EARLY LANGUAGE ACQUISITION AND COGNITION: METHODS AND RESULTS
2.4.1 THE VIOLATION OF EXPECTATION PARADIGM 50
2.4.2 DESCRIPTION OF THE VISUAL EXPECTATION PROCEDURE IN A VIOLATION OF EXPECTATION PARADIGM 51
2.4.3 ANTICIPATORY EYE MOVEMENTS RESPONSE 53
2.4.4 DESCRIPTION OF AN EXPERIMENT USING THE EYE TRACKER IN AN ANTICIPATORY EYE MOVEMENT RESPONSE 58
2.4.5 WHICH QUESTIONS CAN BE ADDRESSED TO INFANTS USING THE VISUAL EXPECTATION PROCEDURE? 62
2.4.6 ADVANTAGES AND DISADVANTAGES OF THE VISUAL EXPECTATION PROCEDURE 62
2.5 THE HEAD-TURN TECHNIQUE 64
2.5.1 THE CONDITIONED HEAD TURN PROCEDURE 64
2.5.2 THE HEAD TURN PREFERENCE PROCEDURE 70
2.6 CHAPTER CONCLUSION: TECHNIQUES AND PROBLEMS 75

3 LANGUAGE AND NATIVISM: THE APPLICATION OF THE TECHNIQUES AND THE STATE OF THE ART 76
3.1 LANGUAGE NATIVISM AND EXPERIENCE: EXPERIMENTS AND METHODS 77
3.2 LANGUAGE ACQUISITION IN THE PRELEXICAL INFANT 78
3.2.1 BOOTSTRAPPING THEORIES AND STATISTICAL LEARNING 78
3.2.2 THE UNIVERSAL CORRELATION OF LANGUAGES 79
3.2.3 THE CLASSIFICATION OF LANGUAGES BY INFANTS 82
3.2.4 THE PERCEPTION OF SYLLABLES BY INFANTS 85
3.2.5 PERCEPTION OF SPEECH SOUNDS AND LINGUISTIC EXPERIENCE 87
3.2.6 WORD ORDER AND THE SETTING OF THE HEAD-COMPLEMENT PARAMETER 93
3.3 STATISTICAL LEARNING: TRANSITIONAL PROBABILITIES AND RULE LEARNING 97
3.3.1 THE SEGMENTATION OF SPEECH BY INFANTS USING TRANSITIONAL PROBABILITIES 97
3.3.2 STATISTICAL COMPUTATIONS IN ADULTS 98
3.3.3 THE ACQUISITION OF LINGUISTIC RULES IN INFANTS 100
3.3.4 USING THE HEAD TURN PREFERENCE PROCEDURE TO CLARIFY HOW INFANTS MIGHT FIND WORDS AND RULES IN A SPEECH STREAM. 104

4 CONCLUSION 115

5 REFERENCES 118
INTRODUCTION
The aim of this thesis is to investigate the earliest stages of language acquisition and describe some of the methods that are used today to study the initial state of development.

Chapter 1 will review some of the most important theories on language acquisition from a nativist perspective. I will consider the theoretical framework behind generative theories and we will see how these assumptions are consistent with the discoveries made by scientists focusing on the biological foundations of language.

The study of the early abilities in newborns and infants can help us understand how language is learnt by unravelling how innate mechanisms and environment interact.

If we think that language acquisition starts when the child begins to utter its first words, it would be unnecessary to study the early stages of development. If, on the contrary, we conjecture that language acquisition begins precisely when we are born, we need to investigate the early mechanisms that assist the infant in the task of acquiring the maternal language. Under this perspective the way infants perceive speech becomes crucial.

The problem that arises when we want to study scientifically the acquisition of language in a pre-lexical infant resides precisely on the subject of our investigation. How can we ask nonverbal infants information of how they perceive or categorize speech or any other physical dimension?

We have two different scientific paths that we can undertake: study the response of the brain through neuroimaging methods or rely on behavioural procedures. Enormous progress has been made today regarding the neuroimaging techniques that allow us to study which areas of the brain are activated when the infant is involved in specific linguistic tasks. Despite the important discoveries and new possibilities that this field of research offers, this thesis will not deal with these methods which require a more specific knowledge of the anatomy and neural activity of the brain and that would therefore demand a specific competence on the field.

My attention will instead concentrate on the behavioural methods that are used in the study of infants’ speech perception and in the investigation of other
cognitive capacities. Chapter 2 will thus describe the most important techniques that allow scientists to study the behavioural responses of newborns and infants: the High Amplitude Sucking Technique, the Visual Fixation technique, the Visual Expectation Procedure, The Eye Tracker Methodology and the Head-Turn Technique. I will describe in detail how each method is performed, what the advantages and disadvantages of using a particular technique are, what research questions can be asked to an infant through these tools. To better understand each methodological procedure I will describe step by step an experiment relying on each of the techniques mentioned above.

In Chapter 3, I will review some of the most important results that have been obtained through the use of these experimental methods with infants. I will try to show how they bear on general debates about the nature of language acquisition.

Many of the studies I will describe have been run at The Language, Cognition and Development Lab at SISSA, Trieste, and some of these studies have been run by the author of this dissertation. The laboratory, directed by Professor Jacques Mehler, studies the cerebral bases of language and of other cognitive abilities on newborns, infants and adults. Every year, the laboratory runs more than a thousand experiments with infants between seven and eighteen months of age, and hundreds of experiments with newborns and adults using behavioural and neuroimaging techniques. Thus, the Laboratory offers a unique perspective in Italy to investigate language acquisition and early cognition in many of its aspects. It is my hope that this dissertation may clarify some of the results found and the relation between the methods used and some exciting discoveries on early cognition.
1 UNDERSTANDING LANGUAGE ACQUISITION
1.1 Nativism and language acquisition

1.1.1 The Poverty of the Stimulus Argument

Learning a language is a very difficult task but every child is able to master it proficiently in a short time.

At first, infants’ use of sounds is a way of expressing their needs or emotions; in the first five months the infant does not produce anything that is linguistically relevant, although, as we will discuss in the second part of this chapter, her interaction with language starts immediately after birth.

Between 5 and 7 months the baby plays with the sounds that she produces showing a more intentional use of them, and shortly after (between 7 and 8 months) begins “babbling”. During this stage, the infant produces syllabic sequences composed of consonant-vowel clusters. These patterns of repeated syllables like ‘pa-pa-pa’ use a set of phonetic units that are the same for all languages.

When the babies reach their first birthday, they start uttering the first words, but their lexical comprehension is three times greater than their production. The duration of the one-word stage varies from child to child and can last 2 or 12 months. The infant’s first words are similar for all languages and are mainly nouns describing everyday simple objects (words for food, parts of the body, clothing, people), some “action words” (up, go) and words used to interact socially like no, bye-bye (Pinker, 1994).

At approximately 18 months, the ability of the baby to produce words grows rapidly and the order in which they combine two words reveals a certain structure underneath it. Although initially the sentences are short, the noun and verb or subject and object respect the position that they would have in more complex and complete grammatical sentences (Brown, 1973; Pinker, 1984).

Between 3 and 4 years of age, a child is able to speak one or more languages and unlike the adult can learn any other language with little effort and achieving a native-like proficiency. The talent that every child has in learning language has no counterpart in other motor or cognitive abilities at this age, a child is born knowing that she will have to communicate linguistically with
other human beings. Language would not be simply learnt from the environment, but would reside somewhere in our brain where a grammar gives an order to an otherwise unsolvable puzzle. But what is actually innate and what is actually learnt by experience?

According to learning-based theories (Skinner, 1957; Tomasello, 2000) all is needed to language acquisition is the input. A child is able to master a language by making random hypotheses and learn from the errors that she inevitably makes. The environment provides the ‘stimulus’ which positively ‘reinforces’ verbal behavior.

Noam Chomsky in ‘A Review of B.F. Skinner’s Verbal Behavior’ (1959) argued that the input is insufficient to guide acquisition.

Language is too creative to be simply imitated:

‘A child will be able to construct and understand utterances which are quite new, and are, at the same time, acceptable in his language’.

Chomsky’s argument was that children are able to understand and to construct an infinite set of sentences, even though the input is of a finite nature. This fact, that a pure imitative/behaviouristic theory of language was unable to explain, called for an alternative model of language learning. Learning a language meant to acquire a certain generative ability, a set of structures, a generative grammar. And for language learning to be successful, the grammar of the target language has to be identified among all the possible rule sets that a child with no explicit instruction of how her native language is structured could make.

Noam Chomsky (1957; Chomsky, 1959) was the first scientist to apply to linguistics the classical induction problem known as the ‘poverty of stimulus’ argument. The input that children receive gives no formal explanation of how the language is structured. Nevertheless all healthy infants master language between 3 and 4 years of age.

Even the mistakes that infants inevitably make reveal that they have a grammar. The most frequent errors arise from an over generalization of certain grammatical rules, a very common example is the use of regular plural suffixes for irregular plurals or the –ed suffix for the past tense of irregular verbs. On the
other hand, there are certain mistakes that children never make, certain generalizations that could be derived from the input, but that would result in an incorrect interpretation of a rule, are not made by the child.

Of course, children do make some linguistic mistakes. Children seem to be naturally biased towards some generalizations and constrained towards others. As an example to our discussion, let us consider the possible generalizations that could be made from sentences like:

a) The girl is pretty
b) Is the girl pretty?
c) The girl who is next to you is pretty

While the rule extraction between a) and b) would be “move the auxiliary of declarative sentences in first position to make interrogative sentences”, the hypothesis would fail with the embedded clause in c) where there are a number of other possible hypotheses and sentences that we could come up with when turning the declarative sentence into an interrogative sentence:

i) Move all auxiliaries to the beginning of the sentence:
   Is is the girl who next to you pretty?

Or also:

ii) Move only the first auxiliary to the beginning of the sentence:
    Is the girl who next to you is pretty?

Or again:

iii) Move the second auxiliary to the first position:
    Is the girl who is next to you pretty?

Of course only this last sentence would be correct in English, but if the child were applying a serial-order depending structure (move the first/second/third...auxiliary) instead of a structure-dependant construction (move the auxiliary belonging to the subject of the main clause to the beginning
of the sentence) she would be making the wrong generalization. Despite the complexity of such sentences, every child by four years of age is able to construct the appropriate generalization. Where does this grammar come from? A child, in order to construct the correct sentences, would need to categorize sentences into declarative and interrogative, she would need to parse them and understand the structural dependencies among words and among clauses. No explicit grammar instruction is given to a child acquiring a language, none of her caretakers would formally instruct the child on what are the principles of English syntax, again the stimulus is poor in this respect. It seems therefore that children learn from the inside what they could not possibly learn from the outside, that is that language is structure-dependent and referring to such notions as ‘subject’, ‘main clause’ as in iii) and not serial-order-dependent as in ii) or category-dependent as in i).

What we have just discussed, by no means wants to be an exhaustive account of the syntactic complexities of languages, but is only an example of the many difficulties and possible wrong generalizations of the rules of the target language that a child could encounter when acquiring a language. There are also more elemental tasks like extracting words from the speech stream which are nonetheless complex for an infant and which we will discuss in more detail in other sections of this thesis.

1.1.2 The Principles and Parameter Theory

To explain the astonishing abilities that every learner demonstrates in analyzing the linguistic input and producing the correct linguistic output, Chomsky theorized the existence of Universal Grammar, which is innately known and conceived as a set of universal linguistic elements. More recently, Universal Grammar has come to be viewed as a set of principles and parameters:

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 or 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 found by experience. The data presented to the child learning
the language must suffice to set 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 the 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

Principles of the universal grammar are therefore all those linguistic
elements that are common for all languages, For example, all languages make a
distinction between content words: nouns, verbs, adjectives and function words:
determiners, pronouns, prepositions etc.

Another principle refers to the general behaviour of a syntactic phrase.
Inside this basic unit there is always a Head, which permeates the syntax of the
whole phrase so that, as Pinker put it, the whole phrase talks of what his head
talks about (Pinker, 1994).
The X-bar structure (Chomsky, 1981) has a Head, X, its complements and specifiers. X can either be a noun, a verb, an adjective, a preposition... The phrase through the specifier and the complement can be extended recursively and to infinity in other syntactic combinations.

This principle can be applied to all languages; all that the child has to do is fix the parameter of the order in which head and complement appear. Indeed, some languages as for example English and Italian, are Head-Complement i.e. the complements follow their head, while other languages like Japanese or Turkish, are Complement-Head. For Chomsky the knowledge of principles and parameters is genetically transmitted and although the input helps to set the direction of the switch, principles and parameters cannot be acquired in the environment.

In the third chapter, we will analyze in more detail those theories that posit that infants are sensitive to the prosodic marking of syntactic units and that they rely on such information in the speech signal to discover the syntactic organization of the target language.

1.1.3 Gold’s Theorem

Another argument that provides strong support for the innateness of language and a further confirmation of the ‘poverty of stimulus’ was given by
Mark Gold. In a seminal article (Gold, 1967), Gold defined a learnability theory which can be applied to language acquisition. Although his model of learning was highly idealized, his results apply to any theory of language learning, or at least, they oblige learning theorists to respect certain formal constraints in theory construction.

Gold’s idealized learning model works as follows. As the child receives no formal information of how her target language is organized, she would have to learn the language only from the set of possible sentences (strings) she hears. The learning strategy that the child would need to apply can be explained as an algorithm that creates hypothesis about the language spoken in the environment of the learner (target language). Every time a new grammatical string is heard in the input, this logical process produces a guess about the grammar of the target language. For learning to occur, a grammar that correctly explains the target language and which cannot be contradicted by any other additional string of the input has to be identified. A class of languages can be defined as identifiable in the limit from text if the algorithm can always succeed in identifying that language.

Gold proved that the algorithm could identify in the limit grammar with complexity close to those of a natural language only if it were provided with both positive and negative evidence about the language.

Having positive and negative evidence closely resembles a Skinnerian view of language. As already discussed, B.F. Skinner’s behaviourist theory states that language learning derives from the parents’ positive and negative reinforcement of their children’s grammatical behaviours. But do children really get such reinforcements?

Roger Brown and Camille Henlon in a 1970 study proved that parents tend to correct their children over the factual errors they make rather than their grammatical mistakes (Brown & Hanlon, 2004). The researchers also studied the answers that parents gave to their children’s questions and found that they were generally inconsistent whether the questions were grammatically correct or not. Children therefore do not get regular and coherent negative evidence and even if they would, they would probably have little benefit from it. Later experimental studies have shown that children do not learn from corrections; when they do not know a specific construction, they cannot produce it correctly even if
repeating after a model (Marcus, 1993; Guasti, 2004).

The fact that children do not seem to learn from negative evidence poses a further problem for a non-nativist view of language acquisition. A child learning a language and making hypotheses over its grammar from the input she receives, would make a very large amount of guesses. If the child only attends to positive evidence, this problem is greatly complicated. A child who would make a mistake about her target language, guessing a rule that makes it a subset of her hypothesized language would have to receive negative evidence to be corrected, that is information of what may not be a possible sentence of the target language. But children do not seem to attend to this information. Indeed, in virtue of Gold’s theorem, disregard of negative evidence in language learning basically makes language unlearnable.

So how can we solve this predicament? We seem to be led to the same conclusion suggested above: Just as a computer without a specific programme could not acquire a grammar as complex as that of a natural language on the basis of positive evidence alone (nor for that matter process external information tout court), so humans could not elaborate complex and elaborate linguistic structures if not endowed with innate abilities for learning language.
1.2 Environment, biology and language acquisition

No one can say that language can emerge without input. Few extreme cases of children reared in total isolation in inhumane conditions have been known. The outcome for these unfortunate human beings was always the same: children did not develop language at all (Gleitman & Newport, 1995). A theory based exclusively on innateness would not take us very far and a theory based exclusively on environment interaction would have the same result.

However, the problem is to characterize what input is needed in order to learn language, on the basis of innate knowledge to be determined. Too much emphasis has been given, for instance, to the special speech that mothers use with their child, coming to regard motherese as a prerequisite to language acquisition. There is no doubt that we talk to a baby in a different way than we do to adults. Infant-directed speech is generally slower, repetitive, high-pitched, with an overdone prosody (Snow & Ferguson, 1979). However, Newport, Gleitman and Gleitman (1977) showed that children whose mothers use motherese do not reach the stages of language development any earlier than those children who have not been exposed to such special linguistic input.

Furthermore not all cultures think that it is necessary to address infants before they have something to say, nevertheless these children develop language anyhow, just by listening to adult-directed speech (Pinker, 1984).

If we want to understand what precise interaction with the environment can lead children to learn language, then it is interesting to take a look at cases in which a limited perception of the environment radically modifies the "normal" linguistic input. If language mostly depends on input, such changes should lead to a deficient acquisition of language. If it does not, invariants across radical input variations will provide us cues for what input can really give to learning.

1.2.1 The deaf child

Profoundly deaf children born in hearing families are faced with the
problem of learning a language without acoustic data. Often their parents do not know sign languages, and actually prefer that their offspring learn to lip-read and utter language rather than learn sign language. Therefore, these children in their early years are deprived of language: they are not exposed to sign language and they cannot hear the language spoken in their environment.

Goldin-Meadow and colleagues (Goldin-Meadow & Mylander, 1984) studied the case of ten of these children from ages 1-4 and found that despite an impoverished environment, these children reached the normal milestones of language acquisition at the same age as their hearing peers.

