between early word types and personality development, she also did not specify word categories. Barrett et al. (1991) focused on the relationship between occurrences of a child's use of words and the mother's use, not on word categories.

I allowed, as a word, any *sound uttered that seems to have meaning and is used for the purpose of communication* (the definition of a word that was discussed in Section 1.1.7). In addition, several mothers who I interviewed said that what they recognized as words were sounds their children had used on several occasions. My definition allows imitative sounds such as *vroom* and *choo-choo* and animal sounds such as *meow, bow-wow*, and *baa*.

4.3 Results

Table 4.1 shows raw percentages of words in each word class for very early vocabularies (10 words or less) and the overall percentages for all data sets. As is evident in the above discussion of word categories, there are cognitive similarities between words

	N	P	E	V	I	D	A	P+E
Proportion in word lists of 10 words or fewer than 10 words	22.8%	23.9%	23.9%	9.2%	8.7%	5.4%	6.0%	47.8 %
Proportion in word lists of more than 10 words	35.8%	22.4%	17.5%	10.8%	4.4%	2.3%	6.7%	39.9 %

Table 4.1: Overall percentages of word types in the collected data.
In this table, N = nouns, P = proper nouns, E = words with social/emotional content, V = verbs, I = imitative sounds, D = deictic pronouns, A = adjectives, P+E = all proper nouns plus all social/emotional words.

that have emotional or social value and the class of proper nouns. Hence, these two groups were combined for additional analysis in addition to being handled separately. See Appendix II for a complete listing of the sample data.

Table 4.1 reveals that the absolute overall proportion of proper nouns and social/emotional words decreases as the children's vocabularies increase while the proportion of nouns increases. For example, word lists of 10 or fewer words contain 47.8% proper nouns plus social/emotional words while later lists contain only 39.9%. On the other hand, the earlier lists contain 22.8% nouns while word lists with over 10 words contain 35.8% nouns. The effect seen here is actually stronger than the table indicates. This is because the word lists for children knowing more than ten words include their first ten words. Thus, the percentages for the larger word lists are affected by the earlier data. In fact, it is reasonable to assume that a child uses an even higher percentage of nouns later than earlier and an even smaller percentage of words with emotional or social content later than earlier. There were not enough data samples to be able to make valid conclusions for verbs, imitative sounds, deictic pronouns, and adjectives. That is, the data sample was too small and the variance too large to be able to analyze or make predictions about these word classes. They were omitted from further analysis.

The analysis asks two questions. First, is there a predictable proportion of each word class in a child's early vocabulary for children of the same nationality and gender? For example, is it possible to predict what percentage of nouns will appear in the vocabularies of American boys who have learned 10 words? Or do similar children vary too much to predict the percentage of nouns that they learn in their early vocabularies? Second, is there any trend in the composition of the vocabularies as the number of words increases? For example, does the proportion of nouns increase with an increase in the size of the vocabulary? These questions reflect the two hypotheses of the thesis. If the proportions of word classes are not predictable, then the first part of the thesis, the claim that early word production does not represent an innate search for nouns, is supported. If first words are not predictable, then it is clearly not possible to claim that they are innately linked to whole objects. If there were a trend in the composition of early vocabularies indicating that an early relatively high proportion of social/emotional words to nouns is reversing and, in fact, nouns are in the process of becoming more important, then this would support the second part of the thesis. Namely, the claim is that early word production is the result of social interaction routines but that later words begin to represent a developing understanding that words link to categories.

Before turning to the formal statistical analysis of these questions, it will be useful view the data graphically. As to the first question, Figure 4.1 graphically demonstrates the large amount of variability in early vocabularies. This figure plots the percentage

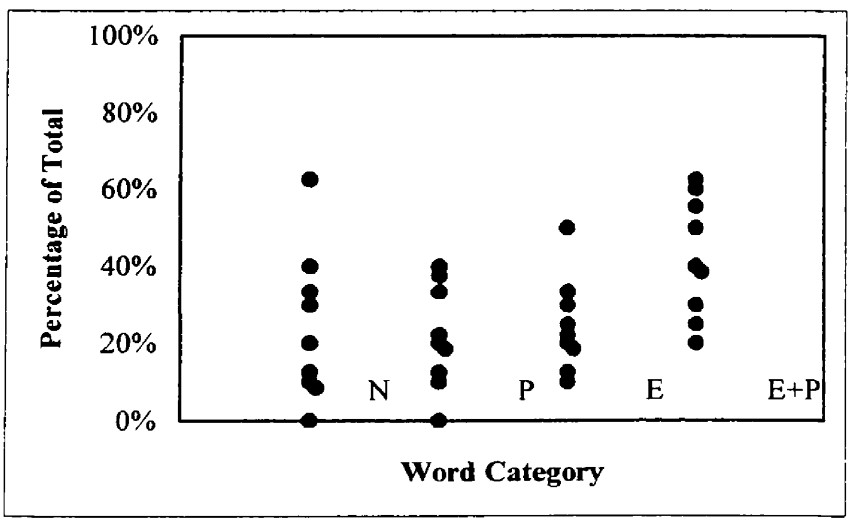

Figure 4.1: Proportions of total vocabulary for children knowing 8-13 words. N = nouns, P = proper nouns, E = words with emotional or social content, E + P = combined proper nouns and social/emotional words.

of nouns, proper nouns, words with social or emotional content, and combined proper nouns and social/emotional words for vocabularies containing eight to thirteen words. For example, in this vocabulary range, nouns may form anywhere from zero to over sixty percent of the total number words that a child knows. The same huge variability is seen to apply to the other word categories as well. At this stage of word production, the graph shows that there is a very large variation in the content of these vocabularies.

As to the second question, Figure 4.2 plots the number of words in each child's vocabulary against the proportion of words in the noun class. It reveals a slight upward

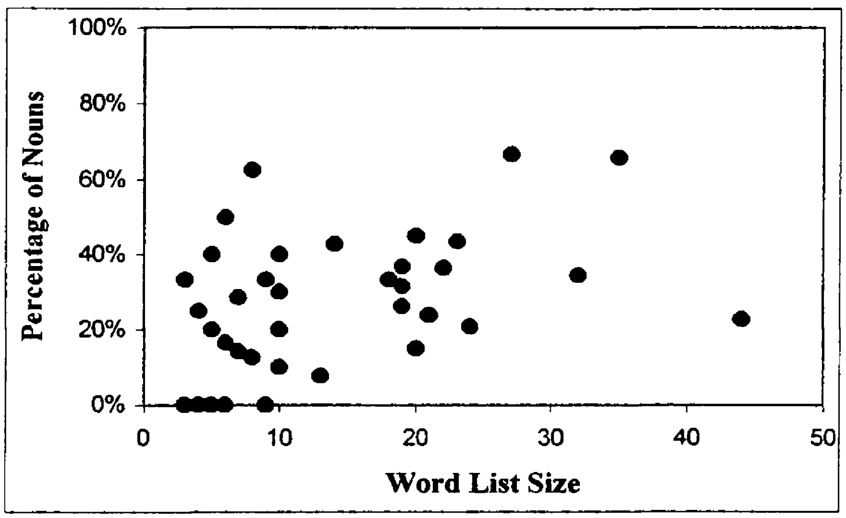

Figure 4.2: Vocabulary size (word list size) plotted against the proportion of nouns in the list.

trend in proportion of nouns as word list size increases. Figure 4.3 shows number of words known plotted against the proportion of personal nouns plus social/emotional words. It shows a slight downward trend in proportion of proper nouns and social/emotional words as word list size increases. These two figures suggest that the proportion of nouns is increasing while the proportion of social/emotion words is decreasing as vocabulary size increases. The two graphs provide a visual understanding of what is happening to very early vocabularies. If the relative proportions of each word

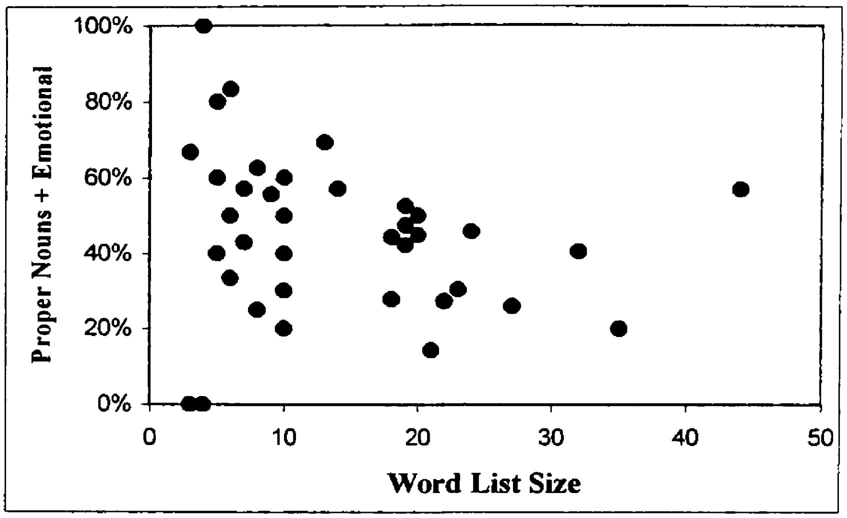

Figure 4.3: Vocabulary size (word list size) plotted against the proportion of Proper Nouns plus Social/emotional words (all words considered to have social or emotional import) in the list

category were constant throughout childhood, then we would expect to see no evidence of upward or downward trends in these graphs.

