Categories, Words and Rules

in Language Acquisition

Jean-Rémy Hochmann

Trieste, 2010

Supervisor: Jacques Mehler

Thesis submitted for the degree of Doctor of Philosophy in
Cognitive Neurosciences

SISSA, via Bonomea, 265 – 34136 Trieste - Italy

Acknowledgement

The content of this thesis owes a lot to the people that I have met during the last five years. I first need to thank my supervisor, Jacques Mehler, for allowing me to come to Trieste, mentoring me and advising me all along. I never came out of Jacques' office without having learned something. Thank you!

I owe a lot to all the people who passed through the Cognition, Language and Development laboratory. I learned from each of them. So thank you to Luca Bonatti, Judit Gervain, Erika Marchetto, Juan Manuel Toro, Ricardo Bion, David Gomez, Amanda Saksida and Marcela Pena. I need especially to thank Marina Nespor for advice and inspiration; Mohinish Shukla, for numerous enlightening discussions, and for the first Italian glass of wine he offered me the night of my arrival; Ansgar Endress for his advice and forcing me to write papers; Erno Teglas, for always being available to answer my questions and provide valuable advice; Agnes Kovacs for inventing the Kovacs paradigm; Silvia Benavides for triggering the consonant and vowel studies; Alan Langus his constant support; Alessio Isaja, for pretending that I understand every single detail of how a computer works, Francesca Gandolfo for providing so many babies, and Marijana Sjekloca for being the pillar of the lab and always ready to help.

I would like to thank my brother Thomas, my mum, my aunt and my dad, who have always been a very supportive family. Finally I thank Liuba Papeo who has tolerated my divagations about consonant, vowels and function words. I can now address other topics!

ABSTRACT

Acquiring language requires learning a set of words (i.e. the lexicon) and abstract rules that combine them to form sentences (i.e. syntax). In this thesis, we show that infants acquiring their mother tongue rely on different speech categories to extract: words and to abstract regularities. We address this issue with a study that investigates how young infants use consonants and vowels, showing that certain computations are tuned to one or the other of these speech categories. Using a Tobii eye-tracker, we developed a paradigm where 12- and 6-month-old infants learned to pair either specific words, or speech sound patterns to the locations where a toy appeared. We show that one-year-old infants rely mainly on consonants to remember words, but are better at generalizing structural relations implemented over vowels. However, we show that this partition of labor between consonants and vowels is not yet present by 6-months of age. Moreover, we propose that six-month-olds, in contrast to twelve-month-olds, may not represent the categories of consonants and vowels. The categories of consonants and vowels and their functional roles should thus emerge later. Consonants may be privileged for word recognition because they are more numerous and perceived categorically. Vowels, instead, carry prosody that marks more abstract constituents, and can thus provide the learner with information about syntactic regularities.

In the second part of this thesis, we study the acquisition of a particular class of words: function words. In contrast to content words (which must be linked to some semantic referents), function words (such as determiners and prepositions) mainly serve syntactic rather than semantic purposes. While the majority of research has focused on

possible phonological cues to explain function word acquisition, we investigate the use of one universal distributional property of function words: their high frequency of occurrence. We show that, after a short familiarization, seventeen-month-old infants prefer to associate the infrequent rather than the frequent part of a label to a novel object. Our results thus suggest a link between a distributional property (i.e. frequency of occurrence) and the functional distinction between content and function words. We propose that the formation of the class of frequent words represent the first step towards the acquisition of function words.

We conclude proposing that core linguistic representations allow infants to bootstrap into language, consenting them to link distributional properties and linguistic functions. Such core representations, however, are not adult-like and must be enriched by various domain-general and possibly language specific learning mechanisms.

Contents

Dieu, on connaît ses limites, ça ne va jamais très loin, mais avec les hommes, c’est illimité, ils sont capables de tout.

God, we know its limits, it never goes very far, but with men, it is unlimited, they are capable of anything.

La Danse de Gengis Cohn

R. Gary, 1967

There is no mode of action, no form of emotion, that we do not share with the lower animals. It is only by language that we rise above them, or above each other – by language, which is the parent, and not the child of thought.

The Critic as Artist

O. Wilde, 1891

CHAPTER 1

Introduction

The ironic quote of Genghis Cohn, Romain Gary's character, expresses what many authors have noticed as the specificity of the human mind. Only we are able to infinitely combine concepts and ideas, in order to create novel concepts and ideas (Fodor, 2002), thus being virtually *capable of anything*. This ability lies at the origin of all human culture, yielding scientific, technological and political progress, with its benefits and its misuses. The ability to compose concepts is intimately related to the language faculty. Allowing humans to produce an infinite number of different sentences (Fodor, 1975; Hauser, Chomsky & Fitch, 2002), language participates to the formation of novel concepts and to their transmissions through communication. Thus, studying the cognitive foundations of language, we may eventually characterize the foundation of Human nature. For the present thesis, our ambition is somewhat smaller. We explore the mechanisms and representations that may underlie the human-unique language faculty.

1.1 Language, a human-specific faculty

There is a large agreement among the scientific community that language is a human-specific ability. However, it is still debated, what exactly is unique to the human language faculty (Fitch, Hauser & Chomsky, 2005; Hauser, Chomsky & Fitch, 2002; Jackendoff & Pinker, 2005; Pinker & Jackendoff, 2005). In fact, many features of the language faculty once claimed to be uniquely human have later been found in other species, particularly in nonhuman primates or birds.

Essential to language is the creation and use of a system of symbols established through social convention, pairing specific signs with intended meanings (Saussure, 1916). The ability to create and use such a system is not unique to humans. The use of vocalization as reference to objects or concepts was observed with monkeys in the wild (Fischer, 1998; Gouzoules, Gouzoules, & Marler, 1984; Seyfarth, Cheney, & Marler, 1980; Zuberbuhler, Noe, & Seyfarth, 1997). Furthermore, chimpanzees can be taught to use and understand arbitrary symbols such as hand signs, colored plastic pieces or lexigrammes (i.e., symbolic pictures) to refer to objects and even abstract concepts (Gardner & Gardner, 1969; Premack, 1976; Rumbaugh, 1977; Terrace, 1979). Furthermore, Kaminski, Call and Fischer (2004) suggested that dogs are able to map words onto objects in a similar fashion as human infants (but see Markman & Abelev, 2004). Rico the dog was able to retrieve the toy asked for by his owner among eight or nine distractors. He was said to understand 200 such commands. Moreover Rico fetched the unfamiliar object when asked for a novel label, e.g. "Fetch the dax!". This behaviour is similar to the indirect word learning seen in mutual exclusivity experiments with eighteen-month-old infants. The processes of word learning may thus not be human-specific

Another particularity of human languages is to use phonemic representations, to code and decode speech productions. Thus, different realizations of the syllable “ta” will yield a common representation abstracted from the variations from one realization to another. This phenomenon, called categorical perception, was thought at some point to be uniquely human (Liberman, 1982; Liberman & Mattingly, 1985). However, similar phenomena were observed with animals for the perception of their own species’ vocal production. This was shown with birds (Kuhl, 1989; Nelson & Marler, 1989), monkeys (May, Moody & Stebbins, 1989; Snowdon, 1987), mice (Ehret & Haack, 1981) and even crickets (Wyttenbach, May & Hoy, 1996). Moreover, categorical perception for speech could be elicited in birds (Dent, Brittan-Powell, Dooling, & Pierce, 1997; Kluender, Diehl & Killeen, 1987), chinchilla (Kuhl & Miller, 1975) and macaques (Kuhl & Padden, 1982). Therefore, if something is special in speech to promote language learning, it may not be categorical perception.

Finally, the development of songbirds’ songs shows similarities with the acquisition of human language such as the existence of dialects and a critical period for acquisition (Gardner, Naef, & Nottebohm, 2005; Marler & Tamura, 1962; Nottebohm, 1993).

Hauser, Chomsky and Fitch (2002) proposed that the unique language-specific and human-specific abilities were to be found in the computational abilities of humans. The statistical computation of transitional probabilities between syllables, an ability thought to play an important role in language learning (see below) was evidenced in monkeys (Hauser, Newport & Aslin, 2001) and rats (Toro & Trobalon, 2005), and there is a current controversy on the ability of non-human species to learn abstract rules (Corballis, 2009; Mondragon, Murphy & Murphy, 2009; Murphy, Mondragon & Murphy, 2008).

One particularity of human language in comparison to other animal modes of communication is the exceptional number of concepts it can express. In fact, mature speakers of a language can understand an infinite number of sentences that have never been pronounced. A simple experiment can exemplify this property. Type the following sentence on the Google search engine "the small cat jumped into the river" (with quotes!). This sentence is pretty simple and should be understood by any English speakers. However, if we believe Google, it has never been written as such on a webpage! Of course many spoken sentences have never been written, but this result exemplifies the infinite expressivity of the human language faculty. In fact, most of the sentences we pronounce every day may never have been uttered before.

The unbound expressivity of human languages comes from discrete infinity (Chomsky, 2000), the property by which a human being can utter an infinite number of sentences from the finite number of words in its lexicon. Hauser, Chomsky and Fitch (2002) proposed that the ability to compute and produce recursive structures, yielding discrete infinity, might be the only component unique to the language faculty. Some experiments conducted by Fitch and Hauser (2004) were presented as a support for the recursion-only hypothesis, suggesting that both humans and tamarin monkeys could learn an $(AB)^n$ artificial grammar that was defined only by adjacent dependencies, whereas only humans could learn the A^nB^n artificial grammar, which required the recursive construction of center-embedded dependencies. However, in a recent paper, we showed that the experiment of Fitch and Hauser lacked crucial controls to support their conclusions (Hochmann, Azadpour & Mehler, 2008). In fact, neither humans nor monkeys learned any recursive rules in these experiments. Rather, they learned approximate linear descriptions of the patterns they were familiarized to, and some human participants learned to count the different types of constituents, a strategy that does not account for language acquisition (Musso et al., 2003). Thus, the debate is still open to understand, what exactly is specific to the human language faculty. To progress in this quest we may focus on acquisition, and ask the corollary question of what is unique to human language acquisition.

1.2 What is unique to language acquisition?

To acquire language, infants need both to learn words and to extract and generalize structural regularities that play a role in learning syntax. For instance, when hearing the utterance "the girl ignored the boy", we need more than just the meaning of each word to understand the whole sentence. The listener has to understand the relation between the verb and both the subject and the object. Learning words requires memorizing specific elements of the input (e.g., *girl*, *boy*) and representing them in a format that allows their recognition and distinction from other words (e.g. *boy* vs. *toy*), whereas learning regularities about syntactic structures implies the ability to extract abstract relations between elements of the input (e.g. whether the verb precedes or follows the object) and generalize them to new sentences. Moreover, syntactic structures are abstract in another sense, as they apply to syntactic categories like nouns and verbs rather than specific words like *girl* and *ignored*.

One dispute in cognitive sciences concerns the extent of prior knowledge that is necessary for infants to develop certain capacities, and thus whether there exists domain-specific learning mechanisms that profit from domain-specific prior knowledge. This dispute reaches its climax in the case of language acquisition.

1.2.1 Language acquisition through general mechanisms

1.2.1.1 Associationism and Connectionism

Opposed to the existence of innate learning mechanisms dedicated to language acquisition, an associationist or connectionist view, successor of an empiricist approach, claims that all the information is present in the input, and that language acquisition mainly results from applying domain general mechanisms, in particular statistical computations, to the speech input (Reddington & Chater, 1997; Seidenberg & McClelland, 1989).

Associationist or connectionist models of cognition are based on the idea that stimuli, which repeatedly occur simultaneously, will eventually elicit the same response. Locke (1690/2004) already considered that most knowledge originates from the sensations provided by the senses, and that learning results from complex associations of ideas. Concepts in this view would thus be decomposable to "simple ideas" such as sensori-motor representations.

The potential of associative learning was famously exemplified by Pavlov (1927), who showed that, when a bell systematically rings before dogs are fed, these eventually start salivating in response to the bell ring even in the absence of food. Such associative learning is called classical conditioning, where an initially neutral stimulus (e.g. the bell ring) gets associated to a meaningful stimulus (e.g. food).

Skinner (1938) proposed that all behaviors resulted from operant conditioning, that is the reinforcement of behaviors through their consequences. If a behavior yields a satisfactory outcome (usually food in animal studies), that behavior will be reinforced and tend to occur more frequently. If, on the other hand, a behavior has unpleasant outcome, it will tend to occur less frequently. Verbal behavior, according to Skinner, should be explainable by operant conditioning (Skinner, 1957).

Behavioralism was thus attractive because it claimed to explain all of human behaviorism on the basis of stimulus-response associations. Furthermore, such mechanism could be modelized at the neuronal level. Hebb (1949) proposed an algorithm that can account for associationist learning at the neural level. The fundamental rule of Hebbian learning is that what fires together, wires together. That is, the connection between neurons that are concurrently (or sequentially) activated is strengthened. Moreover, the activation of one neuron will automatically propagate to those neurons to which it is strongly connected. In the example of Pavlovian dogs, there naturally exists a causal connection between the activation of neurons that respond to food presentation, and those eliciting salivation. The repeated presentation of a bell ring before dogs are fed strengthens connections between the neurons that respond to the bell and those that respond to the food. Thus, when dogs later hear the bell ring, the activation of neurons responding to that ring propagates to the neurons originally responding to food, thus causing the activation of the neurons responsible for salivation.

Hebbian learning thus allow unsupervised learning of various associations. The properties and organization of the network regrouping all neurons can impose certain constraints to the connections between various types of neurons, in order to reflect experimental data. For example, it has been observed that rats easily associate light flashes with electro-shocks, and tastes with visceral sickness, but are unable to associate light flashes to visceral sickness and tastes to electro-shocks (Garcia et al., 1974). Such pattern may reflect the pre-cabling of certain neural populations together (e.g., neurons responding to taste and those responsible for visceral responses), and the complete absence of connections between other neural populations (e.g., visual areas and neurons responsible for visceral responses).

Most popular connectionist models are now based on another kind of models, artificial neural networks, which usually necessitate supervised learning. The most basic version of this model is the Perceptron (Rosenblatt, 1957, 1958). This artificial neural network is especially well designed for categorization tasks. It is composed of one input layer, and one output layer of neurons. The connections between each input and each output neurons are initially all equal, or randomly set. Learning in this model is done through the supervised modification of the connection weights between input and output, in order to force the correct output for each of a series of exemplar stimuli sequentially entered as input. Learning can then be attested by presenting some of the exemplars, or novel stimuli, asking whether the Perceptron correctly classifies them.

More complex (and potentially more powerful) artificial neural networks follow the same principles as the Perceptron, but are equipped with extra *hidden* layers, positioned intermediary to the input and output layers (Rumelhart, Hinton, & Williams, 1986; Rumelhart & McClelland, 1986), or copy layers that model the influence of sequentially presented stimuli as in the Simple Recurrent Network (Altmann, 2002; Christiansen & Curtin, 1999; Elman, 1990, 1993; Rodriguez, 2001; Rodriguez, Wiles & Elman, 1999; Rohde & Plaut, 1999).

1.2.1.2 The role of statistics in language acquisition

The relative success of artificial neural networks in grasping some properties of languages promoted research to verify whether pre-lexical infants could indeed behave as their computer models. The connectionist approach of language acquisition was in fact supported by research in the last fifteen years, showing how much statistical information is contained in the input, and can be accessed by infants. In particular, Saffran and colleagues (1996) showed that infants as young as 8-month-old could compute transitional probabilities between syllables, which would allow them to discover words in their input, and certain regularities in their organization. Saffran and Wilson (2003) further showed that words segmented on the basis of TPs could enter further statistic computations. They showed that 12-month-olds would acquire a simple grammar generated by finite-state automaton, where words are defined by high internal TPs. Gomez (2002) showed that eighteen-month-olds could learn the non-adjacent dependencies between the first and third elements of a three-element string, an ability that could be used to acquire certain morphological rules. Smith and Yu (2007) showed how 12- and 14-month-old infants could compute statistics on the co-occurrence of words and referents, to determine the meaning of novel labels.

Mintz (2003) proposed that distributional bootstrapping could allow adults and infants to form word categories. In fact, a relatively successful strategy to classify words into nouns and verbs is to classify them according to the frames in which they occur, i.e. the sets of words that precede and follow them (Mintz, 2003; Mintz, Newport & Bever, 2002). Moreover this strategy is not just an abstract possibility, but behavioral experiments suggest that at least human adults can use such a mechanism for classifying non-sense words into artificial categories (Mintz, 2002). Mintz (2006) showed that 12-month-olds can use frequent frames made of function words to categorize novel verbs but not novel nouns. Infants were introduced two novel words in typical noun frames (e.g. "I see the *gorp* in the room") and two novel words in typical verb frames (e.g. "She wants to *lonk* it"). Infants further listened to all the novel words in both noun and verb frames. They showed an increased looking time for novel verbs presented in a noun frame (e.g. “Here’s a *lonk* of a dog”), but not for novel nouns presented in a verb frame (e.g. “I *gorp* you now!”). This suggests that the tested infants understood that the novel verbs could appear only in certain frames, thus forming a category. The classification of nouns, however, was not yet effective with the used frames. With a similar method, Höhle, Weissenborn, Kiefer, Schulz & Schmitz (2004) found that German 14- to 16-month-olds but not 12- to 13-month-olds discriminated passages containing a novel word in noun contexts from passages containing that same word in verb contexts; this, if the novel word (e.g."glamm") was previously introduced to the child preceded by a determiner (e.g. "ein glamm"), but not if it was introduced preceded by a pronoun (e.g. "sie glamm"). They concluded that infants of this age can use a frequent determiner to classify a novel word as a noun, but the ability to use pronouns to categorize novel verbs might still be immature. The different results obtained show that these experiments are very dependent on the specific language used, and the specific stimuli. But beyond this variety, infants appear capable of using some function words to form word categories at the beginning of the second year of life.

Other authors insisted on the role of correlations between syntactic category membership and phonological properties such as stress pattern, syllabic structure and vowel reduction (Monaghan, Christiansen & Chater, 2005). Monaghan, Chater and Christiansen (2007) further showed that the combination of distributional cues comparable to the frames discussed above, and phonological cues constituted a good predictor of syntactic category membership. Their approach showed the plausibility of category formation on the basis of perceptual and distributional cues, exemplifying the actual richness of the input provided to infant learners. However, the actual useful phonological cues vary across languages. To give only one example, the number of syllables is useful to discriminate nouns from verbs in English but not in Dutch, French or Japanese (Monaghan, Chater & Christiansen, 2007). Infants must therefore first identify what perceptual cues play a role in their language. Even then, there remains a large overlap between different syntactic categories along any phonological dimensions (Monaghan, Christiansen & Chater, 2005). The combination of multiple perceptual and distributional cues appears therefore necessary for accurate categorization (Monaghan, Chater & Christiansen, 2007).

However, these experiments show only evidence of the formation of categories based on distributional and perceptual analysis, but they do not test for the labeling of those categories (Mintz, 2006). Indeed, they show that in certain frames, only certain words can appear, but the labeling of a category as 'nouns' or 'verbs' should come from other sources of information. In fact, labeling distributional categories represent a fundamental problem of language acquisition, for which solutions are likely to necessitate prior linguistic knowledge. For example, Pinker (1984) proposed that infants are endowed with the knowledge of the syntactic categories nouns and verbs, and the rules to map these categories on the world. We will further discuss this question in Chapter 3 and 4 of this thesis.

1.2.1.3 Structural generalizations

The research on statistical learning in infants reviewed above forwarded the idea that language could be fully acquired by computing statistics and recording dependencies between words or syllables. However, such computations can at most lead to approximate linear descriptions of natural grammars, similar to that provided by finite-state automatons (Gomez & Gerken, 1999; Saffran & Wilson, 2003). The hierarchical nature of natural grammars remains beyond the reach of statistical computations, however complex they might be (Chomsky, 1957; Marcus, 1998).

The unique productivity of human languages indeed requires more than simply encoding sequences of speech input. Learners must also be able to re-combine words in lawful manner, in order to produce novel, never heard sentences. Furthermore, mature speakers must link together sentences that have different surface organizations, such as affirmative sentences, e.g. "The man who is wearing a hat is tall", and corresponding questions, e.g. "Is the man who is wearing a hat tall?" (Chomsky, 1965). This suggests the existence of a hierarchical structure governing to the construction of syntactic sentences. To fill these requirements, learners need to represent abstract rules or abstract structures that guide the formation of sentences.

Thus, a second approach relevant to language acquisition proposed that infants could extract syntactic regularities by learning the rules that generate them. In a seminal work, Marcus and colleagues (1999) habituated 7-month-old infants to exemplars of one of the two structures ABB (instantiated by words like *pukiki*, *mesasa*, etc) or ABA (instantiated by words like *pukipu*, *mesame*, etc). After habituation, infants were tested with items corresponding to both structures. They could discriminate between novel exemplars of the habituation structure and items of the other structure, suggesting that they learned an abstract pattern rather than the specific items they were exposed to. Marcus and colleagues suggested that infants had acquired algebraic-rules manipulating symbols that can generate items of the habituation structure. With a similar paradigm, Gomez and Gerken (1999) showed that twelve-month-old infants could learn a simple artificial grammar generated by a finite-state automaton and generalize it to new sequences generated by the same grammar but using a different vocabulary. They suggested infants achieved this task by learning the pattern of repeated elements. Even though the precise mechanisms involved in these results are debated (see section 2.5.7), these studies show infants' abilities to extract abstract structural relations and generalize them to new tokens. This ability, like statistical computations, is not restricted to the speech domain, but occurs also with tones (Endress, Dehaene-Lambertz & Mehler, 2007) and with visual stimuli (see Appendix A; Saffran, Pollack, Seibel & Shkolnik, 2007; Tyrell, Stauffer & Snowman, 1991; Tyrell, Zingardo & Minard, 1993).

1.2.1.4 Constraints on domain-general mechanisms

Despite their domain-generality, neither of the two mechanisms mentioned above, statistic computations and structure generalization, applies automatically or independently of the nature of the input signal. For example, silent pauses favor the generalization of a pattern to novel stimuli. Peña, Bonatti, Nespor and Mehler (2002) habituated adult participants to a speech stream containing three families of tri-syllabic words defined by non-adjacent dependencies (e.g. puXki, where X could be one of three syllables). These families defined an artificial language. Peña and colleagues showed that adults would accept as part of the artificial language novel sequences respecting the non-adjacent dependency, but only if silent breaks were inserted around the familiarization words. Endress, Scholl and Mehler (2005) showed that adult participants could generalize a repetition structure only if the repetition was located at an edge of exemplar speech sequences, that is, followed by a silent pause.