The one-word stage was reached at one year, by producing iconical gestures. Such gestures were not used by their caretakers, who generally pointed at the objects in the hope of rendering their oral speech more comprehensible. By their second birthday, these children began to produce two- or three-word "sentences". Even more surprisingly, the combination of signs showed to be syntactically structured, with a precise position for verbs and nouns, a structure that was not observed in their caretakers’ gestures.

These facts indicate a biological endowment in humans as regards language acquisition: children who have been deprived of language in some way have developed a form of natural, spontaneous communication nonetheless, a form of communication that shared many crucial aspects of language structure with that of a spoken language.

Another clue to the biological foundation of language comes from an interesting study by L.A. Petitto and P.F. Marentette (Petitto & Marentette, 1991), which showed developmental similarities between manual and vocal babbling.

In normal hearing children, vocal babbling begins between 7 and 10 months. Babbling is universal and has some well defined characteristics:

- Usage of a reduced number of phonetic units;
- Use of syllable patterns with well-formed consonant-vowel clusters;
- No relation between use and meaning;
- Reduplication;
- Continuity between early babbling forms and first words.

Now, Petitto & Marentette collected data from profoundly deaf infants of
deaf parents acquiring American Sign Language (ASL) as a first language and from hearing infants of hearing parents acquiring English as their first language.

Petitto and colleague videotaped the infants at three ages (10, 12 and 14 months) and analyzed the infants’ manual activities. The results showed two distinct kinds of manual activity: gestures and manual babbling. Hearing and deaf children produced both activities, although in very different proportions. While in deaf children manual babbling was between 32 and 71% of the total manual activity, hearing children reached a mere 4 to 15%.

To classify manual activity as "manual babbling", gestures had to have the same properties of vocal babbling. Just as a spoken language is organized into word "subunits" such as phonemes, syllables, and subphonetic features, so a gestural language can be dissected into "visual word" subunits, such as place of gesture articulation, or kind of articulated gesture (Klima & Bellugi, 1979).

Indeed, the analysis of the data collected showed that the deaf infants used a reduced subset of the phonetic units found in ASL, and that each of them had an individual preference regarding the location of the manual babbling, just as each of the hearing infants had her own preferred babble.

The manual babbling of the deaf children used four syllable types: a well formed syllable has a handshape, a location, a path-movement (a change in location) and a secondary movement (a handshape change). Furthermore the manual babbling contained reduplication, and there was continuity between their predominant manual babbling and their first signs.

Thus, both groups of infants went through the same stages of proto-language production at the same time. This study therefore brings in more evidence to the fact that even the very early stages of language development -- way before production begins -- have a biological basis. Language seems to develop its structures independently of the modality of input and output, even in its finest details. The study also shows that babbling is a proper linguistic activity, contradicting those who believe that it is instead a stage of the progressive maturation of the articulatory apparatus.
1.2.2 The blind child

If we thought that learning a language was merely a matter of experience and of contextualizing words and meaning then a child who cannot make such an association visually would have many problems and delays acquiring language. This however is not the case of blind children, who, despite their perceptual limitations, acquire language as easily and rapidly as other children (Landau & Gleitman, 1985).

Parents of sighted children naturally point to objects and repeat a word several times thinking that this will facilitate children acquire the lexicon. But when, for example, we point to a dog and say "dog", how does the child stop over-generalizing (inferring that dog= all furred animals) or under-generalize (inferring that dog= that specific furred animal I saw on that particular occasion)? (Mehler, Dupoux, & Southgate, 1994). Not necessarily does vision clarify word meaning.

No doubt, a blind child cannot determine the meaning of certain semantic categories such as colours and interestingly will stop using these words at the age when other children use them with their correct meaning (age 4). However, by age three, she understands the meaning of abstract words linked to visual experience such as "see" and "look" despite having no possible input about the experience of seeing and how it is different to looking. When you ask a sighted child to look up he will raise his head even if blindfolded: for him, looking has to do with visual inspection. A blind child instead will raise his hand, showing that for him looking necessarily implies a perceptual inspection by hand. For the blind child, touching and looking take different meanings when applied to the different perceptual experiences of the blind and of the sighted: if we ask a blind child to show an object to his mother he will lift it up for a moment and if we ask him to let her touch that object he will bring the object to her and let her touch it. Therefore, he understands what looking and touching mean for sighted persons. When the same words refer to his perceptual experience, the blind child will adapt those abstract words to his way of exploring the world: if we say to the child "You can touch the table but not look at it" he will tap the table and when allowed to look at it he will use his hands to inspect the table (Landau &
Gleitman, 1985).

What these studies have revealed is that learning a language does not simply mean learning by rote the correct association between words and their meaning in the appropriate context. Blind children just like sighted children or deaf children, can rely on some specific mechanisms that are part of the developing brain of humans and that allow the infant to form the correct hypothesis over the target language.

In the next section of this thesis we will survey just how deep in our genetic program language is. We will therefore describe those areas of the brain that are responsible for processing language and we will see whether these tissues are present from birth.
1.3 The language ‘organ’ in the brain

1.3.1 Equipotentiality or lateralization at birth

A way to determine how deep in the biology of the organism specialization for language goes is to investigate whether there is dedicated cerebral tissue for language processing. Paul Broca (1861) was the first to study the possible relationship between the left hemisphere of the brain and speech. Broca sectioned the brain of patients who, subsequent to a cerebral trauma, had lost the faculty of language, but not other abilities. The clinical studies showed that all these aphasic patients presented damage in a particular anterior area of the left hemisphere that took the name of ‘Broca’s area’. Patients with lesions in this region showed impairments with speech production while those with a lesion of the posterior area, the ‘Wernicke’s area’, had difficulties in speech comprehension.

![Figure 1. Broca's and Wernicke's areas](image_url)

These and other clinical studies, together with behavioural and brain-imaging studies, have confirmed a left-hemisphere dominance for speech. An interesting question today is whether the left hemisphere and the right hemisphere are equipotent at birth and then specialize, through experience,
(Lenneberg, 1967) or if at birth the two hemispheres are already disposed to process different types of stimuli. According to this last hypothesis our cognitive system would be conceived as a set of different modules each involved in the process of different kinds of data.

Cerebral structures would therefore be divided into modules each with its specific function and processing a part of the information that is being perceived. For example we think of vision as an inseparable whole, where the perception of colour or movement could not be divided into two different functions in our ability to see objects. However, histories of patients with cortical lesions have shown that these perceptions obey to different mechanisms that can work separately (Mehler et al., 1994).

Possibly, language is also processed by specific modules (Fodor, 1983). One of the markers of the existence of a module is the existence of dedicated brain tissue to process certain stimuli. So, if language is modular, it is also possible that already at birth the left hemisphere is predisposed for language processing.

The study of brain activation is very difficult to accomplish in infancy because of the lack, until recent years, of non-invasive techniques. Hence Bertoncini and colleagues used an indirect method to do research in this field (Bertoncini et al., 1989). The scientist revealed how 2-week-old infants process acoustic stimuli in a different way according to whether they are linguistic or musical in nature. Infants were tested with the dichotic listening technique, and data were collected with the non-nutritive sucking response. Two sounds were presented to the baby simultaneously, one for each ear. When infants listened to syllables in their right ears, they reacted to a change in syllable, showing a recovery of sucking rates following a change of stimuli. Instead, no recovery occurred if they were listening to the linguistic signal in the left ear. Vice versa, a change in musical notes was better perceived by the left ear than by the right ear.

In adults, language processing mostly occurs in the left hemisphere whereas music is mostly processed by the right hemisphere. Because each ear is better connected with the opposite hemisphere, a right-ear advantage is a sign of left hemisphere processing, and vice versa. Bertoncini et al. showed that newborns have a right ear superiority for processing speech, which consequently
signal left hemisphere superiority for language, and hence for language acquisition.

**Figure 2. Dichotic perception and language lateralization in neonates**  
(Bertoncini et al., 1989)

1.3.2 **Lateralization in the newborn: an optical topography study**

Until recently, few non-invasive techniques existed to explore the neonate’s brain. As a result, our knowledge of the infant brain is severely limited. Thanks to the invention of some novel imaging devices, brain imaging techniques suitable for infant studies now enable researchers to explore the cerebral activation from speech stimulation. At the Language and Cognition Laboratory at Sissa, Trieste, Marcela Peña and colleagues (Peña et al., 2003) used optical topography to investigate response to auditory stimulation in twelve full-term neonates.

Optical Topography, also known as near-infrared spectroscopy (NIRS) measures the changes in the concentration of oxygenated and deoxygenated haemoglobin in response to auditory stimulation. The concentration changes were measured in 12 areas of the right hemisphere and 12 areas of the left hemisphere of newborns of 2 to 5 days of age.

Infants were tested while sleeping in a silent room. They listened to
various blocks of infant-directed speech played forward (FW) or backward (BW) by a recorded voice, and to periods of silence (Figure 3). Backward speech is a very good control to compare whether infants specifically react to language, and not to physical features of the stimuli, such as overall energy or waveform. Indeed, a snippet of language played backward has the same physical energy as the same snippet played forward, but all phonetic, phonological syntactical and prosodic transitions are lost in backward speech. Thus, a differential reaction to speech and backward speech signals that the brain is specifically treating some linguistic aspect of the stimuli.

Figure 3. The different blocks of FW speech, BW speech and silence used in the testing protocol by Peña et al. (2003)

Peña and colleagues found that the left hemisphere was activated more during the exposure to normal speech than during backward speech or silence. As Figure 4 shows, the temporal areas in the left hemisphere (the perisylvian areas) were more active during the FW condition than during backward speech or silence, while no area of the right hemisphere was more activated during forward than during backward speech.
Figure 4.: (a) The two sections of the brain (left and right hemisphere) show the positioning of the OT probes: the red dots correspond to the emitting laser diode, the blue dots correspond to the detector. The probes are positioned in specific areas of the skull. The numbers on the dotted line correspond to the recording sites (channels) in which the Hb concentration was measured. (b) The numbers correspond to channel numbers in (a). The plots show the average of the mean of total Hb: red for forward speech, green for backward speech and blue for silence. Channels 1-6 were positioned over the frontoparietal Sylvian fissure; the channels enclosed by the dotted line were placed on the temporal regions below the Sylvian fissure.

While the involvement of the left hemisphere with language is not questioned, there is still debate as to whether the lateralization is the cause or the consequence of language acquisition. Peña and colleagues have shown that already at birth the brain is activated by normal speech stimuli. Thus, it seems
that newborns and adults use the same areas in the brain to process speech. This result gives strong support to the theoretical assumption that the left hemisphere is the ideal ‘language organ’ responsible for language processing and acquisition.

1.3.3 Lateralization and language production

The left hemisphere is not only dominant in language perception, but also in language production. Of course language production cannot be studied at birth, but we can gather some information in the areas responsible for language production when infants begin their very first linguistic emissions.

As we have recalled in Section 1.2.1, the properties of babbling show that this activity is already fully linguistic. Holowka and Petitto (2002) exploiting this fact, conducted a study to test if the specialization of the left hemisphere is already present in babies in the syllabic babbling stage. Ingenuously, they used a fully behavioural technique based on mouth asymmetry to gather information about brain lateralization. Many mouth asymmetry studies have shown how adults present a right asymmetry of the mouth opening when involved in linguistic tasks while no asymmetry is present when the task is non-linguistic. This element would therefore reflect the left hemisphere dominance for language.

The two researchers videotaped English and French babies between 5 and 12 months (the age varied according to when each baby entered the babbling stage) to study possible mouth asymmetries in language production during babbling. More specifically, Holowka and colleague studied the mouth asymmetry of the babies while producing babbles, other vocalizations (nonbabbles) and smiles. While nonbabbles produced no mouth asymmetry, babbling gave rise to right asymmetry and smiles to left asymmetry.
Not only did this study give further proof of the fact that babbling is a linguistic behaviour rather than a consequence of the development of the vocal tract, but it also showed that lateralization is present at a very early stage even in language production. The right hemisphere is directly involved with emotional responses while language production is controlled by the left hemisphere.
1.4 Why it is important to study language acquisition before language production

The first chapter of this thesis has investigated the basic assumptions of a nativist theory for language acquisition. I have briefly reviewed some of the major theories that put forward the arguments for innate linguistic capacities. Generativist theories have not been analyzed in detail but only taken as a starting point. I started with the assumption that infants do not begin the acquisition of language with a blank slate. The brain seems predisposed from birth to language processing and acquisition. This capacity cannot simply be explained with a general ability to learn anything that it is presented to it, or by a ‘language course’ given by a stimulating environment. Indeed, the ability to learn a language is not associated with a general intelligence; there are various examples of syndromes which show a retardation but perfect language abilities or, vice versa language impairments despite a normal cognitive development.

I have also argued that acquisition cannot derive from a ‘language course’ given by a stimulating environment. We have also seen how language could not be entirely learnt from the positive evidence received from the environment: the child’s disregard of negative evidence would make any language unlearnable.

A further proof for the innateness of language has come from the discussion of those cases in which the role of the environment for language learning is greatly compromised because of a limited perception of it. We have found that deaf and blind children reach the normal milestones of language acquisition like any other child (Landau & Gleitman, 1985; Newport & Meier, 1985).

According to nativist theories then, the infant begins the task of acquiring her maternal language with some important mechanisms built in the brain that allow her to represent extract and generalize the structural information of the maternal language that rely on specific cues contained in speech. There is no doubt that language needs to be learned; a child will speak the language that she is exposed to and will speak none if she is barred from it.
Thus, language is a unique mixture of nature and nurture. I hope to have convincingly argued that the study of the pre-linguistic infant is pivotal in order to understand how this mixture occurs.

The problem is, precisely, that prelinguistic infants are prelinguistic. One cannot ask them how they develop language, and therefore, our advancements into this area are strictly linked to the psychological techniques of investigation of early language learning.

Today there are some very effective and reliable testing procedures that can help researchers understand what happens inside the mind of an infant. The next chapter will review the principal behavioural methods that are used today in many laboratories of the world to study speech perception and cognitive abilities in infants.
2 STUDYING THE COGNITIVE INFANT: PROBLEMS AND METHODS
2.1 The problem of methods for the study of early cognition

When conducting experiments with adults and older children researchers are able to ask direct questions and get direct answers to the quest they are positing; such approach however is not possible with infants who are in a preverbal state. Studying these initial stages of development is yet essential to understand what innate mechanisms are present at birth and how these eventually interact with the perceptual experience of the infant and so lead to the acquisition of language.

Over the last forty years, researchers interested in the cognitive development of infants have found answers to their questions either studying, with the aid of experimental techniques, the behavioural responses of infants, or recording the babies’ brain activity while exposing them to particular stimuli.

In this chapter I will describe the most common behavioural procedures that are used in infancy research. They include the study of such behavioural responses as sucking, visual fixations and eye movements. I will illustrate how these methods work, what their advantages or disadvantages might be and how the procedures have sometimes been adapted to best suit the hypothesis they had to test.

All the methods that I will describe (with the exception of the High Amplitude Sucking technique) are currently used in The Language, Cognition and Development Lab at SISSA, Trieste, where thousands of infants have been tested over the last nine years, and where some important results in the study of infants’ cognitive abilities have been found. I will try to give a sense of the issues related to the techniques at use by describing in detail two behavioural experiments that have been run in this laboratory, and that I know directly having participated to the process of experiment setup and data collection.

2.1.1 The basic steps of an experiment

After the researcher has found a question that interests him, he must answer it with the instruments that are accessible to him. In other words, the
problem he wants to solve has to be testable. Transforming the theoretical question into an empirical issue is what in psychology is called *operationalization* of the question. Often, the process of operationalization is far from trivial. Eventually, the scientist has to find a testable question, state a possible solution to the question he has raised, and then verify whether the proposition is probably true or probably false. To do so he must base his conclusions on the collection of data.

In the specific case of newborns and infants, one could collect data by direct observation, as for example Piaget did. But there are many limitations in this procedure. This method is inadequate because based on subjective judgments that are inevitably biased in the case of parents recording the behaviours of their siblings in everyday situations.

Experimentation is the most valuable tool that a researcher has to test his hypothesis because under experimental conditions he is able to control all those external variables that falsify the possible interpretations that can be given to a result.

The scientist dealing with infant research has three important behavioural paradigms that he can use to address his research question:

- the habituation/dishabituation paradigm, which rests on the well-grounded phenomenon that infants lose interest in a repeatedly presented stimulus and show interest to a novel one;
- the violation of expectation paradigm, which tries to elicit longer looking times in the infant by presenting him with visual events that he should considers to be odd or impossible.
- The eye-movement response, which relies on the ability of the infant to anticipate reinforcement.¹

¹ The idea of reinforcement is related to the principles of operant conditioning. B.F. Skinner first applied the concept of operant conditioning to emphasize the link between stimulus and response. He showed how a pigeon in a *Skinner box* would “learn” the contingency between behaviour and reinforcement i.e. if the pigeon pecks at a button inside the cage while a light
The choice of the procedure to use in the experiment rests on a series of considerations that the researcher must carefully assess and which depend on the nature of his research. He has to consider which method is best suitable to the age-range he wants to study and which behavioural response of the infant interests him. Each of the paradigms that I mention above provides one or more dependant variables that can be measured. For example, one can take the sucking rate in a habituation/dishabituation paradigm as the dependent variable. This is what is employed in a High Amplitude sucking technique. In the case of the violation of expectation paradigm, the dependent variable can be the fixation time to a certain stimulus. Or, in the case of eye movements, the dependent variable may be the rapidity of orientation towards a stimulus, or the total amount of saccades in a certain zone of the scene the infant is witnessing.

The researcher has also to consider in great detail the stimuli that he is going to use in the experiment. The question he asks the infant must be very clear and the possible interpretations that could be given to a result should be as univocal as possible.