In order to determine whether there is formal statistical significance to the general observations made about the data, linear regression models were developed for the word classes of *nouns, proper nouns, social/emotional words*, and the sum of *proper nouns* and *social/emotional words.*[32] For each of the samples (vocabularies) the percentage of words in the category under consideration was the response variable for the model. In preliminary tests, nationality, gender, and the number of words in the vocabulary of each child were used as the explanatory variables for the model. It was discovered, however, that the only variable that had a statistically significant effect on the regression model was the number of words in the child's vocabulary so further results will be discussed in

[32] Ordinary least squares linear regression is a common statistical tool used to determine a linear model which seeks to minimize the errors between the observed data and the predicted values. The method finds an equation that best fits the data values. That is, it plots a line through the graph of data points that minimizes the difference between the calculated values and the actual observations.

terms of only vocabulary size unless specified. Least squares linear regression requires two assumptions. First, it requires that each of the samples is independent. That is, the number of proper nouns, for example, that one child knows has no effect on the number of proper nouns that another child might know. With few exceptions, the children who were included in the study didn't even know one another so we can reasonably assume that this assumption is valid. The second assumption is that there is constant variance in the response variable among the samples. However, this was clearly not the case because the response variable is measured as a percentage of the total number of words in a given child's vocabulary. The variance in this response is naturally higher when the size of the word list is small and lower when the size of the word list is higher. This effect is commonplace in percentage data and it was observed in this particular data set. To see this, one may observe that the fan shapes of the data in Figures 4.2 and 4.3 shows earlier vocabularies exhibit greater variation than later ones.

Since the constant variance requirement could not be met, a correction had to be made in the model. Weighted least squares regression is performed the same way as regular regression but uses a weighting factor that is based on the source of the variation. The change in variance was caused by the number of words in a given word list, so the weighting used was $\sqrt{n}$ where n is the word list size. Table 4.2 shows the results of weighted least squares regression using word list size as the explanatory factor.

To apply statistical analysis in answering the first question, as to whether there is a predictable proportion of any word class in early vocabularies, we refer to the R-squared values in Table 4.2. The R-squared value measures the strength of the correlation between the real values and the predicted values of the model. It gives the percentage of variance in the response variable that is explained by the model. If the model fits perfectly, the R-squared value would be 100% or 1.0. In a good model, R-squared is high and should, ideally, be above .80. The highest value in Table 4.2 is for

	Coefficient Value of word list size	t-statistic	R^2	Intercept
Proper nouns	-.0012	.5879	.0067	.2571
Nouns	.0067	.0056	.1618	.1845
Social/emotional words	-.0033	.0491	.0851	.2603
Sum of proper nouns And emotional/social	-.0046	.0893	.0642	.5173

Table 4.2: Results of regression for word class categories.
For each category listed, a regression model was developed to test the hypothesis that the proportion of the given category in a child's vocabulary could be predicted based on the number of words a child knows. If a coefficient value is positive, the proportion is increasing as the number of words increases. If a coefficient value is negative, the proportion decreases as the number of words increases. The number of words a child knows is considered statistically significant in explaining these proportions if the t-statistics are less than .10 (meaning that the errors made by the model are acceptably small). The strength of a model is considered good if the R^2 values are greater than .80. The intercept is the model's prediction of a baseline proportion. For example, the model for social/emotional word proportion (S/E) would be $S/E = .2603 - .0033*W$, where W is the size of the word list.

nouns at .1618. This means that any predictions about the proportions of words in each of the given categories would be extremely weak. There were no word classes that could be predicted with the model. This does not mean that there is not enough data. The R-squared value could not be improved by getting more data. That is, it is not mathematically possible for the R-squared value to approach 1.0 even in a larger sample if it is this low in the smaller sample. Even in models that used the additional variables of child's sex and nationality, the highest R-squared value obtained was .2111. None of the models were able to predict the proportion of words in these word classes successfully to even a very liberal standard. This negative result provides statistical validity to the intuition of Figure 4.1. Because of the variability in a child's early vocabulary, it is not possible to predict the proportion of word types in that vocabulary.

To apply statistical analysis to the second question, as to whether there is any statistically significant trend in the composition of the vocabularies as the size of the vocabulary increases, we turn to the t-statistic in Table 4.2. The t-statistic is a measure of how much the observed data differs from the predicted results. In this analysis, it tests the hypothesis that the number of words has no effect on the proportion of a given word class. A low value of the t-statistic (less than .1) suggests that this hypothesis is false. In all cases except proper nouns, we can conclude that the word list size was a statistically significant ($t < .1$) factor in explaining the proportion of the word class under consideration.

The statistical significance of the t-statistic for nouns and for social/emotional words and the combined proper nouns and social/emotional words, means that one can make conclusions based on the coefficient value. A positive value in the model indicates that a category is less important in early vocabularies than it is later on. A negative coefficient value means that the word category is more important earlier than later. For example, as the size of the word list increases by one word, the proportion of words in that list falling in the noun category increases by 0.67% since the coefficient value of word list size for nouns is .0067. On the other hand, there was a negative value in the social and emotional model, indicating that words with social or emotional content are more important in early vocabularies but less important later. Specifically, if the size of the word list increases by one word, the overall proportion of words in the social or emotional category decreases by 0.46%.

This is the result that could be seen pictorially in Figures 4.2 and 4.3. Having statistically significant positive or negative coefficient values in these models is important to the thesis. Word category proportions, although not predictable, are not static. If these categories were always the same from the very beginning of word production, the coefficients would have been zero.

4.4 Discussion

The statistical analysis of the data demonstrates that word categories cannot be predicted in early vocabularies but the number of nouns is increasing and the number of words with social and emotional content is decreasing. The word list size is significant in the sense that it demonstrates that the nature of a child's vocabulary is changing but that does not mean that the model itself can predict what kinds of words will be in an individual child's first vocabulary.

The main result of this analysis is that it is not possible to predict the proportion of any word class in early vocabularies. Using all available information, mathematical models of early language were extremely weak. There is too much variation among children to be able to predict what kinds of words children will learn first. The inability to make predictions about what word classes will be included in children's first words supports the thesis that children are not looking for object labels when they first begin to produce words.

The other result of the analysis is that words that are emotionally or socially significant are important in very early vocabularies but are less important in larger vocabularies. While nouns are always important, they are not especially important at first. If proper nouns are combined with social/emotional words, we saw in Table 4.1 that they make up 47.8% of very early vocabularies, whereas nouns represent only 22.8% of these vocabularies. Later, social/emotional words decrease to 35.8% and nouns increase to 39.9%. These trends are statistically significant and demonstrate that words with social/emotional content are very important at the beginning of word production but become less important as the number of words in the child's vocabulary increases. The use of words that have social or emotional content implies that social interaction may be of primary importance to infants. This supports the second claim of the thesis, namely that children's early word productions reflect an interest in social communication rather than knowledge of language.

Many researchers have noted that although there are language-specific variations, nouns are the dominant category regardless of language being learned (de Boysson-Bardies 1999; de Boysson-Bardies and Vihman unpublished; Markman 1993; Waxman and Kozowski 1990; Waxman 1994; 1997a,b; Benedict 1979). By the time a child's vocabulary approaches 50 words, de Boysson-Bardies calculates that over 50% of American children's words are objects or animals. Even Japanese children's vocabularies, whose languages regularly use verbs and adjectives for reference, include 45% nouns. However, although these later vocabularies may show statistically significant proportions of word classes, the results of these studies cannot be applied to very early vocabularies. The current study shows that nouns are not as important in early word production as the larger word lists indicate. Early on, the percentages are highly variable – so variable that it is not possible to predict the initial word class proportions. Later vocabularies are not as variable and begin to converge as differences between individuals even out (de Boysson-Bardies 1999; Fenson et al. 1994; Bates and Goodman 1999). Later on, the percentages will be predictable because the variability is decreasing

(Bates and Goodman 1999). However, very early vocabularies are highly variable, are changing during early acquisition, and can be said to be in a *transitional* stage.

The transitional nature of these results suggests that something important is happening to the child's vocabulary. I have suggested that early words are not linguistic in the sense of representing semantic categories. It is possible that the transition children are making is a change from the use of words as a form of social communication to the use of words as categories. The transitional nature of this stage may be a reflection of cognitive development. By the time they know many words, children do link words to categorical meaning. But in the very beginning, words are merely sounds that have social value.

While not statistically analyzed, there are some interesting trends in the contents of the early vocabularies. Examination of my data samples (see Appendix II) reveals that, while *no* occasionally appears in early vocabularies, common use of success/failure words such as *uh-oh, there, no, did it* and disappearance words such as *gone* and *all-gone* rarely appears before the 10-word stage. Gopnik and Meltzoff (1986) correlate means-ends task and object permanence task abilities with the acquisition of disappearance words and success/failure words. If they are correct, then the current study supports the theory that children do not have these cognitive abilities before they begin to speak but that these abilities are developing as they learn to produce words.

4.5 Conclusion

Statistical analysis of early vocabularies supports the conclusion that children are not innately predisposed to look for object labels when they begin producing words. Vocabularies of 50 or more words might have smaller variances and, hence, have predictable word classes but this result cannot be applied to very early word productions. Further, social/emotional words are very important in early vocabularies and become less important later. Conversely, nouns are not especially important in early vocabularies but become more important in later vocabularies.

Chapter 5

Summary and Conclusions

5.1 Summary of Research

In this thesis, I have presented arguments supporting the hypothesis that children are not predisposed to search for object labels when they begin to speak. A small study of 46 children's early vocabularies was conducted to support this hypothesis. The study showed that the proportion of word classes in very early vocabularies (three to less than fifty words) are not statistically predictable. Hence, the claim that nouns make up a large share of early vocabularies was not supported at this stage of word production. Additionally, the makeup of these early vocabularies changes rapidly. The study demonstrated to a high level of significance that the proportion of nouns in their vocabularies increases and the number of words with emotional or social content decreases. This result supports the second part of my hypothesis that children's early words begin as sounds associated with socially and emotionally salient contexts.

5.2 What Can We Learn From Early Vocabularies?

Many studies have shown that two and three-year-old children extend categories (Waxman 1994; Waxman and Kosowski 1990) and attach a novel label to a whole object (Markman 1993). Several authors have suggested that by the time children know fifty words, the largest category of words in their vocabularies will be nouns and there will be language-specific patterns (de Boysson Bardies 1999; Vihman 1996; de Boysson Bardies and Vihman 1991; Gopnik 1981; Gopnik and Choi 1995).[33] The results of my investigation show that these findings cannot be extrapolated to the very first words that infants learn. Earlier investigators worked with children who were already able to communicate and respond linguistically and who usually had already passed the word spurt. The transitional nature of the very early vocabularies that I worked with indicate that, whatever may typify later vocabularies, especially as they approach or pass the typical rapid word acquisition associated with a word spurt, features of these vocabularies cannot be applied to first words. Very early vocabularies may be qualitatively different than later vocabularies because the type of learning is qualitatively different. Infants are primarily concerned with interpersonal functions and the regulation of social behavior (Grieve and Hoogenrad 1986). They can enhance their ability to initiate and maintain social interaction by producing sounds that their parents interpret as word production. Infants quickly realize that words are useful for more than social interaction. Perhaps by the time they can say between 10 and 20 words, children begin to use words as language, with an understanding that words refer to categories of things and actions. This may be associated with a change from a child-centered world to one where people and things interact with each other (Lock 1997).