Conversely, Shukla, Nespor and Mehler (2007) showed that natural prosodic breaks interfere with the use of statistical information for extracting words. They habituated adult participants with an artificial speech stream containing statistically defined words. The pitch of syllables varied following the prosodic contours of natural intonational phrases. Moreover, prosodic boundaries were also marked by duration variations, the initial syllable of each intonational phrase being shorter (100 ms), and the final syllable longer (140 ms) than internal syllables (120 ms). If, in a subsequent test, stimuli were presented acoustically, participants recognized statistically defined words if these occurred within a prosodic contour, but not if they straddled a prosodic boundary. However, participants could also recognize the words that straddle prosodic boundaries if test stimuli were presented in the visual modality. Shukla et al. (2007) thus proposed that statistics are computed automatically, but prosody acts like a filter, suppressing possible statistically defined words that span two prosodic constituents.

Altogether, these experiments suggest that prosodic properties of speech stimuli can trigger or hinder otherwise general mechanisms such as statistical computations and structural generalizations.

1.2.2 Language learning through dedicated mechanisms

1.2.2.1 Arguments for dedicated learning mechanisms

Lenneberg (1967) observed that the acquisition of language fully succeed at a normal path if, and only if learners are exposed to the linguistic input (usually speech) before puberty. He thus proposed the existence of a critical period during which the human-specific faculty of language would develop if it were triggered by appropriate stimuli. This feature is reminiscent of biologically constrained faculties in humans and other animals, such as binocular vision, which requires brain maturation but also light input to develop (Banks, Aslin & Letson, 1975; Hubel & Wiesel, 1970), and the development of bird songs (see Brainard & Doupe, 2002 for review). Thus language acquisition, at least in its structural aspects such as syntax and phonology, appears determined by biological factors.

A further argument for biological bases of the structure of languages came from Greenberg's (1963) observation that unrelated languages do not follow arbitrary rules, but are formed according to certain universal constraints, applying to syntax and morphology. Humboldt (1836) had already observed one century before that all languages are equal in their intellectual procedure, which "makes infinite use of finite means". The fact that unrelated populations developed languages such as Welsh, Italians, Swahili, Berbers, Hebrew, Quechua, Maori, Loritja, etc. that followed common rules (i.e. Greenberg's universals) suggest that common biological constraints may guide the formation of each of these linguistic systems.

Chomsky (1981; 2000) went further, advocating for language to be studied like an organ, as biologists would study the heart or the eye. This view entails that the formation of the language organ should be innately constrained, and does not fully depends from interaction with the environment. He particularly proposed there exists a Universal Grammar, constraining the possible forms of human languages. A corollary of this proposal is that infants should acquire language by the means of a dedicated innate Language Acquisition Device (LAD).

Chomsky based his claim on two important arguments. First, sentences are not linear but have a hierarchical structure. Second, the input lacks the appropriate evidence for infants to learn the hierarchical structure of sentences, the so-called Poverty of Stimulus argument.

The hierarchical nature of language is well illustrated by the following example:

(1)

a. The student is tired.

b Is the student tired?

c The student who is writing his thesis is tired.

d *Is the student who writing his thesis is tired?

e Is the student who is writing his thesis tired?

The problem is the following: given the two sentences 1.a and 1.b, at least two rules can explain the derivation of the question 1.b from the affirmation 1.a. One such rule states to select the first auxiliary and move it to the beginning of the sentence. A second rule, however, states to select the auxiliary of the *main clause* and move it to the beginning of the sentence. From the affirmative sentence 1.c, the first rule would derive the incorrect sentence 1.d, whereas the second rule would derive the correct sentence 1.e. If we assume, following Chomsky (1980) that at least some children never encounter complex examples of question formation like 1.c and 1.e (that is the poverty of the stimulus), so that the only examples they have are the ambiguous ones (1.a and 1.b), how can children know which rule they should generalize? To solve this problem, children must use the notion of *main clause*, or another equivalent notion, and thus know that language is hierarchical. Children appear to do so at least by the age of three (Crain, Nakayama, 1987). The fact that infants acquire linguistic knowledge for which they lack crucial evidence suggests that they are endowed with innate knowledge that guides their learning[1].

This problem is in fact related to the more general induction problem (Goodman, 1955; Hume, 1739/2003; Wittgenstein, 1953), pointing that, given a finite number of lawful exemplars, there are infinite possible generalizations. Goodman (1955) proposed the following example. Let us consider the new concept 'grue' that means "green until the time t, and blue afterwards". Considering all the emeralds we have ever seen fit both the definition of 'green' and 'grue', how do we know which concept correctly applies to them? The intuitive answer to this problem is that the concept 'grue' appears more complex to us, as requiring the notions of both 'green' and 'blue'. We would thus choose the *simpler* concept 'green' over the complex 'grue'. However, our rating of simplicity is dependent on our prior representations. If we were used to reason with the concepts 'grue' and 'bleen' (i.e. blue until the time t, and green afterwards), we would express 'green' as meaning "grue until time t, and bleen afterwards". 'Grue' would now be rated simpler than 'green'.

In fact, the simplicity of a concept, rule or structure cannot be evaluated independently of the organism that must acquire it. To rightly evaluate, for a given organism, the complexity of a representation, one must know the prior representations

[1] In a recent work, Perfors, Tenenbaum and Regier (submitted) proposed that the hierarchical phrase structure of languages does not need to be included in infants' innate knowledge about language. Rather, they propose that a Bayesian learning mechanism is able to determine that the natural child directed input is better described by phrase structured grammars than by linear grammars. This Bayesian learning mechanism, however requires that learners a priori entertain the ability to construct both phrase-structure and linear grammars. Thus, instead of impoverishing learners' innate abilities, Perfors et al. in fact enrich these with linear grammars in addition to phrase-structure grammars that Chomsky (1965, 1980) had already put into Universal Grammar.

it possesses, and the learning mechanisms it is endowed with. A representation will be easier to acquire if it fits the organism's learning mechanisms and existent representations. Conversely, if infants acquire language so effortlessly, it may be because they are endowed with constraints, representations and/or computational abilities that make it easy for them.

1.2.2.2 The Principles & Parameters theory

In the Principles & Parameters approach, Chomsky (1981; Chomsky & Lasnik, 1993) proposed that the LAD should consist in two types of knowledge. First, universal linguistic principles constrain the possible human languages. That is, human infants are endowed with the knowledge of what a language should look like, and certain of its properties. The properties resulting from principles should be observed universally and cross-linguistically. Second, the differences between languages can be understood as the setting of linguistic parameters. Parameters are seen as switches that should be put in one or another position according to the information extracted from speech input (Baker, 2001). It has been proposed that principles involve the definition of certain syntactic categories such as nouns and verbs (Pinker, 1984), or content and function words, and the hierarchical structure of languages (Chomsky, 1970). Parameters represent the properties that can vary from one language to another. For example, all languages have verbs and verb complements, but some like English or Italian place the complement after the verb, whereas others like Japanese or Basque place the complement before the verb. Learners must thus use language-specific information to identify the word order in their language.

The main problem of the Principles & Parameters theory is to explain how parameters are set. Various sources of information, including lexical, syntactic, prosodic and distributional information, have been proposed to inform infants on the correct value of a parameter in their language.

In particular, many researchers focused on the head-complement parameter, which accounts for the ordering of words and constituents in a language. In many, if not all, languages, the position of a sentence constituent is informative of its syntactic role. For example, in English, the subject tends to precede the verb, while complements follow it. Thus the sentence "a boy kisses a girl" indicates that the girl receives a mark of affection, and the boy performs that action. However, in other languages such as Japanese, the order of verbs and complements is reversed and the structure of the Japanese sentence would be subject-object-verb: "the boy the girl kisses". The last sentence would be ill-formed in English. Moreover, the order of verbs and complements strongly correlates with the order of other sentence constituents. In particular Verb-Object languages tend to use prepositions and have determiners preceding the nouns, whereas Object-Verb languages rather use post-positions and markers that follow the nouns. Thus, infants need to learn early the order of verbs and objects in their language to correctly parse sentences. Learning the order of verbs and complements in one's language corresponds to setting the head-complement parameter in the right position, understanding that one's language is head-initial (as English) or head-final (as Japanese).

Pinker (1984) proposed a lexical bootstrapping mechanism to achieve this task. Building on its knowledge of the existence of the lexical categories of nouns and verbs, and having learned a few members of these categories, infants could generalize the order they observe to set the parameter in the right position. For example, knowing that the word “eat” is a verb and that the word “cookie” is a noun referring to an eatable object, infants could generalize that English is a Verb-Object language, i.e. a head-initial language[2].

Others proposed pre-lexical mechanisms for setting the same parameter. These distributional, phonological or prosodic bootstrapping approaches assume the existence of a bond between perceptual properties of speech, and the values of linguistic parameters. One such proposition was recently put forward by Gervain and colleagues (Gervain et al., 2008). Observing that the position of frequent and infrequent words (i.e. respectively function and content words) in a language strongly correlates with the head-complement order, Gervain and colleagues proposed that pre-lexical infants as young as 7-month-olds may profit from their ability to compute relative frequencies of occurrence to set the head-complement parameter. Mazuka (1996) proposed that prosodic variations could be responsible for the setting of the very same parameter. She showed that prosodic boundaries were more pronounced in right-branching clauses than in left-branching clauses. Subordinate clauses being branched to the left in head-final languages, and to the right in head-initial languages, Mazuka (1996) proposed that the observation of clause boundaries might thus inform infants whether their language is right-branching or left-branching. Nespor and colleagues (2008) observed that complements are marked mainly by pitch prominence in head final languages such as Turkish, and mainly by duration prominence in head initial languages such as French (Nespor et al., 2008). Moreover, in languages such as

[2] In contrast with the lexical bootstrapping approach, Gleitman and colleagues (Gillette et al., 1999; Gleitman, 1994; Gleitman & Gleitman, 1992; Landau & Gleitman, 1985) proposed that language acquisition relies on syntactic bootstrapping mechanisms, where syntactic knowledge boosts the acquisition of novel words. We will further discuss these two views in Chapter 3.

German that allows both verb-object and object-verb constructions, the complement-head structures are mainly marked by an initial pitch prominence, while the head-complement structures are mainly marked by a final duration prominence. Pre-lexical infants may therefore set the head-complement parameter on the basis of the analysis of sentence prosodic contours. Distributional and prosodic cues may in fact combine to inform infants on the value of the head-complement parameter in their language. However, until now, the only available pieces of evidence for the role of one or the other sources of information are cross-linguistic correlations, which may or may not be causal.

Alternatively to or in combination with distributional, prosodic and phonological bootstrapping approaches, parameters may be set through probabilistic computations (Pearl, 2007; Pearl & Lidz, 2009; Yang, 2002). In particular, given the space of hypothesis defined by the possible values of linguistic parameters, a Bayesian learning mechanisms may select the values of parameters that best fit the available input data. Precisely, the learner would compute the probability of the observed data according to each set of parameter values, and compute the posterior probability of each set of parameters according to the data she can observe. The learner will then select the most probable set of parameter values as those accounting for her language.

However, in front of ambiguous data, which can fit several hypotheses of the prior hypothesis space, Bayesian probabilistic learning mechanism often fail to reach the conclusion children reach in natural language acquisition (Lidz, Waxman & Freedman, 2003; Pearl & Lidz, 2009; Pearl, 2009). Pearl (2007) proposed that domain-specific filters, which select the data processed by the learning mechanisms, should be added for these to reach the same conclusions as children.

Pearl studied in particular the case of the acquisition of English metrical phonology, which is thought to depend on the value of at least nine parameters, yielding 156 grammars in the prior hypothesis space. Pearl (2007) shows that an unbiased probabilistic learning mechanism is unable to converge to the correct English grammar. A simple reason for these results is that English grammar is not the optimal description of the input, which contains about 27% of exceptional data that does not fit English grammar (Pearl, 2009). However, the performance of probabilistic learning mechanisms improves if they are equipped with a way to evaluate and discard the most ambiguous piece of data (Pearl, 2007; Pearl & Lidz, 2009; Regier & Gahl, 2004). The evaluation of the ambiguity of a word stress pattern relies on language-specific knowledge. Thus, even though the probabilistic model proposed by Pearl (2007) relies on a domain-general learning mechanism, it requires two types of language-specific knowledge, i.e. the prior hypothesis space and domain-specific data filters.

Similar Bayesian learning mechanisms may in fact model the way infants set the head-complement parameter to the correct value, i.e. head-initial or head-final, on the basis of the analysis of distributional and/or prosodic cues. As for bootstrapping mechanisms, such learning however requires that the prior hypotheses make specific predictions about the frequency distribution or the implementation of speech prominence.

1.2.2.3 Special mechanisms for speech processing

Speech is not the only potential carrier of linguistic information. In fact, hand gestures can effectively replace speech in case of muteness or deafness, and give rise to sign language. Nevertheless, it remains true that when speech is available, human communities preferentially develop languages over that modality. It has been proposed that speech represent a better modality than gestures for communication because it allows to communicate in situations where people cannot see each other, i.e. at night or distance. Moreover, speech would allow speakers to communicate while keeping their hands free, a significant advantage for cultural transmissions of tool use, where oral explanations can accompany a demo. Nevertheless, it is not clear how these advantages of speech would be evident to young infants acquiring language. Moreover, a systematic trait in an animal species that spread all over the globe (here, the human species) suggests a biological rather than environmental origin. Several observations may support this hypothesis.

Since Broca's first observations (1861), numerous behavioral studies showing an advantage for the right ear when processing speech stimuli (Broadbent & Gregory, 1964; Kimura, 1961; Shankweiler & Studdert-Kennedy, 1967), neuropsychological and imaging studies have confirmed that speech processing in adult speakers relies on dedicated neural networks, primary located in the left hemisphere (Geschwind & Levitsky, 1968; Perani et al., 1996). Similar lateralization was observed with imaging techniques in three-month-olds (Dehaene-Lambertz, Dehaene & Hertz-Pannier, 2002) and newborns (Peña et al., 2003). Most interestingly, Peña's experiment shows that the left hemisphere is more activated for normally played speech stimuli, than for time-reversed speech stimuli or silence. Time-reversed speech is an optimal control to speech, because it matches the latter in duration, pitch and intensity. Thus, these results suggest that speech contains certain properties, triggering special processing primary located in the left hemisphere.

Benavides-Varela and colleagues (in prep.) brought further evidence for a specificity of speech processing. They studied memory for short speech sequences in newborns. Using the Near-Infrared Spectroscopy, they first presented infants with one unique word that was repeated for three minutes, e.g. *mita*. A two-minute phase followed, during which neonates were either exposed to a second word, e.g. *noke*, music, or silence. In a test phase, infants then heard again the first word, or another different word, e.g. *pelu*. Interestingly, Benavides-Varela and colleagues found different cerebral activations for the same and novel words in the test phase for those infants who were exposed to silence or music during the intermediate phase. This suggests that these neonates had encoded some properties of the fist word they were exposed to. However, those who were exposed to a word in the intermediate phase did not show evidence of memory in the last phase. Thus, the formation, consolidation or maintenance of memory traces for speech sequences appears to be interfered with by the presentation of new speech material, but not of any acoustic material, like music. This study thus strengthens the proposition that speech is processed differently from other acoustic stimuli, from birth onward.

Behaviorally, Vouloumanos and Werker (2004) observed that 2- to 7-month-old infants and even newborns (2007) prefer to listen to normal speech than non-speech analogues matching speech in terms of spectral and temporal properties. However, newborns, in contrast to 3-month-olds, do not prefer to listen to speech than to rhesus monkey vocalizations (Vouloumanos, Hauser, Werker & Martin, 2010). The preference for speech at birth may in fact reflect a general bias for biological productions, similar to that observed in the visual domain where newborns prefer to observe biological than non-biological movements (Simion, Regolin & Bulf, 2008).

Altogether, these results suggest that speech is special, either because it is produced by biological movements as exposed in the motor theory of speech perception (see Liberman, 1996 for a review), or because it contains certain properties that trigger or constrain certain mechanisms, realized by neural populations primary located in the left hemisphere of the brain (Mehler & Dupoux, 1994s).

1.3 Core cognition

1.3.1 Dedicated learning mechanisms in animal cognition

Evolutionary biology has promoted the view that "biological mechanisms are hierarchically nested adaptive specializations, each mechanism constituting a particular solution to a particular problem" (Gallistel, 2000). While this view is widely accepted when considering molecule and organ structures, it has only been recently extended to studies of learning mechanisms. That is, organisms are endowed with a number of instinctive behaviors and learning mechanisms that answer the specific need of their species to survive in their environment.

For example, indigo buntings, a migratory bird species, are endowed with a mechanism that allows them to identify the north from the night sky, to orient their migrations. Emlen (1975) showed that birds register the landmarks (for example the Polaris Star in the current night sky) that correspond to the center of rotation of the night sky. This learning ability is species-specific, as most species would not spontaneously identify the center of rotation of the night sky, and domain-specific, yielding knowledge about the north position, and nothing else.

Honeybees also need an absolute landmark in order to be able to communicate other members of their species the location of a food source. They do so through a complex dance that involves symbolically representing the distance and direction relative to the sun azimuth of the food source. In fact it seems bees represent the solar ephemeris. Bees presented with a sun-substitute at midnight on their internal clock treat it as if it were due north (Lindauer, 1957). Dyer and Dickinson (1994) showed that bees, which had been exposed to the sun only four hours in the late afternoons, when it was west, nevertheless represented the sun east when first released free in a cloudy morning, when the sun was not observable (see also Lindauer, 1957). Moreover, bees' behavior as the day elapsed suggests that they did not represent the sun as moving steadily from east to west. Rather, bees represented the morning sun in a position 180 degrees from the average position they had experienced on previous evenings, and switched to the evening west position from noon onward. Thus, honeybees appear endowed with innate knowledge that the sun should move from somewhere in the east in mornings to somewhere in the west in afternoons. Learning the sun ephemeris involves adjusting some parameters of this a-priori pattern, to makes it fit with the observed movements of the sun.

Similar parameter-setting mechanism seems to apply to the acquisition of some species-specific bird songs (Marler, 1991). Marler & Tamura (1962) observed that different populations of a same species of songbirds, the White-crowned Sparrows, produce songs that are both highly similar and showing some inter-population variety. This suggests that they equipped with a template of their species-specific songs, and adapt it to the songs experienced in their environment (Marler, 1991). Gardner, Naef and Nottebohm (2005) showed that isolated juvenile canaries could imitate artificial songs that do not conform to the structure normally observed in adult canaries. However, once in adulthood, imitated songs were reprogrammed to form typical canary phrasing, thus suggesting that these birds are equipped with innate constraints on the formation of their songs, which mature only in adulthood.

1.3.2 The Core Cognition Hypothesis

The study of such highly domain-specialized learning mechanisms in animal species lead to the proposition that also humans might be endowed with specialized core cognition systems (Carey, 2009; Carey & Spelke, 1996; Gallistel, Brown, Carey, Gelman, & Keil, 1991; Spelke, 1994). The core cognition, or core knowledge, proposal was first stated by Spelke concerning the cognition of objects, and generalized to ecologically important classes of entities (Spelke, 1994).

Spelke (1990) showed that infants divide the world into objects according to the principles of cohesion, boundedness, rigididty, and no action at distance. According to the cohesion principle, two points belong to the same object, if a path of connected points lying on the object surface can link them. The cohesion principle states that all points on an object move on connected path over space and time. If two points cannot be connected over space or time, they belong to two different objects. The boundedness principle defines the boundary of objects, as the ensemble of pairs of points that are two-dimensionally adjacent, but belong to two surfaces that can be separated in depth or over motion. According to the principle of rigidity, two surfaces undergoing different rigid motions are seen as belonging to different objects. Thus objects tend to maintain their size and shape over motion. Finally, the principle of "no action at a distance" states that two separated objects move independently from each other (Spelke, 1990).

Spelke and colleagues (1992) further showed that even though young infants' knowledge of objects is rich, including the above principles, it is not fully identical to adults'. That core knowledge will be enriched and refined with development. For example, four-month olds already have the notion of continuity and solidity of objects, but not that of gravity or inertia, which should emerge later (Spelke, Breinlinger, Macomber, & Jacobson, 1992). Spelke and colleagues (1994) wrote:

"*According to the* core knowledge hypothesis*, knowledge of certain constraints on objects guides the earlier physical reasoning. This knowledge remains central to common sense reasoning throughout development and constitutes the core of adults' physical conceptions. New beliefs emerge with development, amplifying and extending human reasoning and surrounding core physical conceptions with a multitude of further notions. As physical knowledge grows, however, initial conceptions are not abolished, transformed, or displaced.*" (Spelke, Katz, Purcell, Ehrlich, & Breinlinger, 1994).

Spelke (1994) later generalized this statement to initial (core) knowledge, including at least four domains: physics (objects and object motions), psychology (i.e. expectations about human or animal behaviors), number and geometry. "*Humans are endowed with a number of systems of knowledge, each consisting of a limited set of principles capturing highly reliable constraints on a significant class of entities. Over the course of human development, each system of knowledge grows as the principles at its core are enriched by further, generally less reliable notions.*" (Spelke, 1994).

Carey (2009) further systemized the theoretical proposal of core cognition According to her, a core cognition system must satisfy six properties:

a. Core cognition has rich integrated conceptual content, which goes beyond the perceptual or sensori-motor primitives.
b. Core cognition representations are created by innate perceptual input analyzers.
c. These input analyzers operate throughout life.
d. Core cognition systems are domain-specific learning devices.
e. Some core cognition is shared with other animals, and thus have a long evolutionary history.
f. The format of core cognition is iconic rather than involving sentence-like symbol structures.

Not all of these properties have the same importance; the core of core cognition consists in the properties a, b and d. Core cognition systems rely on specialized learning devices selected through evolution, which analyze specific input to yield conceptual representations.

The last decades of developmental studies showed that human infants are indeed equipped with several specialized mechanisms aiming at processing one type of input to extract knowledge that goes beyond the sensory-motor experience. These core cognition systems include the domain of objects, numbers, and agency (see Carey, 2009 for an extensive review). It is however interesting to notice that none of these core cognition domains is human specific, but all are shared with some other animal species.

Above, we have argued that language, a human-specific faculty, may rely on dedicated learning mechanisms, selected through evolution. Just like it is the case for the cognition of objects, numbers or agency, we propose that language acquisition may partly rely on a series of modular perceptual-input analyzers, mediating causal connections between perceptual properties and linguistic representations. As for core cognition, these learning mechanisms should be triggered by certain properties of the linguistic input and yield linguistic representation that directly and specifically inform the linguistic system.

1.4 Conclusion

In sum, there is not doubt that language is a human-unique faculty. The debate lies on the extent to which mechanisms dedicated to its acquisition must be posited. Since the 80s, the use of more and more powerful computers has allowed researchers to study carefully the amount of statistical information contained in the input. Those studies have forwarded the idea that there is sufficient information in the input to learn a great deal about one's language. The demonstrations that pre-lexical infants are able to access at least some of that statistical information strengthen the claim.

