Part I   Linguistic Knowledge

Introduction

In this part, we begin by laying out some basic facts about language acquisition which any theory of linguistic knowledge must be able to explain. We then consider a common-sense answer to the question of how children acquire language, and show that although it is intuitive, it does not begin to explain the facts we have set out. As an alternative, we present the theory of Universal Grammar, which posits innate knowledge of certain linguistic principles that guide the language learner. We will show that this theory fares far better in explaining the remarkable facts about language acquisition with which we start.

There are some important points of terminology which need to be clarified at the outset. In the field of linguistics as a whole, and particularly in this text, we make a distinction between linguistic **competence** and linguistic **performance**. Linguistic performance consists of the utterances which we actually make (or might make). This performance is fraught with limitations of attention, time, and resources – that is, we often make mistakes when we talk, such as forgetting the beginning of our sentence, or letting a new thought interrupt, or failing to complete a sentence due to a cough, sneeze, or headache. Although such performances are "real," they are not the basis for our theory of linguistic knowledge. Instead, we seek to understand our linguistic competence – the rules in our head which govern our speaking attempts – what it means to "know" a language. The full extent of our linguistic competence is never seen, since we are capable of producing an unlimited number of linguistic utterances. The rules which underlie these potential utterances are what we seek.

Don't be fooled by our use of the word "rules," however. Our aim is to produce a **descriptive** grammar: a list of the rules we have internalized during the course of language acquisition. Our grammar *describes* linguistic knowledge. We do not seek to produce a **prescriptive** grammar, listing the rules we *should* follow in order to speak "correctly." We do not judge. Whatever rules the language users employ are the rules of the language, as far as we're concerned.

We turn now to the beginning of our overview of linguistic knowledge.

1 Introduction to Language Acquisition

Introduction

One of the most remarkable characteristics of human beings is that virtually every single one acquires language at a very young age. This fact is even more remarkable considering the full complexity of the system which is acquired. If you doubt any part of these assertions – that everyone acquires language, that we acquire it at a young age, that the system is complex – be assured that we will soon back them up with evidence. For now, assume that they are true, and ask yourself, "How?" How does every single very young child acquire such a complex system? You might think that parents teach their children language, but this is not correct, except in so far as the parents may provide a language model by talking to and around their children. What we will argue in this book is that children acquire language so readily because it is in their genes. That is, we will argue that children are born with innate knowledge which guides them in the language acquisition task. This innate knowledge, known as the **Language Acquisition Device** (LAD), includes principles common to all human languages, called **Universal Grammar** (UG). What is innate does not have to be learned – hence language acquisition consists of learning what is peculiar to the language environment (e.g., the particular words of one language or another), and applying the universal principles.

The theory of innate linguistic knowledge which we will discuss in this book is largely due to the thoughts and influence of Noam Chomsky, a linguist at the Massachusetts Institute of Technology. Chomsky's 1957 book *Syntactic Structures* revolutionized the field of linguistics by introducing new technology and new goals for the study of the properties of language. His book *Aspects of the Theory of Syntax*, published in 1965, laid out the reasoning for a theory of innate knowledge, and since then much of his work and volumes of work inspired by him have added to our understanding of this theory and its implications.

In this chapter, we begin our presentation of this theory of language acquisition. We start by looking at what is known about the course of language development in general. These facts will lead to a more detailed presentation

of the proposed innate language faculty. For now, our discussion will be fairly general, but the arguments for innateness should be clear. We will return to take a more thorough look at very specific aspects of language acquisition in later chapters, where we will discuss the syntax of these constructions in detail. Overall, our goal is to test the tenets of Universal Grammar.

The knowledge that a child brings to the task of language acquisition, the LAD, is the means by which the child/learner analyzes the linguistic input (from parents and others). The linguistic input to the learner is called the **Primary Linguistic Data** (PLD). On the basis of these data, the LAD hypothesizes a series of grammars. The last grammar that is formulated is the adult grammar, or Final State. The process can be represented schematically in the following way:

Input (PLD) → LAD → Final State

We should specify what we mean by "language." When we talk about a language, we are not talking about sentences in a textbook or on a blackboard – or even a list of all the sentences any speaker might produce. Language is not a concrete set of things out in the world that we can point to or measure. Rather, it is something inside our brains or minds. In linguistics, we attempt to describe what we know when we know a language by formulating a **grammar**. A grammar is a set of rules which characterizes all and only the sentences of the language that we as speakers are able to produce and understand. We take it as given that to learn a language is to master the rules of the mental grammar. Sometimes the mental grammar is called a "Generative Grammar," because it is a system of rules that generate (= produce) sentences of the language.