Gopnik and Meltzoff (1986) relate means-ends task and object permanence task abilities with the acquisition of disappearance words such as *gone* and *all-gone* and success/failure words such as *uh-oh, there, no, did it*. Children may not be able to use certain words until they have acquired these cognitive abilities. The data presented in

[33] For instance, 91% of the words produced by children learning American English are nouns while their Japanese counterparts average only 56% (de Boysson Bardies 1999).

this paper are suggestive that such a feature of early speech exists and that further research might confirm or disconfirm this effect. I suggest that children do not have these cognitive abilities before they begin to learn words but that these abilities are developing while they are in their early word-learning stages. The fact that these words rarely occur in vocabularies of fewer than ten words and never in vocabularies under five words may be evidence that some cognitive abilities necessary for language acquisition are not necessary for children to learn their very first words.

On the other hand, actually learning and saying words may help a child develop cognitive abilities. Premack's (1984) work with chimpanzees suggests that certain concepts are impossible without language. While I contend that first words are not truly language, they provide the raw material for a child to begin to learn language. Research with children as well as computer models provide evidence to support the hypothesis that children may need to learn words before they are able to develop linguistic concepts and categories (Bates and Goodman 1999; Allen and Seidenberg 1999; Goldberg 1999; Redington and Chater 1997; Plunkett 1995; Bates et al. 1995; Vygotsky in Rieber and Carton 1987). Computer models show that the language children hear spoken around them contains enough information to extract semantic structure (Redington and Chater 1997; Plunkett 1995). Children's brains may contain pathways similar to neural networks to process this information and they may need to learn words in order to acquire certain concepts.

It is very likely that early words are learned one at a time without any connections between them. Even words that seem referential from the beginning are learned because they are frequently used by caregivers in particular contexts and are emotionally salient

(Barrett et al. 1991). Words gradually become decontextualized and become meaningful categories. Categorization entails symbolic connections among words, their meanings, and their uses. While they can't use syntactic bootstrapping to learn to produce their first words, once children learn the meanings of some words and determine their semantic categorizations, they can use syntactic structure to infer the meanings of new words (Snow 1999; Bloom 1993). Some word meanings, in fact, cannot be learned without syntactic information to acquire them (Gleitman 1993). Whether from innate or emergent knowledge, children quickly become sensitive to word categories, to phrases and clauses, and the ways of arranging those units to express events and relationships and learn where word types can occur (Hirsh-Pasek and Golinkoff 1996). However, children initially seem oblivious to the syntax of a word (Bloom 2000).

5.3 Implications for Future Research

This thesis suggests a number of directions for future research. The study conducted included only a small number of children's vocabularies from a variety of sources. Some results reached statistical significance but more data samples, collected in a more uniform manner, would allow better inferences to be drawn. An ideal data sample would consist of data collected from several hundred children in a situation where a trained observer confirmed mothers' inventories. This data sample was too small, for instance, to be able to say anything statistically significant about children's early use of adjectives, verbs, deictics (pronouns), or imitative sounds. Additionally, this data represented snapshots in time – no attempt was made to follow children as they learned more words. A diachronic study of individual children might confirm or disconfirm the trends that were observed. The immediate direction this study suggests for future research, then, would be to include a larger number of children who could be followed from the time they said their first words to when they begin to put two words together.

Researchers are not currently in agreement on how words become symbolic. When and how does this happen? What can the statistical analysis of the transitional nature of children's vocabularies tell us about children's cognitive development? Does maturation provide the cognitive abilities necessary for children to advance in their word production or does word production provide the catalyst necessary for learning to progress? Perhaps both processes occur. Some researchers emphasize the importance of the age of the child but others think grammar development is related to the number of words a child knows (Bates and Goodman 1999). It is possible that, when a certain number of words have been learned, the one-word stage becomes mathematically unstable, can no longer be used successfully to represent *events,* and some restructuring occurs as "words" naturally break down into grammatical categories (Nowak et al. 2001; Nowak et al. 2000). The number and usage of words in the one-word stage might be used to predict when the two-word stage is about to occur. Could this be determined by studying the contents and statistical trends in children's vocabularies?

This study may also have some implications for parents and educators of young children. It confirms earlier findings that social interaction with caregivers is very important in early language acquisition. There is a correlation between quality of attachment and cognitive skills, and also between cognition and language (Grieve and Hoogenrad 1986). If the quality of attachment is poor, social skills will be underdetermined and language, although not absent, may be hindered. Ramey et al. (1981) found that when mothers respond to infant vocalization with touch rather than with vocalizations, infants are slower in the development of cognitive and language skills. This interaction difference results in small cognitive differences in the first year of life but large differences in the second year. Successful learning depends on having a sympathetic teacher. Fowler (1981) showed that an intervention program emphasizing attachment and increased use of language by the mother is effective in enhancing cognitive and language skills if applied at *any time in the first year* for a period of several weeks. Further research on the relationship between social interaction and early word production is needed to clarify the significance of early mother-child relationships in the development of language. Long-term studies would be needed to determine if there are lasting effects of these early factors that affect language acquisition.

5.4 Conclusion

Learning to say words is clearly a complex process that involves innate social predispositions but also depends heavily on the environment to proceed successfully. Some cognitive abilities are prerequisites and must be in place before word production can occur while others, including knowledge of word categories, cannot develop until words begin to be learned. One of the lessons from this thesis is that what happens in the mind of an older child cannot be used to predict how an infant's mind works. Early one-word utterances reflect knowledge that the child has about the world – knowledge about the social world of people and the cognition that other people are ready and willing to respond to them (Mervis 1987). But these very early utterances do not reflect knowledge of categories. Early vocabularies can, in fact, be characterized as being in a state of transition and this transition reflects a change in the nature of the words they contain. Earliest word productions are connected to emotional and social contexts while later vocabularies reflect true word classes and categories. Evidence from the study reported in this paper supports the conclusion that categories develop gradually as the child learns words.

Appendix I

Neural Networks

Central to emergentist theories is the family of computer models known as Artificial Neural Networks. Defined briefly in Chapter 1, the models will be more fully described in this section, including how they work, how they are applied to language learning, and why some researchers think they describe the learning process.

What is an Artificial Neural Network?

The concept of an artificial neural network is related to the concept of artificial intelligence. The idea that we might be able to create an artificial intelligence has a long history.[34] However, not until the twentieth century had enough work been done on the actual function of the brain to be able to describe how neurons behave. In 1943, Warren McCulloch and Walter Pitts published a description of a neural network with a logical calculus of the sequences of nerve connections based on the fact that nerves fire in an all-or-none manner (Jones 1999).

Associationist theories have also been around for a long time. In the seventeenth century, Thomas Hobbes and John Locke suggested that knowledge, gained through experience with the world, is stored in the mind in the form of *associations*. Like Aristotle, these people were empiricists – they believed that events in the environment give rise to thoughts and ideas. These thoughts and ideas could excite other thoughts and

[34] In the seventeenth and eighteenth centuries, a popular diversion was the creation and use of small *automata* – machines that were specially constructed to emulate particular activities of animals or humans. In 1748, de la Mettrie's *L'Homme Machine* (Man a Machine) was published. He claimed that all human activity had a mechanical explanation. This was not well received at the time and the document was burned as heresy (Jones 1999).

ideas and were interconnected in the mind. *Connectionist* theory views these associations as connections between neuronal units or nodes of information (Payne and Wenger 1998). While in some ways similar to associationism, connectionism emphasizes that learning involves input from both the environment and from internal or *top-down* (higher-level) conceptual processes that rely on knowledge and memory. The neural network theories developed in the 1940's appeared to be the best kinds of models for these theories. As early as 1949, D.O. Hebb suggested a method for designing artificial neurons so that learning could occur. His model included a *Hebbian Learning Principle*, that is, the strength of the connection between neurons is adjusted to reflect its familiarity with an input. The more probable the input to a neuron is, the larger the output will become. Similar models were described in the 1950's and 1960's but could not be implemented because the immense number of connections between nodes required far more computational power than was available at the time. Figures A1a and A1b illustrate how quickly the number of connections increases as more nodes and layers are added to the network. Note that although there is just one additional layer of four nodes, the

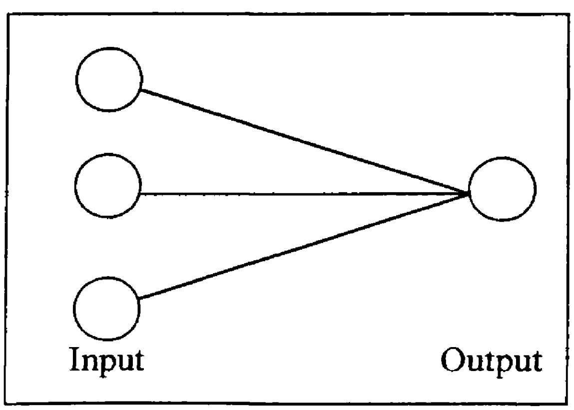

Figure A1a: A simple network

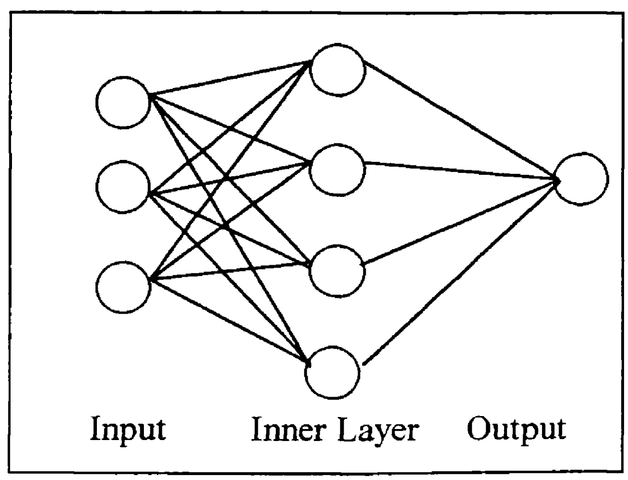

Figure A1b: A more complex network

number of connections increases from three to sixteen.[35] For a functional model that could solve real problems, the computational requirements of such interactive models were well beyond the ability either of an individual to calculate by hand or of the primitive computers available to scientists in the 1940's.

