Language Acquisition

Module 1: First language acquisition, second language acquisition, bilingualism: definitions, differences and similarities

“When we study human language, we are approaching what some might call the human essence, the distinctive qualities of mind that are, so far as we know, unique to [humans].”

(Chomsky, 1968, p. 100)

Language is a sophisticated matter that cannot be studied from one perspective only, but we need to appreciate its complex nature by discussing the contributions of psychology and psycholinguistics, sociology and sociolinguistics, neuroscience, discourse analysis, and education. But first, let us clarify and define our terminology as we dig deep into the complexity of language acquisition.

One of the first terms is ‘Native Language’ (NL). It refers to “the first language that a child learns. It is also known as the primary language, the mother tongue, or the L1 (first language).” (Gass & Salinker, 2008).

Target Language (TL): refers to the language being learned.

First Language Acquisition (Child Language Acquisition): The field that investigates these cases of monolingual language acquisition is known by the generic name of child language acquisition or first language acquisition. A robust empirical research base tells us that, for children who grow up monolingually, the bulk of language is acquired between 18 months and three to four years of age. Child language acquisition happens in a predictable pattern, broadly speaking. First, between the womb and the few first months of life, infants attune themselves to the prosodic and phonological makeup of the language to which they are exposed and they also learn the dynamics of turn taking. During their first year of life they learn to handle one-word utterances. During the second year, two-word utterances and exponential vocabulary growth occur. The third year of life is characterized by syntactic and morphological deployment. Some more pragmatically or syntactically subtle phenomena are learned by five or six years of age. After that point, many more aspects of mature language use are tackled when children are taught how to read and write in school. And as children grow older and their life circumstances diversify, different adolescents and adults will embark on very different kinds of literacy practice and use language for widely differing needs, to the point that neat landmarks of acquisition cannot be demarcated any more. Instead, variability and choice are the most interesting and challenging linguistic phenomena to be explained at those later ages. But the process of acquiring language is essentially completed by all healthy children by age four of life, in terms of most abstract syntax, and by age five or six for most other ‘basics’ of language.
Second language acquisition (SLA): refers to the process of learning another language after the native language has been learned. Sometimes the term refers to the learning of a third or fourth language. The important aspect is that SLA refers to the learning of a nonnative language after the learning of the native language. The second language is commonly referred to as the L2. As with the phrase “second language,” L2 can refer to any language learned after learning the L1, regardless of whether it is the second, third, fourth, or fifth language. By this term, we mean both the acquisition of a second language in a classroom situation, as well as in more “natural” exposure situations.

Foreign Language Learning: Foreign language learning is generally differentiated from second language acquisition in that the former refers to the learning of a nonnative language in the environment of one’s native language (e.g., French speakers learning English in France or Spanish speakers learning French in Spain, Argentina, or Mexico). This is most commonly done within the context of the classroom.

Second language acquisition, on the other hand, generally refers to the learning of a nonnative language in the environment in which that language is spoken (e.g., German speakers learning Japanese in Japan or Punjabi speakers learning English in the United Kingdom). This may or may not take place in a classroom setting. The important point is that learning in a second language environment takes place with considerable access to speakers of the language being learned, whereas learning in a foreign language environment usually does not.

What about Bilingualism?!
Edwards (2006) starts off his article on the foundations of bilingualism by saying “Everyone is bilingual. That is, there is no one in the world (no adult, anyway) who does not know at least a few words in languages other than the maternal variety. If, as an English speaker, you can say c’est la vie or gracias or guten Tag or tovarisch—or even if you only understand them—you clearly have some command of a foreign tongue . . . . The question, of course, is one of degree . . . .” (p. 7). He goes on to say, “it is easy to find definitions of bilingualism that reflect widely divergent responses to the question of degree” (p. 8). Bhatia (2006) states this in an interesting way when he says “the process of second language acquisition—of becoming a bilingual” (p. 5). In other words, the end result of second language acquisition is a bilingual speaker.

Given that bilingualism is seen as the end result and given that we know that native-like competence in a second language is rare, there is some difficulty in discussing bilingualism in this way. Thus, Bhatia and Edwards are referring to two different phenomena. Edwards is saying that one is bilingual at any point in the SL learning process, whereas Bhatia is referring only to the end point and does not deal with whether or not that end point has to be “native” or not.

Valdés (2001a) also discusses the issue of degree when she says “the term bilingual implies not only the ability to use two languages to some degree in everyday life, but also the skilled superior use of both languages at the level of the educated native speaker” (p. 40). She acknowledges that this is a narrow definition, for it considers the bilingual as someone who can “do everything perfectly in two languages and who can pass undetected among monolingual speakers of each of these two languages” (p. 40). This she refers to as the “mythical bilingual.” She argues that there
are, in fact, different types of bilinguals and that it is, therefore, more appropriate to think of bilingualism as a continuum with different amounts of knowledge of the L1 and L2 being represented. In this view, the term *bilingualism* can refer to the process of learning as well as the end result, the product of learning.

Finally, Deuchar and Quay (2000) define bilingual acquisition as “the acquisition of two languages in childhood” (p. 1), although they point to the difficulties involved in this definition given the many situations that can be in place. They point to De Houwer (1995), who talks about *bilingual first language acquisition*, referring to situations when there is regular exposure to two languages within the first month of birth and *bilingual second language acquisition*, referring to situations where exposure begins later than one month after birth but before age two. Wei (2000, pp. 6–7) presents a useful table of various definitions/types of bilinguals.

**SLA Vs. Bilingualism:**
There are some key differences between the two fields. SLA often favours the study of late-starting acquirers, whereas bilingualism favours the study of people who had a very early start with their languages.

Additionally, one can say that bilingualism researchers tend to focus on the products of bilingualism as deployed in already mature bilingual capabilities of children or adults, whereas SLA researchers tend to focus on the pathways towards becoming competent in more languages than one. This in turn means that in SLA the emphasis often is on the incipient stages rather than on ultimate, mature competence.

A third difference is that bilingual research typically maintains a focus on all the languages of an individual, whereas SLA traditionally orients strongly towards the second language, to the point that the first language may be abstracted out of the research picture. In this sense, SLA may be construed as the pure opposite of monolingual (first) child language acquisition. Indeed, in both fields monolingual competence is often taken as the default benchmark of language development.

**Some terms used to describe cases of bilingualism from the literature:**
- balanced bilingual: someone whose mastery of two languages is roughly equivalent
- coordinate bilingual: someone whose two languages are learned in distinctively separate contexts
- dominant bilingual: someone with greater proficiency in one of his or her languages and uses it significantly more than the other language(s)
- early bilingual: someone who has acquired two languages early in childhood
- late bilingual: someone who has become a bilingual later than childhood
- maximal bilingual: someone with near-native control of two or more languages
- minimal bilingual: someone with only a few words and phrases in a second language
- receptive bilingual: someone who understands a second language, in either its spoken or written form, or both, but does not necessarily speak or write it
- simultaneous bilingual: someone whose two languages are present from the onset of speech
- concessive (succesive) bilingual: someone whose second language is added at some stage after the first has begun to develop
Effects of Bilingualism on cognitive Development:
Bilingual and monolingual speakers may develop different patterns of cognitive skills due to the different language environments they experience.

- Control of attention: a lot of work shows that bilingual speakers activate both language systems even when they are only conversing in one language (Brysbaert, 1998; Francis, 1999; Gollan & Kroll, 2001; Smith, 1997). This requires a good deal of attentional control. During linguistic tasks, bilinguals must constantly inhibit one language and activate the other but must be capable of switching quickly from one language to another if required. These sorts of attention switching skills are considered to be under the control of the central executive mechanism in the brain, which is said to control and regulate a large number of cognitive processes such as planning, memory and attention. There is now robust evidence that bilinguals outperform monolinguals on these tasks, much of it provided by Bialystok and her colleagues. For example, Martin-Rhee and Bialystok (2008) tested bilingual and monolingual four- and five-year-old children on the Simon task (Simon, 1969). In this task, a red square and a blue square are presented on a computer screen and the participants have to press a red button in response to the red square and a blue button in response to the blue square. Sneakily, for half of the trials, the squares occur on the opposite side of the screen to the corresponding keys. So for example, the red square might appear on the left side of the screen, requiring the child to press the red button on the right side of the computer keyboard. People’s instinctive reaction in this task is to press the key on the same side as the stimulus, so inhibiting this reaction requires a good level of cognitive control and slows down their reaction times. Martin-Rhee and Bialystok found that bilinguals responded faster than monolinguals because they were quicker at resolving the conflict between the two possible responses. They concluded that bilinguals are better at selectively attending to conflicting cues because “they must constantly control attention between two active and competing language systems” (Martin-Rhee & Bialystok, 2008, p. 91).