This view is however undermined by two arguments (at least). First, the powerful computers that are now available to us are unable to acquire language, even tough their statistical power largely overcomes infants' or adults'. (This failure is unlikely to reflect a lack of efforts if one considers the economic benefit that companies such as Apple or Microsoft would obtain from selling a computer that master language acquisition.) Second, and more importantly, the hierarchical nature of language suggests it is generated by rules or procedures that lie beyond the reach of statistical computations. In fact, we have presented a series of theoretical and experimental arguments in favor of the idea that language acquisition relies on biologically constrained dedicated mechanisms. In particular, speech processing appears to be constrained in a way that may facilitate language acquisition. Given the ubiquity of highly specialized learning mechanisms in the animal cognition, we consider worthy of investigation the hypothesis of dedicated mechanisms underlying the acquisition of a human-unique ability, language.

Showing the existence of such mechanisms requires the demonstration of two properties. First, we need to show that language acquisition mechanisms are triggered by specific perceptual or distributional properties of the linguistic input. Second, we need to show that infants infer from perceptual or distributional information knowledge about their linguistic system. The experimental work presented below intends to address both issues. Studying constraints on the processing of consonants and vowels, we will first show that different learning mechanisms sustaining language acquisition rely on different speech categories, suggesting adequacy between speech and language acquisition mechanisms. Second, adopting a distributional bootstrapping approach of function words acquisition, we will show that infants infer linguistic or functional knowledge from the distributional description of their input.

Even for a word, we will not waste a vowel.

Anglo-Indian proverb

CHAPTER 2

Consonants and Vowels

2.1 The CV Hypothesis

Young language learners could profit from a (partial) “division of labor”, such that one speech category might preferentially support the acquisition of the lexicon, whereas another might be more dedicated to the identification of structural regularities, in particular those signaling relations between constituents. Here, we will present a series of observations that hint at different roles for consonants and vowels in language acquisition. Namely, we propose that consonants are more involved with word identification and encoding, because they are better suited than vowels for categorical perception. Vowels, in contrast, carry prosodic variations and provide cues to determine the boundaries and the organization of syntactic constituents (Nespor & Vogel, 1986; Selkirk, 1984). This functional difference between consonants and vowels constitutes the Consonant-Vowel hypothesis, hereafter referred to as the CV hypothesis (Nespor, Peña & Mehler, 2003). Experimental evidence showed that adult participants (Bonatti, Peña, Nespor & Mehler, 2005) as well as toddlers (Havy & Nazzi, 2009; Nazzi, 2005; Nazzi & Bertoncini, 2009) rely mainly on consonants for lexical processes, and that adult participants generalize a repetition-based regularity over vowels, but not over consonants (Toro, Nespor, Mehler & Bonatti, 2008a; Toro, Shukla, Nespor & Endress, 2008b). It is not clear, however, whether consonants and vowels are readily used for different functions before infants have acquired a sizable lexicon. If the specialization arises only after infants have a lexicon of a few hundred words, it would be difficult to conceive that it could play a role in the early stages of language acquisition. Here, we explore whether 12-month-old infants (who have just begun to build the lexicon) have already assigned different functional roles to consonants and vowels.

Long before Nespor et al. (2003) had formally proposed the CV hypothesis, ancient thinkers already hinted at the categorical distinction between consonants and vowels. In 500 BCE, the Sanskrit grammarian Pānini described phonemic categories based on their mode of articulation, thus hinting at the notion of features and possibly at the categorical distinction of consonants and vowels (Kiparsky, 1993; Staal, 1962). In the *Sophist*, Plato (360 BCE) explicitly distinguished between vowels and consonants: "And the vowels are especially good at combination – a sort of bond pervading them all [the letters], so that without a vowel the others cannot be fitted together". This statement is in agreement with the Ancient Greek etymology that suggests that vowels (φωνηεν, *phonien*) are sounds (φωνη), whereas consonants can come only *with* (συν-) vowels (συμφωνον, *symphonon*).

Probably inspired by the graphic system of Hebrew, whose letters represent only consonants, not vowels, Spinoza (1677) formulated a similar idea, stating that vowels are the soul of the letters and the consonants are bodies without soul (Mehler, Peña, Nespor & Bonatti, 2006). These grammarians and philosophers thus had intuitions about the existence of two broad categories of speech sounds: consonants and vowels. Our present proposal adds to these intuitions that the aforementioned categories have different functions.

2.1.1 The CV Hypothesis Part I: The consonantal bias in lexical processes

The role of consonants in the CV hypothesis originates from the observation that across languages, consonants allow more quality distinctions than vowels, where by quality we mean distinctions in terms of articulatory features[1]. With only a few exceptions, consonants are indeed cross-linguistically more numerous than vowels. For example, in Malay the proportion is 20C: 5V; in Italian 24C: 7V; in Hausa 32C: 5V; in Arabic 29C: 3V; in Igbo 27C: 5V; in Sindhi 46C: 10V. In fact, Maddieson (2008) computed the consonant-vowel ratio as the number of consonants divided by the number

[1] There are systems where consonants may be distinguished for duration (e.g. Italian, *caro-carro;* [káro]-[kár:o]; 'dear' – 'cart') and systems where vowels may be distinguished because of suprasegmentals, such as nasality or tones (e.g. French, *beau-bon*; [bo]-[bɔ̃]; 'beautiful' – 'good)'. Also in these types of systems, by and large, consonant distinctions outnumber vowel distinctions. The role of tones and pitch accents should however be further studied in our framework.

of vowel qualities in each of 563 languages. The ratio ranges from 1.11 for Andoke (isolate; Colombia), with 10 consonants and 9 vowel qualities, to 29 for Abkhaz (Northwest Caucasian; Georgia), with 58 consonants but only 2 vowel qualities. Most importantly, 5 vowel systems are the most common and most systems have over 20 consonants (Nespor et al., 2003). Because consonants are more numerous than vowels, they are relatively more informative than vowels for lexical distinctions and may be at the origin of a lexical specialization of consonants. However, we will see below that the preferential role of consonants for lexical distinctions goes beyond their numerical superiority and persist in languages with a low consonant-vowel ratio.

Nespor et al. (2003, p.206) observe: "Consonants are not only more numerous than vowels, but, unlike vowels, they tend to disharmonize within a word, i.e. to become more distinctive. That is, there is a tendency for the consonants that belong to the same lexical item to alternate in quality. Just to name a few cases, in Japanese the combination of two voiced obstruents within a root is avoided (Itô & Mester 1986); Arabic avoids adjacent root consonants produced by the same articulator (McCarthy 1991); Classical Greek avoids three aspirated consonants within one word, the so-called Grassmann Law.

In contrast, vowels not only have less distinctive power than consonants because of being fewer in number in most systems, but also because of their tendency to loose distinctiveness. For example, vowels do not disharmonize, in general, but rather tend to harmonize throughout a domain in many languages. Because vowel harmony assimilates vowels for certain features, their original distinctive power is reduced. In addition, the domain of vowel harmony is often not lexical, but a signal to syntax. In Turkish for example, it includes, besides all the affixes of a word, also most of the clitic elements that are syntactically attached to it, thus signaling syntactic constituency at the lowest level (Nespor & Vogel, 1986)."

A second observation motivating our belief that consonants have a predominant role to acquire the lexicon is that, in Semitic languages such as Arabic and Hebrew, lexical roots are only represented on the consonantal tier (Berent, Vaknin & Marcus, 2007; Prunet, Beland & Adrissi, 2000). For example, in Arabic, the root *ktb* has the lexical meaning related to "write". The different vowels intervening between and around the consonants serve to form different words and word-forms that are related in meaning (e.g., *katib:* writer, *kataba*: he wrote, *kitab*: book, *maktaba*: library, etc). In contrast, there is no documented language that has lexical roots based only on vowels. Furthermore, the case of consonantal lexical roots may be an extreme case of the situation observed in other languages, where consonants are in general more informative than vowels for lexical distinctions (see Keidel, Jenison, Kluender & Seidenberg, 2007 for the analysis of the French adult lexicon). Below, we analyze the lexicon of infants for French and Italian, two languages that differ in the ratio of consonants and vowels (French has 19 consonants and 13 vowels; Italian has 24 consonants and 7 vowels). We verify that the sequence of consonants is more informative to identify a word than the sequence of vowels. This is true for both languages, suggesting that the lexical role of consonants does not change when the ratio of consonants and vowels varies.

2.1.2 The CV Hypothesis Part II: The role of vowels to signal syntactic organization

The CV hypothesis attributes a specific role to vowels based on three observations. First, vowels are the main carrier of prosodic information (Lehiste, 1970;

Ramus, Nespor & Mehler, 1999). Second, that information provides cues that correlate with some important morphosyntactic properties (Morgan & Demuth, 1996; Nespor, Shukla & Mehler, in press; Nespor & Vogel, 1986; Selkirk, 1984). Third, pre-lexical infants and even neonates are sensitive to rhythm and to phrase boundaries (Christophe, Mehler & Sebastian-Gallés, 2001; Christophe, Nespor, Dupoux, Guasti & Van Ooyen, 2003; Nazzi, Bertoncini & Mehler, 1998; Ramus, Hauser, Miller, Morris & Mehler, 2000). Syntax acquisition may thus start with inferences from prosodic, specifically rhythmic, cues carried by vowels.

In fact, vowels more than consonants, can vary in terms of pitch, intensity and duration in relation to their sentential position. That is, vowels are more affected than consonants by prosody, which provides signals to syntactic constituency (Gleitman & Wanner, 1982; Morgan & Demuth, 1996, Nespor & Vogel, 2008; Selkirk, 1984). In particular, a phonological phrase boundary always coincides with a syntactic boundary. These boundaries are available to infants (Kemler-Nelson, Hirsh-Pasek, Jusczyk & Wright-Cassidy, 1989; Gerken, Jusczyk & Mandel, 1994) as well as newborns (Christophe et al., 2001)

Certain syntactic properties correlate with prosodic cues across languages. For example, complements are marked mainly by pitch prominence in head final languages such as Turkish, and mainly by duration prominence in head initial languages such as French (Nespor et al., 2008). Moreover, those prosodic cues remain valid within languages that allow both types of constructions. For example, in German, the complement-head structures are mainly marked by an initial pitch prominence, while the head-complement structures are mainly marked by a final duration prominence (Nespor et al., 2008; Shukla & Nespor, 2010). Observing this and other correlations, prosodic bootstrapping theories (Morgan & Demuth, 1996) have proposed the existence of a bridge between prosodic and syntactic properties.

Moreover, vowels carry another type of rhythmic information. The proportion of time occupied by vowels in the speech input determines the rhythmic class of languages. Ramus, Nespor & Mehler (1999) showed that vowels occupy about 45% of the speech stream in stress-timed languages (e.g. Dutch, English), about 50% in syllable-timed languages (e.g. French, Italian) and 55% in mora-timed languages (e.g. Japanese). Newborns can use this information to discriminate between two languages that belong to different rhythmic classes (Ramus, Hauser, Miller, Morris & Mehler, 2000). The rhythmic class to which a language belongs correlates with important morphosyntactic properties (Nazzi, Bertoncini & Mehler, 1998; Nespor, Shukla & Mehler, in press; Ramus, Nespor & Mehler, 1999). In particular, the percentage of the speech stream that vowels occupy is indicative of the complexity of the syllabic repertoire of a given language. In turn, typological studies have shown that languages with simple syllabic tend to be verb final, to use post-positions and have a rich case system (Donegan & Stampe, 1983; Gil, 1986; Fenk-Oczlon & Fenk, 2005; Nespor, Shukla & Mehler, in press). Infants may therefore infer certain morphosyntactic properties of their language from the identification of its rhythmic class.

In sum, prosody, in particular rhythmic information, carried by vowels, provides infants with knowledge about the shape of words, signals important syntactic boundaries, and provides cues for fundamental syntactic properties such as the relative order of heads and complements. These observations inspired the second part of the CV hypothesis: the variation of vowel quantities carries syntactic information.

2.1.3 Experimental evidence in favor of the CV hypothesis

A number of experimental results support the CV hypothesis. Specifically, infants are able to use statistical information, such as dips in transition probabilities (TPs) between syllables to identify word boundaries in a continuous speech stream (Saffran, Aslin, & Newport, 1996). Newport & Aslin (2004) argued that adults could compute also TPs both between successive consonants and between successive vowels. However, assuming participants use TPs to identify potential words, the CV hypothesis predicts that they should perform better at computing TPs over consonants than over vowels. Indeed, Bonatti, Peña, Nespor and Mehler (2005) showed that when the statistics were more complex than those in Newport & Aslin (2004), adult participants could use TPs over consonants to segment a continuous speech stream but not over vowels. Moreover, Mehler, Peña, Nespor & Bonatti (2006) showed that when in one stream, TPs between consonants and TPs between vowels predict different segmentations, the consonant statistics are favored. Thus, consonants appear to be a privileged category for discovering words in a continuous speech stream.

Word learning experiments in infants further confirmed the advantage of consonants in encoding lexical items. Nazzi and colleagues (Nazzi, Floccia, Moquet & Butler, 2009) showed that in a word-learning situation where 30-month-olds must ignore either a consonantal one-feature change or a vocalic one-feature change (e.g. match a /duk/ with either a /guk/ or a /dɔk/), both French- and English-learning infants choose to neglect the vocalic change rather than the consonantal change. This preference was observed for word-initial (/guk/-/duk/-/dɔk/), word-final (/pib/-/pid/-/pɛd/) and word-internal consonants (/gito/-/gipo/-/gupo/), and did not depend on an inability to process fine vocalic information. In agreement with these results, 16- to 20-month-old infants could acquire simultaneously two words differing only in one consonant, whereas they could not do so for minimal pairs differing in one vowel (Havy & Nazzi, 2009; Nazzi, 2005; Nazzi & Bertoncini, 2009). Furthermore, these findings are not restricted to one specific consonantal class (Nazzi et al., 2009; Nazzi & New, 2007).

The second prediction of the CV hypothesis is a preference for vowels for extracting and generalizing structural relations. The generalization of a syntactic regularity implies that an observed relation between at least two sentence constituents can be applied to, or recognized in, novel sentences that never occurred before. For example, in English, the noun phrase preceding a verb is usually its subject, while the noun phrase immediately following the verb is interpreted as its complement. Toddlers as young as 23-month-old use this information to interpret a novel verb as either a transitive (e.g. *the duck* is gorping *the rabbit*) or an intransitive action (e.g. *the duck and the rabbit* are gorping; Naigles & Kako, 1993). As we observed above when discussing the role of prosody in the first steps of syntax acquisition, relational patterns such as the relative position of prominent elements may help infants learn about the implementation of syntactic relations in their language.

A particular case of relational pattern is identity, where a structural regularity is defined by the relative positions of identical elements (e.g. ABB, ABA). Indeed, since the seminal work of Marcus and colleagues (Marcus et al., 1999) repetition-based regularities have been extensively used to test generalization abilities in infants and adults (see Endress, Dehaene-Lambertz, & Mehler, 2007; Endress, Scholl & Mehler, 2005; Kovacs & Mehler, 2009). Marcus and colleagues used the following structures: ABB (instantiated by words like *pukiki*, *mesasa*, etc) and ABA (instantiated by words like *pukipu*, *mesame*, etc). They showed that, after being habituated to exemplars of one of the structures, infants could discriminate between novel exemplars of both structures, suggesting that they had extracted and generalized the rules. The CV hypothesis - claiming that vowels are favored to signal structural relations - predicts that repetition-based structures should be easier to detect and generalize when they are implemented over vowels than over consonants. Experimental work corroborated this prediction for adults. Toro, Nespor, Mehler & Bonatti (2008a) showed that while adult participants easily learned the ABA regularity over vowels, they were unable to learn the same regularity over consonants. Adults remain unable to generalize ABA over consonants even when vowel duration was reduced to one third of the duration of consonants, while they could generalize ABA on barely audible vowels (Toro, Shukla, Nespor & Endress, 2008b). Thus, the reliance on vowels for extracting repetition-based regularities is not solely due to a major acoustic salience. Rather, vowels and consonants are involved in different types of processes, as suggested by the existence of a different neural substrate for each category (Caramazza, Chialant, Capasso & Miceli, 2000; Knobel & Caramazza, 2007).

2.1.4 Our experimental studies

In the experimental work presented below, we ask whether the documented functional difference between consonants and vowels can play a role in early steps of language acquisition. Evidence for the lexical role of consonants exists for participants older than 16-month (Havy & Nazzi, 2009), who already have a sizable vocabulary (about one hundred and eighty words according to the MacArthur-Bates Communicative Development Inventory; Dale & Fenson, 1996). Moreover, evidence for the specialized role of vowels was reported for adults (Toro et al., 2008a, 2008b). Recently, Pons and Toro (2010) suggested that 11-month-olds were already better at learning the AAB structure over vowels than over consonants. In a preferential looking paradigm, infants habituated to words respecting the AAB structure over vowels (e.g., *dabale*, *tolodi*, *tibilo*) could discriminate between novel words respecting the same structure (e.g., *nadato*, *lotoba*, *dilite*), and words that did not respect the AAB structure (e.g., *dutone*, *lanude, bitado*). In contrast, infants habituated to the AAB structure instantiated over consonants (e.g., *dadeno, lulabo, nunide*) did no discriminate novel words respecting the structure (e.g., *dedulo, lulina, nunobi*) from words that did not respect the structure (e.g., *dutani, litedo, bilune*). However, Pons and Toro (2010) used the same vowels and consonants in the test phase as in the familiarization phase. In particular, test and familiarization words shared the repeated vowels (e.g. *dabale* as a familiarization word, and *batalo* as a test word). Thus, an alternative explanation of Pons and Toro's results is that infants process mainly vowels and learned the repeated tokens (i.e. *aa, oo, uu, ii, ee*). They did not necessarily generalize the AAB structure. Showing generalization would require using novel vowels and consonants to form novel words in the test phase.

Here, we directly assess the CV hypothesis of a functional difference between consonants and vowels, by testing 12- and 6-month-olds in two very similar paradigms, which vary only in those details that differentiate a word-learning experiment from a situation promoting the discovery of structural relations. Six-month-old infants are considered pre-lexical. They may nevertheless have acquired a certain amount of distributional knowledge about their language (Kuhl, 1993; Maye, Werker & Gerken, 2002), and may recognize a very limited number of words such as "mummy", and their own name (Mandel, Jusczyk & Pisoni, 1995). Twelve-month-olds' word learning abilities are still immature (Lock, 1980; McShane, 1979; Stager & Werker, 1997). Infants begin to learn words in a very fast manner only towards the end of the second year of life (Carey & Bartlett, 1978; Golinkoff, Church Jacquet, Hirsh-Pasek & Nandakumar, 1996; Golinkoff, Hirsh-Pasek, Bailey & Wenger, 1992; Heibeck & Markman, 1987) after the maturation of a series of conceptual constraints that support word learning (Halberda, 2003; Hochmann, Endress, & Mehler, 2010; Mervis, 1987; Markman, 1990; Markman & Hutchinson, 1984; Markman, Wasow & Hansen, 2003; Soja, Carey, & Spelke, 1985). If the lexical role of consonants is already set before this stage, it can constrain and shape infants' vocabulary development. Similarly 12-month-olds' syntactic system is far from being complete. Therefore, showing a preferential role for vowels in the generalization of structural relations may have important implications for theories of syntax acquisition.

2.2 Evidence for a consonantal bias in the lexicon: analysis of infants' vocabulary

We first intend to seek new evidence for a consonantal bias in lexical acquisition, which has previously been reported for French-learning infants as young as 16-months (Havy & Nazzi, 2009) and in 30-month-old English-learning infants (Nazzi et al., 2009). Once a consonantal bias has emerged, infants should learn with more ease words that differ along the consonantal tier than words differing along the vocal tier (Havy & Nazzi, 2009; Nazzi, 2005; Nazzi & Bertoncini, 2009; Nazzi et al., 2009; Nazzi & New, 2007). Thus, we inquire whether infants' vocabulary is constituted of words differing mainly in consonants or in vowels.

We studied the corpus of words in the MacArthur-Bates Communicative Development Inventory (Dale & Fenson, 1996) for two languages: Italian (Caselli et al., 1995), the language of participants in Experiments 1 to 5 (see below), and French, a language that has a more balanced number of consonants and vowels. For both languages, we ask whether consonants are more informative than vowels for identifying words, depending on the size of infants' vocabulary.

2.2.1 Methods

For each word of the MacArthur-Bates Communicative Development Inventory (CDI), we coded words phonetically, using 24 consonants (/j/, /w/, /p/, /b/, /t/, /d/, /k/, /g/, /f/, /v/, /s/, /z/, /ʃ/, /ts/, /dz/ /tʃ/, /dʒ/, /m/, /n/, /ɲ/, /ʒ/, /l/, /ʎ/, /r/) and 7 vowels (/a/, /e/, /ɛ/,

/i/, /o/, /ɔ/, /u/) for Italian, and 19 consonants (/j/, /w/, /p/, /t/, /k/, /b/, /d/, /g/, /f/, /s/, /ʃ/, /v/, /z/, /ʒ/, /l/, /ʁ/, /m/, /n/, /ɲ/) and 13 vowels[2] (/a/, /e/, /ɛ/, /i/, /o/, /ɔ/, /u/, /y/, /œ/, /ø/, /ɑ̃/, /ɛ̃/, /ɔ̃/) for French. Each word yielded one consonant sequence and one vowel sequence, which would correspond to the word representation along the consonantal tier and the vocal tier, respectively. For example, the word "banana", /banana/ yielded /b.n.n./ as consonant sequence and /.a.a.a/ as vowel sequence.

For each language, we obtained the percentage of infants knowing each word of the CDI questionnaire. The data for Italian is available online (http://www.istc.cnr.it/material/tools/macarthur/PVBcomp08_17.pdf) and the data for French was provided by S. Kern (http://www.sci.sdsu.edu/cdi/frencheuropean.htm).

As an approximation, we consider that an infant with a vocabulary of N words understands the N words that are most frequently understood by infants evaluated in the CDI. For the Italian corpus, the data collapsed that information for infants ranging in age between 8- and 17-months. For the French corpus, we considered the percentage of infants ranging in age from 8- to 16-months regarding their understanding of words.

We computed the proportion of different consonant sequences and the proportion of different vowel sequences for vocabulary sizes from 10 words to 400 words.