Implementation of these models was made possible by the invention of inexpensive high-speed computers. Such models have now been used not only for language modeling but are currently being used in a number of applications to solve other real-life problems.[36] Neural networks are used to solve problems that use, as input, large numbers of what may appear to be independent variables. They are able to detect similarities in inputs even though a particular input may never have been seen previously. A network can detect important predictive patterns, not apparent to observers, that allow for excellent interpolation capabilities, especially when the input data is noisy[37] as would certainly be the case for infants and small children who are trying to make sense out of their environment.

Many of the accomplishments of neural networks seem mysterious and almost unbelievable. Neural network programs work by making many very small adjustments to

---

[35] In these and later diagrams, circles represent *nodes* or discrete pieces of information. In a computer model, of course, information must be numeric in order to be processed. The lines connecting these nodes represent operations performed on the contents of nodes. The meeting points of these lines or connections represent the results of calculations done on incoming information. Output nodes represent an answer to the problem presented to the network. Inner layer nodes are both the result of calculations and the input to further calculations.

[36] Applications for this kind of computer program include air traffic control, appraisal and valuation of property, betting on horse races, direct mail advertising, economic models, employee hiring, expert systems, fraud detection, medical research, photo and fingerprint identification, prediction of lake and river levels, scheduling of buses, trains, and airplanes, and weather prediction (Cormac 1999).

[37] In physics, *noise* is defined as any disturbance that obscures or reduces the clarity or quality of a signal (Morris 1970). For psychologists, this kind of noise may be *external* such as the static heard when a radio is not tuned to a station or it may be *internal* interference generated by spontaneous activity in the nervous system (Payne and Wenger 1998).

the formulas that calculate answers, storing those answers as intermediate results, and making many calculations with those intermediate results before passing them on to final processing and outputting an answer. The whole process obviously requires a large number of calculations and a lot of memory storage. The goal is to get the right answer no matter how many repeated approximations it takes. Solutions are possible for these programs, as they are for our brains, because both brains and modern computers are able to make many computations very quickly and have large amounts of storage available (Elman et al. 1996). In the process of finding the right answer, patterns develop in the network that make it possible to predict a correct result when novel data are presented.

A very simple example will demonstrate how these networks operate. Suppose a ballpark hotdog vendor wants to know how many hotdogs and buns to order. He collects data about baseball games and their environment for a period of time and notes how many hotdogs he actually sold on those days. Table A1 presents some possible sample data.

	Tickets sold	Day of the week	Temperature	Hotdogs Sold
Day 1	13	Monday	32	9
Day 2	25	Tuesday	54	16
Day 3	8	Wednesday	10	2
Day 4	54	Saturday	60	35

Table A1: Data from the ballpark

Since computer programs need numerical input for calculations, non-numerical input data must be converted to numbers. For our purposes a number from one to seven can be used to represent the day of the week.[38] Now the vendor creates a model like the one in Figure A2.

Lines or *connections* between circles or *nodes* (labeled T for "Tickets sold," D for "Day of the week," and P for "Daily temperature") in this schematic represent mathematical formulas that multiply the values from the input nodes by a multiplier or *weight* (labeled M_1 through M_3) that represents the strength of the connection. The resulting values are added together ($T \times M_1 + D \times M_2 + P \times M_3$) to give an answer represented by the output node (labeled N). Multipliers may be set randomly to begin with. It doesn't matter what the multipliers or weights are initially set to because

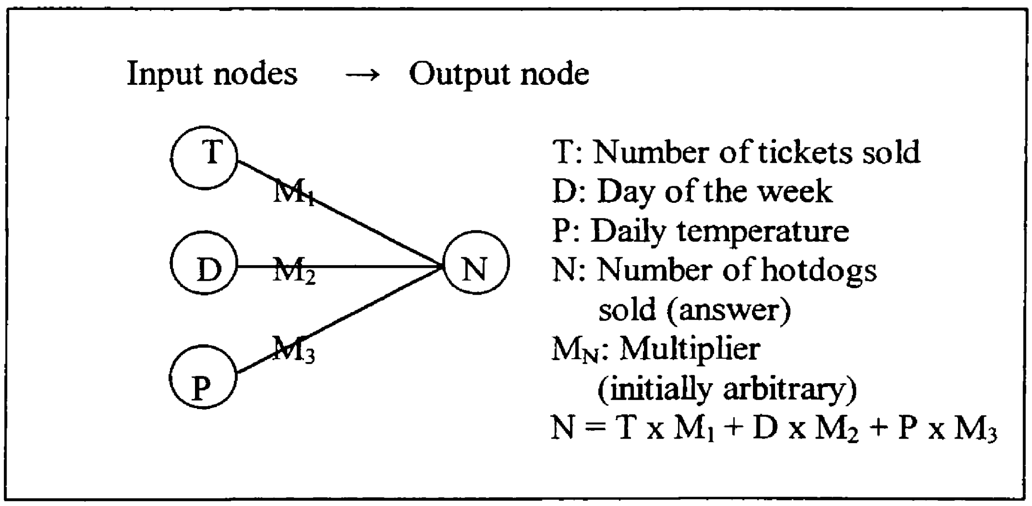

Figure A2: The ballpark vendor's model

the program is designed to adjust them when approximate answers are compared to the desired answer. The programmer does not need to anticipate what the weights are to ensure a solution. It is in these weights that knowledge is gradually built up. This program imitates *architectural* innateness – the weights are not preset. That is, there is

[38] Monday = 1, Tuesday = 2, Wednesday = 3,. . . Sunday = 7.

no information or defaults wired into the system. Rather, the system defines what kinds of problems can be solved. *Representational* innateness would require these weights initially be set so that the "right" answer is inevitable. Emergentism claims that there is no need to burden the human genome with billions of preset weights. They can be assumed to be initially random because the *structure* of the system guarantees a solution, not the initial contents of the system.

The first step in running the example model is to train the system on data that has been previously collected. In any model, the "right" answer depends on what is being modeled. For instance, in a model of phonology acquisition, this step might be analogous to a child hearing adults say words many times. The phonological representations that the child hears are what the system considers the "right" answer. In our ballpark model, the number of hotdogs sold on one day is the "right" answer. For instance, if we arbitrarily start with each of the connection weights set to .5, using data from Table A1, a first approximation of day 1 (Monday) output is:

13 (tickets) x .5 + 1 (Monday) x .5 + 32 (temperature) x .5 = 23

This is too high since our vendor only sold 9 hotdogs that day. This program is written so that each approximate answer is compared to the known answer and the weights (multipliers) are adjusted based on the difference. Since this number is high, the weights should be decreased. In a model of word learning, this might be analogous to processes in the brain that allow production to gradually improve to match expected output. We do not know how the weights are reset in the brain or what kinds of neurological processes accomplish the changes. Designing algorithms for computer programs to alter weights based on output error is still something of an art and depends largely on the programmer's skill. In our example, changing the weights according to the ratio between the correct answer and the answer we arrived at will suffice as a simple demonstration. In this case, since the first approximation was 23 while the known answer was 9, the program might adjust the multiplier to 9/23 of .5 or .2. A second approximation of the answer is:

$$(13 \times .2) + (1 \times .2) + (32 \times .2) = 9.2$$

This is a respectably close second estimate of the expected answer of 9 hotdogs. Of course, every day's data must be run through the system until each data set gives a reasonably close approximation to the observed correct answer. Once the model works to approximate correct answers for our data collection, it could be used to estimate how many hotdogs the vendor could expect to sell in the future. In a phonological model, this might be the equivalent of a child's gradually improving word production. Eventually, a child can reliably produce words.

This extremely simple model is not intended to explain the kind of model that would link words and their meanings but merely to demonstrate that neural network models are based on simple ideas. Early in the development of network models, it became apparent that such systems would work much better and could solve much more complex problems if there were more nodes between the input and output. Figures A1b and A3b illustrate graphically what this kind of network might look like. All nodes between any two layers are interconnected but the answers computed for the middle layer are not normally available as visible output. Hence, they are called *hidden* and the set of inner nodes is called the *hidden layer*. This hidden layer makes it possible to accumulate

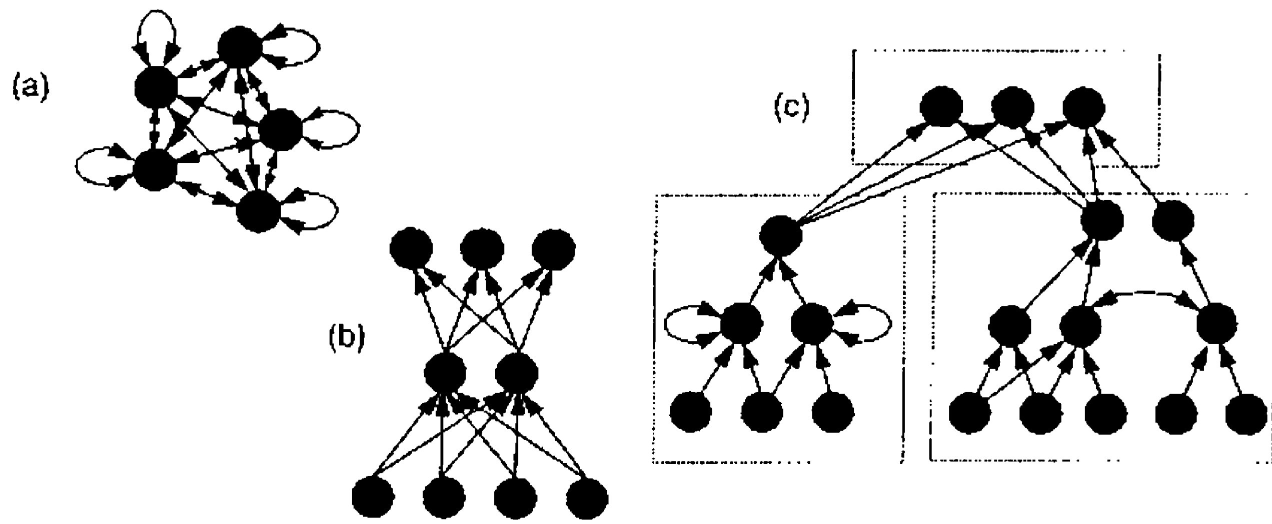

Figure A3: Various types of connectionist networks. (a) A fully recurrent network; (b) a three-layer network with hidden nodes; (c) a complex network consisting of several modules. Arrows indicate flow of information (Elman et al. 1996).

information from earlier data so that the model can be used to predict more complicated kinds of processes.