- Metalinguistic Awareness: refers to the ability to reflect on and think about the nature of language and its functions. For example, work on phonological awareness (awareness of the sound system of a language) has produced contradictory results. Bruck and Genesee (1995) reported that bilingual children showed better performance on an onset-rime segmentation task at age five years (separating words into the onset and rime; swift into sw and ift) although the advantage disappeared a year later. However, monolingual five-year-olds were better than bilinguals on a phoneme-counting task (e.g. identifying that the word run has three phonemes). Similarly, Bialystok, Majumder & Martin (2003) found that Spanish–English, but not Chinese–English bilinguals, were better at a phoneme segmentation task (e.g. segmenting run into /r/, /ʌ/ and /n/). However, they also reported no difference between groups on a phoneme substitution task (e.g. substitute /s/ for /k/ to make sat out of cat). These mixed results suggest that the pattern of performance may not be straightforward. It is probably the case that other factors are equally influential in these tasks; factors such as the child’s age, her ability in her two languages, the task she is performing and even perhaps the nature of the two languages she is learning.

- Language Proficiency and Fluency: there seems to be evidence that bilinguals are disadvantaged compared to monolinguals in some tasks. In particular, studies have shown that bilinguals may have more difficulty accessing words from memory. For example, bilinguals tend to be slower at
rapid picture naming tasks and tend to experience more ‘tip of the tongue’ phenomena, which happen when a speaker just cannot bring the right word to mind (Gollan & Kroll, 2001; Gollan & Silverberg, 2001).

**Bilingualism and code-switching: a common phenomenon:**
A common phenomenon among bilingual speakers is code-switching, which essentially refers to the use of more than one language in the course of a conversation. Sometimes this might happen because of the lack of a concept in one language and its presence in the other; sometimes it might be for humor; and sometimes it might happen simply because of the social context. For example, Grosjean (2001, p. 3) presents the following diagram to illustrate the issue of language mode, which is “the state of activation of the bilingual’s languages and language processing mechanisms at a given point in time” (p. 2). The native language (here called the base language) is always totally activated; it is the language that controls linguistic activities. The guest language, on the other hand, can be in low to high activation depending on the context. Only in bilingual language mode (the right side of the diagram) is there almost equal activation, and it is in these contexts when code-switching occurs.

![Language Mode Diagram](image)

**Figure 2.2** The language mode continuum.


Bilingualism, or at least some form of knowledge of more than one language, is so common throughout the world that Cook has proposed that the “normal” propensity is for humans to know more than one language rather than taking monolingualism as the default position. He refers to this as *multicompetence*, which he defines as the “knowledge of two or more languages in one mind” (Cook, 2003, p. 2; cf. Cook, 1991, 1992). If multicompetence is the “norm,” then there needs to be a re-evaluation of what it means to be a native speaker of a language. Cook (2005) argued that there are effects of multilingualism on how individuals process their native language, even
individuals with a minimal knowledge of a second language. Cook further argues that the monolingual orientation of second language acquisition belies the reality of the context of language learning in much of the world where knowledge of more than one language is the norm.

**Module 2: First/Child Language Acquisition**

**A: Language Acquisition Research**

i. L1 Research: *Behaviorism vs. Language Acquisition Device (Nativism)*

Until the 1960s, the research agenda of language acquisition studies, just like that of psychology and linguistics in general, was strongly determined by behaviorist learning theories. An explanation referring to mental capacities of the learner did not seem to make much sense in that context; it would, indeed, have been regarded as a non-scientific approach to the problem. Only after the constraints and restrictions of behaviorist psychology had been shaken off could the language sciences begin to understand language learning as a mental activity happening in the cognitive system of the individual. Chomsky’s (1959) famous and influential review of Skinner’s (1957) book *Verbal Behavior* is a milestone on this road to the cognitive turn. What this term is meant to convey is that it is the study of human cognition, which is now identified as the major task of linguistics, in close cooperation with other sciences, especially cognitive psychology and philosophy (see Chomsky 1968). With respect to the language faculty, the issues put on the research agenda by this change of perspective include, among other things, the problem of how to characterize the knowledge system represented in the mind of a person who speaks and understands a particular language, as well as to explain how this knowledge is used and, most importantly in the present context, how this linguistic knowledge and the ability to use it are acquired. The Language Acquisition Device, then, represents the initial state of the language faculty, that is, prior to any exposure to the language to be acquired (see Chomsky 1988). This new approach had an enormous impact on L1 research, and as early as in the early 1960s appeared the first of an ever increasing number of publications applying these ideas to the study of first language acquisition.

ii. L2 Research:

L2 research, on the other hand, took somewhat longer to liberate itself from the dominating influence of behaviorism. This is partly due, perhaps, to the fact that for a long time it had exclusively been occupied, and still continues to be primarily concerned, with foreign language learning in classroom settings, rather than with naturalistic L2 acquisition. The idea that learning crucially implies changing previously acquired behavior seems to have been deeply rooted in language teaching. It is therefore not surprising that interference from L1 was, and in part still is, regarded as the major factor determining the shape of L2 speech. The research paradigm which elaborated this idea in considerable detail is Contrastive Analysis (CA).

Contrastive Analysis continued a line of thought which had been expressed quite clearly as early as 1945 by Charles C. Fries in the following frequently quoted statement:
The most efficient materials are those that are based upon a scientific description of the language to be learned, carefully compared with a parallel description of the native language of the learner. (Fries 1945: 9)

The next step was taken by Robert Lado, a former student of Fries, in assuming that ‘individuals tend to transfer forms and meanings, and the distribution of forms and meanings of their native language and culture to the foreign language and culture’ (Lado 1957: 2). This assumption, which Lado as well as many others at the time regarded as an uncontroversial generalization based on empirical observation, was turned into a prediction, the perhaps major theoretical claim of CA, when Lado (1957) and Weinreich (1953) before him argued that ‘those elements that are similar to his [i.e. the learner’s, JMM] native language will be simple for him, and those elements that are different will be difficult’ (Lado 1957: 2).

We can sum up by saying that second language research suffered longer than first language research from its behaviorist heritage. By focusing on the comparison of linguistic structures justified exclusively in grammatical terms rather than with respect to their psycholinguistic plausibility, and, moreover, by defining learning primarily in terms of habit formation and changing of habits, questions relating to the possibility of a common underlying language making capacity for the various types of language acquisition could not even be formulated. As a result, the role of the native language in second language acquisition was seen exclusively as a possible source of transfer.

iii. L2 Research: A Cognitive Turn

An explicitly cognitive orientation of second language acquisition research was initiated in the late 1960s. The change is best illustrated by the seminal paper by Pit Corder (1967). He refers to the child’s ‘innate predisposition to acquire language’ and the ‘internal mechanism’ which makes the acquisition of grammar possible, and then raises the question of whether the child’s language making capacity remains available to second language learners. Although he is careful about the conclusions to be drawn from these assumptions, he leaves no doubt about the fact that he favors a positive answer, postulating ‘the same mechanism’ for both L1 and L2 acquisition, and proposes (p. 164)

as a working hypothesis that some at least of the strategies adopted by the learner of a second language are substantially the same as those by which a first language is acquired. Such a proposal does not imply that the course or sequence of learning is the same in both cases.