The recruitment of subjects, especially in the case of infants, is the next step. In the case of newborn subjects, participants are recruited directly from the maternity wards and normally tested in the hospital itself, where the researchers have set an experimental room suited for experiments with neonates. When the subjects of an experiment are non-neonate infants, recruiting has to follow other ways. Most laboratories recruit infants through notices published on newspapers or magazines. The Sissa lab in Trieste has direct access to the database record of the city’s newborns and thus sends letters to propose parents to participate with inside it is on it would have access to food, if it pecks when the light is off no food is dispensed. The pigeon after several trials will “learn” the behaviours that generate reinforcement and will privilege these behaviours over those that do not obtain reinforcement and that gradually disappear.
their child to the studies that are conducted in the laboratory.

Often, the design of an experiment requires the experimenter to plan two groups of subjects: the experimental group and the control group. The experimental group receives the stimuli that allow the experimenter to verify what he wishes to evaluate. The control group receives another treatment not containing the experimental stimuli, so that it can be used as a milestone to compare the effect of the experimental manipulation.

The setting of the experimental room is also very important. One condition that cannot be overlooked is the general comfort and quiet that the experimenter must provide to the infant. However little he is, an infant is very sensitive to environmental conditions and few apparently unimportant details may be the key for the success or the failure of an experiment. It is very important that there are no possible distractions for the infant; silence must be kept as much as possible. An experimental room must also be bare of any distractions. For this reason, walls are normally covered by black curtains and lights are dimmed, so that the concentration of the infant is not distracted. Such details, that are rarely reported in the literature, turn out to be pivotal for the conduction of good experiments.

It is also very important to instruct parents on how they should behave during the experiment because there must be no interference on their part -- something highly unnatural for caretaker. Often babies have to sit on their parent’s laps (from 6 months on), and controlling parents' interference is a problem. Parents are given specific instruction as to how they should hold the child and it is recommended to them that they must not speak to the baby or caress the baby. Often parents are asked to wear dark opaque glasses or headphones, for the same reasons that I have just underlined.

The outcome of an experiment can be influenced even unconsciously; so all the information and details of the experiment and what the baby is supposed to do must be given to the parent only after the experiment has been run.

In the following pages, I will describe the major behavioural techniques I have mentioned above.
2.2 The High Amplitude Sucking Technique (HAS)

One of the problems that arise when we want to study the perception of newborns is to find a method that will elicit a response that the newborn is able to master and control. Since newborns have difficulty in motor coordination and their visual capacity will only develop in the following months, researchers can barely rely on such responses. There is, however, a behaviour that is closely tied to their survival and which is therefore very well developed: the sucking reflex. Not only is this activity related to feeding needs, but it is also a cause of pleasure and an indication of the neonate’s state of alert: the more babies suck, the more they are aware of their surroundings (1994).

This reflex is the basis of the technique called non-nutritive sucking, which can be used successfully from birth to 4 months of age.

Siqueland and DeLucia were the first scientists to explore the possibilities that this response measure could give if associated with operant conditioning techniques (Siqueland & DeLucia, 1969). Their adaptation of operant conditioning to the study of neonates brought about important advances in the study of the newborn’s perceptions.

Siqueland and DeLucia used their technique to study visual perception in four months old. Their experiments tested whether infants were able to learn the contingency between strength of sucking and the brightness of a picture facing them. The infants’ sucking behaviour was conditioned by a visual reinforcement: the longer the infants sucked, the brighter the light became. The results showed that those infants who received visual stimulation increased their sucking rate (experimental group), while those who did not had lower sucking rates (control group).

Siqueland then used this technique to study visual discrimination in infants. The procedure used the paradigm of habituation/dishabituation. During habituation, the infants in the experimental group were exposed for a fixed period of time (7 minutes) to the same visual stimulus and then received a switch to another stimulus for a fixed period thereafter. Siqueland compared the
sucking rate of the infants in the experimental group with that in the control group who had received the same visual stimuli in both periods. The experimental group showed dishabituation to the novel stimuli, indicated by an increase in sucking rates while the sucking response in the control group continued to decline. The technique thus proved to be successful in testing visual discrimination in infancy.

2.2.1 Description of the High Amplitude Sucking technique in speech perception

The high amplitude sucking procedure was adapted to study speech discrimination by Eimas, Siqueland, Jusczyk and Vigorito (1971).

In the standard procedure, the technique works as follows. Each baby is tested individually and is placed in a reclining seat in a silent room with dim lights to avoid any distractions. The infant is put in front of a blank wall at approximately one meter away from it. A loudspeaker emits the sound stimuli and sometimes a picture is projected on the ceiling or on the wall to attract the attention of the infant and discourage distractions.

A pacifier is put in the baby’s mouth and is held in place either by a mechanical arm or by an experimenter. In the latter case, to avoid any influence on cueing the infant, the experimenter hears music through headphones to mask the stimuli heard by the baby. The pacifier is attached to a plastic tube that is connected to a pressure transducer that sends electric signals to a computer. The electric charges are proportional to the strength of the sucks. The sucks which reach or exceed a predetermined threshold activate the auditory reinforces.
In the first study by Siqueland and DeLucia, there was a fixed criterion for defining a high amplitude suck that was set for all infants to 18 mm-Hg (mm of mercury, a unit of pressure equal to 0.001316 atmosphere). The problem with this procedure was that the intensity of the sucks varied considerably from one baby to another. This meant that if for one baby all sucks exceeded the fixed threshold, for others no sucks reached it. Eimas and colleagues brought an important variation to the procedure in order to reduce individual variability in the sucking rate. They let each infant suck in silence for one or two minutes to obtain a baseline rate of sucking, and then considered only the top 33% of the infant’s sucks as high amplitude sucks.

Once the baseline has been identified for each infant, the pre-shift session begins and the baby’s sucks control the presentation of the auditory stimuli. The duration of this phase depends on how fast the infant loses interest in the pre-shift sequence, consequent to which the post-shift phase begins.

On this point, Eimas introduced another important variation: he individualized the satiation criterion. Satiation is reached when there is a
decrease in sucking rates (which indicate a loss in interest of the stimuli). When the infants in the experimental group reach satiation, the stimulus is changed. The change in stimulus is called the shift. Before the study by Eimas, infants received different amounts of experience with the first stimulus because of difference in their sucking rates. This meant that often for some subjects the shift was presented before any sign of lack of interest in the first stimuli had been reached. Therefore, Eimas established that any infant had to show a decrease in their own sucking rate of 20% or more for two consecutive minutes in order to hear the new stimuli. The decrease was calculated by making a comparison to the minute immediately preceding it. Because during the very first minutes of auditory exposure the sucking rate may have an irregular pattern, the researcher also prescribed that the decrease in sucking had to be measured from the third minute of sound stimulation onwards.

To describe at best the technique it is perhaps useful to look at the study by Eimas (1971) in more detail. In that classic study (whose scientific importance we describe in Section 3.2.5) Eimas studied the ability of infants to categorize synthetic speech sounds classified by adults as /ba/ and /pa/. After the baseline rate had been reached in the first minute, the presentation of the sound stimuli was made contingent to the rate of high amplitude sucking. A recorder played the sound /ba/ and the intensity of the signal was directly proportional to the amplitude of individual sucks. A maximum intensity level was set to 15 dB to avoid frightening the infant. The auditory stimuli acted as reinforcement for the infant and the contingency resulted in an increase of sucking rates. When the baby reached satiation, i.e. when his sucks showed a 20% decrease in strength for at least two minutes, the baby was considered to be habituated. At this point, the subjects in the experimental group received a new stimulus that could have been either a between category shift (another /pa/) or a within category shift (a syllable /ba/). The subjects in the control group received instead no change in stimuli and their sucking was recorded for 4 more minutes.

Discrimination of speech sounds was deduced by a remarkable increase in sucking rate (dishabituation) calculated between the two minutes before the shift and the 4 minutes after the switch. Using the HAS technique, the researchers were able to ask 1 month-old and 4 month-old infants if they could discriminate
speech contrasts differing in voicing properties and also if they accomplished so in a categorical manner.

2.2.2 Variations of the High Amplitude Sucking Technique

Laboratories conducing infant research sometimes have modified the High Amplitude Sucking Technique. In particular, there is no standard for the baseline period and for satiation level. This could be a problem when comparing results by the various laboratories. The choice of the high amplitude sucking rate is very important because on this level will depend the answer to the question whether the baby has learnt the contingency between her high sucks and the sound stimulation. If the range is set too high, it will mean that for some infants few sucks will be able to reach the relevant criteria. Afterwards, the babies will soon grow tired of maintaining such high levels of sucking. As a consequence false results may ensue, because the decrease in sucking rate may have arisen from fatigue rather than from lack of interest. Another problem that could occur is that infants may not be able to learn the contingency between sucks and stimuli because not enough sucks could have taught the baby the mechanism underlying the operant conditioning.

Another dimension that can vary from one laboratory to another is the satiation level, which gives an indication as to when the shift in the stimuli can be performed. In the Eimas et al. study, there had to be a decrease of 20% or more for two consecutive minutes. Other researchers, such as Jusczyk, used higher rates (Jusczyk, 1977). The different criteria adapted by investigators, however, have not given rise to important differences in results.

2.2.3 Which questions can be addressed to infants using the HAS technique?

Over the last 40 years, the High Amplitude Sucking technique has been mainly used to study discrimination abilities in infants. Usually, the question asked to infants in form of an experiment only contrasts two single stimuli, because it is necessary to avoid long experimental sessions. One of the major limitations of this procedure is the high drop-out rate, which can reach even 50%
of the total subjects. Very often, infants have difficulties reaching a change in stimuli because they fall asleep, or start crying, or lose the pacifier, which can only be replaced in the babies’ mouth in the very first minutes of habituation but not afterwards. Therefore the test sessions have to be short, and that means that the researchers can only present babies with one single contrast. As a result of this limitation, the HAS often provides a Yes-No answer to research questions. In other words the technique can indicate whether infants can or cannot discriminate two speech sounds, but it does not reveal how the infant categorizes this information.

A way of overcoming this obstacle is to present infants with multiple tokens of one stimulus, and infer which of the properties that have been contrasted are distinguished and which are not. One of the first studies that exploited this procedure is a study by Kuhl (Kuhl, 1983). The researcher divided the infants in four groups. In the pre-shift phase she presented one group with tokens of the same vowel (/a/) differing in pitch dimension, with some having a flat tone and some a rise-fall pattern. In the post-shift phase the infants in the experimental group were presented with tokens of another vowel (/i/) also differing in pitch contour, whereas the infants in the control group received no shift but continued to hear the same stimuli. The other two groups heard tokens of both vowels with the same pitch (monotone) in the pre-shift phase. In the post-switch phase, the subjects in the experimental group heard tokens of the same vowels with a different pitch (rise-fall). The results the researcher obtained indicated that infants consider more relevant the differences in vowel colour than in pitch contour and that they categorized vowels according to their colour.

The High Amplitude Sucking technique can also be used to test memory of speech sounds in infants. Jusczyk, Jusczyk, Hirsch-Pasek, Kemler Nelson and Kennedy (1992) modified the HAS to test the capacity of infants to remember speech sounds. The researcher introduced a delay period of two minutes, between preshift and postshift phase, during which infants were shown colourful slides without removing the pacifier. The introduction of a temporal lag between pre-shift and post-shift phases can give information about the strength of the memory trace generated by the habituation phase.

The High Amplitude Sucking method can be adapted in other ways to
best suit the research question posited by a study. For example, unlike what happens in the classical HAS technique, Mehler and colleagues (1988) did not use a conditioning sucking procedure to test French newborns’ discrimination of the native language from foreign languages. This choice was necessary to present the stimuli in their entirety and give a natural prosodic pattern to the sentences the infant listened to. Had the baby controlled the presentation or interruption of the sentences through his sucks, she could have heard sentences beginning or ending in an abrupt and unnatural way. Furthermore, in order to give sufficient habituation to the pre-shift phase, the infants’ satiation was calculated only after the presentation of the sixth sentence. Subsequent to a decline by 33% from the highest sucking rate, the infants in the experimental group heard sentences in another language, while those in the control group listened to sentences in the same language.

This experiment gave some interesting results and showed that the HAS procedure can be effectively used not only for discrimination tasks but also for testing infants’ preferences. Subjects had been divided in four groups. One group heard French during Habituation and Russian in the post-shift (group FR). The second group served as a control and listened to French throughout the testing session (group FF). The third group listened to Russian in the first phase and French in the second phase (group RF), while the fourth group was a control and listened to Russian samples during both periods (group RR). The sucking rates of groups FR and FF were particularly high in the first phase, while in the second phase the greatest increase in sucking rate came from the RF group compared to its control group (RR). Such ample gap was not evident in the FR group in comparison to its FF control group. This behavioural response would therefore substantiate that not only do newborns discriminate one language from another but they also have a preference for their native language.

In sum the High Amplitude Sucking technique is the most appropriate behavioural technique for studying speech perception and representation in newborns and in infants up to 4 months of age. Subjects are able to play an active role in this procedure, as they are able to control the sucking reflex.

Researchers have been able to ask infants not only questions regarding their ability to contrast speech sounds but also investigate speech categories,
memory for speech and preference.

The major problem with this method is the high dropout rate because the success of the experiment depends on the state of alert of the subject. Infants have to be attentive during the trial sessions and this is a very difficult condition to maintain with infants, who change their state very rapidly. A second important problem is that this technique provides one single data point per subject and requires control groups. It is therefore highly expensive in terms of the efforts to be deployed by the experimenter.
2.3 The Visual fixation procedure

The visual fixation procedure is one of the most common behavioural methods that are used today in the study of the perceptual and cognitive abilities of the infant.

This procedure, just like the high amplitude sucking technique, uses the habituation paradigm to reveal certain abilities in infants including those regarding speech discrimination.

The habituation procedure rests on the premise that if an infant is placed in an environment where a stimulus is displayed, she will visually orient towards it; but if the same stimulus is shown to the infant over and over again, it will result in a decreased response (habituation) that will eventually bring the infant to stop orienting towards it.

Habituation was applied to research with infants by Robert Fantz (Fantz, 1964), who, with a series of simple experiments, introduced for the first time the paradigm, which proved to be very useful in many techniques that are used today by many laboratories in the world.

Fantz worked with 4 age groups: infants between 1 month and 2 months, 2 and 3 months, 3 and 4 months and 4 and 6 months of age. He repeatedly showed to the infants 10 pairs of stimuli; in each of these pairs, one element was always the same while the other was different from trial to trial. Each pair was displayed for a minute over 10 consecutive minutes.

The experiment showed that while babies younger than 2 months had equal looking times for both stimuli, the babies between 2 and 6 months reached habituation to the repeated stimulus at a rate that was faster as the age grew.

Since Fantz’s studies, habituation has been adapted to best suit different experimental quests. The habituation phase can thus be presented in different ways: the method Fantz used was a fixed-trial procedure but the presentation of the stimuli can also be controlled by the infant.

Horowitz (1974) adapted the visual fixation procedure to study speech discrimination in infants. She showed that a visual stimulus can be paired with
an auditory stimulus until habituation is produced, and that the switch to a novel auditory stimulus produces recovery of looking time therefore indicating a discrimination of the auditory sound.

2.3.1 Description of the Visual Fixation Procedure in Speech Perception

Infants are tested individually in a silent room with bare walls or with black curtains hanging from the ceiling to avoid any distractions.

The baby sits in an infant-chair or in the parent’s lap facing a screen. The caregiver and the experimenter wear headphones playing masking music to prevent them from hearing the stimuli. A loudspeaker is hidden below or above the visual display. A video camera is connected to the monitor allowing the experimenter to watch the child during the trial session from another room.

When the infant is looking towards the screen the visual stimulus, for example a checkerboard, and the auditory stimulus are simultaneously presented.

When the infant looks away from the screen, the experimenter gives a
signal to the computer to interrupt the presentation of both stimuli and when the infant looks again towards the display because attracted by a flashing light on the screen, the same trial is presented again. The trial is presented over and over again until the looking time declines reaching a criterion that has been established previously by the researcher (a common used criterion for measuring habituation is to reach an average 50% decrease of looking time measured over three successive trials as compared to the average of looking times in the first three trials). When the infant has been habituated, the auditory stimulation is changed while the visual one remains the same. If the novel stimulus is discriminated the infant should show longer looking times. The control group will receive no novel trials after habituation has been obtained to exclude any spontaneous recovery.

Figure 8. A schematic drawing of the Visual Habituation Procedure (taken from Werker, Polka, & Pegg, 1997; Patterson & Werker, 1999).
2.3.2 Variations of the Visual Fixation Procedure

In the procedure we have just described the infant controls the presentation of the stimuli through her looking behaviour; in other versions the procedure can be only partially controlled by the subject.

In case the researcher wants the infant to hear a longer speech stream or a list of words, he can make coincide the end of the trial with the end of the segment regardless of the looking behaviour of the child. It may also be the case that the researcher wants all subjects to be habituated to the same amount of trials and he may thus set a number of trials that the infants will listen to regardless of their looking times.

2.3.3 Which questions can be addressed to infants using the Visual Fixation Procedure?

Primarily, this technique has been employed to study the discrimination abilities in infants. Indeed this technique has been used by various researchers to study the perception of non-native contrasts in infants (Polka & Werker, 1994), but also to study preference. For instance, Pegg and colleagues (Pegg, Werker, & McLeod, 1992) adopted this procedure to study two-month-olds preference for infant directed speech over adult directed speech.