As with any computer program, the functioning of these models depends entirely on how the programmer designs them. They vary depending on how the weights are initially set up, what formulas govern error corrections, how fast the weights change, how real-life data are approximated numerically, how many hidden nodes there are relative to the number of inputs and outputs, what directions information can flow, and how neuronal output is represented. Figure A3 illustrates some of the different ways processing can occur.

Designing networks to solve problems is still as much an art as a skill and reflects the theoretical claims of the modeler (Mangrich 2000; Elman et al. 1996). Networks use simple processes to learn, yet yield surprisingly complex results. From what we now know about brains, these networks represent a very simplified, but biologically plausible model of neuronal connections. They have been used to demonstrate that many kinds of learning can occur without preprogramming rules into the system to find desired answers.

How Are Artificial Neural Networks Used in Linguistic Research?

Since they were first described, there have been many different kinds of artificial neural networks designed. The kind of model that Rumelhart and McClelland used in the 1980's to model English past tense was a very simple one called a *pattern associator.* It was like our hotdog vendor's model in that it consisted of just a set of input units and a set of output units with no hidden layers.[39] Phonemes of a word were characterized as a series of patterns of zeroes and ones depending on what features were represented by those phonemes. Output patterns were considered successful if the activated features in the output matched those of the correct past tense (Rumelhart and McClelland 1993). The patterns of learning and overall success rates of production led Rumelhart and McClelland to conclude their model demonstrated that a neural network could learn English past tenses without resorting to rules. However, it represented a very simplistic view of language knowledge and has been heavily criticized for leaving out important features of language. For instance, the model could not discriminate between words that are semantically different but phonologically identical such as *break* and *brake.*

Recent advances in the design of ANN's have made modeling features of language acquisition much more successful and flexible. One is the previously described

[39] Rumelhart and McClelland's model more closely mimicked actual neurons in the brain, which are either *on* or *off.* That is, they either receive enough stimulation to fire or they don't fire at all. Rumelhart and McClelland (1996) designed their model to so that nodes required a threshold value in order to produce any output at all. If the calculated value of input to the node exceeded this threshold value, the value at the output node was set to 1. Lower values resulted in the node being set to 0.

addition of a 'hidden' layer of nodes between the input and output nodes (See Figure A3b). Newer methods of adjusting weights have been developed that more closely approximate what we think happens in the brain.[40] Another modification was the development of the *simple recurrent network.* Since spoken language is processed linearly in time, information spoken in the early part of a sentence must be remembered for a period of time, sometimes until the end of the sentence. To replicate this process, an extra layer of units, called *context* units, stores the contents of the hidden units at one point in time. In the next step, the system adds that stored information back into the hidden units along with the new input. In this way, the network's activity at any point in time reflects whatever external input is being presented plus its own prior state. Hidden units thus reflect not only new information but prior *remembered* information as well (Elman et al. 1996).

Even more complex models have been developed that are able to disambiguate embedded sentences. In the Subsymbolic Parser for Embedded Clauses (SPEC) model, the tasks of segmenting input word sequences into clauses, forming case-role representations, and keeping track of recursive embeddings are separated into different modules (Miikkulainen and Mayberry 1999). The recursive nature of each module allows that module to remember earlier parts of the sentence. The combined system is able to generalize to novel sentences with embedded clauses. For instance, the system

[40] In conjunction with the use of hidden layers, a method of adjustment of the connection weights called *backpropagation* allows programs to adjust the hidden nodes when their values are unknown. Differences between the values of the observed output nodes and the expected values are calculated and the weights leading to the output are adjusted based on those differences. Since there are more layers now than in the simple two-layer model, we assume that each of the inputs to those hidden units are partly responsible for the error and adjust them accordingly. Since the changes are *propagated backward* from the output back into the network, the process is called *backpropagation* (Elman et al. 1996).

can successfully parse a sentence such as "The girl saw the boy who chased the cat" into appropriate case-roles and determine that a) "The girl saw the boy" and b) "The boy chased the cat." These kinds of complex systems have made it possible to successfully model many empirical observations about child language acquisition.

Elman et al. (1996) and Plunkett (1995) found that models programmed to accept words in a sentence, one at a time, are particularly poor at predicting what word might come next. The models, however, can activate a *range* of possible candidates based on their probabilities of occurring in the next position in the sentence. Although the model was deprived of any clues regarding the grammatical category or meaning of the words that it used to make these predictions, the candidate words were found to be grouped by such distinctions as *animacy, human v. non-human, edible*, and *breakable.* That is, candidate words belonged to the appropriate word class for the position in the sentence.[41] Hence, the *distribution* of words from a corpus consisting of thousands of sentences of varying structure allowed the network to carry out its task. The only stimuli were those that are directly observable in the world and so, did not presuppose either an intelligent teacher or previous knowledge as to the type of grammar that was used to generate the training set. The model showed that distributional information that is present in the input

[41] These models represent words as patterns arranged hierarchically in clusters. Similar patterns are arranged close together; more distant patterns are grouped farther away. Analysis of the models revealed that representations of individual lexical items in the sentences reflected both the *lexical* grouping of individual words and their *grammatical* role in the sentence. Hence, it not only correctly predicted potential candidates for the next word in a sentence, it resulted in the kind of categorization necessary to determine word class *and* sentence structure. That is, verbs clustered together and nouns clustered together. Because it classified words by their grammatical role, the models could determine whether a word could play the role of subject, object, or verb in a sentence and what arguments a verb could take. Hence, it correctly differentiated between verbs that required direct objects from those that only optionally took objects and what categories those objects might be. (Elman 1999; Plunkett 1995).

is sufficient to classify words hierarchically and that a relatively simple neural network can learn to predict sentence structure without prior grammatical knowledge.

Redington and Chater (1997) studied a corpus of 2.5 million words, much of which was child-directed speech from the CHILDES databases. They used the two words before and after the target word as context and studied statistical distributions of what words could follow the target. Their study did not partition words into syntactic categories, but produced a hierarchical tree similar to those produced by Elman et al. (1996) and Plunkett (1995) whose structures reflect the syntactic relationship between words. Such analyses do not prove that children use information in the speech stream to categorize words but do show that such information is available. Artificial neural networks have been able to demonstrate many features of language acquisition including the development of a prototype, the lag of production behind comprehension, the famous "spurt" in speed of vocabulary acquisition, the U-shaped learning curves observed in the acquisition of English past tense, and other under- and over-extension errors typical of young children (Plunkett 1995; Rumelhart and McClelland 1993; Elman et al. 1996).

Appendix II

Data collection sorted by number of words known.

Date Collected	Child's Name and age in mo.	# Siblings Nationality	Words	Categories
/38	Larry <12 mos	00 American	03 go, ride, walk	V, V, V
2/17/01	Bailey 08	01 American	03 dada, mama, puppy	P, P, N
/50	Susan 10	01 American	03 hi, kitty, gram	E, N, P
2/15/01	Kelly 10	00 American	03 dada, mama, doggy	P, P, N
2/17/01	Cody 14	01 American	03 mama, dada, tractor	P, P, N
2/15/01	Bridget 08	02 American	04 dada, mama, bye bye, pat-a-cake	P, P, E, E
2/19/01	Cooper 28	4 (twin) American	04 no, thank you, dad, mama	E, E, P, P
/89 (V)[42]	Molly	0 American	04 cracker, vroom, woof, yum	N, I, I, A
/93 (V)	Laurent	05 French	10 allo (hello), lolo (bottle), donne (give), tiens (here), non (no)	E, N, V, D, E
/92 (V)	Deborah	05 English	11 baby, hi, hu-hu-hu, monkey, uh-oh	N, E, I, N, E
/99 (BB)	Marie	05 French	13 ca (that), bravo, poupee (doll), Nono, mama	D, E, N, P, P
/99 (BB)	Noel	05 French	13 manger (eat), papa, poum (boom), wouah-wouah (woof-woof), coucou (peek-a-boo)	V, P, I, I, E
2/15/01	Grace 16	03 American	05 mama, dad, good, hi, Annie	P,P, A, E, P

[42] In this data presentation, V = (Vihman 1996), B = (Barrett 1986), BB = (de Boysson-Bardies 1999)