What exactly Corder means by ‘strategies’ is not entirely clear, nor does he elaborate on the last point, that is, what might cause the emergence of different learning sequences in spite of the claim that the underlying mechanisms are the same. He does, however, list what he sees as differences between the two acquisition types, namely that (1) children acquiring their L1, as opposed to L2 learners, are inevitably successful, (2) L1 development is part of the child’s maturational process, (3) at the onset of second language acquisition, another language is already present, and (4) the motivation for language acquisition is quite different in the two cases. Corder suspects that this last
factor, motivation, is the principal one distinguishing first and second language acquisition. In order to gain insights into the nature of the underlying mechanism and of the strategies used in second language acquisition, Corder suggests studying the errors found in L2 speech. He distinguishes between random mistakes and systematic errors. The latter, he claims (p. 166), ‘reveal his [the learner’s, JMM] underlying knowledge of the language to date, or, as we may call it his transitional competence’. If, for example, learners use the form thinked, this suggests that they have acquired knowledge about tense marking in English, even if this particular form is an error, deviating from the target norm. The truly stimulating ideas in Corder (1967), with respect to the present discussion, are that he explicitly suggested the same underlying mechanism for L1 and L2 acquisition, introduced the notion of ‘transitional competence’, and demanded that the focus of L2 research should be on the learner, rather than on learners’ productions. This can only be achieved if acquisition studies strive for psycholinguistically plausible grammatical analyses of learner utterances. In other words, L2 learners are assumed to acquire systematic knowledge about the L2: a ‘third system in addition to the NL [native language, JMM] of learners and the TL [target language, JMM] to be learned’ is introduced, to use Selinker’s (1992: 18) words. Note, however, that assuming a kind of transitional competence does not oblige us to subscribe to the idea of one and the same mechanism underlying L1 and L2 acquisition.

Suggestions similar to the ‘transitional competence’ were indeed made by a number of authors, proposing ‘approximative systems’ (Nemser 1971), ‘idiosyncratic dialects’ (Corder 1971) or ‘interlanguages’ (Selinker 1972). These terms are not synonymous, but they coincide in so far as they postulate a structured transitional knowledge base in the L2 learner. It contains elements of the target grammar, possibly also elements of the L1 grammar (‘interlingual errors’, Richards 1971), and, most importantly, elements different from both source and target systems, ‘developmental errors’ (‘intralingual errors’, Richards 1971) which prove that the learner is actively and creatively participating in the acquisitional process. The term most generally adopted is Selinker’s (1972) ‘interlanguage’ (IL)

**B: First Language Acquisition:**

The gift for language which manifests itself in the effortless acquisition of language by toddlers can safely be qualified as a species-specific endowment of humans. In fact, it enables children to develop a full grammatical competence of the languages they are exposed to, independently of individual properties like intelligence, personality, strength of memory and so on, or of particularities of the learning environment, for example social settings, whether the child is an only child or has siblings, birth order among siblings, whether the child has one or more primary caregivers, communicative styles of parents or caregivers, and so forth. The theoretical framework adopted here as the theory about the human language making capacity is that of Universal Grammar (UG), as it has been developed by Chomsky (1981a, 1986, 1995, 2000a, among others) and his colleagues.

A central aspect of the theory of UG is that it views the human language faculty as comprising a priori knowledge about the structure of language. Importantly, knowledge of language is understood as being internal to the human mind/brain, and the object of linguistic theory is therefore the mental grammar or competence of the individual which Chomsky (1986) refers to as I-language, an
internal entity of the individual, as opposed to E-language, ‘E’ suggesting ‘external’, that is, the overt products in language use.

The genetically transmitted or innate implicit linguistic knowledge which UG attempts to capture is formulated in this theory in terms of abstract principles determining the set of possible human languages. They are universal in the sense that it is predicted that no grammar of a natural language, that is, no I-language, will violate these principles. Note that this allows for the possibility that principles may simply not be manifest in a given language. If, for example, UG contains principles explaining properties of prepositions, these will not be exhibited in languages without prepositions. They are nevertheless universal in the intended meaning of the term, namely wherever prepositions occur, they are constrained by the relevant UG principles. To sum up this point, UG is designed to capture all and only those properties which human languages have in common and thus to explain the nature of this species-specific faculty, but not every property present in a grammar must conform to principles of UG. In fact, many grammatical phenomena are language-specific, representing particularities of individual languages, and UG has nothing to say about them. Take, for example, the fact that French interrogative words qui/que ‘who/what’ can both refer to objects, but only qui can question subjects. Although this is undoubtedly a property of French grammar which needs to be represented in the mental grammars of speakers of this language, it is rather unlikely that it follows from a constraint imposed by UG. In other words, the grammar of a particular language is characterized by universal as well as language-specific features. Both reflect properties of the language faculty, but the latter must be acquired and will be contingent on properties of the learner and the learning environment. It is thus conceivable that a learner might not cognize that *que vient? ‘what is coming?’ is ungrammatical.

-The role of UG in explaining the human language making capacity.

The basic idea is that since UG is conceived of as representing the initial state of the language faculty, it can also be understood as a crucial component of the LAD, the Language Acquisition Device. The claim that UG indeed represents the initial state of the child’s linguistic development has, in fact, long been a fundamental assumption of generative theorizing and continues to be a defining property of UG in that Universal Grammar is understood as a theory about what the child brings to the task of language acquisition – or ‘growth’, as Chomsky prefers to say, comparing language development to the growth of organs – before any experience with the target language. To quote only one instance where he explains this idea, Chomsky (2000a: 4) suggests that we think of the initial state as a ‘language acquisition device’ that takes experience as ‘input’ and gives the language as an ‘output’ – an ‘output’ that is internally represented in the mind/brain.

- Milestones of First Language Acquisition:

The First Sounds:

Speech perception – which includes, for instance, the ability to segment the speech stream into meaningful units, to recognize one’s own name in the speech stream, or to distinguish between similar sounding vowels (e.g. /ee/ and /oo/) – is a critical skill that infants develop early in life.
These early language skills also involve visual information; for instance, infants as young as two months have been shown to be able to match vowel sounds they hear with the appropriate lip, mouth, and face movements. These early speech perception skills related to the sound structure of language may help infants to bootstrap into more complex language competencies; bootstrapping refers to the possibility that skills in one area of language might help the child to develop competencies in other language areas. For instance, infants’ ability to recognize their own names in the speech stream (which appears around the fifth month) may provide them with a means to recognize novel, adjacent words.

Early research in speech perception demonstrated that during their first few months of life, infants are able to discriminate between similar sounds (for example, between /b/ and /p/) both in their native language(s) as well as in other languages. Over time, however, infants become more attuned to their native language(s) and less able to make sound distinctions in other languages.

Infants are born with the capacity to learn any language in the world, but the capacity to hear like a native fades very early on.

The first sound made by all infants is crying. All infants can do this immediately from birth; although crying may signal distress, discomfort, boredom, or other emotions in the first month of life, it is not an intentional attempt to communicate. From about the second to fifth month, infants engage in cooing. Coos are generally vowel-like sounds which are often interpreted as signs of pleasure and playfulness.

All infants begin to babble anywhere between four and six months and generally continue to do so until they reach around one year of age. Babbling is characterized by vowel or consonant–vowel sounds such as ouw-ouw or ma-ma. At this age, infants’ tongues tend to be relatively large compared to the size of their mouths, and as a result, these sounds will often be palatals, such as [y] or [Ĩ]. Labial sounds such as [b] and [m] are also common. Babbling begins to conform to the sound patterns of the adults’ language between six and ten months of age, with adult native speakers showing the ability to discriminate the babbles of Chinese, Arabic, English, or French infants.

Although there is no meaning (such as a demand for food) associated with this babbling for hearing or for deaf infants, it can be a source of interactive play. In some cultures, infants are encouraged to continue to babble by caregivers’ smiles or touches, or by their own babbling in return. Infants will often stop babbling in order to listen to their interlocutor (sometimes engaging in give-and-take exchanges known as proto-conversations), and around the fifth month, some infants are able to immediately imitate simple sound sequences presented to them.

The First Words:

Sometime around their first birthdays, children begin to assign specific meanings to the sounds they produce. These first words mark the beginning of what is known as the holophrastic stage. Holophrastic means ‘characterized by one-word sentences’; infants tend to use single words to communicate a variety of complex functions.
Children’s words at this stage tend to be concrete objects which are grounded in and central to everyday experiences and interactions (such as light, tree, water), rather than abstract concepts (peace, happiness). These first words tend to be **content words** (bear or bed) rather than **function words** (the, and, on). For children learning English, most first words are nouns. This seems to be related to the fact that sentences in English typically end with nouns, where they are salient, or more noticeable, to learners. This is not the case for children learning all languages, however. For instance, Korean-learning infants’ first words are often verbs; in the Korean language, verbs are sentence-final and sentences may consist of only a verb.