2.2.2 Results

Figure 2.1 presents the percentages of different sequences of consonants and

[2] The vowels [ɑ] and [œ̃] were not considered because they tend to be replaced by [a] and [ɛ̃], respectively, by many French speakers. Nevertheless, as they occur very rarely, the inclusion of these two vowels would not modify the shape of the results.

vowels for vocabulary size from 10 to 400 words, in Italian. The percentage of different consonant sequences was always superior to the percentage of different vowel sequences, thus making consonants more informative than vowels for word identification. This difference was minimal for a vocabulary of 70 words, when the percentage of different consonant sequences was 92.86% and the percentage of different vowel sequences was 64.29%, and was maximal for a vocabulary of 400 words when the percentage of different consonant sequences was 85.75% and the percentage of different vowel sequences was 38%.

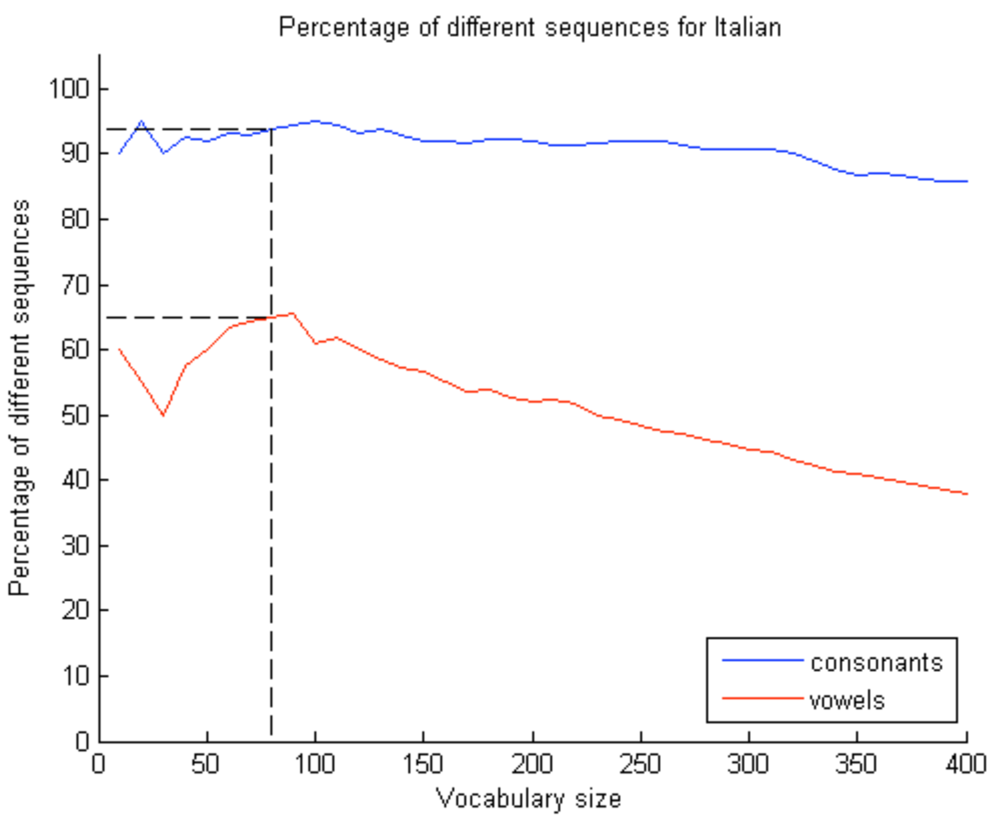

Figure 2.1 - Percentage of different consonant sequences (in blue) and different vowel sequences (in red) for vocabulary sizes from 10 to 400 words, in Italian. The dotted lines indicate the approximate vocabulary of a 12-month-old.

Figure 2.2 presents the percentages of different sequences of consonants and vowels for vocabulary size from 10 to 400 words in French. As for Italian, the percentage of different consonant sequences was always superior to the percentage of different vowel sequences, thus making consonants more informative than vowels for word identification. This difference was minimal for a vocabulary of 10 words, when the percentage of different consonants was 100% and the percentage of different vowel sequences was 90%. The difference quickly increased to reach a plateau for a vocabulary of 170 words, when the percentage of different consonant sequences was 89.41% and the percentage of different vowel sequences was 49.41%. For larger vocabularies, both percentages decreased at a similar rate to reach 76.50% and 35.5%, respectively, for a vocabulary of 400 words.

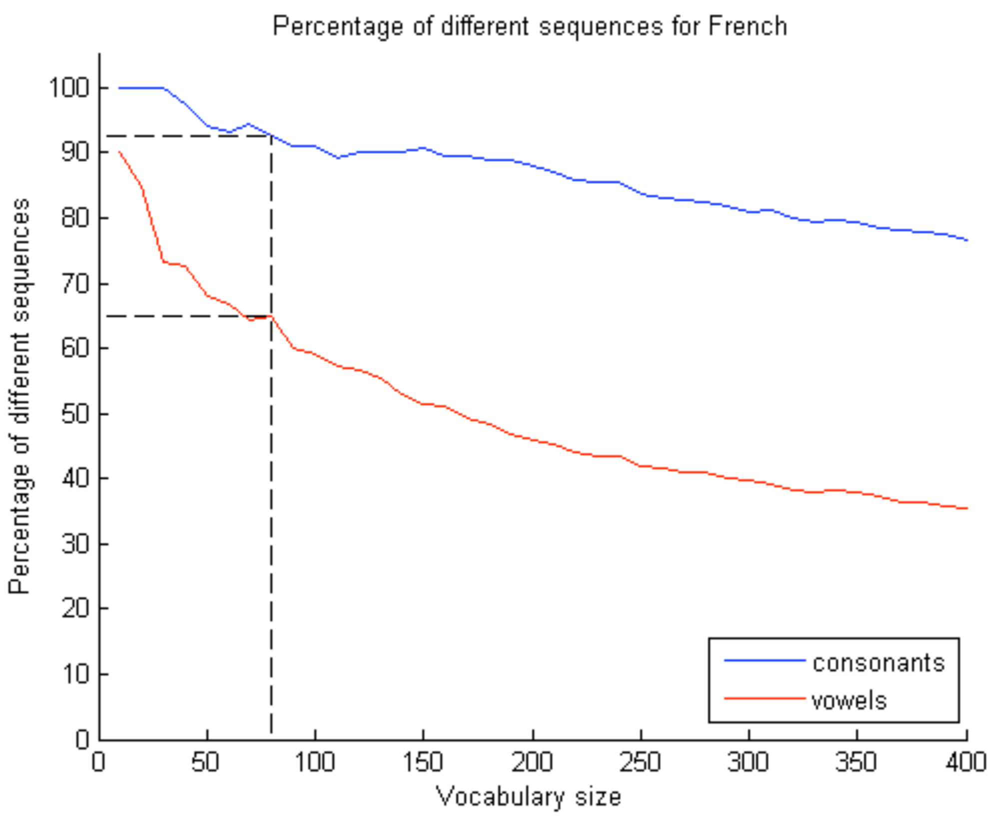

Figure 2.2 - Percentage of different consonant sequences (in blue) and different vowel sequences (in red) for vocabulary sizes from 10 to 400 words, in French. The dotted lines indicate the approximate vocabulary of 12-month-old infants.

We do not know the exact size of the vocabulary of the Italian 12-month-olds we tested. However, the CDI for American English (Dale & Fenson, 1996) reports a mean vocabulary size of 80 words. For that vocabulary size, 93.75% of consonant sequences and 65% of vowel sequences are different in Italian; 92.5% of consonant sequences and 65% of vowel sequences are different in French.

2.2.3 Discussion

We cannot know the exact size of the vocabulary of the Italian 12-month-olds that we tested in Experiment 1 and 3. However, we show here that even in a small vocabulary of 10 words, consonant sequences are more diverse than vowel sequences. Italian twelve-month-olds could identify about 90% of their words relying only on consonants, while they could identify only about 65% relying only on vowels. Moreover, similar information seems also present in French, a language that has a more balanced numbers of consonants and vowels.

These results can be interpreted from two opposite point of views. First, adult speakers teach infants their first words. Thus, the presence of major information on the consonantal tier in infants' vocabulary may simply reflect the major information carried by consonants in adults' vocabulary (Keidel et al., 2007). Furthermore, the consonantal bias in lexical processes observed in infants as young as 16-months may result from an inference of the distribution of information in infants' initial vocabulary. In that view, the larger information held by consonants would be a consequence of the fact that both Italian and French have more consonants than vowels.

Alternatively, our results may suggest that the consonantal bias is already in place when infants first start acquiring the lexicon. That is, if consonants are more important for the lexicon, it follows that it should be easier for infants to learn words that differ in terms of consonants (Havy & Nazzi, 2009; Nazzi, 2005; Nazzi & Bertoncini, 2009), and lexical entries should thus differ more by their consonants than by their vowels.

The lexical role of consonants could be inferred from statistical computations over other sources of information that those which are made available by the lexicon. Several studies reported that by 9 months, infants have acquired general knowledge about word forms in their language. For example, English 9-month-olds but not 6-month-olds prefer strong-weak words to weak-strong words (Jusczyk, Cutler & Redanz, 1993). In a corpus study, Swingley (2005) showed how such knowledge could be generalized from a set of statistically segmented bisyllabic sequences. Statistically segmented syllable sequences tend to correspond to words; thus similar computations may allow pre-lexical infants to discover that consonants are more informative for identifying words. These computations, however, require that infants at least discriminate between the broad categories of consonants and vowels. We will further discuss this point later.

Furthermore, if consonants and vowels were initially equally relevant to the lexicon, many languages should have a larger number of vowels than consonants. However, such languages are very rare. In fact, languages that have a large number of vowels, such as French, have even more consonants. Moreover, one could imagine that languages with a small number of consonants would compensate with a large amount of vowels, in order to increase the variability of syllables and reduce the length of words in the language. Strikingly, this is not the case and languages that have a small number of consonants have even fewer vowels. Hawaiian, for instance, has 8 consonants (/p/, /k ~ t/, /ʔ/, /h/, /m/, /n/, /l/, /w ~ v/) and only 5 vowels (/u/, /i/, /o/, /e/, /a/) (Elbert & Pukui, 1979), thus leading to very long words such as *lauwiliwilinukunuku'oi'oi*, the name of a fish. Such observations would support the proposal that the lexical role of consonants pre-exists to the development of the lexicon, and that a higher number of vowels would not compensate for a reduced number of consonants.

2.3 Constraints on word memory at 12-months: Experiment 1

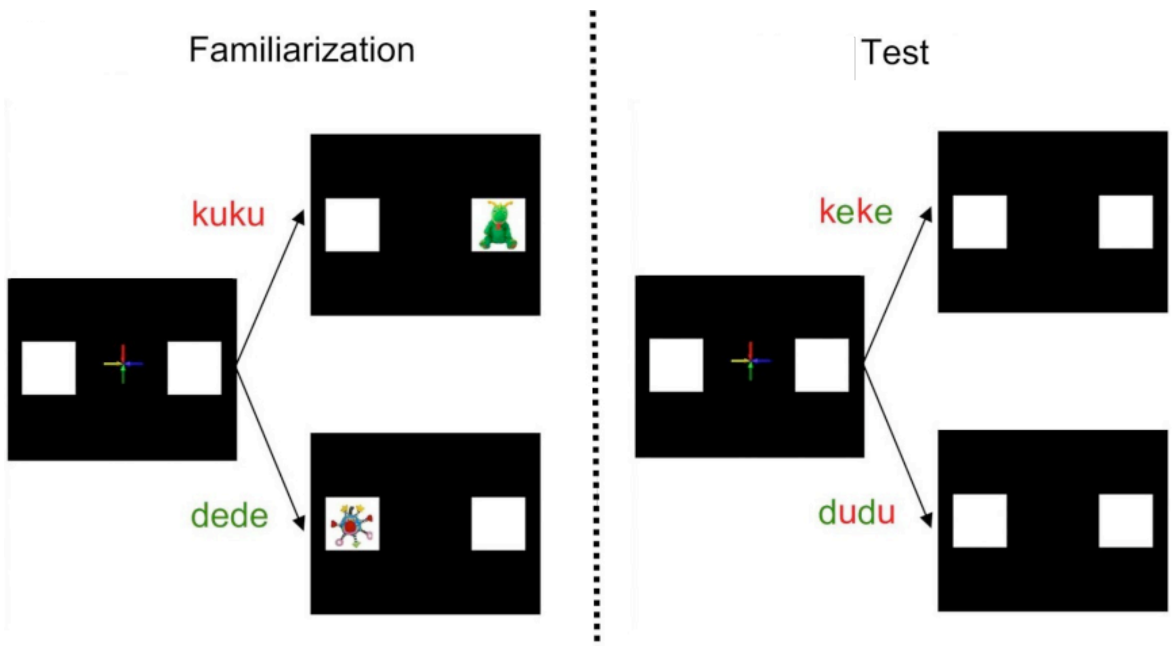

Figure 2.3 – Paradigm of Experiments 1 and 2. Participants took 32 familiarization trials and 8 test trials.

Experiment 1 tests the first part of the CV hypothesis (see Figure 2.3). We ask whether infants rely more on consonants or on vowels when distinguishing among words. Adapting the paradigm developed by Kovács (2008) and Kovács & Mehler (2009b), we teach infants that one word predicts a toy's appearance on one side of the screen (e.g. *dudu*), while another word predicts a toy's appearance on the other side (e.g. *keke*). Infants are then presented with an ambiguous word, composed of the consonants of the former word and the vowels of the latter (e.g. *dede*) (or vice versa, e.g. *kuku*). On this occasion, no toy appears. Reliance on consonants would lead infants to search for the toy on the location predicted by the first word while reliance on vowels would lead infants to search for the toy on the location predicted by the second word. When forming memory representations of novel word, the CV hypothesis predicts that infants should rely more on consonants than on vowels.

2.3.1 Participants

Twenty-six infants were included in the analysis; age range 11 month 15 days to 12 month 15 days. Five other infants participated in the study but were excluded due to fussiness (3) or equipment failure (2). Parents of the infants participating in the 2 experiments signed the informed consent explanation form before the experiments. The Ethics Committee of SISSA (Scuola Internazionale Superiore di Studi Avanzati), where the experiments were conducted, approved the study design.

2.3.2 Stimuli

The words in Experiment 1 consisted of one syllable that duplicates. We used two consonants and two vowels to construct four nonsense words: *kuku*, *dede*, *keke* and *dudu*. Thus, two pairs of words shared the same consonants, but had different vowels: *dede* and *dudu*; *keke* and *kuku*; and two pairs of words shared the same vowels but had different consonants: *dede* and *keke*; *dudu* and *kuku*. Two words sharing neither consonants nor vowels were used in the Familiarization (i.e., *kuku* and *dede*). The two remaining words were used in the Test. Words were synthesized with MBROLA (fr4) with a phoneme duration of 120 ms and a monotonous pitch of 200 Hz. There was no silent pause between two syllables within a word.

The visual stimuli were two pictures of colorful toys. Each appeared inside one of two white squares either on the left or on the right side of the screen. The toys loomed from 4 cm to 7 cm inside the squares for 2 s. The squares had a side-length of 8 cm, positioned at a distance of 13.5 cm. In the Familiarization, each toy was paired with one Familiarization word and one side.

2.3.3 Procedure

The procedure was adapted from Kovács & Mehler (2009b) and is presented in Figure 1. Stimuli were presented via an Apple Dual G5 computer running Psyscope X (http://psy.ck.sissa.it). Infants' gaze was collected with a TOBII 1750 Eye-Tracker (Hofsten, Dahlstrom & Fredrikson, 2005).

The Familiarization phase consisted of 32 Familiarization trials. Familiarization trials started with a display of two white squares on the sides and a central attention-grabber. When the infant looked at the attention-grabber, either one of the two familiarization words was played in a pseudo-random order. We ensured that no word was repeated more than three times in a row. The animated attention-grabber was displayed until the offset of the word, in order to keep the infant's gaze in the middle of the screen. One second after the word offset, a toy appeared in one of the squares, contingent on the word: one word predicted the toy's appearance in one of the squares, while the other word predicted the toy's appearance in the other square. The pairing of the words with toy-locations was counterbalanced across participants.

During test, infants were exposed to 8 trials in a pseudo-random order. Test trials were similar to the familiarization trials, except that infants heard words constituted by the consonants of one of the Familiarization words, and the vowels of the other. Thus if the Familiarization words were *dudu* and *keke*, the Test words were *dede* and *kuku*. If the Familiarization words were *dede* and *kuku*, the Test words were *dudu* and *keke*. No toy ever appeared in the test trials. Two seconds after the word onset, the next trial started.

2.3.4 Analysis

For the analysis, we divided the screen into three equal parts, left, middle and right. In each trial of the familiarization, we measured the proportion of infants anticipating the toy's appearance to the correct side. In the test, we measure each participant's first fixation, and the time spent fixating the left or the right of the screen, after hearing the new word and before the beginning of the next trial. Infants were coded as targeting either the consonant side or the vowel side. The vowel side was the one where the toy appeared after hearing the familiarization word that had the same vowels as the test word. For example, the vowel side for the test word *keke* was the side where during familiarization they learned to turn to after hearing the word *dede*, whereas the consonant side where during familiarization they learned to turn to after hearing the word *kuku*. The consonant side for one of the two test words corresponded to the vowel side for the other test word. We also measured the infants' overall accuracy (Kovacs & Mehler, 2009b; McMurray & Aslin, 2004). That is, for each trial, infants were scored as searching to the consonant-side if the infant looked longer to the consonant side within the 2 s after hearing a new item and before the start of the next trial. Infants were scored as searching to the vowel-side otherwise.

We computed difference scores: (#consonant looks - #vowel looks)/(#consonant looks + #vowel looks) for first and for overall accuracy, and computed a t-test to compare them to the chance level of 0. Positive differences in scores indicate that infants searched for the toys on the consonant side, while negative difference scores indicate infants searched for the toys on the vowel side.

In the absence of the reinforcement due to the absence of puppets in the test phase, extinction effects may be observed. Thus, we also ran our analyses considering only the first 4 test trials.

We also measured the orientation time, as the time that elapsed between the end of the test word (disappearance of the central attractor) and the first fixation produced by infants.

2.3.5 Results

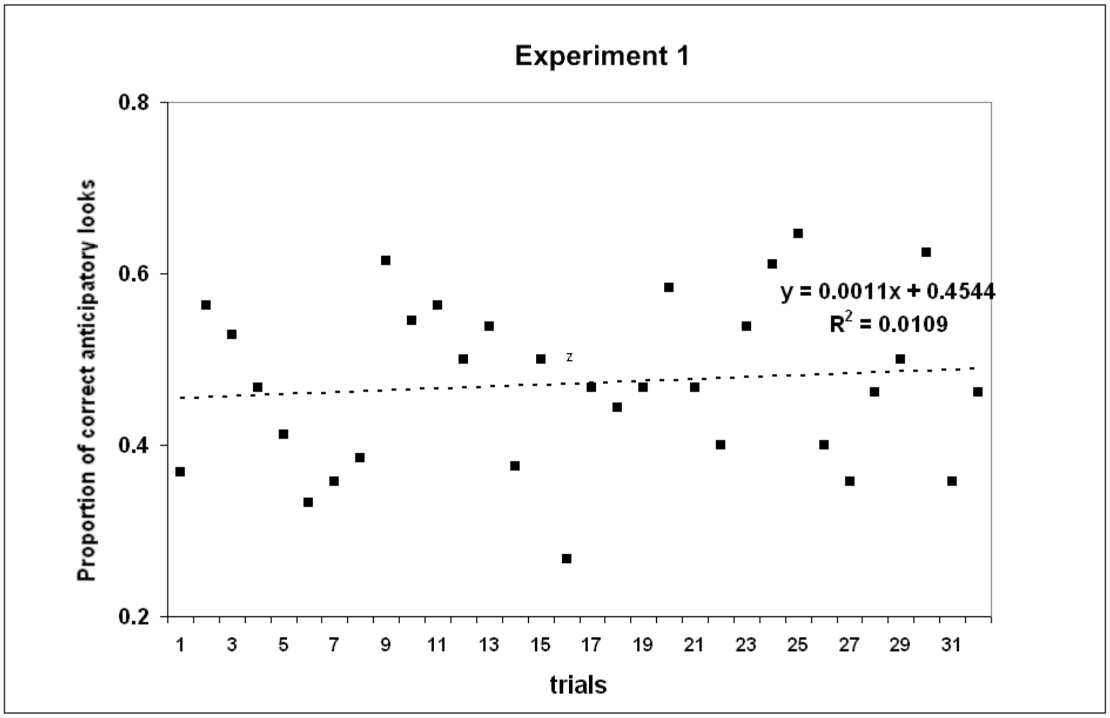

Figure 2.4 - Proportion of correct anticipatory looks for each familiarization trial in Experiment 1. The dotted line depicts the corresponding linear regression.

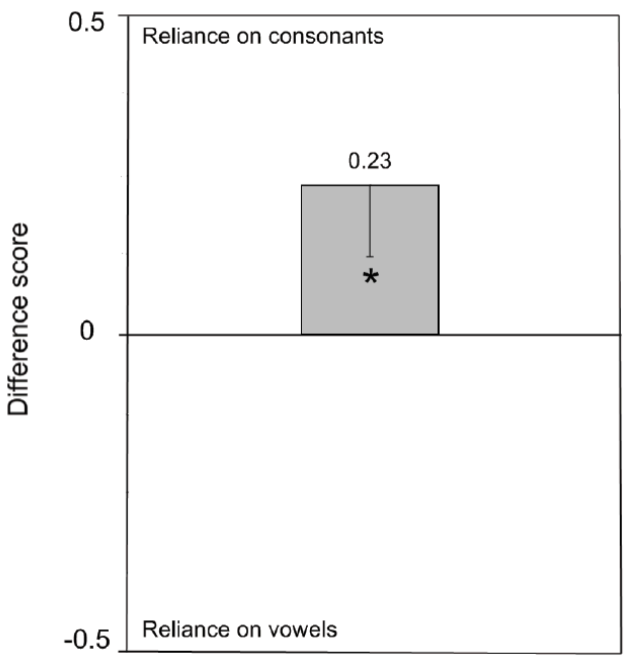

Figure 2.5 - Mean difference score for Experiment 1 considering the first fixations. Infants looked more at the side predicted by consonants. Error bars represent standard error.

The first fixation data for familiarization trials is presented in Figure 2.4. Infants anticipated to one or the other side in 55% of the trials. We computed the proportion of infants showing a correct anticipatory look for each trial. A linear regression analysis gave non-significant results $ß=.0011$, $R^2=.0109$, $t(30)=.57$, $p>.57$.