A very interesting study by Stager and Werker (Stager & Werker, 1997) used the visual fixation procedure to study the ability of 14-month-old and 8-month-old infants to discriminate phonetic contrasts.

The infants sat on their parent’s lap who to prevent any biasing wore a hat with a visor and a cloth blind and listened to vocal music over headphones. The trial began when the infant was looking at the screen. The experimenter sat in a near-by room and watched the baby through a closed circuit TV system. The presentation of the stimuli was contingent to the infant’s looking behaviour.

The researchers ran four different experiments with two age groups: 14 months (in all four experiments) and 8 months (only in experiment 2) following the procedure I have just described.
As we can see from the diagram in Figure 9, in the experiments infants were habituated to two or one word-object pairings (the objects were brightly coloured moving objects). The test trials began after the subjects had shown a decline in looking time. During this phase, two test trials were presented: one trial repeated the same word-object combination ('Same') and on one trial there was a switch in the word-object pairing.

Figure 9 Representations of habituation and test phase in experiments 1-4 in Stager & Werker (1997)
The results show that while 8-month-old infants (experiment 2) were able to tell the difference when two phonetically similar nonsense words (/bih/ and /dih/) switched, 14 month-olds were not (experiment 1 and 2). Fourteen-month-old infants showed to have no difficulty in discriminating two phonetically dissimilar nonsense words (/lif /and /neem/), nor in discriminating the phonetically similar stimuli (/bih/ and /dih/) when the words were not paired with an object (infants were shown a checkerboard pattern which is unlikely to be perceived as an object).

Stager ad Werker came to the conclusion that at this age the ability to make fine-grained discriminations depends on whether infants are involved in a word learning task or not. While at 8 months infants are not relating sound to meaning but are attentive mainly to speech perception, by 14 months they attend to the meaning of words more than to their phonetic information.

2.3.4 Advantages and Disadvantages of the Visual Fixation Procedure

The Visual fixation Procedure is a practical and adaptable technique that can be used to study discrimination abilities or preference in infants as young as
two months of age. It has indeed a very wide age-range and can be used successfully throughout the first year of life as well as with toddlers.

It is easy to set up and the experimenter can be trained quickly.

Being a flexible procedure, the length and number of trials can be adapted to best suit the research question; it is indeed possible to use speech samples of greater duration than a single word.

The main limitation of this procedure is that in most cases it provides no individual data, but only group data. This means that the behaviour of infants can only be judged as a group. A large number of infants are therefore necessary to obtain results also because the dropout rate for fussiness, parental interference or failure to habituate may be high. However, these limitations must be qualified. Practically every infant procedure is subject to the same drawbacks.
2.4 The Visual Expectation Procedure

The Visual Expectation Procedure, like the Visual Fixation Procedure, relies on looking times to visual or auditory stimuli to measure cognitive development or speech perception in infants.

While the procedure I described in the previous section uses the preference and the habituation/dishabituation paradigms, the Visual Expectation Procedure uses the violation of expectation paradigm and the anticipatory eye-movement response to reveal the infant’s cognitive capacities.

2.4.1 The violation of expectation paradigm

The violation of expectation paradigm is one of the most important methods that are used by researchers to study the knowledge of physical events in infants. Research on this field is based on the assumption that infants have an intuitive or perhaps innate understanding of physical laws in respect to space, time, objects and causality (Baillargeon, 1998).

When infants are tested with this paradigm, they are presented with an event that violates their supposed ‘knowledge’ of the physical world; if infants do consider an event as impossible, they will look longer at it, increasing their attention with respect to another scene, as similar as possible to the "violation" scene, but expected from the point of view of the infant's understanding of the world.

Since infants expect a certain outcome to occur, they anticipate what they expect to see and when the event they expect does not occur they will look longer at an alternative outcome.

An experiment by Téglás and colleagues (Téglás, Girotto, Gonzalez, & Bonatti, 2007), using these measures will help us go through the various phases of this method, and will allow us to grasp types of designs and issues on early cognition not necessarily tied to language.
Description of the Visual Expectation Procedure in a violation of expectation paradigm

The experiment I am about to describe was run at The Language, Cognition and Development Laboratory at Sissa, Trieste. Téglás and colleagues (2007) were interested in early numerical and probability reasoning in 12-month-old infants.

Infants were tested in a silent booth, covered by black curtains and with dim lights to avoid distraction. The infants were seated in their parent’s lap in front of a screen at approximately 80 cm away from it. Parents were previously instructed to hold the babies with their hands on the baby’s waist so as to leave some freedom of movement to the child. They were told not to interfere with the experiment by pointing at the screen or by speaking or caressing the baby. The parents wore black opaque glasses that prevented them from looking at the video, thereby influencing the infants’ looking behaviour.

An infrared camera placed above the screen allowed the experimenter to watch the infant from a different screen. The presentation of the stimuli was infant-controlled, because every time the infant looked away from the screen, the experimenter stopped the movie that remained still, pausing on the display and starting again from where it had been interrupted as soon as the infant looked back towards the screen. This procedure was necessary to be sure that all the subjects watched the experimental movies in their entirety.

Importantly, this method requires no habituation, but only a familiarization with the experimental material. So the first two trials consisted of two familiarization movies in which the infants watched four balls (two blue coloured and two yellow) bouncing inside a container with an open exit at its base. One ball of each colour exited in turn in the two familiarization trials. No generalization rule had to be inferred from these neutral movies. After this phase, four test movies were shown to the infant; in them, unlike in the familiarization trials, balls of different colours were not equally distributed: three of the balls bouncing were of one colour (either blue or yellow) and the other one was of a different colour (either blue or yellow). After 13 seconds, an occluder hid the container and one of the objects exited the container. In half of the trials the object was one of the three identical ones (probable outcome) and
on half of the trials it was the different one (improbable outcome). Occlusion was removed after 1 second and all the objects remaining inside the container were shown, so as to avoid excessive memory load. Different lists of test movies were set: half of the movies began with improbable outcomes and half with probable outcomes. This procedure is necessary to counterbalance the trials across the subjects and to make the experimenter blind to experimental conditions. Before each movie, a visual attractor (with four coloured arrows) appeared on the centre of the screen to attract the attention of the infant and orient her to the screen.

After the outcome, if the baby looked away from the screen for 2 or more seconds or looked continuously at it for more than 30 seconds (timeout), the computer automatically presented the next trial.

Looking times were also analyzed off-line: if the baby had 2 or more timeouts she was excluded from analysis. Also looking away exactly in the moment in which the object was exiting the container was a criterion for exclusion. The researcher measured the looking times in seconds scored after the outcome.

The logic of the experiment was as follows: if infants can reason about what event is more or less probable, they should have the expectation that one of the three equally coloured balls should exit the container. Therefore, when the single ball exits, their expectations will be surprised, thus ensuing in longer looking time for the single ball out.
Figure 11. Representation of experiment 1 in Téglás et al.: trial movie with bouncing phase (a) and after occlusion an improbable (b) or probable (d) outcome. Results of looking times over improbable vs. probable outcomes (c)

The results indicated that infants looked longer when the improbable object exited the container: infants had not been habituated to probabilistic events and therefore showed to have a natural intuition on the estimation of probable events.

This study relied on the violation of infants’ expectations in order to make inferences about their reasoning. By measuring the looking times to the visual stimuli it was possible to understand what was considered as surprising and what was instead a predictable outcome for infants of only 12 months of age.

2.4.3 Anticipatory eye movements response

If infants have expectations about how events follow each other, it should be possible to find more online measures of how they anticipate certain
outcomes. Another paradigm that researchers can use with infants in the visual expectation procedure tries to do exactly that, by studying the location and timing of infants’ gaze directions; it is called the anticipatory eye movement paradigm. This method can be particularly useful to make inferences about categorization (McMurray & Aslin, 2004) or to investigate learning and other cognitive abilities in infants.

Research using this paradigm relies on the fact that infants are able to anticipate the outcome of a stimulus either because they have learnt a rule that dictates how the stimulus should appear, or because they have an intuitive knowledge of certain physical events. These experiments can include a training phase which serves the purpose of teaching infants a particular rule and then present them with a series of generalization trials with novel stimuli; in some cases, any training phase is absent and the experimenter tests the infants’ prior spontaneous knowledge of certain events.

To assess eye movements, it is possible today to use automated eye-tracking systems with relative ease. Before describing an experiment using the anticipatory eye movement paradigm, in the next section of this chapter I will describe the evolution of this technical device used to collect data in experiments using the eye movement measure.

2.4.3.1 The Eye Tracking Methodology

Today, most experiments using eye movements simply recur to observation of the eyes, but this technique has many limitations since it is not always easy to determine the infant’s gaze, in particular the vertical gaze position. Off-line coding allows the experimenter to play the recordings of the subjects in slow motion and trained coders are often reliable in their judgments, nonetheless collecting and analyzing data with this technique can take a lot of time and lack precision.

In the early 1960s, electro-oculography (EOG) and corneal reflection photography were applied for the first time in studies of eye-movements in infants.

EOG records the electrical potential between the cornea and the retina.
Electrodes are placed near the orbit of one or both eyes. EOG measures eye position relative to head position, so that any change in head position results in gaze error. Electro-oculography could thus be reliable if head movements were eliminated, and this is impossible with infants who have rapid head movements. Another problem is the use of electrodes that would most probably be cause of fussiness in the infant.

The other technique that was applied to infant studies nearly 40 years ago was the corneal-reflection photography. This method measures eye position relative to fixed light sources. If the eye fixes a small light, this light will form a reflection on the cornea of the eyeball. Initial research with this technique (Haith, 1969) aligned 6 or more infrared lights with the stimulus display. The lights were directed toward the infant’s eye creating a reflection of the corneal surface. An infrared camera recorded the baby’s eye on a videotape. The researchers presumed that the centre of the pupil coincided with the line of sight (gaze axis) and the reflections on the baby’s cornea were used to determine the baby’s fixation position.
However, this technique presented some problems, arising mainly by the assumption that the line of sight and the centre of the pupil coincided. Later studies showed that the fovea (the visual axis) is displaced slightly from the centre of the pupil. Another limitation of this technique was that it was difficult to judge where the position of the centre of the pupil was. Under normal light conditions, the pupil is black, but with subjects with dark eyes it was difficult to verify from the video recordings where the boundary between the iris and the pupil was. The problem was later overcome by creating a ‘red eye effect’ by placing the light that creates the corneal reflection close to the lens of the camera.

The other drawback of this method was that the video camera had a close up lens that filled the video frame with the pupil. As a result, any head movement resulted in a loss of the pupil image. Today, thanks to important technological advances, these problems have been overcome.

The principle that modern eye trackers use today is the same principle developed in the 1960s: an invisible infrared light is directed onto the eye and this enables the eye gaze (the direction towards which one person directs his
visual attention) to be measured. The reflection of the light on the cornea is recorded by a sensor that then calculates the exact point of gaze of each eye, using a geometrical model. The points of gaze are recorded on a computer and can then be visualized when coding the experiment. To prevent the out-of-gaze problem, a camera with a wide-angle lens that can deal with fairly large head movements is used. Some modern eye trackers incorporate the camera and the tracking devices directly into the screen, thus making the setup simpler and more suited to infants' studies.

Before running any experiment using the eye tracker, there is a phase of calibration, to adjust individual eye gazes to fixed point onto the screen. In infant research this phase has to be quick -- the more time lost in calibration, the less cooperative the baby is likely to be during the experimental trials. The infant sits in the parent’s lap in front of the screen where an eye tracker device is mounted (the same screen will be used to play the stimuli during the experiment). The parent wears black opaque glasses in order to be sure that the infrared light source is reflecting the baby’s pupil and not that of the caregiver. A sounding attractor (usually a puppet or a coloured moving object) is played on the screen. The infant follows the attractor, which is moved by the experimenter every 5 seconds to five different parts of the screen. The light level used during calibration has to remain the same during the experiment. Calibration is necessary to provide a metric for every gaze position of the subject within the stimulus field and has to be repeated until satisfactory values are recorded for each calibration point. Normally calibration takes not more than one minute to be accomplished and immediately after it, the experimental movies start being played.
The best infant suited eye tracking device is the desk-based automatic gaze tracker (fig. 22 a). With this device calibration of points of gaze is very rapid and also very accurate. Other types of eye trackers require a head mounted device supplied with a mechanism that is sensible to head movements and that sends information to a camera, which in turn adjusts the focus to maintain the eye image in the field of view of the camera (Fig. 22 b). This type of eye tracker is less suitable to infant studies as it is extremely hard to make a child feel comfortable with the device on her head.

In sum, eye trackers are very useful devices in infant studies because they provide accurate measurements of infants’ gaze direction. The desk-based version is a non-invasive device that allows for a quick calibration and provides a good collection of data. A description of one experiment using eye tracking technology can help the reader to figure out what kinds of questions a researcher can ask with the aid of this technology.

2.4.4 Description of an experiment using the eye tracker in an anticipatory eye movement response

McMurray and Aslin (McMurray & Aslin, 2004) used the anticipatory eye
movement paradigm in experiments with infants aged between six and seven months.

In one of their experiments, after having calibrated the eye tracker for each subject, the researchers presented the babies with a series of training trials.

Infants saw a blue T-shaped occluder. At the base of the occluder a red square and a yellow cross appeared in turn. After shrinking to half of its original size, the red square moved behind the occluder to reappear after 750 msec on the top left corner of the occluder. The yellow cross had exactly the same behavior but it reemerged on the top right corner of the occluder. The scene was repeated for 16 consecutive trials. The 750 msec pause allowed the infant to make anticipatory eye movements to the position where the objects would reappear.

After this phase, during which the infants learnt to anticipate the location of the red square and of the yellow cross, the test trials began. These were divided in blocks of four trials. Each block contained a red square and a yellow cross (trained stimuli) and a yellow square and red cross (generalization stimuli) in random order. Unlike the trained stimuli, the generalization stimuli did not reappear from behind the occluder. The blocks of trials continued until the infant became fussy.
Figure 14. Representation of experiment 3 in McMurray & Aslin (2004) Column A shows the training trials; column B the test trials
The results of this experiment showed that infants learnt to anticipate the location of the visual stimuli during the training phase and that, when they were presented with novel stimuli, they looked for the reappearance of the objects according to their colour membership, but not according to shape membership. The red cross was expected to exit to the left and the yellow square to the right, showing that infants are able to discriminate colours more easily than shapes in forming visual categories.

The anticipatory eye movement paradigm can also be used to study speech perception in infancy. Indeed, in the same study McMurray and Aslin used the occlusion based anticipatory eye movement paradigm to examine speech discrimination abilities in infants. In the training phase infants, of an age between 5 and 7 months, saw the same occluder and at its base a circle. This visual stimulus was paired to one of two synthesized words: lamb or teak. The colour changed randomly across trials and was therefore not related to one word in particular. The only generalization that the infants had to learn was that when they saw the circle and heard the word lamb, the circle would move behind the occluder and reappear on the top left of it, while when they heard the word teak the circle would move behind the occluder and reappear on the top right. The training phase consisted of 20 trials.

Testing was performed in blocks of four trials. The procedure was the same as in the experiment we discussed above. Each block presented one trial identical to the ones seen in the training phase and three trials with a heightened pitch or lengthened duration of the lamb and teak stimuli, all in a random order. In all four trials, the visual stimulus did not reappear from the occluder. The trials stopped when the babies showed to be no longer cooperative, for an average of 23.4 trials.

More than one third of infants showed to be able to learn the general association between stimulus and location. In addition they showed to be more sensitive in recognizing the verbal stimuli when there was a variation in pitch than one in duration.

The two experiments I have described in this section are conceived with a
training phase followed by a series of test trials. However, this procedure is not the only one that researchers adopt when using the anticipatory eye movement paradigm. Indeed, the training phase may not be used at all in experiments that try to evaluate the intuitive representation that pre-verbal infants construct of events in the world, regularities, or linguistic phenomena.

2.4.5 Which questions can be addressed to infants using the Visual Expectation Procedure?

The Visual Expectation Procedure is a very resourceful method that is used today in many laboratories studying infants' cognitive development.

The method achieves important advantages when its use is associated with modern eye tracker techniques.

The method is particularly useful in assessing how infants categorize visual and auditory stimuli. It also helps researchers understand which features of the novel stimuli are more relevant for them.

When this procedure is used in a violation of expectation paradigm it is also possible to get a wider insight into the knowledge that pre-verbal infants have about physical dimensions, or into their naïve intuitions of possible and impossible events.

2.4.6 Advantages and Disadvantages of the Visual Expectation Procedure

The main problem with the visual expectation procedure is the high drop-out rate. Often the training phase can take many trials to be accomplished, and infants tend to get fussy when they have to sit down for a relatively long time. Perhaps this limitation will be overcome in the future, should the experimenter be able to control the duration of the training phase with an eye tracker that can be controlled online. Nowadays, due to problems tied to the computational resources available, most of the data gathered with eye trackers can only be analyzed offline, after the experiment is finished. Indeed, some infants learn more quickly than others the rule that is being presented to them, and could therefore skip certain training trials that in the experimental scripts adopted
today are fixed in number.

Another problem with this technique is that eye trackers, particularly those suitable for infant studies, are expensive and that the need for calibration can sometimes render the infant non-cooperative unless this phase is accomplished rapidly. Thus, the ability of the single experimenter in achieving a good balance between the need of precise calibration and the infant’s limited attentional resources is essential. The fact that the success of the technique is heavily dependent on the quasi-artistic skills of the single experimenter render the replication of eye tracking experiments with infants potentially problematic.