While working to master the vocabulary around them, children often engage in both semantic **overextension** and **underextension**. For instance, a child may overextend the meaning of the word water to include not just drinking water, but also juice, milk, and soda. Underextension, which seems to be less common, refers to the reverse phenomenon: a child, for example, might use baby only to refer to an infant sibling and not to the other babies he/she encounters.

Around age two, children enter the **two-word stage**, characterized by use of phrases which are not more than two words. For English-learning infants, this typically means combining a subject and verb (e.g. baby cry, mama sleep) or a verb and modifier (e.g. eat now, go out). The ordering of these two-word phrases is not fixed, however, and there tends to be limited systematic use of grammatical morphology (for example, the possessive is formed as Miranda bed rather than Miranda’s bed).

As in many other stages of their linguistic development, children’s capacity for comprehending words outpaces their production ability. For instance, around the age of one, children can typically **understand** about seventy different words, but only productively **use** about six. There is about a four- to six-month delay between when children can comprehend a given number of words and when they can produce that many words themselves. Sometime around the end of the second year, children’s productive vocabulary begins to develop rapidly; this is sometimes known as the **vocabulary spurt**. During this period, children begin to add about two hundred words a month to their vocabularies!

At approximately two and half years of age, children begin to produce phrases of three or more words, entering the **multi-word stage** (e.g. Graham go out, Daddy cook dinner, Baby food all gone). Children’s language at this stage has been described as **telegraphic speech** because, like the economical language used in telegraphs, it is seemingly direct and makes only limited use of morphological and syntactic markers.

**First Sentences: Morphological and Syntactic Development**

Many diary, observational, and experimental studies have documented and explored how children become competent users of their language’s system of morphology and syntax. From this research, we know that for all languages, both signed and spoken, this process seems to involve the formation of internal “rules”; in other words, children’s increasingly regular use of grammatical forms (even non-adult-like or “incorrect” usages such as broked or foots) may reflect children’s developing grammatical rule systems.
We also know that children seem to begin to acquire this grammatical competence at a very young age and, as in vocabulary development, comprehension skills outpace production. For instance, children who are only seventeen months of age, and typically still producing only one- or two- word utterances, tend to look longer at video clips that correctly correspond to the grammar of the oral commentary. For instance, children who hear “The bear sat on the bird” and are shown two pictures (one of a bear sitting on a bird and another of a bird sitting on a bear) will look longer at the picture where the bear is sitting on the bird. This research demonstrates that even at very young ages children are tuned into the semantic significance of their language’s grammatical structures.

Research has also demonstrated that morphological and syntactic development is predictable. In other words, all children follow similar patterns and pass through the same developmental sequences as their competence develops. Although there is some variation depending on the language being acquired, many patterns and processes are constant across different language and cultural groups. The development of inflectional and derivational morphology in children’s productive language becomes apparent once the child enters the multiple-word stage and continues through age five. The development of inflectional morphology was the focus of early and intensive investigation. Brown’s investigation of Adam, Sarah, and Eve made important advances in this area. (See Box 6.2.)
Box 6.2 Acquisition of grammatical morphemes by Adam, Sarah, and Eve by MLU stage (I–V) and age (year; month)

This table presents the stages of acquisition of grammatical forms by three children. Their ages at each stage are presented in italicized parentheses. The first number in the parentheses represents the year and the second the month. Thus 2;3 signifies that the child is two years and three months old.

<table>
<thead>
<tr>
<th>Stages</th>
<th>Adam</th>
<th>Sarah</th>
<th>Eve</th>
</tr>
</thead>
<tbody>
<tr>
<td>I (2:3)</td>
<td>Present progressive</td>
<td>I (2:3)</td>
<td>I (1:6)</td>
</tr>
<tr>
<td>II (2:6)</td>
<td>Plural</td>
<td>II (2:10)</td>
<td>II (1:9)</td>
</tr>
<tr>
<td>III (2:11)</td>
<td>Uncontractible copula</td>
<td>III (3:1)</td>
<td>III (1:11)</td>
</tr>
<tr>
<td>IV (3:2)</td>
<td>Past irregular</td>
<td>Articles</td>
<td>Plural</td>
</tr>
<tr>
<td>V (3:6)</td>
<td>Articles</td>
<td>IV (3:8)</td>
<td>Possessive</td>
</tr>
<tr>
<td></td>
<td>Third person irregular</td>
<td>Third person regular</td>
<td>Past regular</td>
</tr>
<tr>
<td></td>
<td>Possessive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V (3:6)</td>
<td>Third person regular</td>
<td>V (4:0)</td>
<td>Uncontractible copula</td>
</tr>
<tr>
<td></td>
<td>Past regular</td>
<td>Uncontractible auxiliary</td>
<td>Past irregular</td>
</tr>
<tr>
<td></td>
<td>Uncontractible auxiliary</td>
<td>Contractable copula</td>
<td>Articles</td>
</tr>
<tr>
<td></td>
<td>Contractable copula</td>
<td>Third person irregular</td>
<td>Third person regular</td>
</tr>
<tr>
<td></td>
<td>Contractable auxiliary</td>
<td>Contractable auxiliary</td>
<td>Third person irregular</td>
</tr>
</tbody>
</table>

Note: Table adapted from Brown (1973).

Stages (I–V) correspond to Mean Length of Utterance (MLU); stage I, MLU = 1.75; stage II, MLU = 2.25; stage III, MLU = 2.75; stage IV, MLU = 3.50; stage V, MLU = 4. Acquisition of a grammatical form was defined as three successive samples in which the morpheme appears at least 90 percent of the time in obligatory contexts.

To calculate the MLU, count the first 100 utterances of a transcribed sample of a child’s language. Next, count the number of morphemes in each utterance and then add up the number of morphemes in all of the 100 utterances. Finally, divide this total by 100 to get the Mean Length of Utterance. If you don’t have 100 utterances, count the number of utterances you have and divide the total number of morphemes by the number of utterances (adapted from Brown 1973). Count categoricals (e.g. wanna, doncha) as one morpheme. Also, don’t count fillers (um, oh) or sounds (aack, etc.).
Through analysis of Adam, Sarah, and Eve’s spontaneous speech, Brown mapped out when different grammatical morphemes consistently appeared in their speech and how this corresponded to other aspects of their language, in particular to **mean length of utterance (MLU)**. MLU is a widely used measurement of the complexity of children’s language and is calculated from the average number of *morphemes* (not words) per utterance. Brown illustrated that: (1) the order of acquisition was similar across the three unacquainted children (with present progressive, plural, and past irregular verb forms appearing first); (2) the age at which children acquired competence in using these forms varied widely (compare, for instance, Eve and Adam at age two years and three months in **Box 6.2**); and (3) the MLU stage served as a good index of the level of development for grammatical morphology (and indeed was much more predictive of grammatical development than age). More recent research has stressed the importance of vocabulary as a predictor of grammatical development.

Another early study which sheds light on when children acquire inflectional morphology was Jean Berko’s famous “wug” study. Rather than recording and analyzing children’s spontaneous speech as Brown did, Berko asked young children of different ages to form the plural of unknown, nonsense creatures, such as “wugs.” (See **Figure 6.2**.) The experimenter pointed to an item and said, “This is a wug.” She then showed a picture with the same two animals and said, “Now here is another one. There are two of them. There are two _____?” Berko found that even preschool children were able to form the plural correctly, demonstrating that they had learned the rule for forming plurals and could apply this rule correctly in novel contexts, and were not just repeating forms which they had previously heard.

In developing these rules, children pass through predictable stages. For instance, children **overgeneralize** in the early phases of acquisition, meaning that they apply the regular rules of
grammar to irregular nouns and verbs. Overgeneralization leads to forms which we sometimes hear in the speech of young children such as goed, eated, foots, and fishes. This process is often described as consisting of three phases:

- **Phase 1:** The child uses the correct past tense of go, for instance, but does not relate this past-tense went to present-tense go. Rather, went is treated as a separate lexical item.