Against this background, the most basic questions about language acquisition are: Who learns language? How and when is it learned? What is the course of language acquisition like? Only when these questions have been answered will we begin to examine the properties of the Primary Linguistic Data, the LAD, and the Final State. Even the answers to the more tractable questions turn out to be quite remarkable if we approach them without any preconceived notions. At this point, it will be important to suppress any preconceptions you may have about how children learn language. Try to hold in abeyance any beliefs you may have, for instance, about the "baby-talk" of toddlers or young children, how they speak in only simple sentences, and use incorrect grammar, such as saying, "I runned home" instead of "I ran home." As we shall see, young children's grammatical successes are vastly more impressive than their mistakes. Also try to erase any preconceptions about the parents' role in teaching children to speak in grammatical sentences. The actual role of parental input to children is probably not what you believe it to be. Perhaps it will be useful to approach the topic of child language development as if we were scientists from another world who have been assigned to study the communication system used by people on

Earth. In preparing to report back the basic facts about how this communication system is acquired, we are simply summarizing what we have observed in our studies of Earth children. At this point we will concentrate on "Just the facts," as Joe Friday from *Dragnet* says. We will collect facts about language acquisition to prepare us to develop a theory, or model, of this process, which is our ultimate goal.

1   Characteristics of Language Acquisition

What would our report say about children's transition to the Final State? Perhaps the most remarkable fact is that, without special training or carefully sequenced language input, every normal child acquires a natural language. We refer to this property of language acquisition as the **universality** of language. The universality of language stands in glaring contrast to the limited attainment of arguably simpler cognitive skills such as the ability to perform basic calculations or to tie knots. A related observation is that every child in a linguistic community succeeds in converging on a grammatical system that is equivalent to everyone else's, despite a considerable latitude in linguistic experience. Finally, we will see that children acquire language quite rapidly and with few wrong turns, considering the number of potential pitfalls that exist.

   Despite the diversity of natural languages, this picture of the course of acquisition has been observed so often across languages that we are certain of its truth. But this in no way renders it less astonishing for a parent watching the process unfold, or less difficult for a scientific theory to explain. What is special about natural language, and about children, that guarantees that almost without fail they will master a rich and intricate system of rules for language production and comprehension by the time they reach school age, i.e., at a time when they are receiving their earliest instruction in other complex cognitive skills? This is one of the puzzles which our theory of language must solve.

   The next observation concerns the primary linguistic data – the environmental experience within which children learn language. Some children spend a lot of time in their first few years with adults, who give them extensive individual attention. If these were the only children we encountered, we might reach the conclusion that children are taught language by their parents or caretakers. But other children receive less individual attention, even if adults are around to care for them, and they nevertheless develop equal facility with language as the children who received much more individual attention. In other situations, children spend most of their time with other children, and not in situations in which an adult speaks directly to them very often. In yet other cases, children are

spoken to in one language by some adults, and in another language by other adults. In all of these circumstances, children manage to acquire a native language – and rather quickly at that. Some of the differences between children's experiences are cultural, some social, and some accidental. After studying many children, our report would conclude that even children growing up in the same linguistic community, and ultimately learning the same language, may have a wide variety of experiences with language during their early years. Despite the variety in experiences, however, the report would note that every child in every linguistic community acquires the language of that community.

A related point is that every child in the same linguistic community learns the same language. This means that the grammars they have internalized are nearly equivalent. Otherwise, mutual understanding would not be possible. Children's uniform convergence on equivalent grammars is all the more striking in view of the fact that children all receive different input (= primary linguistic data). How is it that all children manage to learn to speak and understand a language, despite being raised in environments that vary considerably?