Date Collected	Child's Name and age in mo.	# Siblings Nationality	Words	Categories
/51	Steven <12 mos	02 American	06 baa, bye bye, nite-nite, go potty, go bye-bye, pretty	I, E, E, V, V, A
/92(V)	Alice 10	0 English	06 baby, daddy, hello, hiya, mommy, no	N, P, E, E, P, E
/92 (V)	Charles 12	0 French	06 au revoir (goodbye), beau (beautiful), boom, ca (there), mama, ouah-ouah (woof-woof)	E, A, I, D, P, I
2/15/01	Danielle 16	01 American	06 bye bye, mom, daddy, juice, shoe, toe	E, P, P, N, N, N
/99 (BB)	Leo 10	0 French	07 allo (hello), donne (give), tiens (take it), eau (water), encore (more), papa, mama	E, V, V, N, V, P, P
2/15/01	Caitlin 13	0 American	07 no, mom, dada, what's that, boo, ball, balloon	E, P, P, D, E, N, N
/75	David <12 mos	01 American	08 mamma, daddy, pretty, truck, oh scary, oh boy, oh wow, gramma	P, P, A, N, A, E, E, P
/92	Timmy 11	0 American	08 ball, block, box, car, hi, kitty, quack, mama	N, N, N, N, E, N, I, P
2/14/01	Skyleigh 11	00 American	09 mama, dad, bye-bye, peekaboo, no, kitty, nummy, puppy, bottle	P, P, E, E, E, N, A, N, N
2/14/01	Bailey 11	00 American	09 hi Sara, hot, bye-bye, mama, dad, hi, all gone, all done, bad	P, A, E, P, P, E, A, A, A
/00	Miranda >12 mos	01 American	10 dad, momma, duck, lady, hi, book, ball, that, grandpa, blanky	P, P, N, P, E, N, N, D, P, N
/87 (B)	James 15<20	10 English	mummy, go, quack, there, buzz, moo, boo, teddy, ball, more	P, V, I, D, I, I, E, N, N, V
/87 (B)	Jacqui 15<20	0 English	10 wee, hello, mummy, here, no, down, more, go, Jacqui, bee	V, E, P, D, E, V, V, V, P, N
/87 (B)	Jenny 15<20	0 English	10 choo-choo, bye-bye, there, teddy doggy, moo, shoe, car, mummy, no	I, E, D, N, N, I, N, N, P, E

Date Collected	Child's Name and age in mo.	# Siblings Nationality	Words	Categories
/87 (B)	Madeleine 15<20	0 English	10 there, hello, here, bye-bye, teddy, shoes, vroom, woof, baby, yes	D, E, D, E, N, N, I, I, N, E
/79	Jay 15	01 American	10 hi, hello, wow, uh-oh, thank you, shoe, socks, drink, daddy, bus	E, E, E, E, E, N, N, V, P, N
2/15/01	Cameron 16	00 American	13 scuse me, thank you, mom, dad, kitty, Nanna, please, Clay, moo, bark, wow, uh-oh, smile	E, E, P, P, N, P, E, P, I, I, E, E, V
2/15/01	Jared 15	02 American	14 mommy, daddy, fruitroll, juice, owie, no no, Shesha, puppy, teddy bear, Nanny, Pappa, kitty, fruits, please	P, P, N, N, E, E, P, N, N, P, P, N, N, E
/93 (V)	Laurent 14	0 French	18 bebe (baby), bouton (button), cocotte (hen), coucou (peek-a-boo), Koki, mama, miam (yum), ouah-ouah (bowwow), petits trous (little holes), banane (banana), donne (give), non (no), vroom, allo (hello), voila (here), balle (ball), pas la (not there), parti (gone)	N, N, N, E, P, P, A, I, N, N, V, E, I, E, D, N, D, A
/99 (BB)	Marie 17	0 French	18 attends (wait), bateau (boat), bebe (baby), dodo (sleep), c'est beau (it's nice), c'est beau ca (that's nice), Jacquot, poupee (doll), Tintin, tartine (toast), Ludovic, papa, non, Nono, mama, Mimi chat, papillon (butterfly), voiture (car)	V, N, N, V, A, A, P, N, P, N, P, P, E, P, P, P, N, N
/92 (V)	Deborah 13	0 American	19 baby, ball, daddy, duck, bird, book, byebye, rowrow, patty-cake, Sesame Street, shoe, meow, ah, hi, uh-oh, mama, woof-woof, yay, no	N, N, P, N, N, N, E, V, E, E, E, I, E, V, P, N, I, P, E
2/19/01	Macy 16	00 American	19 puppy, ducky, pepsi, dada, mama, Kudro, cow, truck, no, yes, go, see, pee, poop, shit, ouch, owie, nose, belly button	N, N, N, P, P, P, N, N, E, E, V, V, V, V, E, E, E, N, N

Date Collected	Child's Name and age in mo.	# Siblings Nationality	Words	Categories
2/19/01	Degan 21	01 American	19	P, P, N, P, P, E, E, A,V, N, A, N, E, P, N, E, V, N, P mommy, daddy, doggy, Papa, Nana, no, yeah, yum yum, more, bubble, yellow, bear, uh-oh, Bob, eye, ow, out, duck, Poo
/92 (V) 14	Alice English	0	20	E, N, E, A, E, N, N, N, N, P, N, N, P, N, P E, N, P, P, E bye, eye, hi, clean, whee, bear, nam, baby, blanket, Bonnie, bottle, bunny, daddy, dolly, ernie, hiya, lady, mommy, Oscar, thank you
2/15/01 16	Grace American	00	20	P, P, E, N, V, V, V, E, E, E, P, P, I, I, V, D, D, N, N, P mom, dad, no, pop, don't, don't be mean, shut up, hi, yes, papa, grandma, woof woof, meow, come on, where is he, what's that, baby, ball, Poo
/89 (V) 13	Molly American	0	21	I, I, N, N, A, N, A, V, A, V, N, V, N, P, I, I, E, E, A, A, V bang, burp, cat, dog, good girl, horse, hot, peek, round, squeek, teeth, up, baby, daddy, ho-ho-ho, moo, night-night, no-no, one-two-three, pretty, rockie
/92 (V) 14	Charles French	0	22	A, A, N, N, E, V, I, N, P, D, N, D, E, V, N, N, E, N, N, P, V, E bah (yuck), beau (beautiful), bebe (baby), poupee (doll), bravo, boire (drink), boom, lapin (rabbit), pap, ca (that), gateau (cake), tiens (here), asis (seated), chaussures (shoes), chaussettes (socks), ours (bear), allo (hello), canard (duck), mama, myam (eat/yum), non (no)
/99 (BB) 17	Leo French	0	23	E, V, N, P, P, D, N, E, P, N, V, N, P, A, N, E, D, N, N, N, N, D, N allo (hello), donne (give), de l'eau (some water), papa, mama, la (there), bebe-poupee (babydoll), coucou (peek-a-boo), Koko, bouton (button), manger (eat), ballon (ball), Didier, pas la, parti (not there, gone), petits trous (little holes), non non non (no no no), voila (there it is), cuillerre (spoon), brosse (brush), canard (duck), chapeau (hat), la la dame la (there the lady there), canard dans l'eau (duck in the water)
2/17/01 23	Bradley American	00	24	N, N, P, P, V, V, N, E, E, V, V, E, N, N, P, P, P, P, E, E, V, V, A, A toys, cow, mommy, daddy, poo poo, pee pee, shoes, uh-uh, thank you, nap, hurt, owie, hair, blanky, poppa, mamma, Hugh, Bradley, no, yes, jump, eat, thirsty, yummy

Date Collected	Child's Name and age in mo.	# Siblings Nationality	Words	Categories
/92 (V) 16	Timmy American	0	27	I, N, N, E, N, N, N, N, P, N, I, N, N, N, N, N, N, N, N, N, E, P, P, P, P, N, N baa, ball, bird, bye, cup, girl, kitty, flower, Ruth, eye, moo, moon, nose, baby, bracelet, block, peg, balloon, car, fire, hiya, mummy, Simon, Nana, daddy, eye, fish
2/15/01 20	Selena American	01	32	E, E, P, P, P, V, P, E, V, V, V, E, E, E, P, P, N, E, N, V, V, N, N, N, N, N, N, N, N, N, I, I no, bye, mom, papa, Grandma, eat, sister, nite-nite, peepee, poopoo, go away, excuse me, please, thank you, Barny, Poo, teeth, yes, snow, down, bath, shoes, juice, cup, bottle, cheese, blanky, book, TV, phone, caw-caw, duck-duck
/92 (V) 16	Alice English	0	35	E, N, E, N, N, N, N, N, N, N, N, N, N, A, P, N, N, N, N, N, A, N, N, N, N, E, N, N, V, N, E, P, P, V, A bang, belly, bye, egg, eye, key, meat, plate, shoe, tea, baby, bottle, bunny, clean, dady, elephant, iron, flowers, lady, mommy, shiny, apple, duck, man, milk, no, nose, blanket, down, hat, hello, Oscar, grandpa, up, yum
2/14/01	Rachel 18	00 American	44	P, P, P, P, P, P, P, P, P, P, P, P, N, E, N, N, P, P, P, P, E, N, A, A, P, A, A, A, N, N, N, N, V, E, E, E, N, N, V, P, P, P, A, V mommy, daddy, Grandma, Grandpa, Curtis, Amanda, Randy, Spud, Cami, Chuck, Nanners, Jason, fish, come on, puppy, kitty, Arnie, Tony, Sarah, Mary, no, bath, blue, magenta, Scooby-do, yucky, ick, hungry, cup, baby, diper, wipes, kiss, bye-bye, hi, hello, toes, mouth, drink, Steve, Nessa, Tama, hot, up

Bibliography

Allen, J.P.B. and Paul Van Baren, Ed. 1971. *Chomsky: Selected Readings*. London: Academic Press.

Allen, Joseph and Mark S. Seidenberg. 1999. "The Emergence of Grammaticality in Connectionist Networks," in *The Emergence of Language*, edited by Brian MacWhinney. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers. 115-151.

Aslin, Richard N., Jenny R. Saffran, and Elissa L.Newport. 1999. "Statistical Learning in Linguistic and Nonlinguistic Domains" in *The Emergence of Language*, edited by Brian MacWhinney. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers. 359-380.

Barrett, Martyn. 1995. "Early Lexical Development" in *The Handbook of Child Language*, edited by Paul Fletcher and Brian MacWhinney. Oxford, UK: Basil Blackwell Inc. 362-392.

Barrett, Martyn. 1986. "Early Semantic Representations and Early Word-Usage" in *The Development of Word Meaning* edited by Stan A. Kuczaj, II and Martyn D. Barrett. New York: Springer-Verlag. 39-67.

Barrett, Martyn, Margaret Harris, and Joan Chasin. 1991. "Early Lexical Development and Maternal Speech: A Comparison of Children's Initial and Subsequent Uses of Words," *Journal of Child Language*, 18:21-40.