- **Phase 2:** The child constructs a rule for forming the past tense and begins to overgeneralize this rule to irregular forms such as go (resulting in forms such as goed).

- **Phase 3:** The child learns that there are (many) exceptions to this rule and acquires the ability to apply this rule selectively.

Note that from the observer’s or parents’ perspectives, this development is “U-shaped” – that is, children can appear to be decreasing rather than increasing in their accuracy of past-tense use as they enter phase 2. However, this apparent “back-sliding” is an important sign of linguistic development.

We see similar patterns, known as “developmental sequences,” in other areas of grammar, such as the formation of English negatives and interrogatives. As outlined in Box 6.3, children move through identifiable stages, although these stages are more continuous and overlapping than discrete. Note that from the parents’ perspective, children’s development is also not always straightforward. For instance, a child will likely produce inverted yes/no questions (Did Karalyn eat cake?), while still using normal declarative word order for WH-questions (How Izzy go out?).
These data highlight the extent to which negation and interrogative development are interrelated and also dependent upon development of the necessary vocabulary. For instance, both questions and negatives are dependent upon acquisition of the verb auxiliary system (including, for instance, do, does, did, is, am, have, has) and modals (for example, can, could, may, might). These data also demonstrate the rule-governed nature of children’s language production: the “mistakes” in the examples here are generally not the result of children repeating the language that surrounds them; rather, they reflect the developing rule system of children’s language.

**Specialization of the brain for language:**

One of the first people to find such evidence was Paul Broca, a French surgeon. In 1861, he published a report on a patient who had had great difficulty producing speech. At autopsy, Broca found that he had damage in the lower edge of the left frontal lobe, in an area now called “Broca’s area.” Four years later, Broca published a further report where he showed that damage to areas in the left hemisphere produced aphasia, while damage to the corresponding areas in the right hemisphere did not (Broca 1856, 1861). In 1874, a German doctor, Carl Wernicke, followed up this
research in a monograph where he described patients with deficits in language comprehension, deficits associated with lesions or injury in the left hemisphere but outside Broca’s area.

When people are listening to speech, the auditory signals received through the ears travel first to an area on each side of the brain called Herschl’s area, part of the auditory cortex. Information from the right ear goes directly to the left hemisphere of the brain, and from the left ear goes to the right hemisphere. Then there appears to be a division of labor, with the words of a message going mainly to the left hemisphere, and properties such as intonation, rate of speech, pitch, rhythm, and stress going mainly to the right hemisphere. While these two kinds of information may be stored separately, they are connected by fibers that link the two hemispheres, the corpus callosum.

Evidence of specialization in the brain for language has traditionally been drawn from three main sources: hemispheric dominance, effects of trauma or injury, and the effects on language of hemispherectomy.

**Hemispheric dominance.** Almost everyone shows left hemisphere dominance. When people listen to words played simultaneously into each ear through headphones, they tend to hear words played to the right ear and not those played to the left ear. Words into the right ear go directly to the left or contralateral hemisphere.

**Hemispheric damage.** When someone has a stroke or receives a brain injury, only injuries to the left hemisphere affect language. In particular, only left hemisphere injuries in or near Wernicke’s and Broca’s areas are consistently associated with language loss or disturbance.

**Hemispherectomies.** When there has been extensive trauma to the brain, it is sometimes necessary to remove one or other hemisphere entirely. Whether the removal is partial or complete, the two hemispheres again show that they are not symmetrical. Removal of the left hemisphere typically results in loss of language, whereas removal of the right hemisphere leaves language unaffected (Dennis & Whitaker 1975).

**Hearing vs. Deaf:**

Researchers have used fMRI to examine sentence-processing carried out by native English speakers with normal hearing and by native signers of American Sign Language. The signers were either hearing (and hence bilingual in ASL and English) or congenitally deaf. The areas of the brain activated in the two populations were the classical language areas in the left hemisphere.

In another study of the deaf, researchers looked at the ERPs elicited by anomalies in either signed or spoken sentences. They compared four groups of adults: (a) deaf adults, born of deaf parents, who had learnt ASL at a young age; (b) hearing adults, also born of deaf parents, who had likewise learnt ASL young as a first language; (c) hearing adults, born of hearing parents, who learnt ASL after age seventeen; and, finally, (d) hearing adults with no experience of ASL at all. The left hemisphere showed extensive activity for all four groups. But the right hemisphere was also involved in language processing for the deaf adult children of deaf parents, and, to a lesser extent, the hearing adult children of deaf parents.
A Sensitive Period for Language Acquisition?

Sensitive periods in development have long intrigued biologists. In many species, there are periods when learning of certain kinds can take place more effectively than later in development. Sometimes the sensitive period is critical in that, once it is past, that learning can no longer occur, as in imprinting in chicks, or the acquisition of the pertinent songs in young songbirds. In 1967, Eric Lenneberg marshaled an impressive array of evidence for a critical period for the learning of language. Lenneberg argued that the reason language couldn’t be recovered after puberty was that lateralization (specialization of the left hemisphere for language) was by then complete. And any areas of the right hemisphere that could deal with language were also assigned. After lateralization of both hemispheres, then, it is no longer possible to allocate language functions to other areas of the brain; and if areas already assigned for language become damaged, there is now nowhere else to “put” language. These arguments for a critical period have been challenged by later work. Several studies have shown that lateralization for language is essentially complete earlier than Lenneberg proposed, around age five (e.g., Witelson & Pallie 1973). This in turn raises questions about whether the notion of a critical period for language is correlated with the completion of lateralization. If children found it easier to learn a language prior to lateralization, age five to six ought to mark a point at which all this changes. But children appear highly successful in learning additional languages up to age twelve and even later. Lateralization may therefore offer a less than compelling explanation of difficulties in later learning.

Children with early brain injury appear to attain normal (or nearnormal) language despite damage to areas critical for language in adults. Their language capacities appear to be fairly resilient. But sometimes children may have to have an entire hemisphere removed – for instance, to arrest seizures associated with Sturge-Weber-Dimitri syndrome. Their subsequent language abilities depend on which hemisphere was removed. Language is more affected by loss of the left hemisphere, and visual processing more by loss of the right hemisphere.

Is there an innate language acquisition device?

Many researchers have proposed that people have an innate capacity for language acquisition, a capacity that distinguishes humans from other species. The question is, just what is innate? Are there special areas of the brain dedicated to language? Here, as we have seen, the answer appears to be yes, but the specialization in these areas develops with exposure to language. And lateralization for language, then, only begins during the second year of life. Could there be built-in categories, universal in human languages, such as “noun” and “verb”? Are there any built-in structures? How might built-in categories and structures affect the process of acquisition? Or, instead of this, could humans be endowed with learning mechanisms specialized just for language? These questions have elicited considerable debate. Here we look at some of the proposals that have been made.

In the 1960s, Chomsky proposed that the human capacity for language was innate. The assumptions here are that: (a) natural-language syntax is too complex for children to learn from what they hear around them, because (b) adults offer such a distorted and imperfect source of data. And (c) children learn their first language so fast that they must rely on some innate capacity, specifically for
syntactic categories and syntactic structure. As Chomsky put it, “The grammar has to be discovered by the child on the basis of the data available to him, through the use of the innate capacities with which he is endowed” (1972:183). We will focus on the existence of innate categories and structures.