It is also very important to give precise instruction to the parent accompanying the infant as to not interfere in any way with the experiment.

The advantages of the Visual Expectation Procedure reside on the nature of the questions that can be addressed to the pre-verbal infant using this technique. As discussed in Section 2.2, the habituation paradigms attain a yes-no response because they can tell us whether the infant does or does not discriminate two stimuli; from this information, researchers can conjecture if the subjects include a stimulus in a category or not. The Visual Expectation Procedure, when used in a two-alternative anticipatory eye movement paradigm as in the experiment by McMurray and Aslin (2004), can reveal how infants categorize stimuli rather, than how they fit the categories that have been chosen for them by the researchers. Furthermore, a crucial aspect of this technique is the possibility to test repeated contrasts within the same infant, thus allowing for a great increase in statistical power with respect to simple habituation techniques.
2.5 The Head-Turn Technique

This section will describe two procedures that rely on infants’ head turn responses: the Conditioned Head Turn Procedure and the Head Turn Preference Procedure. Both procedures are used in studies of infant speech perception and are founded on the fact that infants orient towards a sound source and will keep the head turned towards that direction when the visual or auditory stimulation are contingent on their behaviour.

2.5.1 The Conditioned Head Turn Procedure

When researchers use the Conditioned Head Turn Procedure to study the perception of speech stimuli, through operant conditioning they teach infants to turn their heads when they perceive a change in sound stimuli.

This method was originally used as a tool in clinical audiology. It relied on the fact that the baby would naturally orient to a sound source and that this behaviour, together with other responses such as eye-blinking, would therefore be an indication of her hearing faculty.

However, audiologists were not satisfied with this technique because it could not measure auditory thresholds. Furthermore, infants could only be tested for a limited period of time because, in the absence of any reinforcement, they would habituate and no longer turn to the sound source.

It was for these reasons that in the 1960s some clinical audiologists like Suzuki and Ogiba (e.g., Suzuki & Ogiba, 1961) used the head turn response associated with an operant conditioning technique with a visual stimulus as reinforcer. Initially it was used with infants not younger than 12 months, and in the 1970s it was adapted to study the auditory capacities of infants between 5 and 12 months. The audiologists using this method trained infants to turn their head towards a loudspeaker when it played a sound. If the baby turned to the right side, a reinforcer was presented. Since it was not clear which kind of reinforcer
would be most efficient, Moore and collaborators (Moore, Thompson, & Thompson, 1975) examined four types of reinforcement:

- No reinforcement
- Social reinforcement (a smile or a verbal praise)
- Simple visual reinforcement such as a blinking light
- Complex visual reinforcement such as an animated toy.

The results revealed a clear preference for the complex reinforcement that kept the babies to be cooperative for a longer time than the other stimuli.

Other studies by the same authors (Moore, Wilson, & Thompson, 1977) demonstrated that the head turn technique is efficient with infants as young as 5 months of age, while younger subjects were not successfully conditioned.

Let us now examine the basic steps of this procedure in more detail (Figure 15). The infant is tested in a small silent room; she sits on the parent’s lap in front of a small table and opposite the infant sits an experimenter (E1). Both the experimenter and the parent wear headphones with masking music that prevents them from hearing the stimuli. On one side of the infant (at approximately 1.5 metres) there are an audio speaker and (at approximately 1 metre from the baby) a dark plexiglass box and hidden inside it a commercially available battery-operated toy. The experimenter attracts the baby’s attention by quietly showing her some coloured toy, thus maintaining the head of the infant in a midline position. Whenever a sound is played over the audio speaker and the baby turns towards it, the plexiglass box lights up and reveals the animated toy. Wrong head turns are not reinforced. Another experimenter (E2) outside the room watches the infant through a one-way mirror or a video monitor. Whenever the baby makes a head turn, E2 presses a key of a computer that scores the number of head turns.
Eilers (1977) adapted this basic procedure to study speech discrimination in infants between 6 and 8 months and between 12 and 14 months. The basic phases of the procedure are identical to the ones just described, the differences consisting on the fact that there is a continuous background sound during which there may be a brief change in the auditory stimulus. Infants are at first trained to make head turns in response to a sound change. Most infants naturally turn towards the loudspeaker when the stimulus varies and at that moment, immediately, the reinforcement is activated. For those subjects who do not spontaneously orient towards the speaker, the reinforcer is displayed contingent to a sound change. Eventually, after a number of trials the infant will anticipate the appearance of the reinforcer at the change in stimuli. After the infant has learnt the contingency between change in sound, her behaviour and reinforcement, a series of control trials during which there are no changes in
stimuli are introduced.

There were many possible sources of bias in this procedure, depending mainly on the fact that the experimenter outside the room knew before the trial started if the trial was going to be a control trial ('no change') or an experimental trial ('change'). Thus, he/she could initiate the experimental trials when the baby was not paying attention to the other experimenter entertaining her with the toy. In such occasions, the baby has a greater probability of orienting towards the visual reinforcer, independently of a change in sound. Therefore, the experimenter should be blind to the sequence of trials to avoid producing a false behaviour.

Kuhl (1983) introduced a series of variations to the procedure that standardized the technique and introduced a series of measures that reduced the chance of bias on the experimenter’s part.

The procedure is divided in three stages: a training phase, a conditioning phase and testing phase.

The training phase wants to familiarize the baby with the reinforcer, so the latter is activated immediately after the presentation of a new stimulus. This stage can last between 3 and 8 trials.

After the training phase the conditioning phase begins. During this stage, the infant is presented with a series of ‘change trials’. The work of experimenter E2 is very important at this stage because he/she has to gradually condition the infant. At first the experimenter will immediately present the reinforcer as the sound changes then he/she will gradually delay the presentation of the reinforcer to allow the infant to make a head turn and thus control the presentation of the visual reinforcer when detecting a change in sound.

When the baby has scored three correct head turns in a row, the testing phase begins. During this stage experimenter E2 is totally blind to the presentation of the stimuli. He/she has to press a key in a computer when the baby is engaged looking at experimenter E1. The computer then controls the presentation of the trials which can be either control trials during which there is no change in stimuli (for example [a],[a],[a],[a], [a],[a],[a],[a]…) or an experimental trial during which there is a brief change in stimuli (for example [a],[a],[a],[i],[i],[i],[i],[i],[a],[a],[a]…). Experimenter E2 pushes a button whenever
the baby makes a head turn towards the box. If, the head turn is made during a change trial the reinforcer is displayed and the response is scored as correct. On the contrary, making a head turn towards the box during control trials is recorded as incorrect (false head turn), while not turning during this phase is recorded as correct (correct rejections). If the infant does not make a head turn during change trial, no reinforcement is activated and the failure is scored as incorrect (miss).

![Image of a baby tested with the CHT procedure](figure16.jpg)

**Figure 16.** A baby tested with the CHT procedure (from Werker et al., 1997)

2.5.1.1 **Which questions can be addressed to infants using the Conditioned Head Turn Procedure?**

The Conditioned Head Turn Procedure continues to be used for the purpose for which it was originally created, that is, for assessing auditory thresholds in infants.

However, this procedure, with the appropriate modifications, can also be used to study discrimination abilities in speech perception in pre-verbal infants. In Section 2.5.1 we have already seen how phoneme discrimination tasks are realized. The same procedure can be adapted to test whether infants detect, not
only two phonemic contrasts, but also two different frequencies, melodies or voices.

Researchers have also used this procedure in categorization tasks. Adults categorize exemplars of one phoneme differing, for example, in suprasegmental dimensions such as pitch or voice, as belonging to the same category (within category variances), while cross category variances are classified as belonging to different phoneme classes. Kuhl (1983), using the conditioned head turning procedure, demonstrated that infants categorize phonemes in the same way, that is they ignore changes in pitch and voice while maintaining the discrimination of a cross-category membership. The procedure was modified to train infants to respond differently to two speech sounds contrasting two phonemes of different categories (/a/ and /i/) and then testing for generalization of the response to novel stimuli from the two categories.

Since the conditioned head turn procedure can be used with infants as well as children and adults, it can also be a valuable technique to measure the discriminative capacities in various stages of development. Of course, the basic procedure used with infants has to be modified to best suit the age range. With children between 2 and 5 years of age the head turn response can be substituted with a bar or button press, while with older children and adults we can use a hand signal instead of a head turn response. In the study we discuss in Section 3.2.5, Werker and Tees (1984) used the conditioned head turn procedure to test the ability of adults and of infants of three different age groups (6-8 months, 8-10 months, 10-12 months) to discriminate non-native phonetic distinctions. The data collected showed a better perception in the younger group than in all other groups.

2.5.1.2 Advantages and Disadvantages of the Conditioned Head Turn Procedure

The Conditioned Head Turn Procedure is a valuable method that can be used for hearing assessment and speech perception.

However, the procedure has many potential sources of bias which have to be carefully avoided by the experimenters. It is thus important that the parent
and experimenter E1 both wear headphones to prevent them from hearing the stimuli. Also experimenter E2 should not hear the stimuli and, most importantly, should not be aware of the trial succession during the experiment.

Like all behavioural techniques used with infants, the Conditioned Head Turn Procedure can have variable drop out rates that may be influenced, for example, by the correct choice of the reinforcer and of the response, that have to be appropriate to the age under investigation. Another measure that can help reduce the drop out rate is to train the experimenters to be flexible in the presentation of the trials, that can be either delayed or anticipated to respect the different rhythms that each infant is likely to display in learning the contingency between her response and the reinforcement. While not necessary, it is also advisable to use a control group in order to distinguish between wrong performances related to boredom and failure related to perceptual factors.

The Conditioned Head Turn Procedure has among its strengths the fact that it can provide individual data on the perceptual or hearing abilities of single infants, thus identifying those subjects who could have hearing or perceptual difficulties.

This technique is a highly sensitive testing procedure, in which each child can be presented with multiple test trials. Particularly successful are those experiments in which the stimuli used are short speech patterns such as syllables, single and multi-syllabic words or short melodic patterns.

As we have discussed, the procedure can be used with the appropriate variations across the whole developmental span therefore, yielding important results in cross-sectional and longitudinal studies.

It is debated whether the Conditioned Head Turn Procedure can be a valuable tool for assessing categorization in infants, because in these tasks it gives a Yes-No answer to the research question, leaving all other conclusions open to various interpretations.

2.5.2 The Head Turn Preference Procedure

The Head Turn Preference Procedure relies on the fact that infants naturally turn their head towards a sound source and that as long as they are
interested on such a stimulus they will keep their head turned towards it.

In order to assess discrimination or preference, the Head Turn Preference Procedure uses as dependent measure the length of time infants turn their head toward a particular sound. Infants learn that the presentation of a stimulus is contingent on their maintaining a head turn.

In the last 20 years, the technique has evolved considerably. In the following section I will describe how the procedure has changed over the years. In Section 3.3.4, in order to better understand its application today, I will also describe in detail a study that has adopted the technique to study the computational abilities that infants may have, and for which I have direct knowledge as an experimenter.

2.5.2.1 Description of the Head Turn Preference Procedure

Fernald (1985) used the Head Turn Preference Procedure technique for the first time to study infants’ preference between adult- and infant-directed speech. She used as dependent variable only the direction of the infants head turn, and not the length of the temporal interval of preference that is used today (Nelson, Jusczyk, Mandel, & Myers, 1995).

Fernald used a three-sided booth, open on the fourth side. She hung white curtains from the ceiling to the top of the booth so that the infant would not see the rest of the room. The baby sat on the caregiver’s laps in the centre of the booth. On each side of the booth there was a light; the central light was green (the baby faced this light), whereas the two sidelights were red. Loudspeakers were hidden behind the two side panels. A video camera masked in the front panel recorded the baby during the experiment.

In its first uses, the Head Turn Preference Procedure consisted of a series of training trials during which the child learnt that each side was associated with one particular stimulus (the side varied across subjects but not across one experimental session). At the beginning of each trial the green light on the front panel was flashed to get the attention of the subject. Once the baby was in midline position, the green light was turned off and one of the red lights was turned on. The experimenter then played a recording of an 8 second stimuli that
was always associated with one of the sides (for example infant-directed speech on the right and adult-directed speech on the left). The stimuli continued to play and the light to flash regardless of the head turn behaviour of the infant. After the stimuli ended, the red light was extinguished and the green light was switched on. The same procedure was repeated 4 times. During testing, the caregiver and experimenter wore headphones; the green light attracted the attention of the infant to the centre and was then extinguished. The first head turn made by the baby gave way to the playing of the stimuli associated to that side and was accompanied by the blinking red light. The sample had to be played in its entirety regardless of the looking behaviour of the infant. Then the researcher measured the number of head turns to each given side, and found that the side playing the infant-directed speech was preferred.

In subsequent years, the main modification introduced in infant research using this technique was to change the primary dependent measure. Instead of being the simple head turn, the measure became the amount of time that infants orient towards a speech sample on a certain trial.

After these revisions, the Head Turn Preference Procedure has undergone other significant changes.

Today, a typical head turn preference procedure setup looks as follows (Figure 17). The infant is tested inside a three-sided booth; the panels measure approximately 120cm x 180cm. Black curtains hang from the top of the panel to the floor to avoid any distractions. The room is dimly lit for the same purpose.
The infant sits on the parent’s lap and faces a screen or a green light at the infant’s eye level. Above the screen there is a video camera that records the infant during the experiment. The parent wears headphones and listens to masking music in order to avoid cueing the infant towards a response throughout the experiment.

On each side of the infant, at head level, there is a monitor or a red light. Behind each of them a speaker. The parent’s knees have to be aligned with the two sidelights or monitors.

The infant’s attention is drawn to the centre by flashing the green light or in the case of a monitor a blinking light is played on it. Once the infant is looking at the central monitor, the flashing light disappears and one of the sidelights is turned on. When the infant orients towards it, making a 30º head turn, a sound from that speaker begins to be played. The sound does not stop unless the infant turns away from the monitor for at least 2 seconds or, in the case that the infant maintains continuously the head turn, until the entire stimulus of that trial has been played.

There are two phases in the Head Turn Preference Procedure: a training phase and a testing phase.
The training phase can be used in different ways; it can either familiarize infants with the stimuli or train them to learn the contingency between their head turns and the auditory stimulation. In this latter case, the sidelight is turned off as soon as the infant turns towards it and only the sound continues to be played throughout the head turn. This procedure was adopted because it was noticed that otherwise orientation during testing was considerably shorter.

The duration of the head turn proved to be a more sensitive measure than simple orientation. It rendered unnecessary to display all the trials of one stimulus to one side. A further methodological benefit in this new procedure comes from the fact that the computer can control the random presentation of the trials, thus reducing possible sources of bias on behalf of the experimenter.

2.5.2.2 Advantages and Disadvantages using the Head Turn Preference Procedure

The Head Turn Preference Procedure has proved to be a very important instrument in the study of infants’ language perception. It can be used with infants as young as 4 months up to 18 months. Yet, it can have a high drop-out rate, especially with older infants who generally do not like sitting for a relatively long time.

One of the most important strengths of this procedure is that it allows the presentation of long samples of continuous speech. This advantage has permitted researchers to test infants’ sensitivity to many cues present in the speech signal, such as prosodic cues related to clause boundaries (Jusczyk et al., 1992), or native language stress patterns (Jusczyk, Cutler, & Redanz, 1993).

The Head Turn Preference Procedure is a versatile technique that can be modified in diverse ways. For example, the training phase can become a familiarization phase, during which infants are simply presented with speech stimuli, and can then be tested for their retention, or else for how novel speech samples are categorized. Thanks to its versatility, nowadays it is probably the major behavioural technique used for infant language studies.
2.6 Chapter conclusion: techniques and problems

This chapter has reviewed the most important behavioural methods that are used today in the study of speech perception and other cognitive abilities in infants. I have tried to give a short but comprehensive overview of the major techniques used in infant studies. I hope it has become clear that each age, and each research issue, presents different problems and demands different techniques. Nevertheless, undeniably the last twenty years have seen a marked progress both in terms of quality and of quantity of studies dedicated to the cognitive abilities of the pre-linguistic child. The number of laboratories working on such issues has also increased radically.

I now propose to go back to the issue I raised in Chapter 1 -- what can we say on the nature and origin of language if we look at them with the tools provided by the techniques I have presented?

Chapter 3 will explain the theoretical framework that rests behind the experiments I have described so far and this will help us understand the enormous contribution given by these techniques in the study of infant research.
3 LANGUAGE AND NATIVISM: THE APPLICATION OF THE TECHNIQUES AND THE STATE OF THE ART
3.1 Language nativism and experience: experiments and methods

In Chapter 1 I presented evidence that suggests that language has a biological basis. This position, which has been accepted or presupposed by most linguistic literature of the past 30 years, has been recently strongly challenged by a series of results in the psycholinguistic literature. Reviewing the arguments for the innateness of language representations, the philosopher Fiona Cowie concludes:

My aim here has been to sketch the ways in which modern understanding of the mind reveals the inadequacy and implausibility of the claim that humans have innate representations of UG that are responsible for their acquisition of language (Cowie, 2008).

While the debate between nativism and experience goes well beyond this work, in this chapter I intend to recall some of the evidence that seems to me to be of major importance to decide this issue. The evidence I will report also shows several applications of the techniques presented in Chapter 2. Thus, the aim of the current chapter is twofold. On the one hand, I intend to keep an eye on the issues that ensued from Chapter 1 -- is language founded in human biology? Do linguistic representations have an innate basis? On the other hand, I also want to give a sense of the limits and powers that the tools available to experimental psychologists offer researchers to address such questions. Necessarily, my choice of the evidence to present will be partial, although, I hope, representative of the current work psycholinguists are developing to unveil the mysteries of the early bases of language acquisition.
3.2 Language acquisition in the prelexical infant

3.2.1 Bootstrapping theories and statistical learning

As we have discussed in Chapter 1, evidence suggests that humans are born with an innate and abstract linguistic knowledge that guides them in the acquisition of language.