Bates, Elizabeth and Judith C. Goodman. 1999. "On the Emergence of Grammar From the Lexicon," in *The Emergence of Language*, edited by Brian MacWhinney. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers. 29-80.

Bates, Elizabeth, Philip S. Dale, and Donna Thal. 1995. "Individual Differences and their Implications for Theories of Language Development," in *The Handbook of Child Language*, edited by Paul Fletcher and Brian MacWhinney. Oxford, UK: Basil Blackwell Inc. 96-151.

Bates, Elizabeth, Inge Bretherton, and Lynn Snyder. 1988. *From First Words to Grammar*. Cambridge: Cambridge University Press.

Bates, Elizabeth and Virginia Volterra. 1984. "On the Invention of Language: An Alternative View," *Monographs of the Society for Research in Child Development,* 49:30-142.

Bates, Elizabeth, L. Benigni, Inge Bretherton, Luigia Camaioni, and Virginia Volterra. 1979. *The Emergence of Symbols: Cognition and Communication in Infancy.* New York: Academic Press.

Bates, Elizabeth, Luigia Camaioni , and Virginia Volterra. 1975. "The Acquisition of Performatives prior to Speech," *Merrill-Palmer Quarterly*, 21/3:205-226.

Benedict, Helen. 1979. "Early Lexical Development: Comprehension and Production," *Journal of Child Language*, 6:183-200.

Bloom, L. 1973. *One word at a time*. The Hague: Mouton.

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

Bloom, Paul. 1993. "Overview: Controversies in language acquisition" in *Language Acquisition: Core Readings*, edited by Paul Bloom. Cambridge, Massachusetts: MIT Press.

Bowerman, Melissa. 1993. "Learning a Semantic System: What Roles do Cognitive Predispositions Play?" in *Language Acquisition: Core Readings*, edited by Paul Bloom. Cambridge, Massachusetts: MIT Press. 329-363.

de Boysson-Bardies, Bénédicte. 1999. *How Language Comes to Children*. Cambridge, Massachusetts: Massachusetts Institute of Technology.

de Boysson-Bardies, Bénédicte and Marilyn May Vihman. 1991. "Adaptation to Language: Evidence From Babbling and First Words in Four Languages," *Language*, 67:297-319.

de Boysson-Bardies, Bénédicte and Marilyn May Vihman. Unpublished paper. "The First Lexicon: A Comparative Study of Four Languages."

Burling, Robbins. 1993. "Primate Calls, Human Language, and Nonverbal Communication," *Current Anthropology*, 34/1:25-51.

Capirci, Olga, Jana Iverson, Elena Pizzuto, and Virginia Volterra. 1996. "Gestures and Words During the Transition to Two-word Speech," *Journal of Child Language*, 23:645-673.

Chomsky, Noam. 1975. *Reflections on Language*. New York: Pantheon Books.

Chomsky, Noam. 1959. A review of Skinner's "Verbal Behavior," *Language*, 35:26-58.

Cook, V.J. and Mark Newson. 1996. *Chomsky's Universal Grammar, second edition.* Oxford: Blackwell Publishers.

Cormac Technologies, Inc. 1999. "Introduction to NeuNet Pro," [On-Line] Available: http://cormactech.com/neunet/helpfile/4000intr.htm.

Crevier, Daniel. 1993. AI: The Tumultous History of the Search for Artificial Intelligence. New York: Basicbooks.

Cutler, Ann. 1994. "Segmentation Problems, Rhythmic Solutions," *Lingua,* 92:81-104.

Darley, Frederick and Winitz, Harris. 1961. "Age of First Word: Review of Research," *Journal of Speech and Hearing Disorders*, 26:272-290.

Deacon, Terrance. 1997. *The Symbolic Species.* New York: W. W. Norton and Company.

Dunbar, Robin. 1997. *Grooming, Gossip, and the Evolution of Language*. Cambridge, Mass.: Harvard University Press.

Elman, Jeffrey. 1999. "The Emergence of Language: A Conspiracy Theory," in *The Emergence of Language*, edited by Brian MacWhinney, Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers. 1-28.

Elman, Jeffrey, Elizabeth Bates, Mark H. Johnson, Annette Karmiloff-Smith, Domenico Parisi, and Kim Plunkett. 1996. *Rethinking Innateness*. Cambridge, Massachusetts: MIT Press.

Fenson, L. 2000. *The MacArthur Communicative Development Inventories.* [On-Line] Available: http://www.sci.sdsu.edu/cdi/.

Fenson, L., Dale, P., Reznick, J., Bates, E., Thal, D., and Pethick, S. 1994. "Variability in Early Communicative Development," *Monographs of the Society for Research in Child Development*, Serial #242, 59/5:1-107.

Fernald, Anne. 1993. "Human Maternal Vocalizations to Infants as Biologically Relevant Signals: An evolutionary perspective," in *Language Acquisition: Core Readings*, edited by Paul Bloom. Cambridge, Massachusetts: The MIT Press. 51-94.

Finegan, Edward. 1999. *Language, Its Structure and Use*. Fort Worth, Texas: Harcourt Brace.

Fodor, Jerry. 1975. *The Language of Thought*. New York: Crowell.

Fortescue, Michael and Lise Lennert Olsen. 1992. "The Acquisition of West Greenlandic," in *The Crosslinguistic Study of Language Acquisition, Volume 3*, edited by Dan Isaac Slobin. Hillsdale, New Jersey: Lawrence Erlbaum Associates. 111-219.

Fowler, William. 1981. "A Strategy for Stimulating Infant Learning," in *Early Language: Acquisition and Intervention*, edited by Richard L. Schiefelbusch and Diane Bricker. Baltimore: University Park Press. 517-557.

Givón, T. 1999. "Generativity and Variation: The Notion 'Rule of Grammar' Revisited," in *The Emergence of Language*, edited by Brian MacWhinney. Mahwah, New Jersey: Lawrence Erlbaum Associates. 81-114.

Gleitman, Lila. 1993. "The Structural Sources of Verb Meanings," in *Language Acquisition: Core Readings*, edited by Paul Bloom. Cambridge, Massachusetts: The MIT Press. 174-221.

Gold, E.M. 1967. "Language Identification in the Limit," *Information and Control*, 10:447-474.

Goldberg, Adele. 2000. "Learning Argument Structure Generalizations," draft, to be published.

Goldberg, Adele. 1999. "The Emergence of the Semantics of Argument Structure Constructions," in *The Emergence of Language*, edited by Brian MacWhinney. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers. 197-212.

Goldin-Meadow and Mylander, Carolyn. 1993. "Beyond the Input Given: The child's role in the acquisition of language." in *Language Acquisition: Core Readings*, edited by Paul Bloom. Cambridge, Massachusetts: The MIT Press. 507-542.

Golinkoff, Roberta Michnick and Kathy Hirsh-Pasek. 2000. *How Babies Talk*. New York: Penguin.

Golinkoff, Roberta Michnick, Carolyn B. Mervis, and Kathryn Hirsh-Pasek. 1994. "Early Object Labels: the Case for a Developmental Lexical Principles Framework," *Journal of Child Language*, 21:125-155.

Gopnik, Alison. 1981. "Development of Non-nominal Expressions in 1-2-Year-Olds: Why the First Words Aren't About Things," in *Child Language – An International Perspective*, edited by P.S. Dale and D. Ingram. Baltimore: University Park Press. 93-104.

Goodluck, Helen 1991 *Language Acquisition: A Linguistic Introduction*. Blackwell: Oxford.

Gupta, Prahlad and Gary S. Dell. 1999. "The Emergence of Language From Serial Order and Procedural Memory," in *The Emergence of Language*, edited by Brian MacWhinney. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers. 483-496.

Harris, Margaret, Martyn Barrett, David Jones, and Susan Brookes. 1988. "Linguistic Input and Early Word Meaning," *Journal of Child Language,* 15:77-94.

Hirsh-Pasek, Kathy and Roberta Michnik Golinkoff. 1996. *The Origins of Grammar: Evidence from Early Language Comprehension.* Cambridge, Massachusetts: MIT Press.

Hurford, James R. 2000. "The Roles of Expression and Representation in Language Evolution," paper presented to the 3rd International Congress on Evolution of Language, Paris.

Huttenlocher, Janellen and Smiley, Patricia. 1993. "Early Word Meanings: The case of object names," in *Language Acquisition, Core Readings*, edited by Paul Bloom, Cambridge, Massachusetts; The MIT Press. 222-248.

Ingram, David. 1981. "The Transition from Early Symbols to Syntax," in *Early Language: Acquisition and Intervention*, edited by Richard L. Schiefelbusch and Diane Bricker. Baltimore: University Park Press. 259-286.

Jackendoff, Ray. 1992. *Languages of the Mind.* Cambridge, Massachusetts: The MIT Press.

James, Sharon. 1990. *Normal Language Acquisition.* Boston: Little, Brown and Company, Inc.

Jannedy, Stefanie, Robert Poletto, and Tracey L. Weldon, eds. 1994. *The Language Files.* Columbus, Ohio: Ohio State University Press.

Jones, Stephen. 1999. "Neural Networks and the Computational Brain." [On-Line] Available: http://www.culture.com.au/brain_proj/neur_net.htm.

Karmiloff-Smith, Annette. 1993. "Innate Constraints and Developmental Change," in *Language Acquisition: Core Readings*, edited by Paul Bloom. Cambridge, Massachusetts: MIT Press. 563-590.

Kita, Sotaro and Laura Walsh Dickey, Eds. 1998. *Annual Report to Max Planck Institute for Psycholinguistics.* [On-Line] Available: http://www.mpi.nl/world/anrep/98/anrep-98.html

Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Lederer, Anne, Henry Gleitman and Lila Gleitman. 1995. "Verbs of a Feather Flock Together: Semantic Information in the Structure of Maternal Speech," in *Beyond Names for Things: Young Children's Acquisition of Verbs* edited by Michael Tomasello and William E. Merriman. Hillsdale, New Jersey: Lawerence Erlbaum Associates.

Lieberman, Phillip. 1999. *Eve Spoke*. New York: W.W. Norton and Company, Inc.