Current positions on the status of innate categories and structures tend to espouse one of two main views: the continuity view and the maturational view (see O’Grady 1997). Both focus on syntax, which is treated as autonomous, independent of phonology and the lexicon. Morphology is sometimes included with syntax, because it too can be viewed as rule-based and because it marks syntactic distinctions such as “subject of” and “predicate,” as well as parts of speech (by distinguishing nouns from verbs from adjectives, for instance). The continuity view assumes that children use the same notions and relations throughout development; they are present from birth. These researchers generally subscribe to an innate Universal Grammar (UG) common to speakers at every stage of development (e.g., Pinker 1984). The strong version of this approach assumes that children just beginning to speak have the same mental representations for linguistic constructions as adults (e.g., Lust 1994). A weaker version assumes that, although children come with all the categories and operations, they don’t make use of them all immediately. They first have to learn how to instantiate such elements as relativizers (that, who, which), complementizers (to, for, whether), or wh- forms for questions (what, where, why). One result is that the focus in most studies has been limited to the learning of grammatical elements. But languages differ, and researchers have to take that into account. One version of UG that has been invoked contains parameters that allows for variation across languages. Some languages, for example, have complement constructions that are head-initial – the term introducing a complement comes first, followed by the complement, as in English (e.g., They said that he came in at five). Others are head-final and place the head after the complement, as in Japanese. In UG, this difference is captured by a word-order parameter with two values: head-initial and head-final. Upon exposure to a language, children “discover” the value of this parameter and set it accordingly. Languages also differ in whether finite verbs appear with an overt subject in the form of a nominal or pronoun, as in French (e.g., je veux partir ‘I want to-go’), or whether they can omit these, as in Italian (e.g., voglio partire ‘I-want to-go’). This variation is captured by the parameter called subject-drop, which is either permitted or not. Both these parameters are assumed to be present from the start in acquisition, but they can only be set after experience with a language. Children also need to acquire relevant words with which to display their syntactic knowledge. Until they use nouns and verbs, there is no way to tell what underlying grammatical categories or structures they might know. Take two year-olds who haven’t yet learnt I or me in self-reference and usually use a verb alone for their own actions (e.g., Throw ball), or three-year-olds who don’t yet use complementizers like whether or that (e.g., Rod said Nico coming). Have these children set the appropriate parameters yet? The values on some parameters are associated directly with specific lexical items, such as him versus himself, which determine the domain for the antecedent of the pronoun, as in Ken washed him (i.e., someone else) versus Ken washed himself. This parameter either sets the domain as the smallest clause containing the pronoun or as the sentence containing the pronoun. In English, both him and himself are associated with the first setting (the smallest clause), but in other languages, the pronoun him may be associated with the first setting and the reflexive himself with the second, as in Japanese.
The continuity view offers a potentially simple and elegant account of acquisition for syntax (Macnamara 1982; Pinker 1984). If UG in its entirety is present from the start, then children have only to set a certain number of parameters and learn any lexical items with syntactic consequences. There is no need to track changes in learning mechanisms since they remain unchanged, as do children’s representations of linguistic categories and structures.

The maturational view of what is innate differs from the continuity view on the role of experience. In the maturational account, children make progress in syntactic acquisition without much regard to experience. Development is driven instead by a biological timetable. As a result, groups with different kinds of experience adhere to the same timetable. For example, Gleitman (1981) argued that children with normal hearing, children who are blind, and children who are deaf (and not exposed to a sign language) follow much the same timing. All, for instance, produce their first one-word forms around age one; two- and three-word expressions by age two; and some simple grammatical sequences by age three. But notice that all three groups require normal intelligence and normal input to arrive at these milestones.

Proponents of UG often prefer maturation over continuity because biological maturation can be relatively independent of experience. As Felix (1988:371) put it, “The mechanism that ‘pushes’ the child through the sequence of developmental stages is therefore the maturational schedule.” What is innate, then, is UG combined with a biologically based schedule for the emergence of each parameter setting. To take one example, Borer and Wexler (1987) proposed that subject-drop as a parameter in UG is just not available until a particular point in acquisition (also Hyams 1986, but see Ingham 1992). Before this point, exposure to any relevant information in child-directed speech can have no effect. This allows researchers to ignore what happens in acquisition prior to the setting of each parameter. There is no need, in the maturational view, to account for early errors, since they are assumed to play no role in the later emergence of the target construction. Borer and Wexler also proposed that the passive construction (as in The cat was chased by the boy) “matures” only at around age five, thus accounting for its relatively late emergence in English. Here, however, the maturational account runs into distinct difficulty. First, Demuth (1989) showed that, for speakers of Sesotho (a Bantu language), the passive is well established in young children before age three, and, second, several researchers have shown that the passive in English is also acquired considerably before age five (e.g., Clark & Carpenter 1989a; Pinker et al. 1987). They have also shown that emergence of the passive depends heavily on the precise verbs used (e.g., Maratsos et al. 1985).

Another maturational account was put forward by Radford (1990). He proposed that young children go through three stages. The first is pre-grammatical in that any terms used have yet to be categorized as nouns or verbs, say. Next (at around 1;8) comes the lexical stage. This is marked by an increase in vocabulary size, especially for nouns, verbs, prepositions, and adjectives, and by the appearance of word combinations like X + Complement (e.g., open box, in bag), Modifier + X (e.g., nice book, back in, very good), and Possessor + X (e.g., baby cup, daddy gone, doggy down, hand cold). Absent from this lexical stage, according to Radford, are all “functional categories,” such as determiners (a, the, this, that, etc.) and complementizers (whether, that), as well as inflections, such as tense suffixes (wants, jumped), modal auxiliaries (can, must), and the infinitival marker to. These emerge only in the third, functional, stage. Functional categories emerge when they do, according to...
Radford (1990:274), because they “are genetically programmed to come into operation at different biologically determined stages of maturation.” But in other languages, a variety of functional categories emerge much earlier (e.g., determiners in Sesotho: Demuth 1992), and even in English, some members of a specific functional set emerge many months before others (Fletcher 1985; Ingham 1998). These data raise serious questions about the status of a functional stage per se.

An alternative to Radford’s maturational account is that functional categories come in later than some instances of lexical categories because they are semantically more complex and so require more structural knowledge. They are also often pragmatically complex as well. This account is consistent with the data on Korean determiners, a lexical category that emerges relatively late, near age three (O’Grady 1993), and also with the rather long period of acquisition (three years or more) for functional categories in English (e.g., Brown 1973). A major determinant of acquisition there is relative semantic complexity, with less complex meanings mastered before more complex ones. If so, complexity of meaning may offer a more general explanation, crosslinguistically, than the stages proposed by Radford.

Then, the question of what is innate is rather a question of whether we make use of innate learning mechanisms unique to language, and if so, what form these might take. The real debate here should be over the specificity or generality of the learning mechanisms themselves, not the categories or structures to be learnt. As Lenneberg (1967:394) put it:

[N]o features that are characteristics of only certain natural languages, either particulars of syntax, or phonology, or semantics, are assumed…to be innate. However, there are many reasons to believe that the processes by which the realized, outer structure of a natural language comes about are deep-rooted, species-specific, innate properties of man’s biological nature.

What is innate in this view is the manner in which humans process information.

A Learning Mechanism, Just for Syntax?

Some researchers have argued that there is a module in the brain for the rule-governed portion of language. Humans are unique in having developed language; there is strong evidence for specialization of the left hemisphere for language, and part of this specialization consists in a module devoted to the syntactic (rule-based) component of language. As a result, they have argued, whatever mechanisms are used for language acquisition must be specialized and quite distinct from any used for other cognitive acquisitions. In short, language is distinct from cognition, and mechanisms for acquisition can be used only for language.

What is the evidence for this position? First, language is unique to people and is not found in other species. Second, some language disorders show selective impairment of the rule-governed aspects of language (specifically, syntactic rules and regular morphological paradigms). This has been viewed as evidence that language is encapsulated and hence distinct from other capacities (Curtiss 1988). For instance, studies of children with Williams Syndrome initially suggested that WS children’s language was normal but their nonverbal skills were not (e.g., Mervis 1999; Morris & Mervis 1999). This population therefore offered strong evidence for a dissociation between
linguistic skills and general cognitive skills. Children’s cognition could be seriously impaired, in short, but their language remained intact. Yet, as researchers looked more closely at the Williams Syndrome child, that picture became much cloudier.

The findings across tasks and languages show that WS individuals do not after all present evidence for an encapsulated language capacity. The initially rosy picture of their syntactic abilities was exaggerated. Children with Williams Syndrome do not follow the normal course of acquisition for syntax, morphology, or the lexicon. Their syntax and morphology do not, after all, represent a unified, neatly modular syntactic or computational skill, distinct from other linguistic and communicative abilities.
Module 3: The Role of the Native Language in Acquiring a Second Language.

