Maturation and Learning of Language in the First Year of Life

Jacques Mehler and Anne Christophe

Abstract

Language, a species-specific aptitude, seems to be acquired by selection from a set of innate dispositions. In the first part of this chapter, we assess the cortical substrate of speech processing. We show that even though, from birth on, speech is preferentially processed in the left hemisphere, the underlying cortical structures continue to specialize over the period of language acquisition. In the remainder of the chapter, we illustrate a number of innate mechanisms that help the baby to learn language. Infants are able to segregate utterances according to their source language; they are able to segment speech streams into basic units that may be different for contrasting languages. Throughout the chapter, we focus on the processes that enable the infant to go from universal capacities to language-specific ones. The general hypothesis that acquisition proceeds by selection is supported by data.

What advantages can an infant possibly draw from listening to speech? One could imagine that, at birth, the infant neglects speech, until interest in learning words and the syntactic rules linking words in sentences begins at around 12 months. If this is the case, one might as well ignore what infants do during the first few months of life and start to explain language acquisition at the point when the acquisition of the lexicon begins. Alternatively, one might imagine that the infant, at birth, has to pay attention to incoming speech signals. Language acquisition may begin with the establishment of some important properties of the language, which are reflected in the speech signal. This view presupposes that the infant interacts in a predetermined way with speech stimuli, in much the way other animals interact with signals that are specific to their own species.

We hypothesize that during their first encounters with speech infants determine how many linguistic systems coexist in their surroundings, and moreover that they make the necessary adjustments for the acquisition of the maternal language. Chomsky (1988) has proposed a model of how this happens and addresses the issue of how infants can be set to learn different languages:

The principles of universal grammar have certain parameters, which can be fixed by experience in one or another way. We may think of the language faculty as a complex and intricate network of some sort associated with a switch box consisting of an array of switches that can be in one of two positions. Unless the switches are set one way or another, the system does not function. When they are set in one of the permissible ways, then the system functions in accordance with its nature, but differently, depending on how the switches are set. The fixed network is the system of principles of universal grammar; the switches are the parameters to be fixed by experience. The data presented to the child learning the language must suffice to set the switches one way or another. When the switches are set, the child has command of a particular language and knows the facts of that language: that a particular expression has a particular meaning and so on. Language learning, then, is the process of determining the values of the parameters left unspecified by universal grammar, of setting the switches that make the network function. Beyond that, the language learner must discover the lexical items of the language and their properties. Language learning is not really something that the child does; it is something that happens to the child placed in an appropriate environment, much as the child's body grows and matures in a predetermined way when provided with appropriate nutrition and environmental stimulation. (Chomsky, 1988, 62–63, 134)

Chomsky and most of his colleagues, even if they believe that the setting of parameters accounts for most aspects of language acquisition, have focused on how it
applies to syntax. Our own investigation focuses on how initial adjustments apply to the most external aspect of language, namely, its sound structure. Below, we investigate how infants behave when confronted with speech signals and explore when and how they parse and represent language.

If language is a species-specific attribute, which are the specific neurological structures that insure its emergence in every member of the species? As far back as Gall (1835), neurologists have suspected that most species-specific attributes are the expression of dedicated organs. Although we know this to be the case (see Knudsen and Konishi, 1978; Nottebohm, 1981; Geschwind and Galaburda, 1987, for some abilities) after decades of research, it is still difficult to locate the "language organ" precisely (see Ojemann, 1983).

The search for the neurological structures underlying language is still in progress. Pathologies described by Gopnik and her colleagues (Gopnik, 1990; Gopnik and Crago, 1991), show that certain kinds of developmental dysphasia (difficulties or delays in learning language) run in families with a distribution that is consistent with a single autosomal dominant mutation. Williams syndrome children, described by Bellugi and her colleagues (see, e.g., Bellugi et al., 1992), are children who have some cognitive deficits but relatively spared linguistic abilities. Indeed, they have very poor spatial abilities, and their capacity to do even the simplest arithmetical operations is impaired. In contrast, their syntactic and lexical abilities are exceptionally good. Their behaviors contrast markedly with those of children with Down's syndrome, suggesting that the behavioral differences arise because of genetic differences between these populations. Moreover, it is striking that Curtiss and Yamada (1981) also found children who, despite impressive language abilities, had greatly impaired cognitive abilities. These studies show that language acquisition is marginally related to general intelligence, which would be a necessary condition for learning by inductive inferences and hypothesis testing. Likewise, Klima and Bellugi (1979) and Poezner, Bellugi, and Klima (1987) have shown that deaf children acquire sign language in the manner that hearing children learn to speak. Landau and Gleitman (1985) have shown that blind people will acquire a spoken language despite their major sensory deprivation. Both deaf and blind children can acquire language with astonishing ease. If the surroundings are sufficiently congenial to ease the child's predicament, the road to language can be as swift and smooth as that of any other child. The evidence suggests that language arises because of specific neural structures that are part of the species' endowment. Neuropsychology teaches that language is basically lateralized in the left hemisphere. But do these structures arise from the child's acquisition of language, or are they responsible for the acquisition of language in the first place?