These observations lead us to ask whether *explicit* instruction is required at all for language acquisition. In this context, it is worth remarking on the linguistic achievements of the children of immigrant parents. Even in cases where adult immigrants exhibit great difficulty in mastering more than the rudiments of the language of a new community, it is normal for their young children to acquire the new language without any difficulties at all. Clearly, these children are not using their parents as models, yet they learn all the complexities of the language spoken around them just as successfully as children whose parents are native speakers.

Here, then, are the first two facts about language acquisition. First, it is universal (within the human species) and, second, there is a considerable latitude in the kind of environmental inputs that permit children to develop language. A theory of language and mind, to be viable, must be responsive to these observations.

Another observation that would be in our report is that some children learn many languages, although most children learn only one. In communities where more than one language is spoken, children acquire all of the languages of the community. So, in this way at least, language clearly IS a function of the input – if a child's primary linguistic data is limited to one language, she learns just that language. But if a child is exposed to two languages, she learns both; and if exposed to three, the child learns three, and so on. If a polylingual child moves to another speech community, she may no longer be in a position to use all the languages she has been exposed to, and she may quickly forget one or more of them. But children who are exposed to and use more than one language will acquire (and retain) all the languages. As far as we know, *any* child can become polylingual.

Another observation is that every language is learned with equal ease. No particular language is so difficult that it is only acquired by a subset of the children who are exposed to it. It is sometimes said that Russian, or Japanese, or English is hard to learn. And it may be true that English-speaking adults find it more difficult to learn certain languages than others. But it is important to note that every child exposed to Russian acquires Russian, and that it is mastered as rapidly and effortlessly by Russian children as English is acquired by English or American children. The same is true of every other language.

We should note also that children who learn more than one language learn each of them with equal ease. Thus, the language of a child's parents doesn't determine the language of the child except in the most obvious sense that children speak whatever language (or languages) is spoken to them. That is to say, every child exposed to English learns English, and every child exposed to Chinese learns Chinese. One final point reinforces this conclusion. This is the observation that if a child born of English-speaking parents is taken to China, and raised there, the child speaks Chinese, not English. And this will be true, as noted earlier, whatever the level of linguistic achievement of the child's parents turns out to be. The observation that the language of the community is sufficient for language development calls into question one of the most common misconceptions about language – that children need explicit instruction in order to learn a language. To sum up, the facts about the process of language acquisition (with respect to differences in language exposure) all show that language acquisition is *uniform* across languages. Let us call this set of properties, then, the **uniformity** of language acquisition.

The universality and uniformity facts about language acquisition are quite interesting, but they are only the tip of the iceberg. As hypothetical scientific observers, we will have uncovered some other facts that make these facts pale by comparison. First, language acquisition is quite *rapid*. Almost all of the complexities of language are mastered by children before they begin school, by age 3 or 4. Of course, there are many *words* that children don't know – but even college students (and their professors) are still developing vocabulary. Also, children usually only begin learning to read and write when they are 5 or 6, and they may not be particularly eloquent in their speech. But, this should not make us lose track of the fact that most 5-year-olds can produce and understand sentences of considerable length and complexity. By age 5, children have mastered nearly all the different types of structures used in their language. At the same time, children of this age are only beginning to count, many of them can't tie their own shoes or draw human figures, and so on. So the rapid mastery of language by the preschool child is another important fact with which to confront a theory of language and mind.

A fourth important fact about language development is the *sequence* of steps children go through in mastering the language of their community. Children

learning the same language all follow an almost identical pattern. They progress through the same stages of acquisition, and in the same order. It is true that the rate at which individual children pass through these stages may differ, however. In other words, stage and not age is the better indicator of a child's level of language development. This fact, too, must be accounted for by any model of language acquisition.

2   How Does Language Acquisition Proceed?

We have collected a number of facts about language acquisition which we would like a model to be able to account for. Now, let's consider some possible mechanisms which might be proposed as the basis for this remarkable process. We will start with mechanisms which have some intuitive appeal, but we will show that none of them is able to explain the facts we have presented so far. In the following chapters, we will introduce a theory which is up to the task.