The ‘Poverty of the stimulus’ and the absence of ‘negative evidence’ have brought theorists to formulate the hypothesis of the existence of a set of rules contained in the universal grammar which through specific abstract notions (principles and parameters) help the child in the process of acquiring language.

The linguistic input however, does not have specific labels attached to it, nothing that tells explicitly to the child this is the verb, this is the subject or rather this is the head and those the complements.

The so called ‘linking’ problem (Pinker, 1984) has brought many researchers to study which possible cues exist in the linguistic input that help the child find in the concrete speech signal those abstract notions that are part of the universal grammar.

The theories that have tried to bridge the gap and solve the linking problem go under the name of ‘bootstrapping theories’ and take different approaches since each theory sees in one particular aspect of the signal a cue that triggers other mechanisms.

While the syntactic bootstrapping (Gleitman, 1990; Gleitman & Gleitman, 1994) and the semantic bootstrapping (Pinker, 1984) assume a prior knowledge of language, the prosodic bootstrap (Morgan & Demuth, 1996; Christophe, Mehlner, & Sebastian-Galles, 2001; Christophe, Nespor, Guasti, & Van Ooyen, 2003b) posits that infants start the acquisition of lexicon and syntax of their native language by picking up the phonological cues that are present in the speech signal and which are correlated to certain structural parameter.

Since the prosodic bootstrapping theory focuses on the mechanisms that are at work in the earliest stages of acquisition, in the next chapters I will explore
some of the results that have risen from such assumption.

Another line of research has also analysed the speech signal and has found that certain statistical regularities which are contained in it trigger some specific computational mechanisms that help the young learner to acquire the lexicon and the grammar of her native language (Saffran, Newport, & Aslin, 1996c; Marcus, Vijayan, Rao, & Vishton, 1999). The theories known as ‘statistical learning’ and ‘rule extraction’, will also be discussed in the following sections.

3.2.2 The universal correlation of languages

Before we discuss some of the abilities that infants have in detecting some cues which are present in language and which allow them to set some of the parameters necessary to learn the native language, we need to consider in more detail if there is in fact a correlation between the acoustical properties and the abstract grammatical properties of a language, and if such correlation can be extended to different language classes.

Languages according to phonologic classifications (Abercrombie, 1967; Ladefoged, 1975) can be divided in syllable-timed (French, Spanish and Italian belong to this class), stress-timed (for example English and Dutch) and mora-timed (Japanese and Tamil are part of this class).²

A typological study (Ramus, Nespor, & Mehler, 2000) showed how those languages which belong to the same rhythmic class not only have the same rhythm but many other similar characteristics involving their syllabic structure, morphology, syntax, and prosodic prominence.

According to the researchers, the rhythm of a language is reflected in certain acoustic characteristics. Ramus and colleagues measured the syllabic complexity ($\Delta C$) and vocalic ratios ($\%V$) of several languages and thus classified them according to such parameter. The syllabic complexity has been described as

² In phonological theory a mora is a unit of sound which distinguishes light from heavy syllables, The word “Tokyo”, for example, has four morae /to-o-kyo-o/. 
the amount of time we spend on consonants in the speech stream, while %V is
the time we spend on vocals.

Languages that have complex consonant clusters, such as English and
Polish, will have high $\Delta C$ and a low vocalic ratio, while a language with a
simple syllabic structure like Japanese will have a low $\Delta C$ and a high %V.

![Figure 18. Distribution of languages over the $\Delta C$ and %V plane
(Ramus et al., 2000).](image)

To the initial eight languages studied (Figure 18), several others were added in
later research as can be seen in Figure 19:
Figure 19. The figure shows 16 languages plotted in the %V and ΔC. The diagonal line separates the languages in two groups: on the left languages with low %V and head-complement order, on the right languages with high %V and complement-head order. (Dutch and Hungarian can have both orders but are predominantly H-C and C-H respectively.)

The two groups of languages separated by the diagonal line in Figure 19 have similar rhythms and syllabic ratios and also share similar morphosyntactic structures. The languages on the left tend to be inflecting and with a basic VO (verb-object) order, while the group on the right includes languages that tend to be agglutinative and with an OV order.

The proposal is particularly interesting for the prosodic bootstrapping theories because the infants’ ability to perceive and categorise languages by relying on prosodic information would allow for many syntactic parameters correlated with rhythm to be set.
3.2.3 The classification of languages by infants

As we have seen in Section 1.3, interactions between the brain and speech start immediately after birth. The results obtained by Peña and colleagues point in a specific direction, demonstrating that speech is processed in the same regions of the brain used by adults and suggesting that infants are able to tell apart linguistic input from other acoustic stimuli. In fact many studies have indicated that infants discriminate between non-linguistic and linguistic stimuli and show a preference for the latter (Colombo & Bundy, 1983). Infants prefer to listen to their mother’s voice than to a stranger’s (Mehler, Segui, & Carey, 1978) and are particularly attracted to a natural way of speaking, while they lose interest if a text is read backwards.

To see whether infants are able to sort utterances belonging to two different languages shortly after birth, Mehler and colleagues (Mehler et al., 1988) tested the ability of French newborns to discriminate between their maternal language and Russian using the high amplitude sucking technique (HAS) and of 2-month-old American infants to discriminate between English and Italian using looking times as the dependent variable.

The study yielded the following results: 4-day-old newborns are able to discriminate between French and Russian showing a preference in listening to their maternal language, two-month-olds while still discriminating between their maternal language and a foreign one showed no preference in listening to either one (Figure 20).

The same study also conducted experiments to see whether at these two different stages of language acquisition there is the ability to discriminate two foreign languages. The results indicated that while a few days after birth infants can discriminate two unfamiliar languages this capacity disappears at two months. Newborns are better than older infants because, as we will see when discussing other studies, linguistic experience changes the ability to discriminate speech sounds. In fact, the discrimination abilities of infants specialize for those acoustic contrasts that are present in the maternal language while non-native contrasts stop being distinguished (Werker & Tees, 1984).
Figure 20. Sucking rate of 4-day-olds when exposed to French or Russian speech samples pronounced by a bilingual speaker (Mehler et al., 1988). The Y axis represents sucking rate per minute.

The question that arises is what helps infants tell apart the mother tongue from other languages only shortly after birth?

Languages are different in many ways: they use different phonemes and the phonemes can be ordered in certain ways that are specific to every language, so that every language has a particular phonotactic structure that allows for certain sequences while constraining others (for example in English /m/ cannot be followed by /l/, in the same way in Italian /n/ cannot be followed by /b/). Languages also differ in their prosody: they have a different rhythm, intonation and stress.

Since 4-day-olds showed the ability to discriminate between languages it is highly improbable that they had knowledge of the phonemic range not only of their maternal language but also of foreign languages. Studies on prenatal exposure have shown that some prosodic information may pass to the foetus while in uterus, while no phonetic segments reach the foetus because the signal is greatly attenuated (DeCasper & Spence, 1986). The hypothesis that Mehler and colleagues advanced was therefore based on the assumption that infants rely on prosodic cues to classify languages. To test this hypothesis, the researchers filtered the speech sequences, so as to leave the intonation untouched but
eliminating phonetic information. The obtained signal could be compared to our listening to speech with our head underwater.

Once more the infants at both ages were able to make a distinction between the languages indicating that what helps infants in this task is prosody.

This study was only the first step in a series of investigations that tried to understand the precise information contained in the signal that infants use to distinguish languages.

Moved by this quest, Nazzi, Bertoncini and Mehler (1998) conducted a series of experiments with French newborns using the HAS technique. Their investigation was based on the hypothesis that infants are able to discriminate languages according to the rhythmic class they belong to.

In the first experiment the infants were familiarized, for at least 5 minutes, with a set of filtered English sentences (stress-timed) then the stimuli was changed to a set of filtered sentences in Japanese (mora-timed). The postshift period lasted for 4 minutes. In the control group there was no change in language stimuli between the familiarization and postswitch phase.

The results confirmed the hypothesis: the babies in the experimental group increased the sucking rate in the postshift period, while the infants in the control group did not.

The next step was to familiarize the infants with a set of filtered English sentences and see if, in the postshift phase after a switch in the stimuli using filtered Dutch sentences, the babies would show a recovery in sucking. As expected the neonates did not notice a change in stimuli because the languages, being both stress-timed, belong to the same rhythmic class.

Infants are interested in language from birth and are able to differentiate their native language from other languages. Not only does this capacity help the infant to attend specifically to what will be her native language, but in the case of a bilingual surrounding it will allow the child to distinguish precociously between the two languages by relying on rhythmic cues. However, infants could be exposed to two languages belonging to the same rhythmic class, so that this prosodic property could not help them to build separate representations for the two languages. However, other behavioural studies have shown that at around 4 months infants possess another ability that allows them to discriminate two
languages belonging to the same rhythmic class by relying on other differences such as their phonotactic properties (Bosch & Sebastián-Gallés, 1997).

3.2.4 The perception of syllables by infants

Speech stream is mostly continuous. There are no perceptive boundaries between words and when we listen to speech our task is to segment this flow into lexical units (morphemes). Prior knowledge of lexicon helps adults in the segmentation task. Unlike adults, infants do not have access to words but, nevertheless, have the ability to process speech and segment it in word-sized units. How infants represent at a sublexical level such units has been discussed by several studies: some claim a segmentation at phoneme level others at syllable, morae and feet.

A linear perception and categorization of phonemes would certainly be quite problematic as shown by a series of studies conducted over the last sixty years.

In the 1950s, with the intent of projecting reading machines for the blind, Cooper and Liberman experimented with synthetic speech sounds and tried to identify the invariant acoustic properties of phonemes. The task showed to be impossible to accomplish. The researchers used simple consonant-vowel (CV) syllables and tried to divide the stop consonant from the vowel without success. What they found was that coarticulation influences the acoustic sound of consonants, which are influenced by the vowel context. No acoustic invariant for certain consonants has been found until today.

If we think that phonemes (and other aspects of speech at the segmental and suprasegmental level) not only vary according to the vowel setting, but also to who (man, woman, child) pronounces them or to how fast the speakers talk, the task of accessing the same, invariant sound to build a lexicon, becomes all the more difficult for the young learner. We ask again the same question: what are infants relying on to process speech?

Bertoncini and Mehler (Bertoncini & Mehler, 1981) posited that the representation of speech in infants is the syllable. The researchers conducted an experiment with French newborns, testing their ability to detect a change in CVC
type stimulus and CCC type stimulus using the habituation paradigm with the non-nutritive sucking technique.

When the syllable changed, for example, from tap to pat, the babies noticed the change while no distinction was made between the consonant clusters pst and tsp. When at the beginning and end of each CCC sequence, a vowel was added, obtaining upstu and utspu, the infants were able to discriminate the stimuli hearing the sequence as up-stu and ut-spu. What was proposed was the consideration of syllables as gestalt-like perceptions so that blocks of consonants and vowels would be naturally perceived by the perceptual system. Gestalts would therefore be operant not only at a visual level (according to Gestaltists, we see differences in size of two identical shapes by considering them as a whole not by separating their constituent parts) but also at the auditory level. Infants identify and process syllables more easily than other segmental units; syllables can be retained through time and thus infants can discriminate changes in the phoneme chain.

Another study (Bijeljac-Babic, Bertoncini, & Mehler, 1993) found that infants are able to ‘count’ the syllables that are present in the acoustic sequence. In the experiment, 4-days-old newborns were familiarized with a multisyllabic flow of syllables (pa-pa). During the trial session the stimuli changed in the number of syllables (pa-pa-pa). Infants detected the change but it was not clear whether the subjects were using as a distinctive trait the number of syllables or the different duration in time of the sequences. To rule out such possible interpretation the researchers compressed the trisyllabic sequences so as to have duration equalized for bisyllables and trisyllables. The result was that infants could still discriminate between the two stimuli. Infants therefore, are able to make a classification of CV constituents according to their number.

Syllables are processed more automatically by the perceptual system and play a very important part in the acquisition of language; even adults under experimental conditions have shown to have less difficulty in identifying syllables compared to single phonemes. Surprisingly longer and more complicated strings are spotted with less effort by our perceptual system.

As we have discussed in Section 3.2.2, the study by Ramus showed how the syllable structure is related to rhythm and infants are particularly sensitive to
this prosodic property that allows the discrimination of different languages.

Different studies (Cutler, Mehler, Norris, & Segui, 1992) have shown that the segmentation procedure that adults adopt is correlated with the rhythmic class to which their native language belongs to, so that speakers of a syllable-timed language will segment utterances at syllable level, speakers of stress-timed languages will segment at the onset of strong syllables while speakers of mora-timed languages in morae (units shorter than the syllable but longer than the phoneme). Not only do adults segment their native language, but all languages by relying on such units.

At the earliest stages of language acquisition, the rhythmic information contained in the speech signal helps the infant to find the unit that helps them in segmentation procedures and which become more refined and specific in time to suit the phonologic structure of the native language.

3.2.5 Perception of speech sounds and linguistic experience

Our perceptual system detects and discerns through the senses a number of visual/auditory stimuli coming from the environment and then organizes such input in categories (Mehler et al., 1994).

To understand how our perceptual system categorizes the input it receives we can take as an example that of colour perception. A colour naturally changes if the wavelength of a luminous source is changed. While we are perfectly able to tell the difference between the various shades of colours, our perceptual system divides colours into four main categories. From a whorfian approach, it could be argued that such categorization is arbitrary because it has been shaped by linguistic and cultural conventions. According to this formulation, newborns could not possibly have perceptual categories as these would emerge consequent to education in a particular community.

Contrary to this proposal, through behavioural studies Bornstein and collaborators showed that newborns like adults have visual perceptual categories (Bornstein, Kessen, & Weiskopf, 1976), and that they also classify colours into four broad classes that correspond to blue, green, red and yellow. The researchers habituated the newborns by presenting them with visual stimuli
which adults classify as blue, then changed the stimuli to another shade of blue or to green. Newborns reacted when the change spanned a chromatic category, while their reaction was much less evident when the change occurred within a category.

![Figure 21](image)

**Figure 21.** How babies habituated to a chromatic category of 480 nm wavelength or 450 nm wavelength react when presented with a different stimuli. The Y axis represents the average time of visual fixation. The left part of the panel shows that infants dishabituate when the dishabituating stimulus crosses a category boundary (from 480 nm to 510 nm) but not when the same physical change stays within the same perceptual category (from 480 to 450) *(taken from Mehler et al., 1994)*

The categorization of colour then, is not something that we have learned to do, but it is something that we are born with and is the same regardless of age or culture.

Our perceptual system not only categorizes visual input but also auditory stimuli. Sounds and in particular consonants, are divided by adults into broad categories. If, using a synthesizer, we modify the voice-onset time (VOT: the interval between the release of the closure of the vocal tract and the vibration of the vocal fold) we can obtain a series of acoustic differences between two consonants such as *p* and *b*. Adults however do not perceive a difference in all the possible sounds obtained but only recognize them as belonging to one of the
two categories.

Figure 22. Idealized pattern of identification and discrimination results. Although between the first and second arrow there is the same physical distance in sound as between the third and fourth arrow, we hear two different sounds.

In the 1970s Eimas and collaborators (Eimas et al., 1971) conducted a series of experiments to see whether infants behave in the same way regarding to phoneme categorization.

They tested the ability of infants at two ages, 1 and 4 months, to discriminate the VOT in the syllables *ba* and *pa* and found that the distinction, as for adults, is categorical.
The study by Eimas marked the beginning of a series of experiments to understand the abilities of infants in discriminating speech sounds. A well-specified trend emerges from the data collected so far: while in the first few months infants are able to discern all phonetic contrasts of their native language and of foreign languages, they then discriminate only those phonemes which are relevant for their maternal language.

Werker and Tees (1984) tested the ability to discriminate non-native phonetic contrasts in adults, and in babies with an age ranging between 6 and 8 months, 8 and 10 months and 10 and 12 months using the conditioned head turn procedure.

The stimuli contrasted two glottalized voiceless stop consonants differing in place of articulation /kɪ/ (velar) and /qɪ/ (uvular) which have a phonemic
distinction in the Thompson language, a Salish language spoken by the native Indians in south central British Columbia.

Figure 23. The diagram shows the results of Thompson speaking adults, of English speaking adults and English monolingual infants correctly identifying the Thompson glottalized velar/uvular contrast

The study showed that infants are able to perceive foreign language contrasts. The ability declines with age and is a consequence of the perceptual reorganization that follows the specialization of the discrimination abilities to the acoustic contrasts that are relevant for the native language.

Further experiments (using also other non-native contrasts) established that while infants aged between 6 and 8 months have fine-grained discrimination abilities, the percentage gradually lessens and at 12 months the infants’ performances resemble those of adults.
The results obtained by Werker and Tees (1984). Babies gradually lose the discriminating ability in detecting foreign language contrasts.

The loss of such capacity is an example of what has been seen as ‘learning by forgetting’ (Mehler et al., 1994). Initially, infants have the ability to detect all the phonemes that can be found in all languages, then such capacity specializes so as to discriminate the contrasts that are relevant only for the target language. These modifications are common across other areas of language acquisition, reflecting a maturation of the cognitive system rather than a regression of our capacity to hear sounds. Indeed, a study by Best et al. (1988) showed that those contrasts that have no possible counterparts in our language are clearly perceived regardless of age. The researchers analysed the discrimination of Zulu ‘clicks’, a particular linguistic sound very different from any signal found in Western languages. Being such sound unlike any other, the ability to
discriminate it is not lost.