Lightbown, Patsy M. and Nina Spada. 1999. *How Languages are Learned.* Revised Edition. New York: Oxford University Press.

Lock, Andrew. 1997. "The role of gesture in the establishment of symbolic abilities: continuities and discontinuities in early language development," *Evolution of Communication*, 1:159-192.

Lust, Barbara, and Wakayama, Tatsuko Kaneda. 1981. "Word Order in First Language Acquisition of Japanese," in *Child Language – An International Perspective*, edited by P.S. Dale and D. Ingram. Baltimore, MD: University Park Press. 73-90.

Macnamara, John. 1977. "On the Relation Between Language Learning and Thought," in *Language Learning and Thought*, edited by John Macnamara. New York: Academic Press. 1-10.

MacWhinney, Brian. 1999. Preface to *The Emergence of Language*, edited by Brian MacWhinney. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers. ix-xvii.

MacWhinney, Brian. 1998. "Models of the Emergence of Language," *Annual Review of Psychology*, 49: 199-227.

Mangrich, Mark. 2000. University of Montana Computer Science graduate student, personal interview.

Markman, Ellen M. 1993. "Constraints Children Place on Word Meanings," in *Language Acquisition, Core Readings,* edited by Paul Bloom. Cambridge, Massachusetts; The MIT Press. 154-173.

Markowitz, Judith. 1988. "An Exploration into Graded Set Membership," in *Relational Models of the Lexicon*, edited by Martha Evens. Cambridge; New York: Cambridge University Press.

Marler, Peter. 1993. "The Instinct to Learn," in *Language Acquisition: Core Readings*, edited by Paul Bloom. Cambridge, Massachusetts; The MIT Press. 591-617.

McDonald, Janet L. 1997. "Language Acquisition: The Acquisition of Linguistic Structure in Normal and Special Populations," *Annual Review of Psychology*, 48: 215-41.

Meisel, Jürgen. 1995. "Parameters in Acquisition," in *The Handbook of Child Language*, edited by Paul Fletcher and Brian MacWhinney. Oxford, UK: Basil Blackwell Inc. 10-35.

Mervis, Carolyn B. 1998. "Child-basic Object Categories and Early Lexical Development," in *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*, edited by Ulric Neisser. New York: Cambridge University Press.

Miikkulainen, Risto and Marshall R. Mayberry, III. 1999. "Disambiguation and Grammar as Emergent Soft Constraints," in *The Emergence of Language*, edited by Brian MacWhinney. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers. 153-176.

Moeser, Shannon Dawn. 1977. "Semantics and Miniature Artificial Languages," in *Language Learning and Thought,* edited by John MacNamara. New York: Academic Press. 227-250.

Morris, William, ed. 1970. *The American Heritage Dictionary*. New York: American Heritage Publishing Co.

Nelson, Katherine. 1977a. *Lexical Acquisition*. Paper presented to the Society for Research in Child Development, New Orleans.

Nelson, Katherine. 1977b. "The Conceptual Basis for Naming," in *Language Learning and Thought*, edited by John MacNamara. New York: Academic Press. 117-136.

Neville, Helen J, Debra L. Mills, and Donald S. Lawson. 1992. "Fractionating Language: Different Neural Subsystems with Different Sensitive Periods," *Cerebral Cortex*, 2/3:244-258.

Ninio, Anat. 1992. "The relation of children's single word utterances to single word utterances in the input," *Journal of Child Language*, 19:87-110.

Nowak, Martin A., Natalia L. Komarova, Partha Niyogi. 2001. "Evolution of Universal Grammar," *Science*, 291:114-118.

Nowak, Martin A., Plotkin, Joshua B., and Jansen, Vincent A.A. 2000. "The Evolution of Syntactic Communication," *Nature*, 404:495-498.

Ochs, Elinor and Bambi Schieffelin. 1995. "The Impact of Language Socialization on Grammatical Development," in *The Handbook of Child Language*, edited by Paul Fletcher and Brian MacWhinney. Oxford: Blackwell Publishers. 73-93.

Owens, Robert, Jr. 1984. *Language Development*. Columbus, Ohio: Charles E. Merrill Publishing Company.

Pääbo, Svante. 2001a. "The Human Genome and Our View of Ourselves," *Science* 291:1219-1220.

Pääbo, Svante. 2001b. "A Comparative Approach to Human Origins." Address to the Human Genome Meeting, April 19-22, 2001.

Payne, David G. and Michael J. Wenger. 1998. *Cognitive Psychology*. Boston: Houghton-Mifflin Company.

Peters, Ann M. 1986. "Early Syntax," in *Language Acquisition, Second Edition*, edited by Paul Fletcher and Michael Garman. Cambridge; New York: Cambridge University Press. 307-325.

Pettito, Laura Ann. 1993. "Modularity and Constraints in Early Lexical Acquisition: Evidence From Children's Early Language and Gestures," in *Language Acquisition: Core Readings*, edited by Paul Bloom. Cambridge, Massachusetts: MIT Press. 95-126.

Pinker, Steven. 1993. "Rules of Language," in *Language Acquisition: Core Readings*, edited by Paul Bloom. Cambridge, Massachusetts: The MIT Press. 472-484.

Pinker, Steven. 1994. *The Language Instinct: How the Mind Creates Language*. New York: Morrow.

Pinker, Steven. 1987. "The Bootstrapping Problem in Language Acquisition," in *Mechanisms in Language Acquisition*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc. 400-441.

Pinker, Steven and Alan Prince. 1988. "On Language and Connectionism: Analysis of a Parallel Distributed Processing Model of Language Acquisition," *Cognition* 28 (1-2) 73-193.

Plaut, David C. and Christopher T. Kello. 1999. "The Emergence of Phonology From the Interplay of Speech Comprehension and Production: A Distributed connectionist Approach," in *The Emergence of Language*, edited by Brian MacWhinney. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers. 381-416.

Plunkett, Kim. 1995. "Connectionist Approaches to Language Acquisition," in *The Handbook of Child Language*, edited by Paul Fletcher and Brian MacWhinney. Oxford: Blackwell Publishers. 36-72.

Premack, David. 1984. "Possible General Effects of Language Training on the Chimpanzee," *Human Development*, 27:268-281.

Ramey, Craig, Sparling, Joseph, and Wasik, Barara. 1981. "Creating Social Environments to Facilitate Language Development." in *Early Language: Acquisition and Intervention*, edited by Richard L. Schiefelbusch and Diane Bricker. Baltimore: University Park Press. 447-476.

Redington, Martin and Nick Chater. 1997. "Probabilistic and Distributional Approaches to language Acquisition." [On-Line] Available: http://www.ucl.ac.uk/~ucjtfmr/TICS/TICS.html (draft - to appear in *Trends in Cognitive Sciences*).

Rieber, Robert W. and Aaron S. Carton, ed. 1987. *The Collected Works of L.S. Vygotsky, Volume 1 – Problems of General Psychology.* Plenum Press: New York.

Rumelhart, D.E. and J.L. McClelland. 1993. "On Learning the Past Tenses of English Verbs," in *Language Acquisition, Core Readings,* edited by Paul Bloom. Cambridge, Massachusetts: The MIT Press. 423-471.

Savage-Rumbaugh, Sue, Stuart G. Shanker, and Talbot J Taylor. 1998. *Apes, Language, and the Human Mind.* New York: Oxford University Press.

Slobin, Dan Isaac. 1992. *The Crosslinguistic Study of Language Acquisition, Volume 3.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Smiley, Patricia and Huttenlocher, Janellen. 1995. "Conceptual Development and the Child's Early Words for Events, Objects, and Persons," in *Beyond Names for Things: Young Children's Acquisition of Verbs*, edited by Michael Tomasello and William E Merriman. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers. 21-61.

Smith, Linda B. 1996. "Children's Noun Learning: How General Learning Processes Make Specialized Learning Mechanisms," in *The Emergence of Language*, edited by Brian MacWhinney, Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers. 277-303.

Snow, Catherine. 1999. "Social Perspectives on the Emergence of Language" in *The Emergence of Language*, edited by Brian MacWhinney. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers. 257-276.

Snow, Catherine. 1986. "Conversations With Children," in *Language Acquisition, Second Edition*, edited by Paul Fletcher and Michael Garman. Cambridge; New York: Cambridge University Press. 69-89.

Snow, Catherine. 1981. "Social Interaction and Language Acquisition," in *Child Language – An International Perspective*, edited by P.S. Dale and D. Ingram. Baltimore: University Park Press. 195-214.

Thoman, Evelyn B. 1981. "Affective Communication as the Prelude and Context for Language Learning," in *Early Language: Acquisition and Intervention*, edited by Richard L. Schiefelbusch and Diane Bricker. Baltimore: University Park Press. 181-200.

Trask, R.L. 1997. *A Student's Dictionary of Language and Linguistics*. London: Arnold.

Ulbaek, Ib. 1998. "The Origin of Language and Cognition," in *Approaches to the Evolution of Language*, edited by James R. Hurford, Michael Studdert-Kennedy and Chris Knight. Cambridge: Cambridge University Press. 30-43.

Uzgiris, Ina. 1981. "Experience in the Social Context" in *Early Language: Acquisition and Intervention*, edited by Richard L. Schiefelbusch and Diane Bricker. Baltimore: University Park Press. 139-168.

Vihman, Marilyn May. 1996. *Phonological Development*. Oxford: Blackwell Publishers.

Waxman, Sandra R. 1994. "Development of an Appreciation of the Linkages Between Linguistic and Conceptual Organization," *Lingua*, 92:229-257.

Waxman, Sandra R. and Kosowski, Toby D. 1990. "Nouns Mark Category Relations: Toddlers' and Preschoolers' Word-Learning Biases," *Child Development*, 61:1461-1473.

Werker, Janet F., Valerie L. Lloyd, Judith E. Pegg, and Linda Polka. 1996. "Putting the Baby in the Bootstraps: Toward a More Complete Understanding of the Role of the Input in Infant Speech Processing," in *From Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, edited by James L. Morgan and Katherine Demuth. Mahwah, New Jersey: Lawrence Erlbaum. 427-446.