Trial and error

The fact that children progress through similar stages of language acquisition suggests that language acquisition is not a trial-and-error process. If language acquisition was done simply by trial and error, then we wouldn't expect all children to go through the same steps, that is, make the same trials and errors in the same order. We should have anticipated this conclusion already, even if we hadn't observed that children pass through an invariant sequence of stages. It is unlikely that language is acquired in a trial-and-error process, because, if it were, the children learning the same language wouldn't all converge on the same grammatical rules. Children who reached their conclusions by trial and error would surely sometimes fail to communicate with each other, because they would have reached different conclusions. But, as far as we know, this does not happen. As we noted earlier, children in the same linguistic community all learn the same grammar. (There may be small differences, but these too are systematic.) This uniform convergence on the same Final State would be mysterious if children were free to entertain different hypotheses about the processes underlying language.

  It is worth noting that we have already ruled out two ways in which trial-and-error learning procedures could be used successfully by children to come up with the same set of hypotheses. First, this could happen if the input to children were fixed, that is, if every child received the same experience. But we have

already seen that this is not the case. We emphasized that, to the contrary, language acquisition takes place with diverse environmental experiences – children receiving quite varied input all settle on the same hypotheses.

A second logical possibility is that the Final State is so simple that it can be reached universally even using a trial-and-error approach. This, too, is incorrect. We will see abundant evidence of this in the chapters to follow.

Another source of evidence against trial-and-error learning is the fact that children make so few grammatical errors in language acquisition, at least considering the number of logically possible wrong turns that they could conceivably take – that is, considering the number of hypotheses that are compatible with the data at some stage of acquisition, but which turn out to be incorrect in the light of further data. Somehow, children know to avoid most potential pitfalls. Later we will discuss several examples of linguistic hypotheses that children systematically pick out from among a host of logically equivalent alternatives.

One of the main goals of Universal Grammar is to explain why so few errors occur. Of course, the errors that do occur are more prominent than ones that do not occur, because we simply don't take note when children say exactly what we say. But it is important not to let the errors children make play too prominent a role in the theory. For one thing, we all make errors, even as adults. But these mistakes, too, are not random. Our errors follow from the nature of the mental apparatus we have for learning and using language. The theory must explain both what does and what does not occur. As we will see, in many cases it is more difficult to explain the latter. The kind of errors that we have in mind are subtle grammatical errors, ones that you probably would never expect children to make. They are not the kind of errors we often hear, like saying "pasketti" for "spaghetti."

For now, the point to remember is that, from the standpoint of grammar, language acquisition is relatively *error-free*. It does not conform to a scenario in which each child is free to consider a large and varied set of alternative hypotheses. By all indications, then, it seems that children's hypotheses are not determined by trial and error.

Corrective feedback (negative evidence)

Many people believe that children learn to speak correctly because their parents correct them when they say something wrong. This cannot be the basis of language acquisition for several reasons. First, we have just observed that acquisition is relatively error-free. If children do not make certain kinds of mistakes, we can hardly expect their parents to correct them!

What about the errors that children do make? Researchers have found that parents do not systematically correct children for their grammatical errors

(despite your recollections otherwise). Since the early 1970s, dozens of children have been observed, tape-recorded, and videotaped, for hundreds of hours by many researchers within the field of psycholinguistics. All of these data have been combed through, at great effort, looking for facts to support or contradict particular hypotheses, such as the hypothesis that parents reward children for improving their language skills.

One study of the availability of corrective feedback (or negative evidence) to children, by the psychologists Roger Brown and Camille Hanlon, looked at parental responses to children's utterances. First, these researchers recorded all the cases where the parents' response revealed a misunderstanding of what the child was saying. They found that there was *no* correlation between the parents' understanding and how grammatically well formed the child's utterance was. Parents seem to be able to understand the language of their child perfectly well, however babyish and oversimplified it is.

Brown and Hanlon also recorded all the instances of approval or disapproval in the parents' responses. What they found was that the approval responses didn't distinguish between well-formed sentences and incorrect baby-talk. Parental feedback did distinguish between true sentences and false sentences, regardless of whether these sentences were grammatically well formed or not. For example, one child studied by Brown's group said, "Mama isn't boy, he a girl." Despite the ungrammatical utterance, her mother replied, "That's right." On the other hand, when another child said, "Walt Disney comes on Tuesday," his mother said, "No, he does not" (since, as we know, Walt Disney comes on Sunday). In short, parents do give negative feedback, when children do not tell the truth, but they *don't* distribute positive and negative reinforcement in a way that could be causally related to children's language learning.