What comes first, language or lateralization?

Lenneberg (1967) argues that at birth the brain is equipotential and that lateralization arises as a by-product of language acquisition. Lenneberg proposes equipotentiality to accommodate the early findings that in younger children, both lesions to the left hemisphere and lesions to the right hemisphere may result in aphasia that is transitory unless the lesions are really very great. When they are not, language tends to be acquired or reacquired a few weeks after it was affected. It is only after puberty that impairments from small focal lesions to the left hemisphere become irreversible, and also that aphasias almost never arise after focal lesions to the right hemisphere. Today, new investigations make it difficult to accept that the infant's brain is nonlateralized at birth. Before presenting infant data it is necessary to certify that speech comprehension is mostly a left-hemisphere aptitude. A new technique, positron emission tomography (PET), makes it possible to study how the cortex processes not only sentences in one's maternal language but also sentences in a language one does not understand.

Most of the classical investigations with PET have focused on cortical activation while listening to a list of words (see Howard et al., 1992) and little, if any, research has focused on the cortical activation that arises while listening to connected speech sounds. The situation has been corrected by Mazoyer et al. (1993), who have used PET combined with sophisticated individual anatomical and brain-imaging techniques to study the activation of cortical areas in normal subjects who listen to connected speech signals. The activation in the left hemisphere increases as the signals get progressively closer to a normal story in the subjects' mother tongue. Color plate 21 reproduces the activation in subjects who were listening to one of five conditions: a story in their native language (French), a list of French words, a story that was syntactically adequate but had no meaning whatsoever (the nouns and verbs of the
original story were replaced by randomly chosen words from the same category), a "story" in which the nouns and verbs were nonexistent (but that were phonologically legal in French), and a story in an unknown language, Tamil (recorded by a very proficient French-Tamil bilingual who also recorded the story in French). By and large, this study endorses the view that the left hemisphere is more actively engaged when processing standard connected speech than the right. Of course, our findings suggest that this is so only when one listens to a familiar language. The activation in the left hemisphere observed for French but not for Tamil could reflect the fact that the subjects understand French or the fact that French is their mother tongue. These results address a classical concern, namely, whether left-hemisphere superiority emerges through the learning of language. The brain-imaging results show that adults who listen to speech in a language they do not understand, have a left hemisphere that remains unactivated, as compared to the resting baseline.

Consider what happens in the brains of infants who do not understand either French, Tamil, or, for that matter, any other language. Does this mean that at birth, or shortly thereafter, no functional asymmetry will be observable? Is the brain of the neonate activated in the same way by any speech sound that reaches its ears, and if so, does this activation show a pattern that is characteristic for speech stimuli and only for them? Or could it be that speech sounds, regardless of language, mostly activate the neonate's left hemisphere?

Entus (1977) adapted the dichotic presentation technique (i.e., the simultaneous presentation of two stimuli, one to each ear) and found a right-ear advantage for syllables and a left-ear advantage for musical stimuli in 4-month-old infants. Vargha-Khadem and Corballis (1979) were unable to replicate Entus's finding. Bertoncini et al. (1989) carried out a new investigation with younger infants (4-day-olds) and better stimuli. As can be seen in figure 61.1, Bertoncini et al. found that neonates have a right-ear advantage for syllables and a left-ear advantage for musiclike noises of otherwise similar acoustic characteristics. This result is comparable to that reported by Best (1988). However, Best also reports that 2-month-old infants show a left-ear advantage for musiclike stimuli but no right-ear advantage for syllables. This minor discrepancy is likely to reflect the sensitivity of the methods or the response to the contrasting stimuli used by each group. In any case, both studies agree that very young infants display ear advantages that are asymmetric as a function of stimuli. Today the consensus seems to be that the greater engagement of the left hemisphere with speech originates in our genetic endowment rather than occurring as a side effect of language learning. These findings clarify one of the central issues concerning the biological foundations of language. However, we can go one step further.