In the test phase, infants looked to the left or the right in 65% of the trials. Their first fixations occurred slightly earlier when they were located to the consonant side (915 ms after the end of the word) than when they targeted the vowel side (965 ms after the end of the word). However this 50 ms difference was not significant, $t(20) = 1.22$; $P = .24$; d' = .27

Considering first fixations (see Figure 2.5), infants' mean difference score was .23, which was significantly greater than 0, $t(25) = 2.1077$; $P = 045$; d' = .41. Seventeen infants obtained a positive difference score, six infants a negative difference score, and three infants a null difference score. A binomial test showed that significantly more infants obtained a positive difference score than a negative difference score, $P = .035$. If we consider only the 4 first test trials, infants' mean difference score was .25, which was significantly greater than 0, $t(23) = 2.18$; $P = 039$; d' = .45. Fourteen infants obtained a positive difference score, six infants a negative difference score, and six infants a null difference score. A comparison of the number of infants who obtained a positive difference score and the number of infants who obtained a negative difference score was not significant, as evaluated by a binomial test, $P = 0.12$.

Considering the overall accuracy, infants' mean difference score was .13, which was not significantly different from 0; $t(25) = 1.33$; $P = .20$; d' = .26. Fifteen infants obtained a positive difference score, eight infants a negative difference score, and three infants a null difference score. A comparison of the number of infants who obtained a positive difference score and the number of infants who obtained a negative difference score was not significant, as evaluated by a binomial test, $P = 0.21$. If we consider only the 4 first test trials, infants' mean difference score was .22, which was marginally significant, $t(24) = 1.81$; $P = .072$; $d' = .36$ Fifteen infants obtained a positive difference score, five infants a negative difference score, and five infants a null difference score. A binomial test showed that significantly more infants obtained a positive difference score than a negative difference score, $P = .041$.

Thus, altogether, infants privileged the prediction made by consonants rather than that made by vowels.

2.3.6 Discussion

In this experiment, infants needed to learn that one word predicts a toy's appearance in one location, while another word predicts a different toy's appearance in another location. We further asked what prediction infants would make when presented with ambiguous words formed with the consonants of one of the previous words and the vowels of the other.

The observation of anticipatory looks in the familiarization did not show evidence of an increase in the number of correct anticipations. However, this absence of evidence should not be interpreted as infant's failure to learn the associations, as is shown by the test results. In fact, our paradigm is not designed to evaluate learning in the familiarization phase; it is designed to assess the participants' performance in the test phase. Due to limitations in executive functions, most infants probably learned to predict the location of the puppet's appearance for only one of the familiarization words (Kovacs & Mehler, 2009a, 2009b). Thus, the anticipations of infants for the word they have not learned, that is for half of the trials, are directed randomly. This results in noisy data in the familiarization phase. Moreover, infants may have learned not only to predict the location where a toy would appear, but also the timing of these appearances, as is suggested by the latency between the disappearance of the attention grabber and the occurrence of infants first fixation in test trials.

Nevertheless, the observation of first fixations and the overall accuracy in the test phase suggests that infants consider two words sharing consonants as more similar than two words sharing vowels. Namely, 12-month-olds found *kuku* more similar to *keke* than to *dudu*. A second interpretation is that infants solely relied on consonants when associating the familiarization words to either the side or the specific toy that appeared there. Our paradigm does not allow us to understand whether infants associated words to toys or to locations. In both cases, nevertheless they needed to store at least one specific word in memory. When encoding specific words in memory, 12-month-old infants appear to give a higher weight to consonants than to vowels.

2.3.7 CV hypothesis, Part I: Consonants to build the lexicon

The end of the first year of life coincides with the time when infants begin to develop their vocabulary. According to the MacArthur-Bates Communicative Development Inventory (CDI; Dale & Fenson, 1996) questionnaire studies, infants of that age can understand about 80 words. Fenson et al. (1994) evaluated that 12-month-olds learn about 2 words a week. This rate improves in the following months yielding 6-year-olds with a vocabulary of about 10000 words (Bloom & Markson, 1998; Miller, 1996). Showing that infants at 12 months already rely more on consonants for learning new words is therefore far from being anecdotal, as it is going to play a role in the acquisition of more than 99% of a child's vocabulary. Above, we showed that the words in French and Italian infants' vocabulary can be better discriminated on the basis of the information carried by consonants than of that carried by vowels. In fact, Keidel et al. (2007) attributed the origin of the lexical role of consonants to the comparative distribution of information carried by vowels and consonants in the lexicon, which would solely result from the larger number of consonants. However, militating against this hypothesis, evidence for the consonantal bias in the lexical processes are found in languages where consonants largely outnumber vowels, as well as in languages where the numbers of consonants and vowels are balanced (see Mehler et al., 2006; Toro et al., 2008a and Bonatti, Peña, Nespor & Mehler, 2007 for Italian and French). For instance, Cutler, Sebastián-Gallés, Soler-Vilageliu, & Van Ooijen (2000) showed that, when asked to change one phoneme to turn a non-word (e.g. *kebra*) into a known word, participants altered more often the vowel (thus generating *cobra*) than the consonant (generating *zebra*). These results hold both for speakers of Spanish, which has many more consonants than vowels, and for speakers of Dutch, which has a similar number of consonants and vowels. Furthermore, in reading tasks, consonants appear to be privileged for lexical access in various languages such as French (New, Araujo & Nazzi, 2008), English (Berent & Perfetti, 1995; Lee, Rayner & Pollastek, 2001) and Spanish (Carreiras, Gillon-

Dowens, Vergara & Perea, 2009). If the lexical statistics hypothesis fully explained the specialization of consonants for lexical access, a more important effect should have been observed in languages that have many more consonants than vowels. In contrast, we interpret the distribution of information in consonants and vowels as a consequence of the consonantal bias for lexical acquisition (see Bonatti et al., 2007).

The functional difference between consonants and vowels may originate from the fact that different processes have different requirements, and may consequently rely on different categories. Specifically, lexical memory may require more stable, thus reliable, categories that allow the learner to identify words and distinguish each word from other lexical entries. The lexical role of consonants may therefore be due to the categorical mode in which consonants are perceived. In fact, speech perception abilities change in the course of the first year of life. Infants are initially sensitive to all phonemic contrast that can be found in the languages of the world (Jusczyk, 1997; Mehler & Dupoux, 1994). By six months of age, however, infants display a perceptual magnet effect for vowel perception, which is due to the formation of a prototype for each of the vowels of the language of exposure (Kuhl, 1991; Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992). By the end of the first year of life, infants converge to the consonantal categories of the language of exposure, and have lost the sensitivity to non-native consonantal contrasts (Werker & Tees, 1984). Interestingly, even in adulthood, the sensitivity to within-category consonantal contrasts is practically lost, whereas the sensitivity to within-category vocalic contrasts is only diminished (Pisoni, 1973). As a consequence, vowel variations due to sentence prosody, speaker and dialectal accents may hinder word recognition to a greater extent than consonant variations. Therefore, lexical distinctions are better instantiated by consonantal contrasts than by vocalic contrasts. This view predicts that the formation of consonantal phonological categories may be necessary for the lexical role of consonants to emerge. In Experiment 2, we test younger infants to tackle this issue.

2.4 Constraints on word memory at 6-months: Experiment 2

In Experiment 1, we showed that the consonantal bias for word memory is already present by the end of the first year of life. In Experiment 2, we ask whether younger infants would show the same bias. Experiment 2 was thus identical to Experiment 1, except for the age of participants.

2.4.1 Participants

Sixteen infants were included in the analysis; age range 05 months 28 days to 06 months 24 days (average 6 months 09 days). Four other infants participated in the study but were excluded due to fussiness.

2.4.2 Results

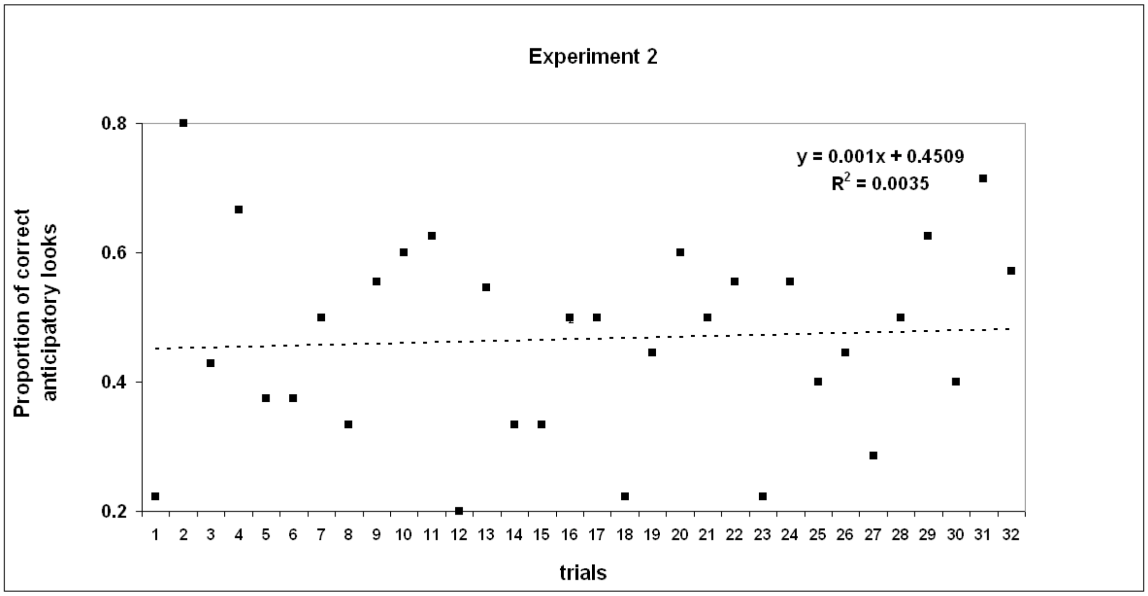

Figure 2.6 - Proportion of correct anticipatory looks for each familiarization trial in Experiment 2. The dotted line depicts the corresponding linear regression.

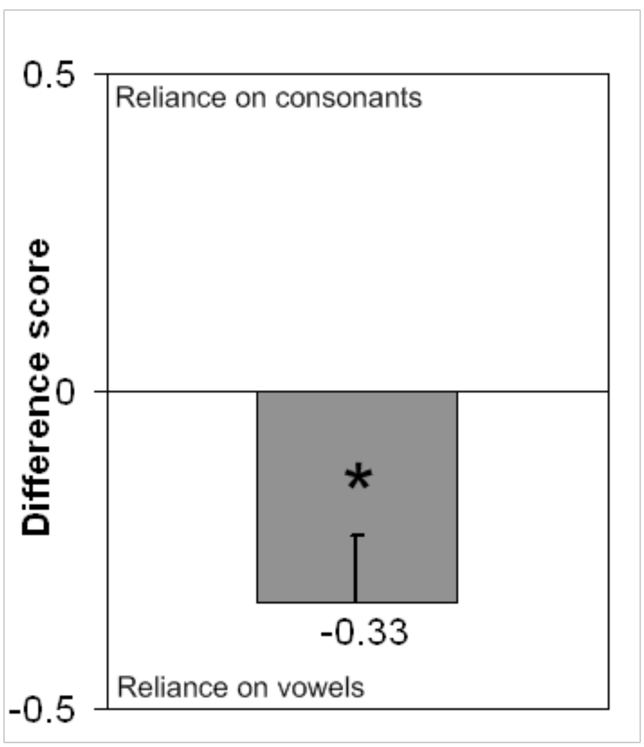

Figure 2.7 - Mean difference score for Experiment 2 considering the first fixations. Infants looked more at the side predicted by vowels. Error bars represent standard error.

The first fixation data for familiarization trials is presented in Figure 2.6. Infants anticipated to one or the other side in 54% of the trials. We computed the proportion of infants showing a correct anticipatory look for each trial. A linear regression analysis gave non-significant results ß=.0010, R^2=.0035, t(30)=. 33, p>.74.

In the test phase, infants looked to the left or the right in 63% of the trials. Their first fixations occurred slightly earlier when they were located to the consonant side (872 ms after the end of the word) than when they targeted the vowel side (880 ms after the end of the word). However this difference was not significant, t(11) = .84; *P* = .42.

Considering first fixations (see Figure 2.7), infants' mean difference score was -.31, which was significantly lower than 0, t(15) = -2.49; *P* = .025; d’ = .622. Four infants obtained a positive difference score, eight infants a negative difference score, and four infants a null difference score. A comparison of the number of infants who obtained a positive difference score and the number of infants who obtained a negative difference score was not significant, as evaluated by a binomial test, *P* = 0.39. If we consider only the 4 first test trials, infants' mean difference score was -.33, which was significantly lower than 0, t(14) = -3.13; *P* = .007; d’ = .807. One infant did not provide data in the first four tests. One infant obtained a positive difference score, nine infants a negative difference score, and five infants a null difference score. A binomial test showed that significantly more infants obtained a negative difference score than a positive difference score, *P* = .0215.

Considering the overall accuracy, infants’ mean difference score was -.07, which was not significantly different from 0; t(15) = -.50; *P* = .62; d’ = .126. Seven infants obtained a positive difference score, six infants a negative difference score, and three infants a null difference score. A comparison of the number of infants who obtained a positive difference score and the number of infants who obtained a negative difference score was not significant, as evaluated by a binomial test, $P = 1$. If we consider only the 4 first test trials, infants' mean difference score was -.13, which was not significantly different from 0, $t(14) = -.88$; $P = .39$; $d' = .228$. Three infants obtained a positive difference score, eight infants a negative difference score, and three infants a null difference score. A comparison of the number of infants who obtained a positive difference score and the number of infants who obtained a negative difference score was not significant, as evaluated by a binomial test, $P = .23$.

2.4.3 Discussion

Experiment 2 was identical to Experiment 1, except that participants were aged 6-months instead of 12-months. As in Experiment 1, we did not observe a significant increase of correct anticipations in the familiarization phase. However, in the test phase, infants' first fixations suggest that 6-month-old infants relied on vowels to predict the location of toy's appearance. That is, when hearing "keke", they searched for the puppet on the side previously associated with "dede", rather than that previously associated with "kuku". This behavior, however, was not confirmed by the analysis of overall accuracy, thus suggesting that infants' prediction on the location of the puppets' appearance is not very strong.

These results contrast with the behavior of 12-month-olds in Experiment 1, who relied on consonants. The behavior of 6-month-olds is, however, congruent with that of younger infants. Indeed, newborns are able to extract the common vowels of a series of syllables (*bi, si, li, mi*), but not the common consonant (*bi, ba, be, bu*) (Bertoncini et al., 1988). Furthermore, newborns can also discriminate between bi-syllabic and tri-syllabic words, e.g. *kepa* vs. *kesopa*, but not between bi-syllabic words varying in the number of consonants, e.g. *rifu* vs. *suldri*, (Bijeljac-Babic et al., 1993) or in the number of moras (Bertoncini et al., 1995), thus suggesting they track the number of syllables or vowels in a speech sequence. More recently, Benavides-Varela and colleagues showed that the information neonates encode when exposed to a repeated bisyllabic word is mainly carried by vowels (Benavides-Varela, Hochmann, Macagno, Nespor & Mehler, in prep.). Using Near-Infrared Spectroscopy, they first exposed infants to one repeated bisyllabic word (e.g. "titi") for six minutes. After a silent pause of two minutes, half of the neonates heard a test word differing from the familiarization word in its consonants ("sisi") whereas the other half heard a test word differing in its vowels ("tata"). The NIRS technique allows to observe the variation of oxy- and desoxy- hemoglobin, signaling cortical activation. In previous studies, Benavides-Varela and colleagues had shown that the presentation in the test of the same word as in the exposure phase results in a decrease of oxy-hemoglobin, while the presentation of a novel word results in an increase of oxy-hemoglobin (Benavides-Varela, Gomez, Bion, Macagno & Mehler, submitted). The presentation of the test word with the same consonants as the exposure word resulted in an increase of oxy-hemoglobin in the right frontal areas, and the presentation of the test word with the same vowel as the exposure word resulted in a decrease of oxy-hemoglobin in the right frontal areas. Thus, these results suggest that newborns consider *titi* as similar to *sisi*, but different from *tata*. They thus rely on the information carried by vowels when encoding a word in memory.

The pattern of results observed in newborn studies, in our experiments with 6- and 12-month olds, and in adult studies suggests continuity in the mode of representation of speech sequences between neonates and 6-month-olds on one hand, and between 12-month-olds and adults, on the other hand. A discontinuity is observed between 6- and 12-month-olds.

The emergence of the consonantal bias in word representation may depend on brain maturation or some learning processes. Havy, Bertoncini and Nazzi (submitted) explored the respective roles of input and brain maturation testing the presence of a consonantal bias in young children with cochlear implants. Interestingly, they found that children who had just been implanted (on average at 47-months) privileged vocalic information when encoding a word in memory, just like neonates and six-month-olds. Following implantation, however, children's performance in using consonantal contrasts correlated positively with the duration of implant, while the performance with vowels did not correlate significantly (Havy, 2009; Havy, Bertoncini & Nazzi, submitted). This suggests that the emergence of the consonantal bias results at least in part from exposure to the speech input.

As discussed in the previous chapter, word learning in the final state requires a stable and reliable format of representation that should be both abstract enough to recognize words across voices, speech rates and intonations; and precise enough to distinguish between minimal pairs of words (e.g. *boy, toy*). Consonantal phonemic categories appear to fill these requirements better than vowel categories; they are more stable, being less contaminated by prosodic variations, and more reliable as the sensitivity to within-category consonantal contrasts is practically lost in adulthood, whereas the sensitivity to within-category vocalic contrasts is only diminished (Pisoni, 1973).

Speech carries the information necessary to acquire the consonantal categories of the surrounding language. Maye, Werker & Gerken (2002) showed that these categories might be acquired by observing the distribution of different consonantal tokens in the input. They exposed two groups of 6- to 8-month-old infants to different tokens along a continuum between the prototypes of the syllables "da" and "ta". One group of infants was exposed to a unimodal representation, hearing more instances of tokens halfway between "ta" and "da". The other group was exposed to a bimodal distribution, hearing more tokens corresponding to tokens closer to one or the other of the prototypes. Importantly, all infants heard as often the tokens corresponding to the prototypes. After 2.3 minutes of exposure, only those infants exposed to the bimodal distribution could discriminate between the tokens corresponding to the prototypes of "da" and "ta" (see also Maye, Weiss and Aslin 2003; Yoshida, Pons & Werker, 2006). This suggests that the unimodal distribution leads infants to form one category that include both "da" and "ta" tokens, while the bimodal distribution leads to the constitution of two different categories.

Thus, we propose that the consonantal-bias in word representations observed from 12-months onwards results from the acquisition of phonemic consonantal categories, which is completed around that age (Werker & Tees, 1984). Younger infants are probably more influenced by the perceptual saliency of different speech sound categories, relying on the sounds that carry more energy, i.e. vowels. This advantage for vowels in word representation may be reinforced by the acquisition of prototypes for native vowel categories around 5-to 6-months of age (Kuhl, 1991; Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992). The magnet effect resulting from these prototypes may give higher reliability to vowel representations.

2.5 Generalizing structural relations at 12-months: Experiment 3

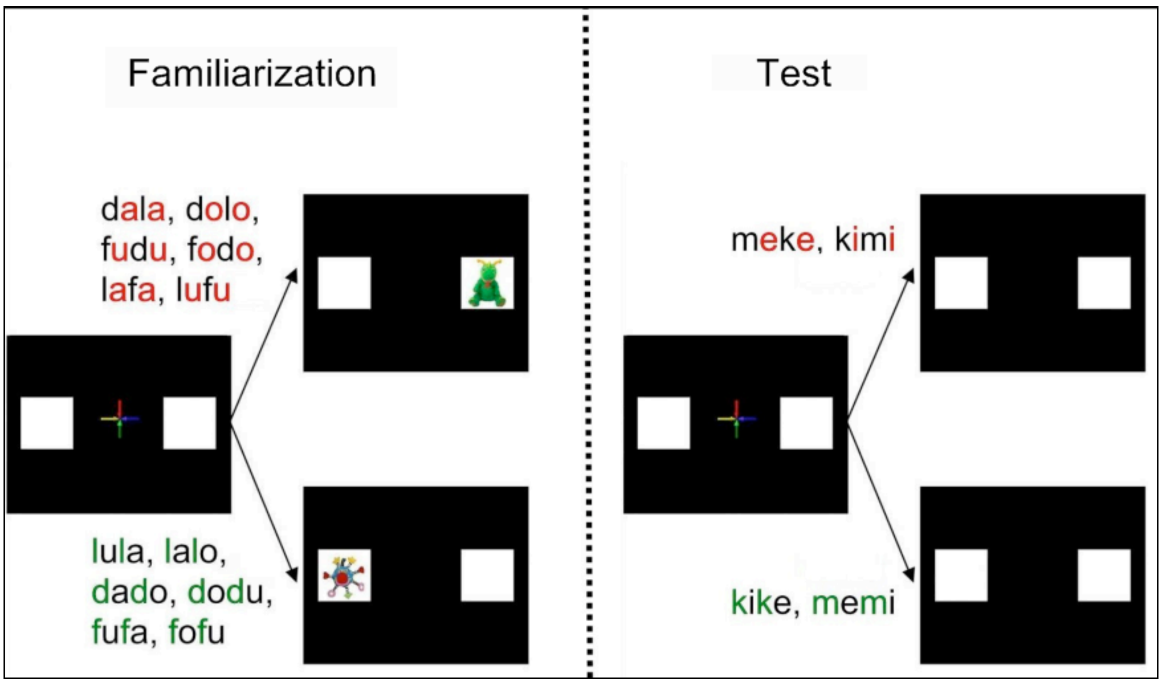

Figure 2.8– Paradigm of Experiments 3 and 4. Participants took 32 familiarization trials and 8 test trials.

Experiment 3 tests the second part of the CV hypothesis (see Figure 2.8), asking whether infants find it easier to learn and generalize regularities defined over vowels or over consonants. We use the paradigm developed by Kovács (2008) and Kovács & Mehler (2009b). Monolingual 12-month-olds have trouble learning simultaneously two regularities in that paradigm, probably due to limitations in executive functions (Kovács & Mehler, 2009b). Indeed, they just learn the simpler one (e.g., adjacent rather than non-adjacent syllable repetition, AAB vs. ABA; repetition rather than absence of repetition, AA vs. AB; Kovács, 2008). Thus, our paradigm allows us to test in a within-subject design which of two structural relations is easier for infants to detect and generalize. In our experiment, we teach infants two regularities, consonant repetition and vowel repetition in order to determine which is easier for 12-month-olds to learn. Each of six words containing a consonant repetition (i.e., *lula*, *lalo*, *dado*, *dodu*, *fufa* and *fofu*) were followed by a toy, which appeared on one side of the screen, whereas each of six words containing a vowel repetition (i.e., *dala*, *dolo*, *fodo*, *fudu*, *lafa* and *lufu*) were followed by a toy, which appeared on the other side of the screen. We then tested for generalization, asking where infants would search for the toy when hearing novel words respecting either the consonant repetition regularity (i.e., *kike* and *memi*) or the vowel repetition regularity (i.e., *meke* and *kimi*), using novel vowels and consonants that did not appear during the familiarization. These two regularities are strictly equivalent in terms of complexity (a simple repetition), varying only in the category that carries the repetition. The CV hypothesis predicts that the generalization of a repetition regularity should be easier if it is implemented over vowels than if it is implemented over consonants. Therefore, when learning only one regularity, infants should learn and generalize the vowel repetition rather than the consonant repetition.