Lado, in his early and influential book Linguistics Across Cultures (1957), stated this clearly:

individuals tend to transfer the forms and meanings, and the distribution of forms and meanings of their native language and culture to the foreign language and culture—both productively when attempting to speak the language and to act in the culture, and receptively when attempting to grasp and understand the language and the culture as practiced by natives.

(p. 2)

Lado’s work and much of the work of that time was based on the need to produce pedagogically relevant materials. To produce these native language-based materials, it was necessary to do a contrastive analysis of the native language and the target language. This entailed making detailed comparisons between the two languages in order to determine similarities and differences.

To understand why language transfer was accepted as the mainstream view of language learning, it is necessary to understand the psychological and linguistic thinking at the time Lado was writing.

It is important at this juncture to clarify one important aspect of our understanding of the term transfer. Although the original term used in the classical literature on transfer did not imply a separation into two processes, negative versus positive transfer, there has been some confusion in the use of the terms in the second language literature. Implicit in the use of these terms is that there are two different underlying learning processes, one of positive transfer and another of negative transfer. But the actual determination of whether or not a learner has positively or negatively transferred is based on the output, as analyzed by the researcher, teacher, native speaker/hearer, when compared and contrasted with target language output. In other words, these terms refer to the product, although their use implies a process. There is a process of transfer; there is not a process of negative or positive transfer. Thus, one must be careful when using terminology of this sort because the terminology suggests a confusion between product and process.

1. Behaviorism:

1.1. The Behaviorist account of L1 learning:

Bloomfield’s classic work, Language (1933), provides the most elaborate description of the behaviorist position with regard to language.

The typical behaviorist position is that language is speech rather than writing. Furthermore, speech is a precondition for writing. The justification for this position came from the facts that (a) children without cognitive impairment learn to speak before they learn to write and (b) many societies have no written language, although all societies have oral language; there are no societies with only written but no spoken language systems.
Within the behaviorist framework speaking consists of mimicking and analogizing. We say or hear something and analogize from it. Basic to this view is the concept of habits. We establish a set of habits as children and continue our linguistic growth by analogizing from what we already know or by mimicking the speech of others.

Bloomfield divides a situation like this into three parts:

1. Practical events before the act of speech (e.g., hungry feeling, sight of apple).
2. Speech event (making sound with larynx, tongue, and lips).
3. Hearer’s response (Ahmad’s leaping over the fence, fetching the apple, placing it in Sarah’s hand).

Thus, in this view, speech is the practical reaction (response) to some stimulus.

1.2. Psychological Background:

One of the key concepts in behaviorist theory was the notion of transfer. The main claim with regard to transfer is that the learning of task A will affect the subsequent learning of task B. What is of interest is how fast and how well you learn something after having learned something else.

In a transfer experiment related to verbal learning, consider a study by Sleight (1911) in which he was concerned with the ability to memorize prose more easily if one has had “prior experience” in memorizing poetry. He compared four groups of 12-year-old children on their ability to memorize prose. Three groups had prior training on the memorization of (a) poetry, (b) tables of measures, or (c) content of prose passages. A fourth group had no prior training in any type of memorization. Following training, the groups were given tests that tested their ability to memorize prose. The question was: “To what extent does poetry memorization, or more precisely, the skills used in poetry memorization, transfer to memorization of prose?” (The results were nonsignificant.)

Let’s consider an example from the area of language learning. According to the initial view of language transfer, if speakers of a particular language (in this case, Italian) form questions by saying:

(4-1) Mangia bene il bambino? eats well the baby “Does the baby eat well?”

then those same (Italian) speakers learning English would be expected to say

(4-2) Eats well the baby?

when asking a question in English. A behaviorist notion underlying this expectation is that of habits and cumulative learning. According to behaviorist learning theory:

Learning is a cumulative process. The more knowledge and skills an individual acquires, the more likely it becomes that his new learning will be shaped by his past experiences and activities. An adult rarely, if ever, learns
anything completely new; however unfamiliar the task that confronts him, the information and habits he has built up in the past will be his point of departure. Thus transfer of training from old to new situations is part and parcel of most, if not all, learning. In this sense the study of transfer is coextensive with the investigation of learning.

(Postman, 1971, p. 1019)

While this statement is not specifically intended as a description of language learning, we can see how the concepts were applied to second language learning. A distinction noted above that is commonly made in the literature is between positive transfer (also known as facilitation) and negative transfer (also known as interference). These terms refer respectively to whether transfer results in something correct or something incorrect, and, to repeat a point made earlier, do not imply two distinct cognitive processes. As an example with relation to second language learning, if a Spanish speaker is learning Italian, when asking a question that speaker might correctly produce

(4-3) Mangia bene il bambino? eats well the baby

because in Spanish one uses the same word order to form questions.

(4-4) ¿Come bien el niño? eats well the baby

This is known as positive transfer. But if that same speaker is learning English and produces

(4-5) Eats well the baby?

the incorrect utterance is known as negative transfer.

With regard to interference, there are two types noted in the literature: (a) retroactive inhibition—where learning acts back on previously learned material, causing someone to forget (language loss) —and (b) proactive inhibition—where a series of responses already learned tends to appear in situations where a new set is required. This is more akin to the phenomenon of second language learning because the first language in this framework influences/inhibits/modifies the learning of the L2.

We turn now to the work on second language learning that was based on these behaviorist positions. As noted earlier, Lado’s work made these theoretical underpinnings explicit. Recall also that the major impetus for this work was pedagogical. In his foreword to Lado’s book, Fries noted:

Before any of the questions of how to teach a foreign language must come the much more important preliminary work of finding the special problems arising out of any effort to develop a new set of language habits against a background of different native language habits . . .

Learning a second language, therefore, constitutes a very different task from learning the first language. The basic problems arise not out of any essential
difficulty in the features of the new language themselves but primarily out of the special “set” created by the first language habits. (Fries, 1957)

Thus, underlying much work in the 1950s and 1960s was the notion of language as habit. Second language learning was seen as the development of a new set of habits. The role of the native language, then, took on great significance, because, in this view of language learning, it was the major cause for lack of success in learning the L2. The habits established in childhood interfered with the establishment of a different set of habits.

From this framework emerged contrastive analysis, because if one is to talk about replacing a set of habits (let’s say, the habits of English) with another set of habits (let’s say, those of Italian), valid descriptions are needed comparing the “rules” of the two languages. It would be misleading, however, to consider contrastive analysis in a monolithic fashion. In fact, there are two distinct traditions of contrastive analysis that emerged. In the North American tradition, the emphasis was on language teaching and, by implication, language learning. Contrastive analyses were conducted with the ultimate goal of improving classroom materials.

As Fisiak (1991) noted, this is more appropriately considered “applied contrastive analysis.” In the European tradition, the goal of contrastive analysis was not pedagogical. Rather, the goal of language comparison was to gain a greater understanding of language.

**A: Contrastive Analysis Hypothesis:**

What are the tenets of contrastive analysis? Contrastive analysis is a way of comparing languages in order to determine potential errors for the ultimate purpose of isolating what needs to be learned and what does not need to be learned in a second-language-learning situation. As Lado detailed, one does a structure-by-structure comparison of the sound system, morphological system, syntactic system, and even the cultural system of two languages for the purpose of discovering similarities and differences. The ultimate goal is to predict areas that will be either easy or difficult for learners.

The pedagogical materials that resulted from contrastive analyses were based on a number of assumptions:

1. Contrastive analysis is based on a theory of language that claims that language is habit and that language learning involves the establishment of a new set of habits.

2. The major source of error in the production and/or reception of a second language is the native language.

3. One can account for errors by considering differences between the L1 and the L2.

4. A corollary to item 3 is that the greater the differences, the more errors will occur.

5. What one has to do in learning a second language is learn the differences. Similarities can be safely ignored as no new learning is involved. In other words, what is dissimilar between two languages is what must be learned.
6. Difficulty and ease in learning is determined respectively by differences and similarities between the two languages in contrast.