We are considering the hypothesis that, at birth, speech stimuli, regardless of the language from which they stem, activate the left hemisphere, whereas later in life, the only speech stimuli that can activate the left hemisphere must either belong to the maternal language or, possibly, to a second language. This raises the possibility that the mother tongue has a special status and that language learning consists not only in the dismissal of acoustic stimuli that are nonlinguistic in nature, but also, possibly, in the dismissal of all speech stimuli that do not belong to one's mother tongue. If so, a related question arises in relation to the representation of speech in bilinguals. This hypothesis, as we shall see later, is compatible with the sparse data now available and seems to mesh well with evidence about the behavioral changes one observes at the end of the first year of life.

In our review of behavioral data about the tuning of infants toward their maternal language phonology, we will examine in turn how infants treat and represent speech at different levels (the utterance, prosodic, pre-
lexical, and segmental levels; and for each level, investigate how infants' performance depends on their maternal language. The evidence we will present suggests that during acquisition the infant not only privileges some properties (possibly setting some phonological parameters) but also learns how to neglect aspects that are not relevant to the language he or she is trying to master.

Discovering the languages of the world

In order to acquire language, a first prerequisite that infants have to meet is that they should be able to sort linguistic stimuli from nonlinguistic ones: They have to attribute meaning and linguistic function to words, not to dog barks or telephone rings. Above, we showed that the infant processes speech-stimuli asymmetrically and in areas that are classically related to speech-processing centers in adults. This suggests that somehow the human brain is capable from birth of sorting stimuli according to whether they are speechlike or not.

Moreover, Colombo and Bundy (1983) showed that 4-month-old infants prefer to listen to words than to other sounds. This experiment is suggestive, though not conclusive; indeed, infants' preferences may depend on the choice of the nonspeech stimuli; for example, animal noises may generate more interest than speech. Sturdier evidence has recently been reported by Eimas and Miller (1992): They showed that 4-month-old infants behave as if they experienced duplex perception, a phenomenon that in adults has been taken to mean that speech is processed in specific ways (Liberman, 1982; Whalen and Liberman, 1987). Eimas and Miller suggest that there is a speech mode that is engaged when processing speech stimuli even in 3- to 4-month-old infants.

However, even though there is good evidence that infants can distinguish speech from nonlinguistic stimuli, they still face a momentous problem. We argued in the introduction that the only way infants could possibly acquire language is by making some choices among a set of finite, innately determined possibilities. Indeed, the input with which they are presented is not rich enough to allow them to build the language structure from scratch. But, according to current statistics, most children listen to more than a single language in their immediate surroundings. How can infants successfully converge toward any given language, that is, correctly select the options adequate for this language, if they indiscriminately use sentences from different languages with incompatible option settings? Yet, there seems to be little ground for worry. Indeed, there is strong converging evidence that the child has predispositions that allow it to cope with this problem in a smooth and efficient fashion.

Bahrick and Pickens (1988) have shown that 4-month-olds react to a change in language. When infants were habituated to an English sentence and then switched to a novel English utterance, they showed no recovery of interest, compared to infants that were switched to a (novel) Spanish utterance. Mehler and his colleagues (1988), using many different short utterances instead of only one, found that 4-day-olds tested in a French environment prefer to listen to their maternal language: They suck more when listening to French than when listening to Russian. Moreover, 4-day-olds were able to discriminate the set of Russian utterances from the set of French utterances. Two-month-old American infants also discriminated English from Italian sentences, but in contrast to newborns, they displayed no preference for either English or Italian. Interestingly, the 2-month-old American infants failed to discriminate the Russian and French sentences that French newborns discriminated.

New results show that 4-day-old infants discriminate French from English sentences, corroborating the discrimination part of the previously reported results. However, the preference of newborn infants for their maternal language was not replicated. Mehler et al. (1988) also made the claim that 4-day-olds could not discriminate Italian from English nor English from Italian, two languages unfamiliar to the infants born to French parents. However, our published results show that both experimental groups have a tendency to increase their sucking rate during the postshift phase. The original claim was based upon the fact that neither experimental group exhibited a significant difference to its corresponding control. However, because there is no significant interaction between the groups and their effects, both groups can be considered together. When one assesses jointly the results for both groups, it appears that 4-day-olds discriminate the two unfamiliar languages without any difficulty. In contrast, at two months, American infants do not discriminate two unfamiliar languages (French and Russian), even when one combines the results for both experimental groups. This suggests that while younger infants still work on all the utterances to which they are
exposed, the older ones already concentrate on utterances that share a structure corresponding to the maternal language and neglect utterances that do not. If so, one has to conclude that by the age of two months the infant has set the first values to individuate the structure of the maternal language.