The transition from a general phonetic discrimination to a specific phonemic one, is in line with the results on the rhythmic discriminations of infants, that are stronger in younger infants than in older infants in their effort to find a regularity of the maternal language.

3.2.6 Word order and the setting of the Head-Complement parameter

While discussing the X-bar structure, we said that while the hierarchical relations in the syntactic phrase are universal, the order of its elements has to be set in the appropriate parameter.

We have also argued that there are no explicit references in the speech signal that tell the child which is the ‘head’ or which are the ‘complements’.

From previous studies (see Brown, 1973; Pinker, 1994) we know that the parameter relative to word order is fixed at an early stage of language acquisition. When the production of two words-combination starts (at an average age of 18 months) the order is already properly set, and when babies are still in the one-word stage they make no mistakes in understanding sentences which could be ambiguous if the babies were not using syntax to help them with the comprehension task (Hirsh-Pasek & Golinkoff, 1996).

Various proposals (Nespor, Guasti, & Christophe, 1996) posit that infants use the prosodic information of phonological phrase boundaries to set the value of the head-complement parameter.

Before we analyse in more detail such proposals we need to establish whether prelexical infants are sensitive to phonological phrase boundaries.

Phonological phrases are the units just below the intonational phrase characterised by containing one or two content words and some function words. Prosodically they are distinguished by a preboundary lengthening.

Various studies (Christophe, Dupoux, Bertoncini, & Mehler, 1994; Christophe et al., 2001) have shown how newborns perceive the cues that are correlated with phonological phrase boundaries and Cristophe and colleagues (Christophe, Gout, Peperkamp, & Morgan, 2003) using the head turning paradigm, revealed how adults and prelexical infants at age 13 months, actually
use phonological phrase boundaries to segment speech and thus access lexicon.

Regarding the bootstrapping of syntactic acquisition, an interesting study by Christophe, Nespor, Guasti and Van Ooyen (Christophe et al., 2003b) revealed how the acquisition of word order might actually be prior to word learning.

The researchers argued that a cue to the syntactic structure of a language is present in the prosodic structure, and is correlated with a prominence within the phonological phrase.
In head-complement languages such as English, French and Greek, the prosodic prominence is phrase-final:

```
[Le grand orang-outang] [était enervé]
```

In complement-head languages, such as Turkish, Japanese and Bengali, the prosodic prominence is phrase-initial:

```
[Yeni kitabimi] [almak istiyor]
```

The researchers chose the stimuli to be used in the experiments with particular attention. The matched pairs of sentences had to differ only in the prominence within the phonological phrase, while the number of syllables, the position of word boundaries, word stress, phonological and intonational phrase boundaries had to be the same. To eliminate all phonetic information, the sentences were resynthesized to contain only Dutch phonemes while leaving the original prosodic structure intact.

The subjects were French infants aged between 6 and 12 weeks, and they were tested using the High Amplitude Sucking technique.

The results showed that the infants were able to discriminate between the phrase-final prominence marked by increased duration and phrase initial prominence marked cross-linguistically by increased pitch and intensity. Whether infants link these prosodic cues to the syntactic properties has not been proved. What has been observed is that infants are sensitive to the prominence within phonological phrases which are correlated with word order. If the head-

---

3 Phonological phrases boundaries are signalled by square brackets. Sentences in French and Turkish are taken from Cristophe et al. (2003).
direction parameter were set prelexically, when listening to a whole sentence the child at around 12 months would be able to grasp more of its meaning than just the few words that she understands.

In sum, prosodic bootstrapping theories posit that the speech signal contains cues to the lexical, morphological and syntactic structure of the target language and that the infant is able to perceive these cues. We have seen how the information contained in the signal is correlated to other levels of linguistic organization and how these correlations are true for all those languages that have been studied so far. Undoubtedly infants interact with language from birth and, according to the studies we have discussed, they possess a specialized mechanism for processing speech that conduces the infant from a general ability to discern all prosodic-phonetic contrasts to a language specific ability that will ultimately allow them to acquire the maternal language.
3.3 Statistical learning: Transitional Probabilities and Rule Learning

3.3.1 The Segmentation of speech by infants using transitional probabilities

As already mentioned, in order to learn a language we need to discover its lexicon by segmenting the linguistic flow in words, and we also need to learn its grammar. According to the bootstrapping theories the infant is able to pick up those cues which are present in the speech signal and which in turn trigger some specific mechanisms at different levels of the linguistic structure.

According to another line of research, the speech signal is informative in another way: statistically.

The task of segmentation, necessary for word learning, would be helped by the statistical relations between syllables. In particular, it has been argued that a certain statistical relation would be especially helpful. By listening to an unsegmented flow of speech, the learner would be able to compute statistically what is more likely to follow one sound and thus calculate the transitional probabilities (TPs) between syllables. Indeed, the transitional probability of B given A is the probability that, if there is A, there is B. For example if we listen to the sentence ‘thevaseisgreen’ we can see how the TP between, for example, the syllables the-va or va-se is 1. However, when other sentences are added, such as ‘thesmallvasefell’ or ‘bringmeavasewithwater’, the TP between va-se remains 1 with the addition of the second sentence and again stays to 1 with addition of the third sentence, while the TP between the-va dips to 0.5 with the addition of the second sentence (va occurs once out of 2 occasions after the) and remains so when the third sentence comes in (va occurs once after the, in the three sentences, which have two occurrences of the). By learning the statistical regularities that are found in recurrent patterns, infants (and adult learners) would be able to find word boundaries.

The study by Saffran, Aslin and Newport (Saffran, Aslin, & Newport, 1996a) had the aim of testing the hypothesis we have just discussed. The researchers, using the head turn preference procedure, familiarized 8-month-
olds with a flow of four three-syllable nonsense "words" (from now on words). These were chained and repeated in a random order, and had no pauses or prosodic information, creating a monotonous synthesized continuous flow like bidakupadotigolabubidaku... The trisyllabic sequences were "words" in this artificial language because, due to the way they were chained together, the TP between their first and second element, and their second and third element, remained high (TP 1). After familiarization, infants then heard trisyllabic strings, some of which were words taken from the familiarization stream (for example padoti), while others were non-words constructed by taking three syllables heard during familiarization, but never appearing in the same order (for example padago). Infants showed increased attention to non-words with respect to words, suggesting that for them non-words were novel items and words were familiar ones. This is a sign that infants learned words during the familiarization period. Thus babies, using the TPs as the only cues to word segmentation, were able to parse the continuous flow using a dip in transition probability as a cue to word boundaries.

Those scientists that believe in this proposal also argue that statistical learning is a domain-general ability that can be applied also to other stimuli besides the linguistic one (Fiser & Aslin, 2002) and that it is not specific to humans. Indeed, tamarin monkeys (Hauser, Newport, & Aslin, 2001) and rats (Toro & Trobalòn, 2005) are able to use the same statistical information.

Infants have a great ability at paying attention to the language that surrounds them and, according to Saffran and colleagues, they are able to make statistical computations to segment the speech flow in its components. However, while the role of statistical learning is quite well defined for what concerns word segmentation, it is still not clear whether this ability allows learners to make generalizations and hence to attain rules. We will discuss these issues in the sections that follow.

3.3.2 Statistical computations in adults

The ability to exploit the statistical information contained in the signal is not age-specific. Indeed, computational abilities have been observed in adults as
well as in infants. In this section I will discuss a study which has been conducted with adult subjects, but that can be interesting in understanding how computational mechanisms work and what the implications for language acquisition might be.

Peña and colleagues (Peña, Bonatti, Nespor, & Mehler, 2002) showed that adults use statistical regularities to segment speech, but that they are not able to extract generalization rules defined over nonadjacent syllables (such as ‘if A occurs, then C occurs’ in an ABC item) if the signal is unsegmented. This is important because the ability to detect such long-distance relations is necessary to acquire those morphosyntactic regularities that imply a dependency between particular items (as it happens, for example, in gender and number agreements).

The researchers created a language with an artificial grammar whereby the highest transitional probability of 1 was non-adjacent: every “word” has an ‘AxC’ structure so that syllable A always predicts syllable C in trisyllabic (ABC) strings. For example, three words were PUraKI, PUfoKI, PUiLiKI. The transitional probability between A and the adjacent x is 0.33 and between x and the adjacent C is also 0.33. The transitional probability between the last syllable of a ‘word’ and the first syllable of the following word is 0.5.

In the first experiment, the subjects heard the continuous flow: PURAkiBeligaTAFODupufokiTALIDUberaga....(alternation between small and capital letters is used to put in evidence the “words” and underlining is used to show “part words”). After 10 minutes of listening, the subjects were asked to compare a word to a part word and say which resembled more to the language they had heard. The adult subjects showed a preference to words than to part words.

In another experiment a new group of subjects was familiarized to the same continuous string and in the test trial they were asked to compare a part word taken from the sequence (for example RAKIBE) and a novel word which did not appear in the stream but had the same AxC’structure of the words heard in the listening period: the ‘rule-word’ (for example PUbeKI, PUgaKI). This time the subjects preferred the part word to the rule word and increasing the familiarization period to 30 minutes did not help subjects recognize the rule-words.
In another experiment, very short pauses (25ms gaps of silence) were added to the flow so as to segment the flow into three-syllable words. The pauses were in the signal, but were not consciously perceived by the subjects: they remained subliminal. After a very short familiarization period of only 2 minutes, a new group of adult subjects was able to extract the structural regularity and preferred rule-words to part words.

Peña and colleagues inferred that while statistical computations are used to segment speech, another kind of computation is needed to pick up non-adjacent regularities and thus acquire the grammatical structure of a language. The ability to extract such patterns depends on the difference in the signal and not on the amount of time needed to consolidate the process, indeed prolonging familiarization from 10 to 30 minutes did not change the outcome of the experiment. Inserting silent gaps to the stream, instead, allowed the subjects to discover the structural regularities. Natural language has a prosodic structure, and in the absence of such cues as rhythm and intonation, the subjects relied on pauses that made the artificial language slightly similar to natural language.

3.3.3 The acquisition of linguistic rules in infants

From the results that we have discussed so far, it seems that statistical learning can be a useful mechanism that infants may use for word learning. The mechanism would allow the child to recognize the information about the frequency distribution of items present in the speech signal and would thus help her in the segmentation task. However, the mechanism cannot make generalizations for items that have never been encountered before and can only work with long segments of speech as inputs, long enough to render the distribution representative.

To learn certain aspects of grammar, however, the infant needs to find rules and make the correct generalizations. As we said before she needs to go from a finite input to an infinite set of sentences; given the properties of statistical learning it seems that this mechanism cannot be used to such end.

Marcus, Vijayan, Bandi Rao and Vishton (1999) showed that statistical learning alone cannot account for extracting and generalizing rules, and that
other learning mechanisms are available to the child during language acquisition.

The researchers tested 7-month-old infants with the head turn preference procedure, using strings of an artificial grammar having an ABB structure or an ABA structure. Infants were divided in two groups. One group was familiarized with a sequence of 16 three-word sentences that had an ABA structure (‘ga ti ga’ ‘li na li’). The other group received an ABB grammar (‘ga ti ti’ ‘li na na’). During the trial phase, infants of both groups listened to sentences which contained words made up of syllables that had not been presented in habituation. Half of the sentences were consistent with the ABA or ABB structure heard during familiarization, and the other half was inconsistent because it was constructed by using the other grammar. All stimuli in habituation and test phase were synthesized to eliminate any prosodic information.

The results indicated that infants looked longer to sentences with the grammar that was not presented during familiarization. A longer looking preference is an indication that infants found "odd" those sentences that did not comply to the structural description imposed by the familiarization grammar.

Two more experiments were run to rule out certain possible interpretations of the results. The second experiment chose the stimuli more carefully, so as to exclude the possibility that the infants could rely on certain phonetic aspects instead of abstract structures. The third experiment aimed at eliminating the possibility that infants could look for adjacent repetitions to discriminate between the two sets of stimuli. Thus, stimuli were presented with two types of grammars containing both adjacent repetitions, but in the opposite order -- ABB vs. AAB. The results for both control experiments confirmed longer looking time for inconsistent sentences, just as in the first experiment.
Figure 25. The diagram shows the results of the three experiments conducted by Marcus et al. (1999). The blue bars denote looking times to stimuli consistent with the grammar structure presented during habituation. The red bars show the looking times of the infants when presented with an item inconsistent with those heard during habituation.

Marcus and colleagues came to the conclusion that the infants had learned the rule, and that the mechanism that allowed them to do so could not be based on transitional probabilities because none of the words in the habituation phase was present during the trial session. The type of generalization that the child was able to learn closely resembles the generalization made by the adult subjects of Peña et al. (2002), an AxC algebra-like rule, where the presence of A predicts the outcome of C, a rule which is useful to learn and generalize certain grammatical and morphological regularities found in languages.

In a more recent study (Marcus, Fernandes, & Johnson, 2007), Marcus revealed how infants can perceive the structural regularities of non-linguistic stimuli, if they first hear those regularities in sequences of speech.

The procedure was the same used in the previously discussed study, only the stimuli consisted in a sequence of tones or sung syllables (experiment 1), and timbres and animal sounds (experiment 2). The results showed that 7-month-old infants did not show a preference for either structure using timbres, tones or
animal sounds, the only discrimination was shown for sung syllables.

In a third experiment, the researchers familiarized three different groups of infants to a sequence of sentences with an ABB or ABA grammar and for half of the trials in the test phase they reproduced the same sequence using either tones, timbres or animal sounds, while for the other half they used the other grammar. This time the infants showed a preference for those sequences of tones, timbres and animal sounds that were inconsistent with the speech pattern used for habituation.

![Graph showing looking times for experiments 1, 2, and 3.](image)

**Figure 26.** Results of the three experiments of Marcus et al. 2007. Black bars indicate mean looking times to inconsistent items after an exposure to speech stimuli. Light grey bars indicate mean looking times to stimuli consistent with items heard in familiarization. White bars show looking times to inconsistent items but with no exposure to speech stimuli.

It seems that infants can extract symbolic abstract rules in linguistic stimuli with more ease than in non-linguistic ones. In this series of studies, speech has been shown to channel infants into the grammar, and this allowed them to transfer their knowledge to other domains.

According to Marcus, in order to extract the rules contained in the artificial grammar presented to them, infants need to have a mental algebra, an unconscious and innate mechanism that tells the infant how the variables of the grammatical system can be manipulated and generalized.
3.3.4 Using the Head Turn Preference Procedure to clarify how infants might find words and rules in a speech stream.

In this section I am going to describe a series of experiments that have been run recently at the Language, Cognition and Development Laboratory at SISSA, Trieste. The purpose of this section is to show how a question related to language and rule learning, which raises several theoretical issues raised in this chapter and in Chapter 1, ties together with the methodological considerations raised in Chapter 3. Having direct knowledge of this procedure as an experimenter, I will try to convey the relation between actual practice and theoretical issues in a well-documented case.

The explanation of this study will be quite detailed. Not only will I describe the procedure of the technique we are at present revising, but I will give a brief account of the theoretical background that is behind the hypothesis that has been advanced by the researchers. We will also see how the results of an experiment are never univocal (particularly when studying infants) and how many possible interpretations have to be ruled out by other control experiments.

Marchetto and Bonatti (2009) used the Head Turn Preference Procedure to test hypothesis on the different mechanisms that infants might use in identifying words and extracting syntactical relations from a speech stream.

In Section 3.3.2, I described the experiment that was run with adults by Peña and colleagues (Peña et al., 2002). The study suggested that adults possess two computational abilities and that can exploit the information present in the speech signal. One of these mechanisms tracks statistical relations among syllables to find words in a continuous speech stream, while the other mechanism can extract generalization such as rules with an ‘A then C’ structure, but only if the signal is segmented. Finding the relations among nonadjacent items in words is potentially important, because it is a necessary prerequisite for acquiring the morphosyntactic regularities of languages.

In the study Marchetto and Bonatti hypothesized that this dual mechanism is also activated in language acquisition procedures in infants.

The researchers created an artificial language with trisyllabic consonant-vowel strings (hereafter words). The words had an internal AxC structure
whereby syllable A always foretold the outcome of syllable C with a transitional probability (TP) of 1 between nonadjacent syllables. There were two word couples: one couple had /ba/ as syllable A and /so/ as syllable C, and one couple had /li/ as syllable A and /fe/ as syllable C. The middle syllable had little variability: it could either be syllable /mu/ or /ga/ reaching a TP of 0.5 between adjacent syllables. The four words of the artificial language were therefore: /bamuso/ /bagaso/ /limufe/ /ligafe/.

The subjects of this series of experiments were 7 and 12-month-olds.

In experiment 1A, 12-month-olds were tested individually in a three-sided booth in a dimly lit room. The baby sat on the parent's lap facing, at approximately 80 cm distance, a monitor. The booth was covered with black curtains that only allowed to be visible the three monitors placed in front and sides of the infant, and the lens of a video camera above the central monitor. Thus, Marchetto and Bonatti tried to reproduce a Head Turn Preference Procedure setting completely computerized, in which monitors replaced the "physical" pulsating lights typical of the standard procedure. While this choice allows the experimenter to better control the visual stimuli, finding the right stimuli, eliciting the right level of interest in babies -- not too little, not too much -- require a significant amount of piloting. Finally, the authors decided to use a novel attractor instead of the most common lights: a recorded movie of a hand making a circular slow motion, presented either on the centre or on the side monitors according to the needs of the experiment design. This change shows how any novel modification of a technique requires a prior work of calibration on which the success of the experiment may depend.