Experiment 3, contrary to Experiment 1 was not designed as a word learning experiment but as an experiment to elicit the generalization of a pattern. In that type of experiments, a variety of generalizations are possible from the limited set of examples that are provided to the participants. To avoid the position of specific phonemes providing a cue for generalizations, the same consonants and vowels were used to create the six consonant-repetition and the six vowel-repetition items used in the familiarization. Moreover, three puppets could appear in each location and were randomly paired with the items, so that the attention of infants could not be attracted by a particular puppet and an item associated to it. In the test phase, we asked whether infants could generalize these associations to novel items formed with novel consonants and vowels, instantiating the repetition structures. Prediction about the location of a toy's appearance in the test trials could be done only if one focused on the identity relation between consonants or vowels. Due to limitations in executive functions, monolingual 12-month-olds usually learn to generalize only one structure in this paradigm, the easier one. We thus ask whether vowel-repetition or consonant-repetition is easier for 12-month-olds to generalize.

2.5.1 Participants

Twenty-four infants were included in the analysis; age range 11 month 15 days to 12 month 15 days. Six other infants participated in the study but were excluded due to fussiness (3) or equipment failure (3).

2.5.2 Stimuli

The stimuli in Experiment 3 consisted in bisyllabic items. These items could have either repeated consonants or repeated vowels. For the familiarization, six items containing a consonant repetition (*lula*, *lalo*, *dado*, *dodu*, *fufa* and *fofu*) and six items containing a vowel repetition were created (*dala*, *dolo*, *fodo*, *fudu*, *lafa* and *lufu*). The same three consonants and three vowels were used to generate both sets of items. For the test, four novel items were generated with novel consonants and novel vowels. Two test items had a consonant repetition (*kike* and *memi*) and two test items had a vowel repetition (*meke* and *kimi*). Items were synthesized with MBROLA (fr4) with phoneme durations of 120 ms and a monotonous pitch of 200 Hz. There was no silent pause between two syllables within an item.

Visual stimuli were three pictures of colorful toys. These appeared inside one of two white squares on the left or right side of the screen. The toys loomed from 4 cm to 7 cm inside the squares for 2 s. The squares had a side-length of 8 cm, positioned at a distance of 13.5 cm. Toys and items were paired randomly.

2.5.3 Procedure

The procedure was identical to that of Experiment 1, except for the speech items infants heard (see Figure 2.8) and the toys they saw. In Experiment 3, a consonant repetition predicted the toys' appearance in one of the squares, while a vowel repetition predicted the toys' appearance in the other square. The pairing of structures with toy-locations was counterbalanced across participants.

2.5.4 Analysis

For the analysis, we divided the screen into three equal parts, left, middle and right. In each trial of the familiarization, we measured the proportion of infants anticipating the toy's appearance to the correct side. In the test, we measure each participant's first fixation and the time spent fixating the left or the right of the screen, after hearing the novel item and before the beginning of the next trial. We also computed the infants' overall accuracy (Kovacs & Mehler, 2009b; McMurray & Aslin, 2004). That is, trials were scored as correct if the infant looked longer to the correct side within the 2 s after hearing a noel item and before the start of the next trial. We computed difference scores: (#correct looks - #incorrect looks)/(#correct looks + #incorrect looks) for first fixations and for overall accuracy, and computed a t-test to compare them to the chance level of 0. Significantly positive difference scores would indicate that infants learned and generalized the regularity.

2.5.5 Results

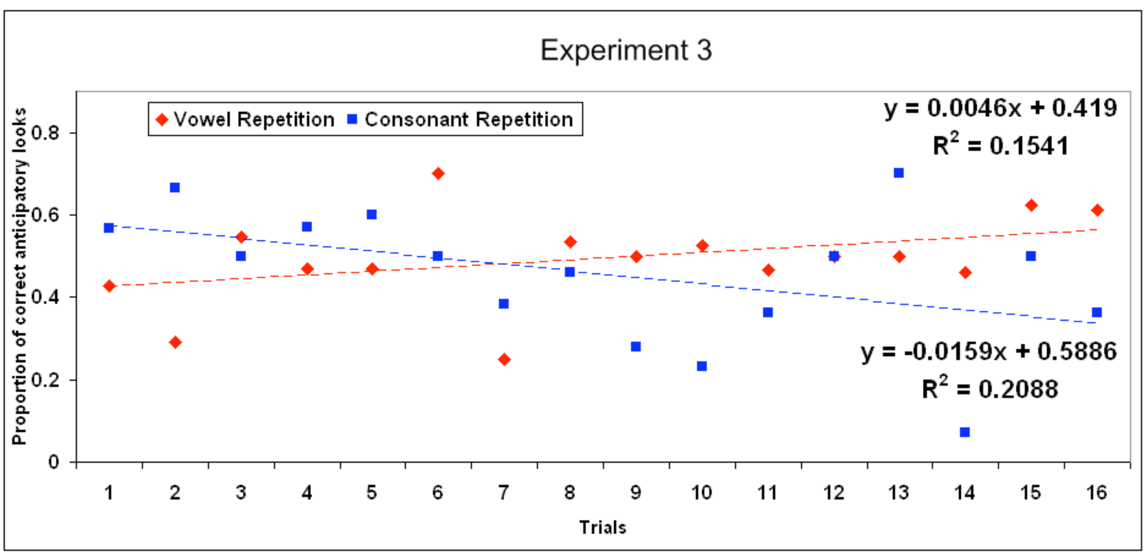

Figure 2.9 - Proportion of correct anticipatory looks for each familiarization trial in Experiment 3. The dotted lines depict the linear regression for vowel repetition (red) and for consonant repetitions (blue), respectively.

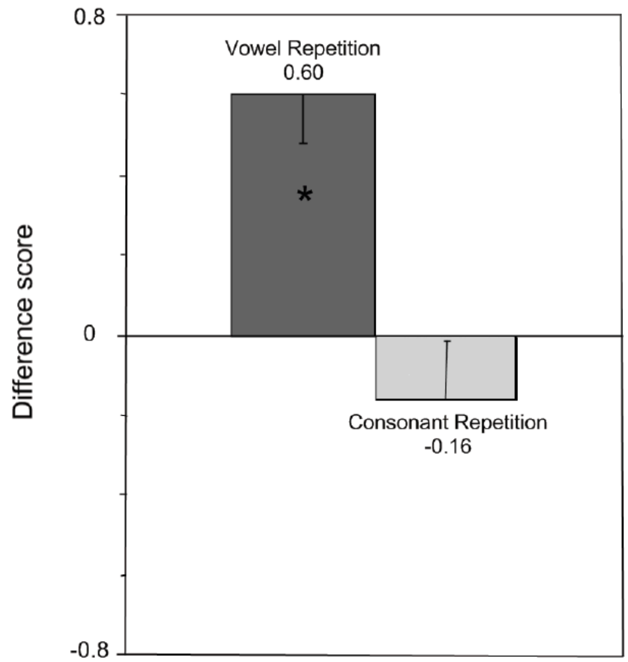

Figure 2.10 - Mean difference scores for Experiment 3 considering the first fixations. Infants looked more at the correct side predicted by the regularity for the vowel repetition, but not for the consonant repetition. Error bars represent standard errors.

Familiarization results are presented in Figure 2.9. Infants anticipated to one or the other side in 56% of the trials. We computed the proportion of infants showing a correct anticipatory look for each trial. A linear regression analysis yielded marginally significant results for the Consonant-repetition, ß=-.0159, R^2=.21, t(14)=-1.92, p=.075, and a non significant trend for the Vowel-repetition, ß=.0046, R^2=.15, t(14)=1.60, p=.133. Qualitatively, the proportion of correct anticipatory looks increased during familiarization for the vowel repetition and decreased for the consonant repetition. Thus, this suggests that infants learned the association between a Vowel repetition and the predicted location but did not learn the association between a Consonant repetition and the predicted location.

The test results are presented in Figure 2.10. Infants looked to the left or the right in 62% of vowel repetition test trials, and in 61% of the consonant repetition test trials. Two infants did not provide data in the vowel repetition tests, so that 24 infants were included in the analysis of the consonant repetition tests and only 22 in the analysis of the vowel repetition tests.

Paired t-test showed that infants obtained significantly higher difference scores for the vowel repetition than for the consonant repetition, considering the first fixations, $t(21) = 4.56$, $P < .0002$, d' = .97; and the overall accuracy, $t(21) = 4.29$, $P < .0004$, d' = .91.

Considering first fixations in vowel repetition test trials, Infants' mean difference score was .60 for the vowel repetition, which was significantly greater than 0, $t(21) = 4.92$; $P < .0001$; d' = 1.05. Seventeen infants obtained a positive difference score, 2 infants a negative difference score, and 3 a null difference score. A binomial test showed that significantly more infants obtained a positive difference score than a negative difference score, $P < .001$. For the consonant repetition, infants' mean difference score was -0.16, which did no differ from chance, $t(23) = -1.06$; $P = .30$; d' = .22. Eight infants obtained a positive difference score, 11 infants a negative difference score, and 5 a null difference score. This distribution did not differ from chance, $P = .144$.

Finally, considering the overall accuracy, infants' mean difference score was .53 for the vowel repetition, which was significantly greater than 0; $t(21) = 5.23$; $P < .0001$; d' = .1.11. Fifteen infants obtained a positive difference score, one infant a negative difference score, and six a null difference score. A binomial test showed that significantly more infants obtained a positive difference score than a negative difference score, $P < .001$. Infants' mean difference score was -.07 for the consonant repetition, which did not significantly differ from chance, t(23) = -.52; *P* = .61; d' = .11. Six infants obtained a positive difference score, nine infants a negative difference score, and nine a null difference score. This distribution did not differ from chance, *P* = .61.

2.5.6 Discussion

The CV hypothesis predicts that generalizing structural relations should be easier over vowels than over consonants. Experiment 3 directly tested this claim by confronting infants with two competing regularities. Given the limited cognitive capacities of 12-month-olds (Kovacs & Mehler, 2009a, 2009b), participants were expected to learn and generalize only one of the patterns. The results suggest that infants learned the association between the vowel repetition regularity and the predicted location of a toy's appearance. Moreover, they could extend this association to new, never heard, items formed with novel consonants and vowels, thus showing that infants extracted an abstract property from the familiarization. In contrast, they showed no evidence of learning and generalizing the consonant repetition regularity. Thus, the same regularity (i.e., a repetition[3]) is easier for 12-month-old infants to learn over vowels than over consonants. Vowels rather than consonants appear to be a privileged category for extracting and generalizing structural relations.

[3] Note that, in our stimuli, an alternative description of the vowel repetition regularity would be a NON-repetition of consonants. Indeed, rather than generalizing the repetition of the vowel, generalizing the pattern that the first and second consonants should differ would yield similar results. However, we consider this possibility unlikely, as Kovács (2008) showed how both 7- and 12-month-olds found easier to learn and generalize a syllable repetition regularity rather than a non-repetition regularity.

Inspired by the prosodic bootstrapping accounts of syntax acquisition (Morgan & Demuth, 1996), the CV hypothesis predicted that structural relations should be easier to generalize over vowels than over consonants. In Experiment 3, we show in a within-subject design that 12-month-old infants are better at extracting a repetition-based regularity over vowels than over consonants. Moreover, ours is the first experiment to show that infants can generalize the vowel-repetition regularity to completely novel words, formed with vowels and consonants that did not appear in the familiarization. This result indicates a special role for vowels to signal structural regularities. Obviously, we are not claiming that all linguistic regularities are acquired through the detection and generalization of repetition patterns. However, the acquisition of syntax, like the repetition structure, requires the ability to generalize structural relations (e.g. the relations between verbs and objects) that cannot reduce to a statistical regularity and thus require computations that go beyond the reach of memory and statistical computations (Chomsky, 1957; Marcus, 1998).

Experiment 3 illustrates that infants are better at generalizing structural relations over vowels. Consequently, they ought to be capable of extracting the structural information carried by vowels, including prosodic information. For example, they may quickly learn and generalize prominence alternations signaled by pitch (Bion, Benavides Varela & Nespor, in press) or by duration, which may allow them to learn the order of heads and complements in their language (Nespor et al., 2008). In addition to syntactic structural relations, infants must also extract structural regularities for other components of the linguistic systems, such as phonology, phonotactics and morphology. Thus, the special role of vowels for signaling structural relations may be useful in several domains of language acquisition.

2.5.7 A word on the mechanisms involved in repetition generalization

Our results showing that 12-month-old infants generalize better repetition structures over vowels than over consonants may inform us on the mechanisms underlying repetition generalization. Two views have been proposed to explain such generalization. Marcus and colleagues (1999) initially proposed that generalizing the structures AAB, ABB and ABA requires the use of symbolic computations, thus learning the algebraic-like rules that can generate such structures. Alternatively, Endress and colleagues (Endress et al., 2007; Endress, Nespor & Mehler, 2009) proposed that these results could be explained if one assumes the existence of a perceptual primitive that is sensitive to identity relation. That later interpretation was supported by the demonstration that adult participants easily learn and generalize repetition-based structures over tones (i.e. ABB and ABA), but perform poorly on an equally complex melodic structure, which does not rely on repetition, i.e. middle-high-low or low-high-middle tone sequences.

The two models thus contrast in the nature of the representations over which a repetition pattern is generalized. While the rule model considers that generalization operates over symbolic representations, the perceptual primitive model views generalization as operating at a perceptual level. To disentangle between the two models, we may thus ask what type of representations are involved in these computations.

To our knowledge, only one experiment has addressed the question of the level of representation involved in repetition generalization. Kovacs (2008) showed that 7-month-old infants could generalize an AAB structure even when the A syllables where pronounced at different pitch (200Hz and 100Hz) and where therefore not acoustically identical. These syllables could be found equal only at a higher level of representation. However, this does not mean either that the generalization is computed over a phonological representation of the input; first because 7-month-old infants have not yet acquired the consonantal phonemic categories (Werker & Tees, 1984), second because humans and non-human species perceive the similarity between sounds separated by one octave (Blackwell & Schlosberg, 1943; Demany & Armand, 1984) and are able to generalize their response to an acoustic tune, to the same tune played one or two octaves above (Deutsch, 1972; Wright, Rivera, Hulse, Shyan, & Neiworth, 2000).

Our results further inform us on the type of representations involved in two ways. First, Experiment 1 shows that consonants are better represented than vowels in word representations. If the generalization acted upon word representation, it should therefore privilege consonants instead of vowels. Second, by 12-months of age, infants have acquired phonemic representations, which are symbolic representations of consonantal categories. The representation of vowels on the other hand appears to rely on prototypes rather than abstract symbolic representations. Algebraic-like symbolic computations should therefore rely on consonant rather than vowel representations.

Therefore, even though these experiments were not designed to assess the type of computations and representations involved in repetition generalization, their results suggest that such generalization does not result from abstract symbolic computations, and do not act at the word representation level, or at the phonemic level. Kovacs (2008) showed that it does not act either at a purely acoustic level. Altogether, these results favor the perceptual primitive model (Endress, Nespor & Mehler, 2009), suggesting that repetition structures are generalized over low-level perceptual representations.

2.6 Generalizing structural relations at 6-months: Experiments 4

In Experiment 3, we showed that the specialization of vowel for implementing structural relations is already in place by end of the first year of life. In Experiment 4, we ask whether is it in place earlier, testing 6-month-olds. Experiment 4 is thus identical to Experiment 3, except for the age of participants.

2.6.1 Participants

Seventeen infants were included in the analysis; age range 5 month 25 days to 6 month 30 days (average 6 months 17 days). Ten other infants participated in the study but were excluded due to fussiness (6), equipment failure (3), or the mother not following instructions (1).

2.6.2 Results

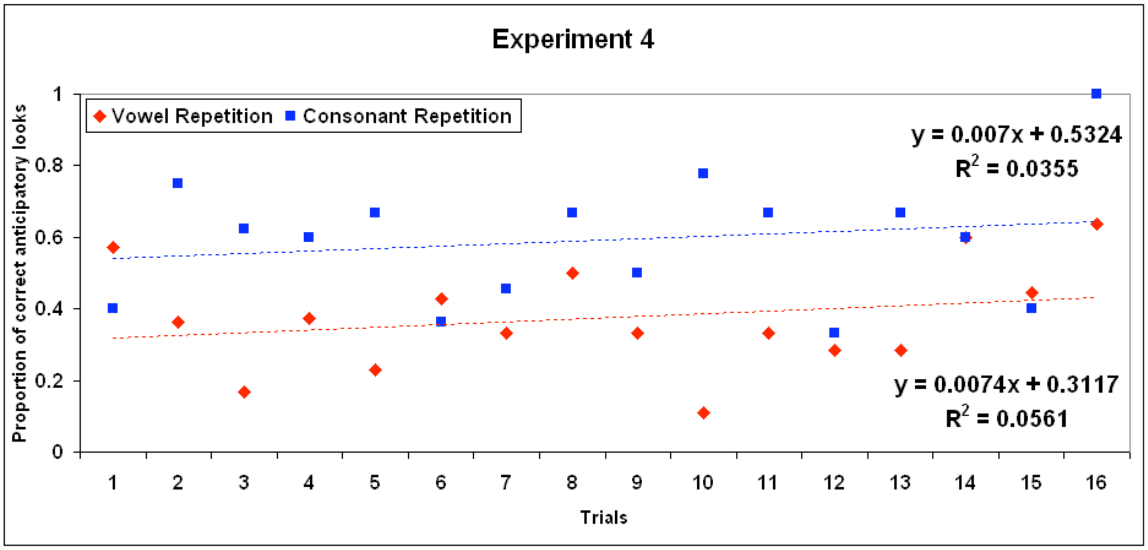

Figure 2.11 - Proportion of correct anticipatory looks for each familiarization trial in Experiment 4. The dotted lines depict the linear regression for vowel repetition (red) and for consonant repetitions (blue), respectively.

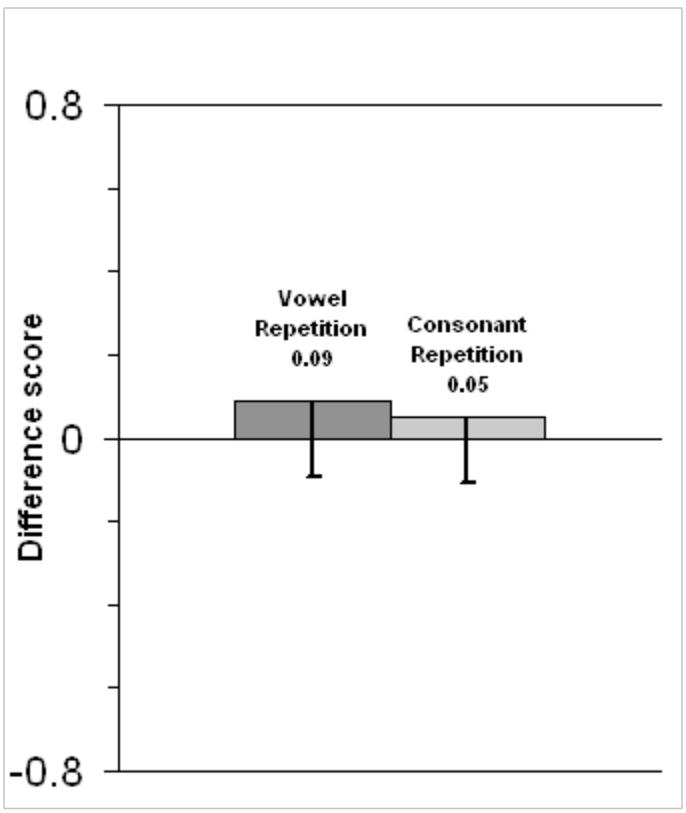

Figure 2.12 - Mean difference scores for Experiment 4 considering the first fixations. Infants showed no evidence of generalizing either the vowel repetition or the consonant repetition structures. Error bars represent standard errors.

Familiarization results are presented in Figure 2.11. Infants anticipated to one or the other side in 47% of the trials. We computed the proportion of infants showing a correct anticipatory look for each trial. A linear regression analysis gave non-significant results for the Consonant-repetition, ß=.007, R^2<.04, t(14)=.72, *P* = .48, and for the Vowel-repetition, ß=.0074, R^2<.06, t(14)=.91, *P* = .38.

The test results are presented in Figure 2.12. No evidence was found that infants generalized one or the other structure. Infants looked to the left or the right in 73% of vowel repetition test trials, and in 66% of the consonant repetition test trials. Two infants did not provide data in the vowel repetition tests, so that 17 infants were included in the analysis of the consonant repetition tests and only 15 in the analysis of the vowel repetition tests.

Paired t-test showed that infants' difference scores did not differ significantly for the consonant-repetition and the vowel-repetition test items, considering either the first fixations, t(14) = .34, *P* = .74, d' = .09; or the overall accuracy, t(14) = .25, *P* < .80, d' = .07.

Considering first fixations in vowel repetition test trials, Infants' mean difference score was .09 for the vowel repetition, which did not differ significantly from 0, t(14) = .49; *P* = .63; d' = .13. Six infants obtained a positive difference score, 6 infants a negative difference score, and 3 a null difference score. For the consonant repetition, infants' mean difference score was .05, which did no differ from chance, t(16) = .34; *P* = .74; d' = .08. Six infants obtained a positive difference score, five infants a negative difference score, and six a null difference score.

Finally, considering the overall accuracy, infants' mean difference score was .07

for the vowel repetition, which did not significantly differ from 0; $t(14) = .37$; $P = .72$; d' = .095. Six infants obtained a positive difference score, six infants a negative difference score, and three a null difference score. Infants' mean difference score was .05 for the consonant repetition, which did not significantly differ from chance, $t(16) = .33$; $P = .74$; d' = .08. Six infants obtained a positive difference score, six infants a negative difference score, and five a null difference score.

2.6.3 Discussion

In this experiment, we found no evidence that six-month-old infants could learn either of the two regularities that were presented to them. This pattern of results is ambiguous. As we explained in the previous chapters, probably due to limitations in executive functions (see Kovacs, 2008; Kovacs & Mehler, 2009b), infants usually learn to predict only one location in our paradigm, the one associated to the simplest regularity. This allowed us to conclude that vowel-repetition is simpler than consonant-repetition for 12-month-olds. However, a global failure in this paradigm may have three explanations.