What might these first adjustments rely upon? The above studies have shown that the main properties that infants pay attention to are carried by the lower 400 Hz of the spectrum. Indeed, when the utterances are reversed, infants do not respond to the change in language, although they still do when they are exposed to filtered speech that contains only the frequencies below 400 Hz. These observations allow us to conjecture that infants begin by paying attention mainly to the prosodic properties of speech (that is, overall properties of utterances such as intonation and rhythm). This allows them to classify inputs according to whether they are drawn from one or another natural language. Rapidly they extract a representation that captures the prosodic properties that characterize their maternal language. Recently, Moon, Panneton-Cooper, and Fifer (in press) have begun investigating the same question with a new experimental method that also relies on the sucking response, and is explicitly designed to test preference. They observed that infants of Spanish-speaking parents preferred Spanish over English, whereas infants of English-speaking parents showed the reverse pattern of preference. These results suggest that by two days of age infants already have acquired a sensitivity for some regularity of the maternal language.

Lambertz (in preparation) has recently adapted a method originally used to study visual perception (Johnson, Posner, and Rothbart, 1991) to assess 2-month-olds' preference for auditory stimuli. After attracting the infants' gaze to the center of a display by switching on a multicolor moving spiral pattern, Lambertz presents an auditory stimulus through either one of two loudspeakers that are located 30° to the right and to the left of the moving spiral. The experimental measure is the time the infant needs to initiate a visual saccade toward the source of the sound. Lambertz used as auditory stimuli French and English utterances that were less than 3 seconds long and were either intact or low-pass filtered. The results, illustrated in figure 61.2, show that American-born infants start reacting to English utterances significantly faster than they do to French utterances. This result is obtained regardless of whether the utterances are filtered or not. Lambertz's results thus show that 2-month-old infants, like newborns, display a preference for their maternal language when tested under appropriate conditions. Moreover, the fact that the preference holds for filtered speech confirms the fact that infants probably rely on some gross properties of utterances to categorize them (their prosody).

The picture that seems to arise from all these studies is that a few days after birth, infants are able to tell apart two different languages, even when neither of them is present in their environment; moreover, they already show a preference for their maternal language. Two months after birth, infants still show a preference for their maternal language (not surprisingly), but
they seem to have lost the ability to distinguish two languages that are not familiar to them. These results seem to hold with low-pass-filtered speech stimuli, which indicates that infants rely on prosodic properties of speech to perform this task. This picture, which needs to be made more precise and confirmed, raises a number of captivating questions. For instance, what happens with children raised in bilingual environments? Do they show a preference for the two languages in their environment, or does one of them take precedence over the other? Also, when and how does left-hemisphere lateralization restrict itself to speech stimuli belonging to the maternal language? Last, what happens to deaf children who are not exposed to speech stimuli? Is it the case that signs play a role exactly equivalent to speech, benefiting from a special treatment from birth on, or is it only later, because of the absence of speech, that deaf infants realize that signs are the vehicle of language in their environment? Even though we still have no idea how to answer these questions, our original concern in starting these investigations is satisfactorily addressed: Infants do not get confused when exposed to several languages, because they know how to tell them apart. Even better, they seem already to know, a few days after birth, which language is going to be their maternal language.

Knowing how to distinguish languages is, of course, only the very first step in language acquisition. In the next section, we will turn to the question of how infants segment the continuous speech stream into linguistically relevant units.