There were two positions that developed with regard to the Contrastive Analysis Hypothesis (CAH) framework. These were variously known as the *a priori* versus the *a posteriori* view, the *strong* versus *weak* view and the *predictive* versus *explanatory* view. In the strong view, it was maintained that one could make predictions about learning and hence about the success of language-teaching materials based on a comparison between two languages. The weak version starts with an analysis of learners’ recurring errors. In other words, it begins with what learners do and then attempts to account for those errors on the basis of NL–TL differences. The weak version, which came to be part of error analysis, gained credence largely due to the failure of predictive contrastive analysis. The important contribution of the former approach to learner data (i.e., error analysis) was the emphasis it placed on learners themselves, the forms they produced, and the strategies they used to arrive at their IL forms. Those arguing against the strong version of contrastive analysis were quick to point out the many areas where the predictions made were not borne out in actual learner production.

But there were other criticisms as well. Perhaps the most serious difficulty and one that ultimately led to the demise of contrastive analysis, a hypothesis that assumed that the native language was the driving force of second language learning, was its theoretical underpinnings. In the 1960s, the behaviorist theory of language and language learning was challenged. Language came to be seen in terms of structured rules instead of habits. Learning was seen not as imitation but as active rule formation.

The recognition of the inadequacies of a behaviorist theory of language had important implications for second language acquisition, for if children were not imitators and were not influenced in a significant way by reinforcement as they learned language, then perhaps second language learners were not either. For example, it is not uncommon for beginning second language learners to produce an utterance such as 4-6:

(4-6) He comed yesterday.

in which the learner attempts to impose regularity on an irregular verb. There was no way to account for this fact within a theory based primarily on a learner transferring forms from the NL to the TL.

Not only did errors occur that had not been predicted by the theory, but also there was evidence that predicted errors did not occur. That is, the theory did not accurately predict what was happening in nonnative speech.

Yet another criticism of the role of contrastive analysis had to do with the concept of difficulty. Recall that a fundamental tenet of the CAH was that differences signified difficulty and that similarity signified ease. Difficulty in this view was equated with errors. If a learner produced an
error, or errors, this was a signal that the learner was having difficulty with a particular structure or sound. But what actually constitutes a sign of difficulty?

Recognition of the complexity of comparing languages became apparent quite early, particularly in works such as that of Stockwell, Bowen, and Martin (1965a, 1965b), who, rather than dichotomize the results of language comparison into easy and difficult and therefore dichotomize the needs of learning into a yes/no position, established a hierarchy of difficulty and, by implication, a hierarchy of learning. Included in this hierarchy are different ways in which languages can differ.

<table>
<thead>
<tr>
<th>Table 4.1 Hierarchy of difficulty</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Category</strong></td>
</tr>
<tr>
<td>Differentiation</td>
</tr>
<tr>
<td>New category</td>
</tr>
<tr>
<td>Absent category</td>
</tr>
<tr>
<td>Coalescing</td>
</tr>
<tr>
<td>Correspondence</td>
</tr>
</tbody>
</table>

**B: Error Analysis:**

It is a type of linguistic analysis that focuses on the errors learners make. Unlike contrastive analysis (in either its weak or strong form), the comparison made is between the errors a learner makes in producing the TL and the TL form itself. It is similar to the weak version of contrastive analysis in that both start from learner production data; however, in contrastive analysis the comparison is made with the native language, whereas in error analysis it is made with the TL.

Even though the main emphasis in second language studies during the 1950s and 1960s was on pedagogical issues, a shift in interests began to emerge. The conceptualization and significance of errors took on a different role with the publication of an article by Corder (1967) titled “The significance of learners’ errors.” Unlike the typical view held at the time by teachers, errors, in Corder’s view, are not just to be seen as something to be eradicated, but rather can be important in and of themselves.

Errors can be taken as red flags; they provide windows onto a system— that is, evidence of the state of a learner’s knowledge of the L2. They are not to be viewed solely as a product of imperfect learning; hence, they are not something for teachers to throw their hands up in the air about.

It has been found that second language errors are not a reflection of faulty imitation. Rather, they are to be viewed as indications of a learner’s attempt to figure out some system, that is, to impose regularity on the language the learner is exposed to. As such, they are evidence of an underlying rule- governed system. In some sense, the focus on errors is the beginning of the field of second
language acquisition, which at this point is beginning to emerge as a field of interest not only for the pedagogical implications that may result from knowing about second language learning, but also because of the theoretical implications for fields such as psychology (in particular learning theory) and linguistics.

In the same article, Corder was careful to distinguish between errors and mistakes. Mistakes are akin to slips of the tongue. That is, they are generally one-time-only events. The speaker who makes a mistake is able to recognize it as a mistake and correct it if necessary. An error, on the other hand, is systematic. That is, it is likely to occur repeatedly and is not recognized by the learner as an error.

Taken from the perspective of a learner who has created a grammatical system (an interlanguage), everything that forms part of that interlanguage system by definition belongs there. Hence, there can be no errors in that system. Errors are only errors with reference to some external norm (in this case the TL).

A great deal of the work on error analysis was carried out within the context of the classroom. The goal was clearly one of pedagogical remediation. There are a number of steps taken in conducting an error analysis.

1 Collect data. Although this is typically done with written data, oral data can also serve as a base.

2 Identify errors. What is the error (e.g., incorrect sequence of tenses, wrong verb form, singular verb form with plural subject)?

3 Classify errors. Is it an error of agreement? Is it an error in irregular verbs?

4 Quantify errors. How many errors of agreement occur? How many irregular verb form errors occur?

5 Analyze source.

6 Remediate. Based on the kind and frequency of an error type, pedagogical intervention is carried out.

Error analysis provides a broader range of possible explanations than contrastive analysis for researchers/teachers to use to account for errors, as the latter only attributed errors to the NL. In comparison, there are two main error types within an error analysis framework: interlingual and intralingual. Interlingual errors are those which can be attributed to the NL (i.e., they involve cross-linguistic comparisons). Intralingual errors are those that are due to the language being learned, independent of the NL.

Error analysis was not without its detractors. One of the major criticisms of error analysis was directed at its total reliance on errors to the exclusion of other information.
Perhaps the most serious attempt at showing the inadequacies of error analysis comes from a 1974 article by Schachter. She collected 50 compositions from each of four groups of learners of English: native speakers of Persian, Arabic, Chinese, and Japanese. Her research focused on the use of English restrictive relative clauses (RC) by each of these four groups. The findings in terms of errors are given in Table 4.3 (taken from Table 4.4).

**Table 4.3 Number of relative clause errors**

<table>
<thead>
<tr>
<th>NL group</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persian</td>
<td>43</td>
</tr>
<tr>
<td>Arabic</td>
<td>31</td>
</tr>
<tr>
<td>Chinese</td>
<td>9</td>
</tr>
<tr>
<td>Japanese</td>
<td>5</td>
</tr>
<tr>
<td>American</td>
<td>0</td>
</tr>
</tbody>
</table>

**Table 4.4 Relative clause production**

<table>
<thead>
<tr>
<th>NL group</th>
<th>Correct</th>
<th>Error</th>
<th>Total</th>
<th>% errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persian</td>
<td>131</td>
<td>43</td>
<td>174</td>
<td>25</td>
</tr>
<tr>
<td>Arabic</td>
<td>123</td>
<td>31</td>
<td>154</td>
<td>20</td>
</tr>
<tr>
<td>Chinese</td>
<td>67</td>
<td>9</td>
<td>76</td>
<td>12</td>
</tr>
<tr>
<td>Japanese</td>
<td>58</td>
<td>5</td>
<td>63</td>
<td>8</td>
</tr>
<tr>
<td>American</td>
<td>173</td>
<td>0</td>
<td>173</td>
<td>—</td>
</tr>
</tbody>
</table>


*Compare the construction of relative clauses in (Chinese and Japanese) and (Arabic and Persian).*

(4-13) Japanese
Watashi-wa Eigo-o hanasu josei-o mimashita.
I subj English obj. talks woman obj. saw
“I saw the woman who speaks English.”

(4-14) Chinese
Wo kandao nei ge shuo yingyu de nuren.
I saw the CL speaks English language RM woman
“I saw the woman who speaks English.”
CL = classifier; RM = relative marker
(4-15) Arabic
ana raait Al Emraah allely tatakalem Al-Englizy.
I saw the woman who speaks the English
“I saw the woman who speaks English.”

(4-16) Persian
an zaenra ke inglisi haerfmizaene didaem.
that woman that Eng. speaks I saw
“I saw the woman who speaks English.”