There is further evidence against corrective feedback as a major contributing factor in language development. Evidence of this kind is quite convincing, even if it is anecdotal. It is often noted that children fail to respond on the occasions when their parents do try to correct them. For example, one child was heard to say, "Nobody don't like me." Since her parent was a linguist, the parent of this child attempted to respond to the grammaticality of the child's sentence rather than its truth-value. The parent tried to correct the child, saying, "No, say, 'Nobody likes me'." The child responded saying, "Nobody don't like me." This went on for eight repetitions, until finally the child thought she understood the correction. She said, "Oh! Nobody don't lik*es* me."

This anecdote offers considerable insight into the problem children would confront if they were to use parental input as the basis for correcting their own grammatical mistakes. As the example shows, it is one thing for the child to know that she has made a grammatical mistake and quite another to know how to go about instilling changes in her grammar so as to eliminate the mistake. The example points out a distinction that needs to be drawn, between **input** and

intake. Even if corrective feedback, say, were present in the input to children, what they make of the input, their intake, is another matter.

Imitation

It is often suggested that children develop grammatical knowledge through imitation. That is, children learn language by imitating their parents' talk. While there might be a grain of truth to this proposal, it cannot be the whole story. One source of evidence against the claim that children learn by imitation comes from children's **overregularizations**. As one example, children produce incorrect past tense and plural forms such as "goed" and "foots." These forms are clearly not the product of imitation, since they do not appear in the input from adults. These non-adult forms are evidence that children form generalizations (i.e., rules) that enable them to produce linguistic utterances they have not encountered. In this case, children have clearly attained the adult rules for past or plural, but have incorrectly applied them to stems which irregularly do not take them in the adult grammar.

It is interesting to note that non-adult utterances such as "goed" and "foots" continue to appear in children's spontaneous speech for several months, sometimes even years. Apparently they are expunged from children's grammars only after children encounter a number of occurrences of the adult irregular alternative. Somehow, these alternative forms are in competition with the errant forms emanating from the child's grammar. It has been suggested that children's errant forms are jettisoned because their grammars contain a **blocking principle**, or a **uniqueness principle**. According to this type of principle, the presence of the irregular form blocks the rule-generated one. Both forms of expression are permitted simultaneously only if there is abundant evidence (in the input) for each of the alternative forms of expression. Clearly these principles apply in the adult grammar as well as the child's, accounting for the adult acceptance of both "dived" and "dove", but not "singed".

Later in this book, we will discuss in more detail the creative aspect of the child's emerging linguistic system, which is called **productivity**. As we will discuss, children not only produce novel forms that are not part of the adult language, but they develop a system of linguistic rules that is capable of producing an unlimited number of sentences that are acceptable in the adult language. This means that the non-adult utterances children produce pale in comparison to the adult utterances they produce, but which have not previously been encountered in the linguistic environment.

Another example of a construction used by children, but not by adults, concerns their questions. Children sometimes ask ill-formed questions such as, "Where he is going?" What makes it interesting is that children who ask questions like

this also ask correct questions of other types, such as, "Is he going?" This means that they have internalized the rules that are needed in forming these questions. Yet, they often fail to apply the rules correctly in forming questions. These forms are a product of the child's internal grammar (and possibly their performance limitations) and do not reflect their attempts to imitate what they encounter in the input.

It has also been found that some English-speaking children around the age of 3 or 4 consistently insert an "extra" word in some of their questions, such as, "What do you think what pigs eat?" The appearance of the extra word in the language of children learning English cannot be explained as children's response to the input. Although these constructions are not grammatically well formed in English, structures like this are found in certain dialects of German and other languages. It is worth noting that, if the innateness hypothesis is correct, then there is no *a priori* reason to expect children's grammatical hypotheses to be closely tied to their linguistic experience. However, if children make non-adult hypotheses, these should be consistent with the possibilities allowed by UG. This is true whenever the difference between children's productions and the target language mirror differences that occur across languages.