**How to find the assembling units of language**

While speech is a continuous signal, language is discrete; indeed, we compute the meaning of sentences by operating on individual words. There is thus no doubt that upon hearing utterances, speakers of a language have to access individual words. If we now think about the task facing the infant hearing a language, it is clear that acquiring the lexicon is a necessary stage of learning. Indeed, the sounds languages use to refer to things are extremely idiosyncratic. There is nothing kelef-ish, perro-ish, or hund-ish about a dog, no more that there is something very chien-ish about it. Moreover, infants should be able to access the same representation regardless of who pronounces a word, the rate with which the word is spoken, whether it is whispered or screamed, said in a happy or sad intonation, and so forth. Mehler, Dupoux, and Segui (1990) have argued, on the basis of the preceding arguments, that infants should have access to a prelexical, normalized representation of speech. In the next section, we will examine what we currently know about such a representation. Moreover, Mehler, Dupoux, and Segui argue that infants should also be able to segment the continuous speech stream into word-sized units, in order to be able to memorize the word forms that will constitute their future lexicon. Two alternatives to this solution have been offered in the literature, both considered implausible by Christophe, Dupoux, and Mehler (1993). The first one states that all words to be learned are first heard in isolation. This implies that speech directed to infants has to be special, since adults rarely utter words in isolation. But then, mothers who—out of belief or distraction—neglected to speak in this special way would be responsible for their child's failure to learn language. We prefer to avoid such an implausible claim and explore alternative solutions. The second alternative (Hayes and Clark, 1970) assumes that infants are endowed with a mechanism that computes transition probabilities between successive prelexical units. In this view, word boundaries appear as troughs in the transitional probability function. However, if such a clustering mechanism might plausibly work on short strings of prelexical units, it is quite implausible for segmenting whole sentences into words, given the size of the lexicon.

Most parties accept that words have to be recovered from the speech stream. The same must be claimed for syntactic constituents. Indeed, infants do at some point have to acquire syntax. Mazuka (1993) has recently made a principled argument to support the view that even syntactic parameters should be set through prosody. She argues that syntactic parameters cannot be set by using the syntactic category of words, because parameters are supposed to be useful precisely to find categories. Instead, syntactic parameters should be set independently of syntactic analysis. Mazuka makes the assumption that prosody may inform the infant not only about where the units are, but also about their dominance relationships. More generally, Mazuka suggests that some syntactic parameters that are relevant for analysis of the fine-grained structure of language might be set by focusing on information available at the utterance level.

Christophe et al. (1994) attempted to assess the hypothesis that infants are able to segment the speech
stream into word-sized units. They reasoned that if this hypothesis is correct, then infants should be able to discriminate bisyllabic contexts that contain a word boundary from bisyllabic contexts that do not. Indeed, they showed that four-day-old French infants were able to discriminate consonant-vowel-consonant-vowel strings (CVCVs) spliced out from the middle of a long word (e.g., *mati* in "mathématicien") from CVCVs spliced out from between two words (e.g., *mati* in "panorama tibétain"). What could have been the basis of discrimination? Christophe and her colleagues measured a significant word-final vowel lengthening, which is not surprising since French is an accent-final language. They also measured a significant word-initial consonant lengthening, which has been observed in many different languages (for English, Umeda, 1976; for Dutch, Quené, 1991; for Czech, Lehiste, 1965; for Estonian, Lehiste, 1966; and for Italian and Swedish, see Vaissière, 1983). The fact that infants are sensitive to word boundaries should be true universally (for all languages of the world), even though different languages may use different boundary cues—indeed, in order to realize which cues signal which boundaries in their particular language, infants should at least perceive all potential cues. Other languages thus have to be investigated before we can conclude that infants are sensitive to word-boundary cues in general. But we already know that infants have the capacity to perceive potential word boundary cues in French.

Umeda, 1976; for Dutch, Quené, 1991; for Czech, Lehiste, 1965; for Estonian, Lehiste, 1966; and for Italian and Swedish, see Vaissière, 1983). The fact that infants are sensitive to word boundaries should be true universally (for all languages of the world), even though different languages may use different boundary cues—indeed, in order to realize which cues signal which boundaries in their particular language, infants should at least perceive all potential cues. Other languages thus have to be investigated before we can conclude that infants are sensitive to word-boundary cues in general. But we already know that infants have the capacity to perceive potential word boundary cues in French.

Jusczyk and his colleagues have conducted a number of experiments directly aimed at evaluating which kind of units infants perceive in the signal. They presented infants with continuous speech samples that had been interrupted either at a boundary between two constituents or within a constituent. They used a modified preference-looking paradigm much like the one developed by Fernald (1985). Studies with 9-month-old, 6-month-old, and even 4½-month-old American infants show that they prefer listening to speech that has been artificially interrupted between two clauses rather than within a clause (see Hirsh-Pasek et al., 1987). Moreover, 4½-month-old American infants who are presented with Polish sentences also show a preference for the speech samples that preserve the clauses (Jusczyk, Kemler Nelson et al., in preparation). The authors suggest that the clause's integrity is signaled by both temporal and frequency cues, and that these cues may well be universal. Indeed, in languages as different as English and Japanese, clause boundaries are marked by a fall in pitch and a lengthening of the last segments of the clause (Fisher and Tokura, in press).