First, both structures may be two complex for 6-month-olds to generalize. Second, infants may be able to generalize both structures, but none is rated simpler than the other. Some infants may thus generalize the vowel-repetition, while others generalize the consonant repetition. This may yield non-significant results when averaging infants' scores together. Finally, infants may be able to generalize both structures, but may not represent speech sequences in a format allowing them to distinguish between the two structures. For example, if infants do not represent vowels and consonants as different speech sound categories, they might consider all items in our experiment as instantiating repetitions of the same kind, and would observe no consistency in the location of the toys' appearances.

Kovacs (2008) showed that 7-month-olds could generalize syllable-repetition, when contrasted with bisyllabic words exhibiting no repetition. Even though her participants were in average 4 weeks older than ours (i.e. 7 months and 22 days), her results suggest that infants that young are able to generalize a repetition. However, they may not be able to do so on subsyllabic units, such as consonants and vowels. Experiment 5 aims at testing whether 6-month-old infants are able to generalize the vowel-repetition structure, when contrasted in the familiarization with bisyllabic words exhibiting no repetition at all. In Experiment 6, we ask whether 6-month-old infants are able to discriminate between vowel-repetition and consonant-repetition exemplars.

2.7 Generalizing vowel-repetition at 6-months: Experiments 5 and 6

2.7.1 Experiment 5

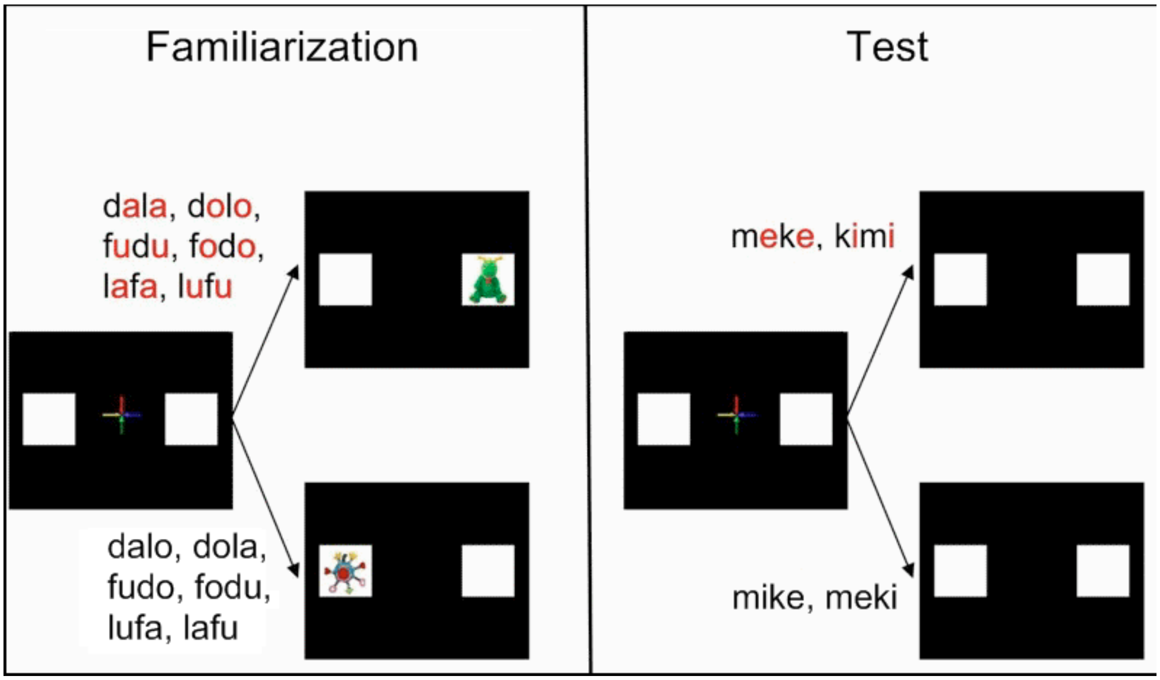

Figure 2.13 – Paradigm of Experiment 5. Participants took 32 familiarization trials and 8 test trials.

Experiment 5 aims at testing whether 6-month-old infants can generalize the repetition structure when it is instantiated over a subsyllabic unit, i.e. vowel-repetition. The paradigm of Experiment 5 (see Figure 2.13) was the same as that of Experiment 4, except for some of the stimuli used in familiarization and test. To facilitate infants' task in comparison to Experiment 4, infants participating in Experiment 5 are presented with two types of items in the familiarization. Items exemplifying the vowel-repetition structure predict the appearance of a toy in one location of the screen, and items that do not instantiate *any* repetition (neither vowel not consonant repetition) predict the appearance of a toy in the other location. In the test, we ask whether infants predict the location where a toy should appear when hearing novel vowel-repetition items and novel items without repetition, built with vowels and consonants that were not used in familiarization.

2.7.1.1 Participants

Twenty infants were included in the analysis; age range 5 month 24 days to 6 month 11 days (average 6 months and 0 day). Three other infants participated in the study but were excluded due to fussiness (3).

2.7.1.2 Stimuli

The stimuli in Experiment 5 consisted in bisyllabic items. In the familiarization, these items could have either repeated vowels as in Experiments 3 and 4 or not. Six items containing a vowel repetition (*dala*, *dolo*, *fodo*, *fudu*, *lafa* and *lufu*) and six items containing no repetition at all (*lufa*, *lafu*, *dalo*, *dola*, *fudo* and *fodu*) were created. The same three consonants and three vowels were used to generate both sets of items. For the test, four novel items were generated with novel consonants and novel vowels. In the test, we presented novel words with a vowel repetition (i.e. *meke* and *kimi*) as in Experiments 3 and 4, or without any repetition (i.e. *meki* and *mike*). Items were synthesized with MBROLA (fr4) with phoneme durations of 120 ms and a monotonous pitch of 200 Hz. There was no silent pause between two syllables within an item.

2.7.1.3 Results and discussion

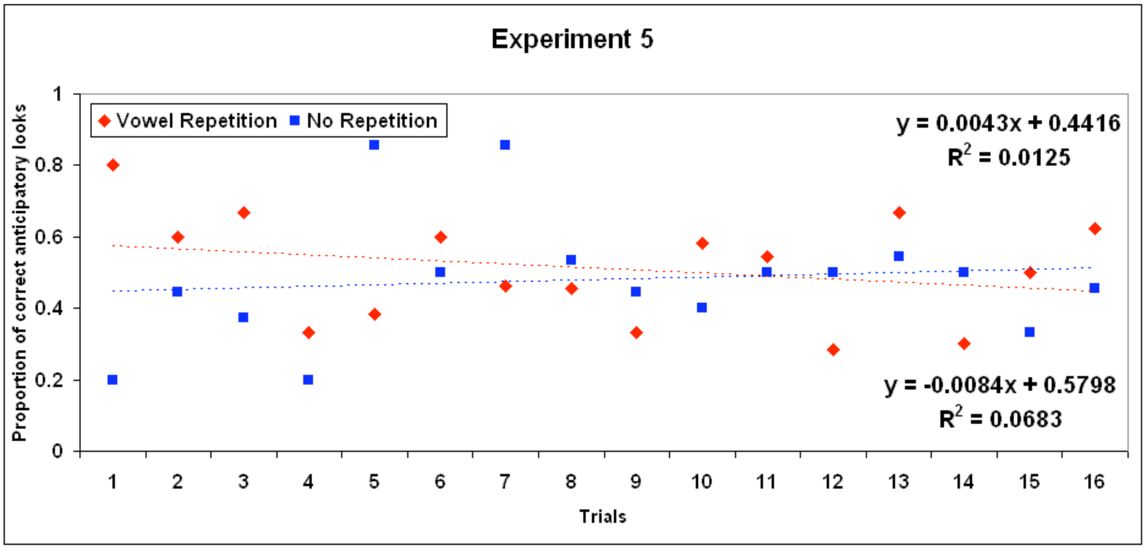

Figure 2.14 - Proportion of correct anticipatory looks for each familiarization trial in Experiment 5. The dotted lines depict the linear regression for vowel repetition (red) and for no-repetition (blue) trials, respectively.

Familiarization results are presented in Figure 2.14. Infants anticipated to one or the other side in 52% of the trials. We computed the proportion of infants showing a correct anticipatory look for each trial. A linear regression analysis gave non-significant results for the Vowel-repetition trials, ß=-.008, $R^2<.07$, t(14)=-1.01, *P* = .33, and for the no-repetition trials, ß=.004, $R^2<.013$, t(14)=.42, *P* = ..68.

The test results are presented in Figure 2.15. Infants looked to the left or the right in 71% of vowel repetition test trials, and in 64% of the consonant repetition test trials. Two infants did not provide data in the vowel repetition tests, so that 18 infants were included in the analysis of the vowel repetition tests and 20 in the analysis of the no-repetition tests.

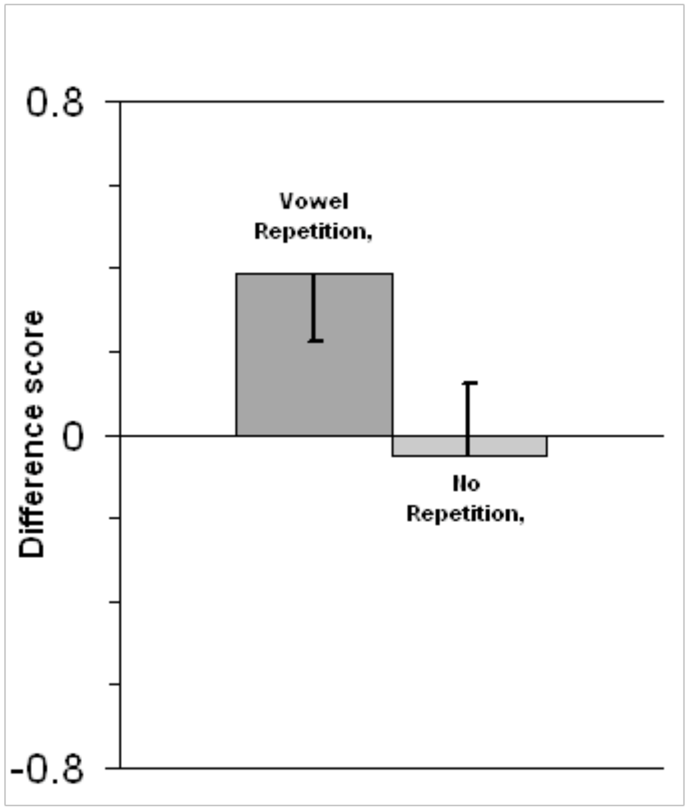

Figure 2.15 - Mean difference scores for Experiment 5 considering the first fixations. Infants looked more at the correct side predicted by the regularity for the vowel repetition, but not for the no-repetition Error bars represent standard errors.

Considering first fixations in vowel repetition test trials, Infants' mean difference score was .39 for the vowel repetition, which was significantly greater than 0, t(17) = 2.36; *P* = .03; d’ = .56. Ten infants obtained a positive difference score, 4 infants a negative difference score, and 4 a null difference score. This distribution did not differ from chance, *P* = .18. For the no-repetition trials, infants' mean difference score was -0.05, which did not significantly differ from 0, t(19) = -2.69; *P* = .014; d’ = .59. Seven infants obtained a positive difference score, nine infants a negative difference score, and four a null difference score. This distribution did not differ from chance, *P* = .80.

Finally, considering the overall accuracy, infants’ mean difference score was .27

for the vowel repetition, which was not significantly greater than 0; $t(17) = 1.61$; $P = .13$; d' = .38. Nine infants obtained a positive difference score, four infants a negative difference score, and five a null difference score. This distribution did not differ from chance, $P = .27$. Infants' mean difference score was -.06 for the no-repetition trials, which did not significantly differ from 0, $t(19) = -.34$; $P = .74$; d' = .08. Seven infants obtained a positive difference score, 10 infants a negative difference score, and 3 a null difference score. This distribution did not differ from chance, $P = .63$.

The analysis of first fixations suggests that 6-month-old infants are able to generalize the vowel-repetition structure. The overall accuracy sowed qualitatively similar results, but those were not significant. Thus, there remain two possible accounts for the results of Experiment 4. First, infants may be able to generalize both vowel-repetition and consonant-repetition, but none is rated simpler than the other. Second, infants may be able to generalize both structures, but may not represent speech sequences in a format allowing them to distinguish between the two structures.

In Experiment 6, we ask whether infants discriminate vowel-repetition and consonant-repetition items. We administer participants the same familiarization as in Experiment 5, which should allow them to generalize the vowel-repetition structure. In the test phase, in addition to the generalization test with novel vowel-repetition words, we present consonant-repetition words, a structure they have not previously been exposed to. If infants' failure in Experiment 4 resulted from their inability to discriminate vowel-repetition and consonant-repetition items, participants in Experiment 6 should expect the toys to appear on the side associated to the vowel-repetition both for novel vowel-repetition test items, and for consonant-repetition test items.

2.7.2 Experiment 6

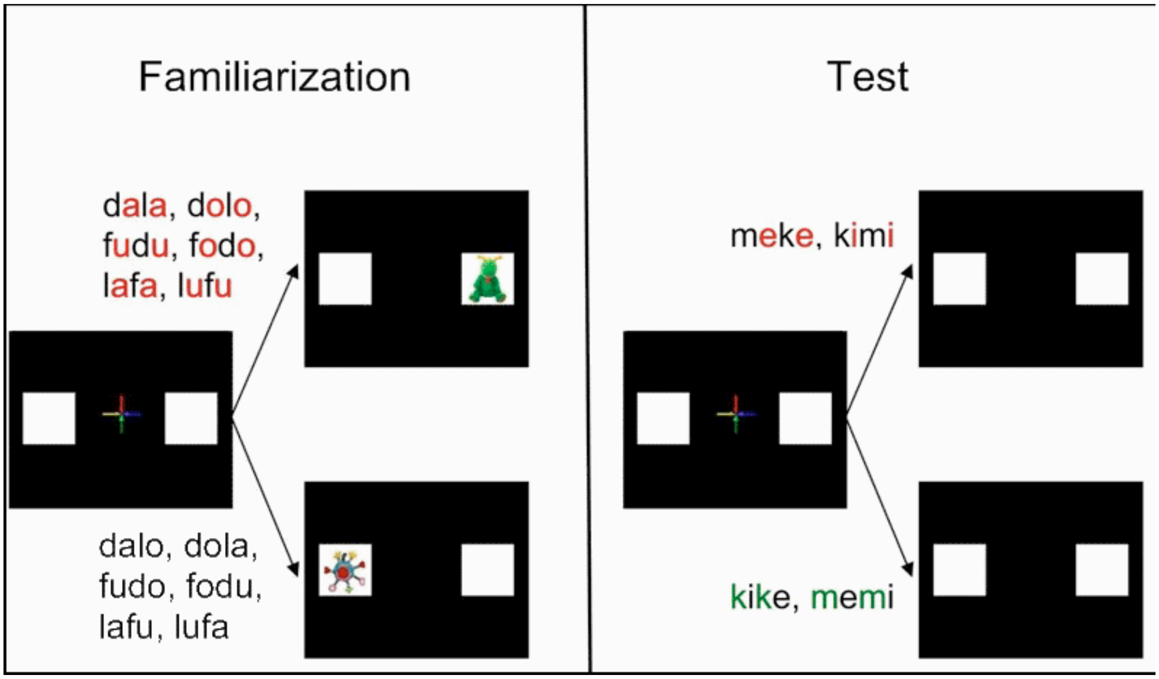

Figure 2.16 – Paradigm of Experiment 6. Participants took 32 familiarization trials and 8 test trials.

In the discussion of Experiment 4, we propose that infants may have failed in that experiment because they could not discriminate between vowel-repetition and consonant-repetition items. We directly test this claim in Experiment 6. To this aim, the paradigm of Experiment 6 (see Figure 2.16) was the same as that of Experiment 5, except for the test stimuli. In the test of Experiment 6, instead of items without any repetition, we present consonant-repetition items. If participants do not discriminate consonant-repetition items from vowel-repetition items, they should expect the toy to appear in the same location for all test items, i.e. the location predicted by vowel-repetition items.

2.7.2.1 Participants

Twenty-four infants were included in the analysis; age range 5 month 19 days to 6 month 14 days (average 6 months and 1 day). Nine other infants participated in the study but were excluded due to fussiness (7) or equipment failure (2).

2.7.2.2 Stimuli

The stimuli in Experiment 6 consisted in bisyllabic items. In the familiarization, these items could have either repeated vowels as in Experiments 3 and 4 or not. Six items containing a vowel repetition (*dala*, *dolo*, *fodo*, *fudu*, *lafa* and *lufu*) and six items containing no repetition at all (*lufa*, *lafu*, *dalo*, *dola*, *fudo* and *fodu*) were created. The same three consonants and three vowels were used to generate both sets of items. For the test, four novel items were generated with novel consonants and novel vowels. The test items were the same as in Experiments 3 and 4. Two test items had a consonant repetition (*kike* and *memi*) and two test items had a vowel repetition (*meke* and *kimi*). Items were synthesized with MBROLA (fr4) with phoneme durations of 120 ms and a monotonous pitch of 200 Hz. There was no silent pause between two syllables within an item.

2.7.2.3 Results and discussion

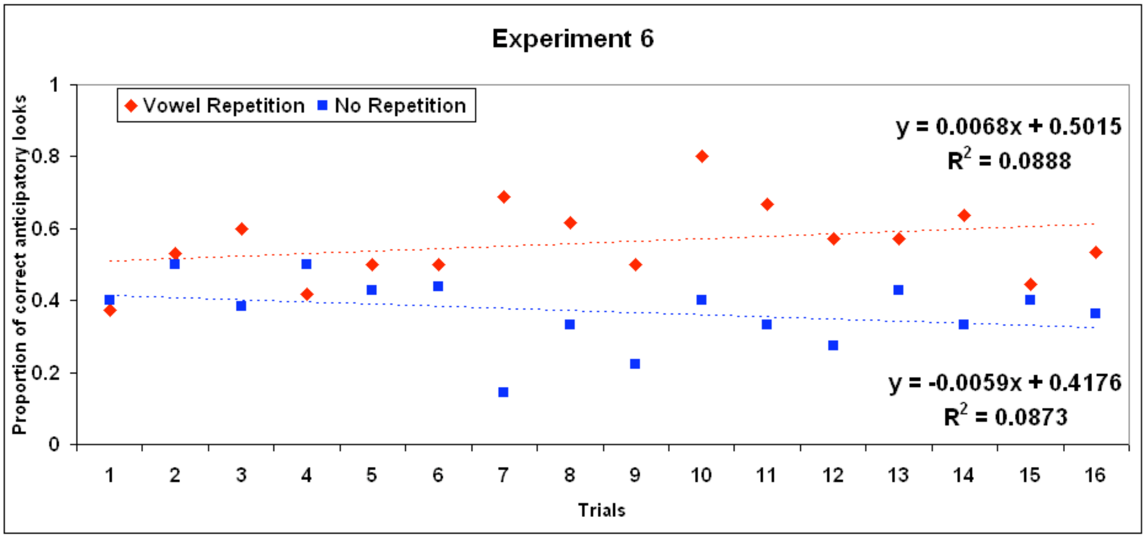

Figure 2.17 - Proportion of correct anticipatory looks for each familiarization trial in Experiment 6. The dotted lines depict the linear regression for vowel repetition (red) and for no-repetition (blue) trials, respectively.

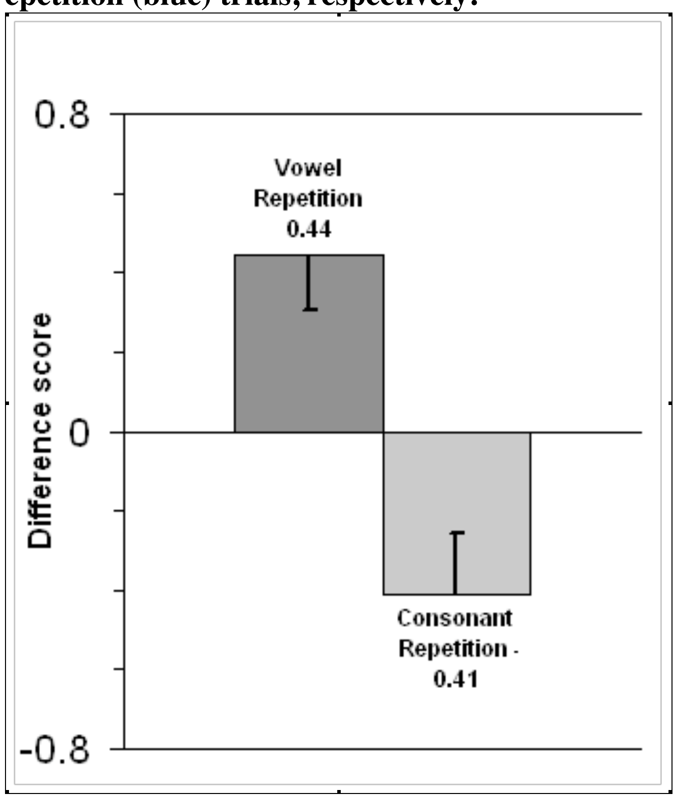

Figure 2.18 - Mean difference scores for Experiment 6 considering the first fixations. Infants looked more at the correct side predicted by the vowel repetition exemplars, both for novel vowel repetition and for consonant repetition items. Error bars represent standard errors.

Familiarization results are presented in Figure 2.17. Infants anticipated to one or the other side in 45% of the trials. We computed the proportion of infants showing a correct anticipatory look for each trial. A linear regression analysis gave non-significant results for the Vowel-repetition trials, ß=.007, R^2<.09, t(14)=1.17, *P* = .26, and for the no-repetition trials, ß=-.006, R^2<.09, t(14)=-1.16, *P* = ..27.

The test results are presented in Figure 2.18. Infants looked to the left or the right in 57% of vowel repetition test trials, and in 55% of the consonant repetition test trials. Three infants did not provide data in the consonant repetition tests, so that 24 infants were included in the analysis of the vowel repetition tests and only 21 in the analysis of the consonant repetition tests.

Paired t-test showed that infants obtained significantly higher difference scores for the vowel repetition than for the consonant repetition, considering the first fixations, t(20) = 3.56, *P* < .002, d' = 1.11; and the overall accuracy, t(20) = 2.87, *P* < .001, d' = 92.

Considering first fixations in vowel repetition test trials, Infants' mean difference score was .44 for the vowel repetition, which was significantly greater than 0, t(23) = 3.19; *P* = .004; d' = .65. Eighteen infants obtained a positive difference score, 4 infants a negative difference score, and 2 a null difference score. A binomial test showed that significantly more infants obtained a positive difference score than a negative difference score, *P* = .0043. For the consonant repetition, infants' mean difference score was -0.41, which was significantly lower than 0, t(20) = -2.69; *P* = .014; d' = .59. Four infants obtained a positive difference score, 16 infants a negative difference score, and 1 a null difference score. A binomial test showed that significantly more infants obtained a negative difference score than a positive difference score, P = .012.