Expansion

Another possible explanation of how children develop language is that the parents *expand* their child's talk, turning the child's simple forms into proper sentences. For example, if the child said "Daddy chair," the parent might respond, "Yes, that's right dear, you're sitting on Daddy's chair." Although parents sometimes do this, it cannot be a mechanism that is crucial for language development. This was shown in a study with preschool children. One group of adults was asked to speak to their children by regularly expanding what the children said. They did this five days a week for three months. Another group of parents served as a control group. For this group, expansions were not given. The main finding was this: the experimental group of children showed no improvement in their language relative to the control group that did not receive expansions of their utterances. Thus, even if some parents do sometimes expand on their children's utterances, this does not seem to be a primary factor in language development.

Motherese

A related suggestion is that parents **simplify** their speech to their children. The idea is that parents use simple forms at the early stages, and gradually increase

the complexity of their own sentences to match the child's development. In this fashion, the child moves ahead a little at a time. Adults do talk differently to children than to other adults, using what is sometimes called "Motherese," or "baby-talk." Adults mumble less to children, they use fewer incorrectly formed sentences, they use shorter sentences, and they frequently use different intonation patterns with young children. However, in many studies it has been found that they do not *invariably* use grammatically simpler sentences. Also, in studies comparing children whose parents used Motherese to those whose parents didn't use Motherese, no difference was found in language development. So it doesn't seem that Motherese serves to pace the information presented to the child, in order to help her learn the language in easy steps.

There is a more fundamental problem with this proposal. By simplifying the linguistic input, a parent is actually making the child's task harder, not easier. Simple input is consistent with a greater variety of grammatical hypotheses than is complex input. We will see an example of this later, in our discussion of structure dependence in chapter 3. It will be shown there that the formation of simple yes/no questions, such as, "Is Bill going to the show?", from their declarative counterparts, "Bill is going to the show," is compatible with both the correct, structure-dependent principle and the simple, but ultimately incorrect, structure-independent operation: move the first "is," "will," "can" etc. to the front of the sentence. The structure-independent operation produces the wrong yes/no questions with more complex input. For example, the application of the structure-independent operation to declarative sentences with a restrictive relative clause, such as, "The man who is running will win," yields the question, "Is the man who running will win?", not, "Will the man who is running win?" This example shows that more complex input can help the learner eliminate erroneous hypotheses. It should be clear that Motherese would not be helpful, and could be detrimental to language development in certain cases.

Conclusion

What can we conclude, then, about *why* and *how* language is acquired? The main conclusion we would draw is that language development seems to be internally driven. It comes from within the child, rather than from anything parents provide from the outside. It seems that children are just built to learn – they don't respond in any obvious way to external pressure or rewards. Rather, they are internally motivated to pick up information from what they hear, and record it and use it in their own behavior. We will see more evidence for this conclusion when we present the results from experimental studies of child language

development. We will argue that the success at language acquisition which children's spontaneous productions reveal is actually an underestimate of children's abilities. For example, experimental procedures enable us to ascertain that children attain knowledge of grammatical principles that are not attested in their linguistic environment.

We will discuss in some detail the solution to this puzzle advanced by the theory of Universal Grammar. According to this theory, children's hypotheses about language are taken to be constrained in advance, in the sense that hypotheses are limited by biological characteristics of the human species. According to this viewpoint, there are innate restrictions on the hypotheses that children may entertain in forming their grammatical system. In this way, the remarkable linguistic feats that children accomplish in their first few years can be accounted for. We will consider further evidence for this position throughout this text.

Bibliographical Comments

Chomsky (1957) and Chomsky (1965) are two of Chomsky's most influential works laying out the theory of language presented here and in the rest of the text. There are many books reporting data on the process of language acquisition; among them are P. Bloom (1993), Brown (1973), Goodluck (1991), Slobin (1985). The original study on negative evidence was presented by Brown and Hanlon (1970); a more recent review of negative evidence is presented by Morgan and Travis (1989). The anecdote about children's failure to be corrected is from McNeil (1970). The proposal that a blocking principle applies to children's overregularizations is made in Pinker (1995). Children's non-adult questions mentioned here will be discussed in detail in chapters 20 and 22. One study of the potential effects of expansion on children's language acquisition is reported by Brown, Cazden, and Bellugi (1969). The phenomenon of Motherese and the limitations of its effects on language acquisition mentioned here are discussed by Newport, Gleitman, and Gleitman (1977); other studies of this phenomenon are collected in Snow and Ferguson (1977).