Jusczyk et al. (1992) investigated whether infants were also sensitive to units below the clause. They placed the pause either at the major syntactic boundary, namely between the subject and the verb, or at a minor syntactic boundary, generally between the verb and its complement. They observed that 9-month-olds but not 6-month-olds prefer listening to stories with the pause at the major syntactic break. However, in some conditions this major syntactic break, between subject and verb, does not correspond to a major prosodic break. In particular, when the subject is a pronoun, it tends to cliticize with the following verb (unless it bears emphatic stress); when this happens, subject and verb together form a clitic group, a very cohesive prosodic unit. Recently, Gerken, Jusczyk, and Mandel (1994) carried out an experiment explicitly designed for contrasting syntactic and prosodic units. They used the same stories where either the full noun phrase was repeated for each sentence (e.g., *the caterpillar*), or where it was replaced by a pronoun (*it*). Only in the first case did the major syntactic break between subject and verb correspond to a major prosodic boundary (a phonological phrase boundary). Unsurprisingly, the researchers found that 9-month-old infants were sensitive to prosodic units, but not to syntactic units per se. Thus, 9-month-old infants who heard sentences with lexical noun phrases preferred sentences with a pause inserted between the subject and the verb to the same sentences with the pause inserted between the verb and its complement. However, when the infants were tested with pronoun-subject sentences, they did not show a marked preference for either segmentation.

From this series of studies, one is tempted to conclude that infants perceive sentences as strings of clauses from a very early age. Moreover, they should do so in all languages worldwide, without any need for adaptation or tuning. Only at the age of 9 months is there positive evidence that infants hear phonological phrases. We still do not know at what age infants display sensitivity to clitic groups (formed by the grouping of a content word and its adjacent clitics, such as articles, pronouns, auxiliaries, etc.), but a reasonable bet given the present data would be even later than 9 months. But, as we will see now, there is evidence from other paradigms, using lists of words instead of long
samples of continuous speech, that infants are probably sensitive to prosodic units smaller than the phonological phrase before 9 months.

Jusczyk, Friederici, and colleagues (1993) report that 6-month-old American infants show a preference for lists of unfamiliar English words as compared to Norwegian words uttered by a single bilingual speaker. However, these infants show no preference for lists of English as opposed to Dutch words. Moreover, these results hold when the words are low-pass filtered, indicating that the infants relied on the prosodic properties of the words. The authors stress the fact that the prosodic properties of English and Norwegian words are very different, while English and Dutch words are quite similar in their prosody. This study implies that 6-month-old infants already have a notion of what is a legal prosodic unit in their language, even though these prosodic units are only two syllables long.

Jusczyk, Cutler, and Redanz (1993) have raised a similar question in relation to stress. As Cutler and her colleagues have argued, most content words in English start with a strong syllable (Cutler and Carter, 1987), and adult English-speaking listeners make use of this regularity in their processing of continuous speech. Jusczyk, Cutler, and Redanz showed that American 9-month-olds, but not 6-month-olds, show a preference for lists of unfamiliar strong-weak words, as opposed to lists of unfamiliar weak-strong words. Again, the result holds when the words are low-pass filtered. This study may be taken as evidence that at the age of 9 months, infants have an idea of what is the most frequent prosodic structure in their native language. The implication is that before the age of 9 months, infants have some means of segmenting speech into word-sized units, which allows them to make statistics about a sufficient number of such units. Another implication, of course, is that from the age of 9 months infants may take advantage of this regularity in English.

It appears that the age at which sensitivity to smaller-sized prosodic units arises depends on the paradigm chosen. This should not be too surprising. When presented with long samples of continuous speech, infants' attention is likely to be focused on the larger-sized prosodic units (the intonational phrase, the phonological phrase). This would explain why they do not react to the disruption of smaller units such as clitic groups or phonological words. In contrast, presenting infants with lists of words focuses their attention on the smaller units. This may be why the results of the two word-list studies reported suggest sensitivity to smaller prosodic units before the age of 9 months, while the continuous-speech studies did not.

The studies reviewed in this section relate to how the child recovers specific cues to parse sentences, and recovers prosodic, syntactic, and semantic units. However, as we mentioned in the introduction of this section, the infant needs some prelexical, normalized representation of speech to work on. In the next section, we review what we currently know about such representations, with special emphasis on the question of their language specificity.

How is speech represented?