The caretaker wore headphones playing music and was instructed to not interfere with the baby during the experiment, he/she should not speak to the infant or point or orient towards any of the screens but only hold the baby and eventually reorient her towards the centre should the baby move persistently.

During familiarization, the visual stimulus was shown in the central screen while the infant listened to the artificial speech stream. The sequence was made up of the four words mentioned earlier: /bamuso/bagaso/limufe/ligafe/ and was synthesized to obtain a flat prosody. The words were separated by a pause of 200ms and were repeated in random order but never appeared twice in
a row. The infant listened to the familiarization stream for 2 minutes and 52 seconds. During this phase the experimenter (who is not visible to the child) simply watched the infant through a monitor and eventually stopped the experiment if the baby cried or became very fussy. No other intervention was needed on the part of the experimenter.

Immediately after familiarization the infant was presented with 16 test trials in random order.

Half of the trials presented the infant with novel instances of “rule words”: /baliso/, /bafeso/, /libafe/, /lisofe/ (bold font is used to put in evidence the AxC rule), while half of the trials consisted of non-words: /sogali/, /femuba/, /bafemu/, /lisoga/; the first two words had a CBA’ structure where C and B were respectively the third and second syllable of one of the two families while A’ always belonged to the other family and the other two non-words had an A’CB structure (Table 1). Non-words had a TP of 0, for both adjacent and nonadjacent syllables while the novel rule words had 0 frequency for adjacent syllable but a TP of 1 for nonadjacent syllables. Each one of the four rule words and each one of the four non-words was presented in two trials. Thus, the material patterned as closely as possible the stimuli presented by Peña et al. (2002), but reducing families from three to two and family member variations from three to two. Even for this aspect, the moral is that in order to study the same phenomenon in adults and infants, any experiment requires a significant amount of adjustments.

In each trial, to attract the infants’ attention in midline position, the movie of a flashing light was displayed in the central monitor. After the participant had fixated continuously at it for 1.5 seconds, the experimenter pressed a key in an Apple G5 computer that ran the experiment using PsyScope X (http://psy.ck.sissa.it). The computer gave input to one of the two side monitors that displayed the same visual attractor that was played during familiarization (the moving hand). As infants made a clear head turn towards it, the experimenter pressed another key that started playing from the loudspeakers concealed behind the monitors, a repeated sequence of either a rule word or non-word. The sequence together with the visual attractor continued to play until the infant turned away from the monitor for at least 2 seconds or ended after 65
seconds of continuous looking.

At the end of each trial, the flashing light attracted again the infant towards the centre and, after a 1.5 seconds fixation the computer played another trial. The same procedure was repeated for 16 trials. The camera recorded the looking behaviour of the infant and the researchers then coded looking times offline.

The data collected showed that infants looked longer while listening to non-words than to rule words, suggesting that, as in the study by Peña with adult subjects, infants when familiarized to a segmented stream can extract generalizations.

With infant subjects, however, the interpretation of results is always problematic. While it was clear that infants preferred to listen to non-words, there could be two alternative reasons to account for this preference. One was the reason hypothesized by the researchers i.e. that infants were surprised by the absence of nonadjacent relations in the non-words. The other reason could have been that infants noticed that the syllables heard in familiarization were mismatched in some of the items heard during the test phase. According to this possible interpretation, in the familiarization stage infants would only have extracted words from the speech stream, but never extracted the structural rule that governed them, and were only surprised by the fact that some stimuli in the test phase did not look like the words in familiarization.

The material used in this experiment can help us to exclude partially this alternative explanation. Indeed two of the non-words (/bafemu/ and /lisoga/) matched the first syllable position of the words heard in familiarization, so if infants were looking for partial matches in the test items, those trials that presented these words should have shorter looking times. Likewise, if infants were monitoring the middle syllable they should have found a match in the two other non-words /sogali/ and /femuba/. Indeed, the researchers analyzed if there was any difference between those trials that displayed a match in the first or second syllable and those that did not, and there was no hint of it. The only possible doubt existed regarding the position of the last syllable, which had no match in any of the non-words. This possibility was tested in the control Experiment 1B.
In this experiment, the familiarization stream and the rule words were the same as in Experiment 1A, while two of the non-words were replaced so as to be identical to the words in familiarization in their last syllable.

As in the first experiment, infants looked longer to non-words than to rule words therefore excluding the alternative interpretation that arose from the first experiment.

The results, however, needed yet another confirmation. It could be that infants in Experiments 1A and 1B paid attention to some phonetic or phonological pattern used in the stimuli. To exclude this other possible explanation, the researchers ran Experiment 1C. They created a new familiarization stream, new novel rule-words and used as non-words the rule words that were used in test trials in Experiment 1A and 1B. Table 1 summarizes the stimuli used in Experiments 1A, 1B, and Table 2 presents the stimuli used in Experiment 1C.
### Table 1. The stimuli used to construct the experimental material by Marchetto and Bonatti in Experiments 1A and 1B (from Marchetto & Bonatti, 2009)

<table>
<thead>
<tr>
<th>WORDS</th>
<th>RULE-WORDS</th>
<th>NON-WORDS</th>
<th>Exp. 1A</th>
<th>Exp.1B</th>
</tr>
</thead>
<tbody>
<tr>
<td>/bamuso/</td>
<td>/baliso/</td>
<td>/sogali/</td>
<td>/sogali/</td>
<td></td>
</tr>
<tr>
<td>/bagaso/</td>
<td>/bafeso/</td>
<td>/femuba/</td>
<td>/femuba/</td>
<td></td>
</tr>
<tr>
<td>/limufe/</td>
<td>/libafe/</td>
<td>/bafemu/</td>
<td>/mubafe/</td>
<td></td>
</tr>
<tr>
<td>/ligafe/</td>
<td>/lisofe/</td>
<td>/lisoga/</td>
<td>/galiso/</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2. The stimuli used to construct the experimental material by Marchetto and Bonatti in Experiment 1C

<table>
<thead>
<tr>
<th>WORDS</th>
<th>RULE-WORDS</th>
<th>NON-WORDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>/feliga/</td>
<td>/fesoga/</td>
<td>/baliso/</td>
</tr>
<tr>
<td>/febaga/</td>
<td>/femuga/</td>
<td>/bafeso/</td>
</tr>
<tr>
<td>/solimu/</td>
<td>/sogamu/</td>
<td>/libafe/</td>
</tr>
<tr>
<td>/sobamu/</td>
<td>/sofemu/</td>
<td>/lisofe/</td>
</tr>
</tbody>
</table>
Also in this experiment, 12-month-olds looked longer at non-words than to rule words, changing the stimuli therefore had no effect on the results that had been obtained to that point.

The study by Peña et al. (2002) suggested a further conclusion. It also showed that segmentation is crucial for rule learning: adults were unable to extract structural regularities when the speech stream was unsegmented. Thus Marchetto and Bonatti tested this prediction with 12-month-old infants in Experiment 2. They exposed infants to the same familiarization stream used in experiments 1A and 1B, with the difference that there were no pauses between the items. Thus the familiarization phase was slightly shorter, lasting 2 minutes and 14 seconds. The procedure and the test stimuli were otherwise identical to the ones described in the first experiment.

Because this time there was no difference in looking times between rule words and non-words, the researchers inferred that infants like adults need to be exposed to a segmented stream in order to capture structural rules.

Yet, another control was needed. Perhaps infants failed to find rules in the continuous speech stream of Experiment 2 because they could not extract any
information from the stream -- whether rule-information, or information of any other kind. To exclude this possibility, and show that infants are able to compute transitional probabilities to extract words from a continuous speech stream, Experiment 3 was run. Infants were familiarized to the same stream used in Experiment 2 and were then tested with words and non-words identical to those used in experiment 1B. In words the adjacent TP was 0.5, while in non-words it was 0.

Twelve-month-old infants looked longer to non-words than to words confirming the hypothesis that was formulated by the researchers: infants at this age are able to generalize abstract rules if the stream is segmented (Experiment 1); yet if the signal is continuous they can find lexical items (Experiment 3), but not capture generalizations (Experiment 2).

These results provided stronger evidence to the argument of the existence of two distinct mechanisms governing statistical computations and generalization. The specific linguistic data analyzers in the brain would be triggered by particular properties of the signal.

![Figure 28. Results of Experiments 1 (summary of 1A-C), 2 and 3 with 12-month-old babies (from Marchetto & Bonatti, 2009). Asterisks indicate significant differences as detected by the Head Turn Procedure implemented by the researchers.](image-url)
The researchers then advanced the hypothesis that the mechanism that is able to extract the within-word structure develops after the infant has solved the problem of extracting lexicon from the speech signal. Thus, the ability would not be present in earlier stages of development. To test this hypothesis, Marchetto and Bonatti ran experiments with 7-month-olds.

As recalled in Section 3.3.3, researchers have explored the computational abilities of infants at this age. Marcus et al. (1999) had investigated the ability of 7-month-olds to find structural regularities in discrete units, showing that at this age infants are able to extract strict identities with words that have identical repetitions (see Section 3.3.3 for a discussion), while Saffran et al. (1996c) had found that 8-month-old infants are able to detect adjacent TPs in a continuous stream (see Section 3.3.1 for a discussion). The intent of Marchetto and Bonatti was now to verify whether 7-month-olds can extract structural regularities of the AxC type.

Experiment 4 was a repetition of Experiment 1B, but this time subjects were 7-month-old infants. The data collected showed that the infants displayed no difference in looking times between rule words and non-words. The results indicated that at this age infants do not possess yet the mechanism that allows them to find within-words structural relations.

But again, there were alternative interpretations. An alternative to this interpretation could have been that the infants would not respond at all to this experimental design. To test this possibility, the researchers ran Experiment 5. In this experiment the familiarization stream was the same as in the previous experiment, but the test items were words and non-words. Thus, if infants have a specific difficulty to extract rules, but not to extract other type of information such as words from a segmented stream, they should now succeed. Infants showed a difference in looking times between words and non-words showing a preference for the latter.

Which information in the signal did the infants use to identify words? Two possible cues could have been exploited: nonadjacent syllable TPs or segmentation marks between words. To test which of these cues the infants had used Experiment 6 was run. This time the familiarization stream had no pauses
at word boundaries, but the test items were the same as in Experiment 5 (words and non-words). The infants showed no difference in looking times between the test items. Thus, infants were not able to identify the nonadjacent TPs, and this result suggests that they had been successful in the previous experiment because they were helped by the segmentation marks in familiarization.

Figure 29. Results of Experiments 4, 5 and 6 with 7-month-old babies (from Marchetto & Bonatti, 2009). Asterisks indicate significant differences. Infants succeeded only to find words from a segmented stream, not rules nor words from a continuous stream.

The result is not incompatible with previous findings. Saffran and colleagues (Saffran et al., 1996a) had indeed proved the computational abilities of this age range but the sensitivity was evident over adjacent TPs, while in the present study the task was that of identifying nonadjacent statistics.

The purpose of this long discussion was to show that the details of the procedure, the kinds of items, the exact age tested and the methods selected by the researchers all count to elaborate a clearer picture of early language abilities. When looked at this level of details, the results suggest that infants are able to find words and rules in a speech stream using a dual mechanism that becomes available to them at different ages. At seven months infants are able to detect
relations involving identities (Marcus et al., 1999), or can compute statistics over adjacent elements (Saffran et al., 1996a), but cannot compute more general rules that may be involved in morphological processes, nor can they compute more complex statistics such as nonadjacent relations between syllables. Few months later, at 12 months, infants are able to extract words by capturing the associations among nonadjacent elements and then reveal abstract generalizations to build the morphological structure together with lexicon. These mechanisms are not only age-related but are also triggered by specific signal properties.

Thus, both an ability to learn by experience (i.e., statistical computations) as well as an ability to "grasp" structural relations are both necessary ingredients of early language acquisition. The hope to get rid of one of them and rely uniquely on the other -- an ability that has inflamed most recent literature -- is not grounded in the data available.
4 CONCLUSION
This thesis has begun by asserting the importance of studying the initial state of language acquisition. The theoretical framework behind nativist theories, which reject the possibility that the brain at birth be a *tabula rasa*, has brought us to explore the methodologies that allow us to study the cognitive and linguistic abilities of infants before the onset of language production.

The abilities that infants have in processing speech must not be viewed as the result of the capacity of the human brain to learn anything. Language acquisition is known to be independent from general intelligence. One of the issues associated with this theory is that the brain has a modular structure and that specific modules process language. The ability to acquire language would therefore be part of our genetic endowment and the presence at birth of specific brain tissue to process language would provide evidence for these hypotheses. To this end, we have recalled several lines of research, and in particular recent experiments using optical topography, suggesting that at birth the left hemisphere already processes speech while disregarding other acoustic signals such as backward speech.

We have also discussed the thesis that because children receive no information about the rules that generate the linguistic input, a part of this knowledge and the representations it requires must be innate.

The assertion behind the ‘poverty of stimulus’ argument, or the claims supported by Gold’s theorem, are not in contradiction with the fact that the input that infants receive is indeed rich in cues. We have argued that the speech signal is rich in prosodic or phonologic information. Through the aid of behavioural methods, researchers have discovered that newborns and infants are able to perceive these cues. Typological studies have demonstrated that these prosodic (non-structural) characteristics are correlated with structural properties. The hypothesis that has been advanced is that infants are able to rely on these cues to solve the segmentation task and even to set certain parameters.

We have also seen that the speech signal can be rich in statistical information. Again, behavioural experiments using artificial grammar have shown that infants are able to pick up adjacent regularities (TPs) in unsegmented speech streams to extract words, while they are able to find nonadjacent dependencies when the signal is segmented.
The fact that the input is rich in prosodic and statistical cues does not exclude the existence of mechanisms that are able to pick up these cues and correlate them with structural properties.

A large part of this thesis has been dedicated to the study of the behavioural methods that allow researchers to get closer in understanding how innate mechanisms and input interact. To this purpose, I have described the most important procedures that are used today in laboratories studying the cognitive and speech perception abilities of newborns and infants.

Pre-verbal infants have a universal language that researchers can rely on to obtain answers to their quests: their behaviour. Although newborns and infants can only control a relatively small number of behaviours, such as sucking a pacifier, looking towards an object or turning their heads, researchers are able to exploit these dependent variables under controlled laboratory conditions.

Many of the important results that have been obtained could not have been possible without these experimental techniques. Working with infants opens up exciting possibilities, offering researchers a window into the human mind. The progress is only limited by their fantasy and ingenuity.
5 REFERENCES


Peña, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F. et al. (2003). Sounds and silence: an optical topography study of language...


Ringraziamenti

Ringrazio la professoressa Giusti, per la sua disponibilità e curiosità scientifica che hanno dato le basi per la stesura di questa tesi.

Ringrazio il professor Mehler per il suo incoraggiamento e disponibilità, e tutti i suoi studenti che si sono resi sempre disponibili a darmi una mano, grazie quindi a Agnes Kovacs, Erno Teglas, Erika Marchetto, Jean Remy Hochmann, Silvia Benavides e David Gomez. Grazie anche ad Alessio per l’aiuto tecnico.

Ringrazio Marijana per il suo supporto e per avermi aiutata a trovare del prezioso materiale bibliografico.

Ringrazio la mia famiglia, in particolare la mia mamma per l’aiuto con i bambini.

Ringrazio Silvio per avermi sempre incoraggiata, per il suo aiuto e la sua pazienza in questi mesi.

Grazie anche ai miei bambini: Giulia, Andrea e Paolo che hanno fatto ciò che potevano perché ci fosse un po’ di silenzio in casa in questo periodo.

Tutta la mia gratitudine va infine a Luca Bonatti per la sua guida scientifica, il suo sostegno e la sua infinita disponibilità. Senza di lui non sarei riuscita a finire ciò che avevo lasciato in sospeso per troppo tempo.
Il sottoscritto/a  FRANCESCA GANDOLFO
Matricola n.  742119
Facoltà  LINGUE E LETT. STRANIERE
iscritto al corso di laurea/diploma in  LINGUE E LETTERATURE STRANIERE
Titolo della tesi (*)  EARLY LANGUAGE ACQUISITION AND COGNITION: METHODS AND RESULTS

Dichiara che la sua Tesi è: Consultabile da subito.
Dichiara che la sua Tesi è: TOTALMENTE RIPRODUCIBILE.

Venezia,  17/12/2008
Firma dello Studente  Francesca Gando 1fo

Estratto della Tesi:  This thesis investigates the early stages of language acquisition from a nativist perspective and describes the most important behavioural methods that are used today to study the initial state of development. After discussing whether language is founded in human biology and whether linguistic representations have an innate basis, I describe thoroughly those procedures that allow to study pre-lexical infants and which can give researchers an insight on language acquisition and early cognition. The methods described are: The High Amplitude Sucking Technique, The Visual Fixation Technique, The Visual Expectation Technique and The Head Turn Paradigm. The thesis then presents some of the results that have been obtained in psycholinguistics with the application of the procedures mentioned above and describes in detail a series of experiments that have been run at The Language, Cognition and Development Lab at Sissa, Trieste.

(*) Il titolo deve essere quello definitivo uguale a quello che risulta stampato sulla copertina dell’elaborato consegnato al Presidente della Commissione di Laurea (*) Da inserire come ultima pagina della tesi. L’estratto non deve superare le miglia battute