Psychologists and linguists have combined the study of adult competence for language with that of the properties of the human brain that are responsible for its emergence. In other words, they pursue the study of infants' abilities before learning has had the opportunity to leave an imprint, that is the initial state, and also the ability of adults who share a grammatical system, namely, a stable state (see Chomsky, 1965; Mehler and Bever, 1968). One of the major advantages of jointly considering the initial state and the stable state is best illustrated by focusing on how infants discover contrasting language phonologies and on how adults use them. Numerous studies have established that speakers of different languages process speech in ways that are suited to the phonology and prosodic structure of their maternal language. Mehler et al. (1981), Cutler et al. (1983, 1986), and Segui, Dupoux, and Mehler (1990) have shown that native speakers of French and of English rely upon different prelexical units to access the lexicon: French-speakers rely on syllables, whereas English-speakers rely on stressed syllables (see also Cutler and Norris, 1988; Cutler and Butterfield, 1992). More recently, Otake and colleagues (1993) have shown that the mora acts as the segmentation unit, or prelexical atom, for speakers of Japanese. The aforementioned work is distinct from more standard conceptions in that it does not regard the phoneme as the necessary and unique device to represent the speech stream, and in its reflection of the notion that languages may use differing prelexical devices to segment and represent the speech stream. Furthermore, these authors argue that natural languages have
rhythmic structures that are related to the "timing units": French would be syllable-timed, English stress-timed, and Japanese mora-timed. Do these rhythmic structures play any role during language acquisition?

This is precisely the question that Bertoncini and her colleagues have started to investigate. Bijeljac-Babic, Bertoncini, and Mehler (1993) investigated whether infants rely on a syllable-like notion to organize speech stimuli. Could the infant detect the difference between words with different numbers of syllables? The researchers showed that French 4-day-olds discriminated between bi- and trisyllabic items. Moreover, when the length of the items was equated by means of a speech editor and a special algorithm that leaves the spectral templates unmodified except for duration, the infants were still capable of discriminating the lists. Last but not least, when the number of phonemes changed from four to six, keeping the number of syllables constant, infants did not react to the change in stimulation. These results confirm the hypothesis that the syllable, or a covariant structure, is used to represent speech stimuli. Bertoncini et al. (1988) had already noticed a vowel-centered representation of speech sounds in very young infants. This representation becomes more phoneme-like for the 2-month-olds.

The above results are difficult to assess in the absence of cross-linguistic validation. French, the language of the stimuli used, but also the surrounding language for most of the tested infants, is a syllable-based language (Mehler et al., 1981) but this is not the case for other languages (see Cutler, Norris, and Segui, 1983; Otake et al., 1993). What would infants do if they were tested with Japanese utterances? As mentioned above, Japanese is a mora-based language. Items like *toki* and *tooki* are both bisyllabic, but the first one is bimoraic, while the second one is trimoraic. Bertoncini and her colleagues (in preparation) have explored this issue in detail. In a first experiment it was shown that French 4-day-olds discriminate bi- and trisyllabic items (that were also bi- and trimoraic) spoken by a native Japanese monolingual speaker. Second, the experimenters used lists of stimuli that differed only by the number of moras, but not by the number of syllables (e.g., *toki* and *tooki*). The results suggest that French 4-day-olds are insensitive to this contrast. If this observation is confirmed, one may conclude that the syllable (or something correlated with it) is the universal structure in terms of which speech is represented at birth. Another possible interpretation would be that French newborns are incapable of discriminating morae because they have already realized that French is syllable-based. Given the young age (3 days) of the infants tested, we consider this alternative interpretation implausible. But to reject it definitely, we will have to test Japanese newborns under the same conditions, and show that they behave like French newborns.

In this section we have raised the question of how infants converge toward the rhythmic organization of their maternal language. In one hypothesis, the syllable is the primitive core structure for representing any language. If so, according to the particular language to which one is exposed, one learns to correct this representation and specify whether the syllable is accented or not, whether it has vowel reduction or not, and so forth. In another hypothesis, all of the many organizations displayed in speech signals are computed in parallel by infants. Other components of language determine which organization should be preserved, given the constraints arising from production and lexical considerations.

**From prosody to segments**

In a series of classic studies, Werker and Tees (1984) showed that younger infants are able to discriminate consonant contrasts that are not valid in their maternal language, even though older infants and adults are not. This decline in the ability to distinguish foreign contrasts starts at 8 months of age. By the time infants are one year old they behave just like adults; that is, they do not react to foreign contrasts. This behavior suggests a critical period and a related neurological rearrangement. Best, McRoberts, and Sithole’s (1988) study of click discrimination by English-speaking adults and American-born infants leads to a different interpretation. Zulu clicks are ingressive consonants that are absent from English. Consequently, these speech sounds cannot be assimilated to a contrast used in English. English adults without any experience with Zulu clicks process these as natives do, suggesting that adults do not lose an early ability when they acquire the phonology of their mother tongue. Indeed, Best, McRoberts, and Sithole studied American infants ranging from 6 to 14 months in age with Zulu clicks and with contrasts that do not belong to English but
which may be assimilated to phonemes used in that language. Best (1990) replicated Werker's reports of decreasing performance with foreign contrasts that are similar to one in the maternal language, but Best, McRoberts, and Sithole's subjects behaved very differently with Zulu clicks. Zulu clicks are so distinguishable from any native English category that they are not assimilated to categories in English. Werker and Tees's results are then due to assimilation of novel sounds to the categories of one's native language whenever assimilation remains possible.

This new view about the convergence toward the phonology of the maternal language does mean the process is less biological and ultimately unrelated to a biological clock. Whether infants assimilate foreign contrasts to familiar categories or use routines that are compiled for the purpose of production to code the surrounding contrasts, the fact of the matter remains that around 8 to 10 months of age, infants change from a phonetic to a phonemic mode of perceiving consonants (that is, they stop treating speech as consisting of mere sounds, and instead begin to represent only the linguistically relevant contrasts). More recently, Kuhl and colleagues (1992) have shown that 6-month-olds have already acquired some properties that characterize the vowels in their maternal language. From these studies, and from related studies by Werker and Polka (1993), it appears that infants might specify vocalic properties before they specify the properties of the consonants that figure in their native languages.

From the above studies one can draw the obvious conclusion that learning to speak requires the discovery of the segments that are part of the language. But as we saw before, this is far from being all there is to learning one's mother tongue. We saw that the infant has the ability to establish natural boundaries for prosodic structures. One can propose that infants learn the prosodic structure of their language before they can represent the segmental information about the vowels and consonants in that language.

Conclusions

Our investigations have led us to propose a general hypothesis about the relationship between certain areas of the brain and language. We have argued that the child is already functionally lateralized at birth. Later in life, as PET studies show, only utterances that belong to one's own language will increase the activity of the left hemisphere. Hence, one has to conclude that while the brain is initially prepared to be charitable about what will be processed linguistically, it becomes more and more demanding until it ends up accepting as linguistic only those stimuli that are part of the mother tongue (or possibly of another language that has been learned).

In the years to come we need to investigate in greater detail the nature of the units that are used to represent the speech stream in contrasting languages. We have proposed that languages can be classified into families according to whether they rely on the syllable, the foot, the mora, or other units. The data we presented, however, are also compatible with the view that the syllable is the universal atom for representing speech, regardless of language. In learning more about one's language—in particular, its lexicon and the production routines to build sentences—other segments or structures tend to be discovered. Indeed, our preliminary results point in this direction. However, research at this juncture and on this point is so scarce that we can say little before more data become available.

Acknowledgments

The preparation of this chapter was supported by a grant from the Human Frontier Scientific Program, as well as by a grant from the Human Capital Mobility Project. It was also carried out with the help of CNET (Convention 837BD28 00790 9245 LAA/TSS/CMC), and CNRS (ATP “Aspects Cognitifs et Neurobiologiques du Langage”). We would like to thank Luca Bonalli, Teresa Guasti, and Christophe Pallier for their help, comments, and fruitful discussion, as well as Steve Pinker for many thoughtful suggestions on a first version of this chapter.

Notes

1. Of course, the primary acoustic cortical areas will be bilaterally activated by noise or any other acoustic stimulation.
2. Work by Petitto (1993), and by Poizner, Bellugi, and Klima (1987) suggests that deaf, native speakers of sign language have cortical asymmetries homologous to those of hearing subjects. It seems unlikely to us that signs taken from, say, American Sign Language would spontaneously activate the left hemisphere of newborns. Could it be that, in the absence of speech, the left cortical areas devoted to its processing are taken over by other highly structured inputs?
3. The same bilingual speaker pronounced the utterances in both languages, so this result cannot be attributed to potential differences in timbre.
4. A mora is a subsyllabic unit that may consist of either a consonant-vowel or vowel syllable, or of the nasal or stop coda of a syllable.