Finally, considering the overall accuracy, infants' mean difference score was .33 for the vowel repetition, which was significantly greater than 0; t(23) = 2.34; P = .028; d' = .48. Fourteen infants obtained a positive difference score, four infant a negative difference score, and six a null difference score. A binomial test showed that significantly more infants obtained a positive difference score than a negative difference score, P = .031. Infants' mean difference score was -.42 for the consonant repetition, which was significantly lower than 0, t(20) = -2.61; P = .017; d' = .57. Four infants obtained a positive difference score, 14 infants a negative difference score, and 3 a null difference score. A binomial test showed that significantly more infants obtained a negative difference score than a positive difference score, P = .031.

Altogether, these results confirm that 6-month-old infants are able to generalize the vowel-repetition structure. Moreover, they appear to generalize that pattern to consonant-repetition items, which they had not heard during the familiarization. This suggests that infants' failure to generalize any of structure in Experiment 4 resulted from their inability to discriminate between vowel-repetition and consonant-repetition exemplars.

2.7.3 Discussion

The results observed in Experiment 5 and 6 show that 6-month-old infants are able to generalize the association between a series of items instantiating a vowel-repetition and the location where a toy should appear, to novel vowel-repetition items.

This shows that the vowel-repetition structure is not too complex for 6-month-olds to learn. We replicated these results in Experiment 6. However, in that experiment, 6-month-olds appear to have the same expectations for consonant-repetition items, a structure to which they were not previously exposed, as for novel vowel-repetition items. The results of Experiment 6 alone could be accounted by young infants' limited abilities to inhibit a learned response. As they learned to look to one location for vowel-repetition items, they may not be able to inhibit that response to any stimulus. This interpretation is however excluded by Experiment 5, where infants did not look to the location predicted by vowel-repetition when hearing novel words containing no repetition at all. Rather, it seems that 6-month-olds extend the generalization of vowel-repetition to consonant-repetition. This suggests that these infants have trouble discriminating vowel-repetition from consonant-repetition items.

In fact, our results suggest that 6-month-old infants represent speech in a very different way than that of 12-month-olds and older human beings. In particular, younger infants appear not to represent consonants and vowels as different speech categories. From birth, speech is perceived as a series of syllables (Bertoncini, Floccia, Nazzi & Mehler, 1995; Bijeljac-Babic, Bertoncini, & Mehler, 1993; van Ooyen, Bertoncini, Sansavini, & Mehler, 1997). At six-months, infants may still represent syllables in an analogical holistic way, like adults represent meaningless sounds. In this view, no subsyllabic unit is represented at that stage of development. Both vowel-repetition and consonant-repetition are therefore perceived as approximate or partial syllable repetitions. Alternatively, the first stage may consist in the representation of syllables as a series of segments, without differentiation between consonants and vowels, all being included in a larger category. Both vowel repetition and consonant repetition items would therefore exemplify the segment-repetition.

2.8 Changes in the speech format of representation between 6- and 12-months.

The results of Experiments 1-6 suggest that the way infants represent and use speech sequences radically changes between the ages of 6- and 12-months. Indeed, we showed that, by the end of the first year of life, infants exhibit a pattern conforming with that of adults, relying mainly on consonants when forming word representations (see Bonatti et al., 2005; Cutler et al., 2000; Mehler et al., 2006 for adult data), and generalizing with more ease a structural relation if it is carried by vowels, than if it is carried by consonants (see Toro et al., 2008a, 2008b for adult data). Six-month-olds, however, show a quite different pattern. They appear to rely mainly on information carried by vowels when forming a word representation. Moreover, while they are able to generalize repetition on a sub-syllabic unit, they do not find it easier to generalize such structural relation over vowels or consonants. Thus, the functional specialization characterizing consonant and vowel processing must emerge during the second half of the first year of life. We even suggest that 6-month-olds do not represent anything like the categories consonants and vowels.

2.8.1 The holistic syllable

The difficulty of 6-month-olds to discriminate between vowel-repetition and consonant-repetition exemplars suggests that infants at that age do not dispose of the broad categories of consonants and vowels, or do not use them. In contrast, the success of 12-month-olds in discriminating vowel-repetition and consonant-repetition, independently from the specific consonants and vowels used (all Italian vowels were used, and consonants included stop consonants, fricatives and liquids), suggests that they do represent vowels and consonants as different speech objects.

If 6-month-olds do not discriminate between consonants and vowels, how do they represent speech sequences? One possibility is that they have a unique segment format, and represent speech sequences as a series of undifferentiated segments. However, we consider this possibility ought to be rejected on the basis of experimental work suggesting that syllable rather than segment is the favored format of representation for adults (Liberman & Streeter, 1978; Massaro, 1974, 1987; Mehler, 1981; Mehler, Dommergues, Frauenfelder, & Segui, 1981; Pallier, 1994; Segui, 1984; Segui, Dupoux, & Mehler, 1990) and infants (Bertoncini, Floccia, Nazzi & Mehler, 1995; Mehler & Bertoncini, 1981; Bijeljac-Babic, Bertoncini & Mehler, 1993; van Ooyen, Bertoncini, Sansavini, & Mehler, 1997). In particular, Bijeljac-Babic, Bertoncini and Mehler (1993) showed that neonates could discriminate between bi-syllabic and tri-syllabic sequences, but not between bi-syllabic sequences varying in the number of segments. This suggests that infants can use the syllabic unit from birth, and there is no reason to consider they have lost that ability six months later, to regain it in adulthood. Mehler and Bertoncini (1981) further showed that 2-month-old French-learning infants could not discriminate two sequences consisting in three consonants, i.e. *pst* and *tsp*, but could discriminate two well-formed syllables, i.e. *pat* and *tap*. If these sequences were represented as a series of three segments, infants should behave similarly for both pairs of stimuli. Infants could however discriminate the two sequences of three consonants when vowels were added so that syllabification was possible, i.e. *upstu* vs. *utspu*. Thus, infants appear to discriminate between the two stimuli of a pair only in the situations where adult perceive syllables. These results strongly suggest that two-month-olds represent speech as a series of syllables, rather than as a series of segments.

Six-month-olds may still represent speech as a series of syllables. But contrary to 12-month-olds, 6-month-olds appear to represent syllables in a holistic format of representation, which does not allow to differentiate between the information carried by consonants and that carried by vowels. Nevertheless, vowels carrying more energy than consonants, the syllabic representation will therefore consist mainly in the information brought by its vowel. This view is in agreement with our results that 6-month-olds judged two words more similar if they share the same vowels. Future experiments should however directly test this interpretation.

Alternatively, infants may well be able to represent subsyllabic units, such as consonants and vowels but do not use them, because of a precedence of the syllable unit. This view is reminiscent of the "externality effect" observed in the visual domain, when infants younger than 2-months attend to the external frame rather than to the internal content of visual stimuli (Bushnell, 1979; Milewski, 1976; Pascalis et al., 1995). Moreover, Ghim and Eimas (1988) tested infants' sensibility and attention to global and local aspects of geometrical arrays. They habituated 3- and 4-month-olds to global squares and diamonds made of local squares and diamonds, or with global crosses and Xs made of local crosses and Xs. They showed that infants could acquire and remember information about both the global and local forms, but also found evidence for a global precedence. Indeed, if habituated to a global square formed of local squares, infants showed a global novelty preference for a global diamond formed of local diamonds, compared to a global square formed of local diamonds. However, they showed no preference when a global diamond formed of local diamonds was contrasted with a global diamond formed of local squares. Thus the novel global shape in both stimuli of the later test condition attracted infants' attention, and caused the neglect of the local novelty. In sum, young infants appear to process the global aspects of visual stimuli in priority, even though they are able to represent local aspects. We may ask whether a similar phenomenon is at play in the speech domain. Six-month-old and younger infants may well represent consonants and vowels as different sources of information, but privilege the syllabic unit when processing speech, especially when generalizing structural relations.

2.8.2 Acquiring the consonant and vowel categories

A few phoneticians have argued against any role for consonant and vowel categories in speech perception or production (Faber, 1992; Linell, 2005), proposing that these categories were solely consequences of the writing system (Port, 2006, 2007, 2008; Port & Leary, 2005). The results we obtained with 6-month-old and 12-month-old infants in Experiments 3-5 would argue against such an extreme conclusion. Twelve-month-olds but not 6-month-olds could generalize the vowel-repetition structure when it was confronted to a consonant-repetition structure. Still, 6-month-olds could generalize the vowel-repetition structure when it was confronted to tokens showing no structure at all (i.e. no repetition).

If one-year-old infants do not represent categories like consonants and vowels, how can we explain that they systematically generalize the vowel-repetition rather than the consonant repetition? An answer based on the acoustic signal should probably call upon vowels' major energy and saliency, and propose that vowel-repetitions attract more attention and were thus more likely to be generalized. However, given that the low-level features available to 12-month-olds are also available to 6-month-olds, and having shown 6-month-olds' ability to generalize the vowel-repetition structure, this answer would predict the same behavior for 6- and 12-month-olds. Moreover, the saliency hypothesis would predict that 12-month-olds should always rely on vowels, also when encoding a word in memory. Experiment 1 showed that this was not the case. Our results thus reject any explanation based on low-level acoustic features.

Therefore, the minimal assumption we must make to explain 12-month-olds' behavior in our experiments is to grant them the ability to discriminate between consonant-repetition and vowel-repetition items, on other basis than low-level acoustic features. It follows that 12-month-olds must represent consonants and vowels separately at a higher level of cognition. In contrast, 6-month-olds do not represent these categories yet (or do not use them), and must rely on another mode of representation such as the holistic syllable. Below, we discuss how infants may switch from the holistic syllable format of representation, to a format using consonant and vowel categories.

Three (at least) sources of information may be considered by infants to form the large categories of consonants and vowels. First, as we discussed and experimentally tested in the previous chapters, consonants and vowels have different functions. In particular, consonants are more important than vowels for the identification of words. Infants may use such functional information and form the class of sounds that are informative for meaning, and the class of sounds that are less informative. However, in most languages the functional distinction is only relative. Vowels also carry lexical information, e.g. *ball*, *bell*, *bull* in English; *pasta* (pasta), *pista* (track), *posta* (mail) in Italian. It is therefore unclear how infants would draw a threshold between the lexically informative and lexically non-informative speech sounds, especially before having acquired a substantial vocabulary.

In fact, consonants and vowels may rather be defined according to acoustic properties. Even though there is no definite criteria to discriminate between consonants and vowels on the basis of a purely acoustic analysis, certain consonants, i.e. affricates, fricatives and stops, have very different acoustic properties than other speech sounds, i.e. nasals, liquids, semi-vowels and vowels (Singhvi, personal communication). The latter are actually more sonorant, carry more energy and exhibit spectral peculiarity such as periodicity and lower frequencies. Similarly, Stevens (2002) proposed that speech sounds could be categorized according to *acoustic landmarks*, with peaks, valleys and discontinuities in low-frequency energy corresponding respectively to vowels, glides and consonants. The partition between sonorants and non-sonorants, or low-frequency energy peak and discontinuities, may thus constitute a first approximate of the categories of consonants and vowels. In fact, in some languages, certain sonorant consonants can occupy the position of vowels in the nucleus of a syllable; in English, consider the final

[l] in *little* and the final [m] in *rhythm*; in Croatian consider the [r] in *Trst*.

These acoustic differences may be completed by a distributional analysis. Conducting a distributional analysis of different speech sounds however first requires representing these sounds as units that can be tracked. If infants, at some stage of development, dispose only of the syllabic unit to represent speech sequences, how can they learn the distribution of consonants and vowels? Recalling Plato's definition of consonants and vowels, we observe that a fundamental difference between the two speech sound categories is the almost exclusive one-to-one relation between vowels and syllables. One vowel alone can indeed constitute a syllable, and one syllable cannot comport more than one vowel. If syllabic consonants such as [r] in Serbo-Croatian are not easily discriminable from vowels on the basis of acoustic features, and can form a syllable by themselves, e.g. [r] in *rvanje* (Croatian, wrestling) they can however appear with another vowel in one syllable, e.g. *tri* (Croatian, three).

Infants may thus discover the vowel category by tracking those syllables that appear inside other syllables; e.g. *a* in *ra*. Vowels would be atomic syllables that can appear inside other syllables and cannot be further divided. The extraction of atomic syllables may correspond to the formation of vowel prototypes that starts around 5- to 6-months of age (Kuhl, 1991; Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992). This may represent the initial step in the change from holistic representations of syllables to structured representations relying on the formats of consonants and vowels.

2.8.3 The emergence of the functional specialization of Consonants and Vowels

Whether 6-month-old infants still form holistic representations of syllables, or do represent the broad categories of consonants and vowels, our results show that by six months, neither the lexical role of consonants nor the structural role of vowels are in place. Both these roles emerge during the second half of the first year of life.

We have discussed earlier the role that the formation of the phonemic consonantal categories in the emergence of the lexical role of consonants. Categorical perception is indeed stronger for consonants than for vowels (Pisoni, 1973), thus resulting in more reliable information carried over the consonantal tier than over the vocalic tier. We therefore predict that the lexical role of consonants should not emerge before 10-months of age, when infants start forming consonantal phonemic categories.

The role of vowels in the CV-hypothesis was inspired by prosodic bootstrapping theories. Because vowels carry prosody, and because prosody informs syntax, vowels should in turn inform syntax. That motivation clearly states that the structural role of vowels would be a consequence of the structural role of prosody, which we may assume to derive from possibly innate perceptual biases such as the iambic-trochaic law (Hayes, 1995). Indeed, several experiments have exemplified how prosody leads participants to chunk continuous streams (Shukla et al., 2007) and group independent constituents (Bion et al., in press). In consequence, prosody can participate in solving global or local syntactic ambiguities (see Cutler, Dahan & van Donselaar, 1997 for a review), and discovering hierarchical structures in speech streams (Langus, Marchetto, Bion & Nespor, submitted; Mueller, Bahlmann, & Friederici, 2010).

The structural role of vowels may therefore not be innate, but require infants to identify vowels as the category carrying the essential of prosodic information, and consequently transfer the structural role of prosody to vowels. This transfer may happen as soon as infants dispose of enough information, i.e. the vowel category and sensitivity to prosodic grouping and/or chunking. There is evidence that even newborns are sensitive to certain prosodic boundaries (Christophe, Mehler & Sebastian-Galles, 2001) and to the rhythm prosodic variations carry (Nazzi, Bertoncini & Mehler, 1998). Thus, if the structural role of vowels, which we evidenced in 12-month-olds, has not yet emerged by 6-months of age, it may be because they have not yet formed the vowel category. If this scenario is correct, we should predict the specialization of vowels shortly after the vowel category is formed. Alternatively, even if 6-month-olds already dispose of the vowel category, the transfer of role from prosody to vowels may require more information that is not yet accessible to 6-month-olds.

2.8.4 Conclusion

The experimental work discussed in this chapter showed that 12-month-old infants represent the consonant and vowel categories, and have assigned specialized functions to each category. Specifically, they rely more on consonants than on vowels when encoding a word in memory, but extract and generalize structural relations over vowels rather than consonants. Our results further suggest that these categorical representations and functional specializations are not yet available to 6-month-olds. Younger infants indeed relied more on the information carried by vowels than that carried by consonant when encoding words in memory. Moreover, they did not show a preference for generalizing a structural relation over vowels rather than consonants, even though they are able to generalize such structural relations.

We propose that these discrepancies can be explained by a change in the format of representation during the second half of the first year of life. Our hypothesis, which should be thoroughly tested, is that younger infants perceive speech as a sequence of syllables represented in a holistic format, where no distinction is made between consonants and vowels. By one-year of age, however, infants now dispose of more abstract representations, and in particular divide speech sounds into two broad categories, consonants and vowels, to which different functions are associated. The exact mechanisms responsible for this change in format of representation should be further inquired, but we propose that acoustic and distributional information may play a role. However, the role of consonants, we suggest, follows from the categorical perception of consonantal phonemic categories, whereas the structural role of vowels follows from its carrying prosody.

‘Twas brilling, and the slithy toves
Did gyre and gimble in the wabe;
All mimsy were the borogoves,
And the mome raths outgrabe.

Through the Looking-Glass, and What Alice Found There
Lewis Carroll, 1872

CHAPTER 3

Frequency and function words

3.1 Function words, Syntactic Bootstrapping and The Second Gavagai Problem

3.1.1 The Second Gavagai Problem

Quine (1960) famously enounced the “gavagai” problem, according to which a child acquiring language is facing a hard problem when trying to understand the meaning of a novel word. In Quine’s metaphor, children are in the same situation as an adventurer witnessing an island native that would point to a running rabbit and say “gavagai”. Does “gavagai” refer to the rabbit, its ears, its color or its running? This problem has since received several solutions, involving the use of conceptual biases (Markman, 1990;

Markman, Wasow & Hansen, 2003), socio-pragmatic (Baldwin, 1991, 1993; Csibra, 2003, Sperber & Wilson, 2004; Hirsh-Pasek, Golinkoff, & Hollich, 2000; Akhtar & Tomasello, 1996) and syntactic cues (Bernal et al., 2009; Brown, 1957; Gleitman, 1994; Naigles and Kako, 1993; Waxman & Booth, 2001, 2003). But Quine's situation contains a second problem that has hardly been addressed. Assuming that the learner solved the classical gavagai problem and identified the referent, what part of speech actually refers to the referent? In other words, given that the island native intended to name the rabbit, is rabbit said "gavagai" in his language? Or is it "gava", "vagai" or "gai"? This problem is not anecdotal as words are rarely pronounced in isolation. Only 7 to 12% of child directed speech utterances consist in isolated words (Brent & Siskind, 2001; Christiansen, Allen & Seidenberg, 1998; Fernald & Morikawa, 1993; Fernald & Simon, 1984). In particular, nouns in languages like English or Italian are usually associated to a determiner. Instead of just "rabbit", "gavagai" may well mean "*the* rabbit" or "*a* rabbit". Learners would then need to strip the part that corresponds to the determiner, or more generally to function words, before pairing the noun and its referent. Are infants equipped with a mechanism to identify function words, which should not be associated to referents?

3.1.2 Identifying function words

Function words (such as determiners and prepositions) lie at the intersection of syntax and the lexicon. They clearly are words that have to be acquired. However, in contrast to content words (which must be linked to some semantic referent), function words mainly serve syntactic rather than semantic purposes. On the one hand, function words might thus impair word learning – because they are words that children might try to learn, and yet they have no clear meaning they could be mapped onto. On the other hand, function words might facilitate word learning – by providing syntactic cues that might then be used for learning other (content) words. To use the syntactic cues associated with function words, however, infants need to *identify* them in the first place.

While different authors have uncovered different cues that tend to distinguish content and function words (Cutler, 1993; Shi, Morgan & Allopenna, 1998), such proposals meet with two problems. First, to be useful for language acquisition, the cues must be available in any language a child might end up learning, and cannot be specific to a particular language (e.g., English). Second, early in life, infants need to be able to *use* such cues to identify function word candidates. Here, we start assessing these issues, asking whether infants can attribute different properties to potential content words and function words based on a language-independent distributional property of function words, namely their high frequency of occurrence.

3.1.3 Words, syntax, the chicken and the egg

Children acquire both the syntax and the lexicon of their native language. However, different theories disagree on the relation between the development of syntax and that of the lexicon. Specifically, proponents of semantic bootstrapping (Pinker, 1984) and usage-based theories of language acquisition (Dabrowska, 2001; Tomasello, 2003) hold that vocabulary acquisition facilitates syntax learning, while proponents of syntactic bootstrapping accounts (Gillette et al., 1999; Gleitman, 1994; Gleitman & Gleitman, 1992; Landau & Gleitman, 1985) propose that syntax boosts vocabulary acquisition. We will now briefly review both kinds of theories.

Semantic bootstrapping theories describe how infants can bootstrap the initial steps of syntax acquisition based on their knowledge of (a limited number of) words. For example, semantic categories such as objects and actions might initially be used to discover how syntactic categories such as nouns and verbs are implemented in the language. Specifically, infants might first acquire a few words related to the objects and actions they observe. Then, they might use these words to learn the corresponding syntactic categories. For example, object names might be mapped onto nouns, and words describing actions onto verbs. Based on such a mapping, infants might discover crucial aspects of the syntactic organization of their native language. For instance, knowing the verb 'eat' and the noun 'cookie' might be sufficient to decide whether the object comes after the verb (e.g. "eat cookies", corresponding to the canonical English word order), or whether the object precedes the verb (e.g. "cookies eat", corresponding to the canonical Japanese word order; Pinker, 1994, p.112). On this view, infants can start acquiring syntax only after having learned a minimal set of words, because knowledge of these words is crucially required to bootstrap grammar acquisition.

Semantic information might help grammar acquisition in yet another way. According to usage-based theories of syntax acquisition (see e.g. Dabrowska, 2001; Tomasello, 2003), infants and children first learn specific word sequences, with very limited knowledge of their underlying structure. That is, they might remember words only in specific contexts, and assign meaning to words only within this context.

Crucially, however, as they do not analyze sentences in terms of their underlying structure, they should be unable to use words in contexts that differ from those they have heard. For example, if they have heard the word "broke" only in the sentence "The window broke", they should be unable to use the word in new contexts such as "He broke it" or "The windows got broken" (e.g., Savage, Lieven, Theakston & Tomasello, 2003; Tomasello, 2000; but see Lidz, Gleitman & Gleitman, 2003; Thothathiri & Snedeker, 2008)[1]. As they get older, children should gradually discover that the sentences they have heard have in fact an underlying structure, eventually leading to the kind of abstract syntactic knowledge observed in mature, adult speakers. According to this theory, children thus need to acquire a substantial vocabulary before learning any syntactic regularity.

In contrast to such views, syntactic bootstrapping models hold that syntactic knowledge facilitates vocabulary acquisition (Gillette et al., 1999; Gleitman, 1994; Gleitman & Gleitman, 1992; Landau & Gleitman, 1985). For example, upon hearing a

---

[1] Evidence in favor of usage-based theories of language acquisition usually points to the unbalanced use of different syntactic constructions with specific words, thus questioning the existence of productive rules (Pizutto & Caselli, 1994; Tomasello, 1992). For example, Pine and Lieven (1997) showed that, when children start using determiners, they happen to use most nouns with either *the* or *a*, but not with both determiners. Yang (unpublished) however showed that such pattern in fact follows from the ubiquity of Zipf's laws in linguistic production. If one considers that few rules and few words should be used very frequently, while many words and rules are used with low frequency, children's production data actually supports rather than contradicts the existence of productive rules guiding children's (and adults') production (Yang, unpublished).

sentence like "the duck and the bunny are gorping", listeners as young as two-year-olds are likely to conclude that "to gorp" must have an intransitive meaning, since it has no object. Upon hearing the sentence "the duck is gorping the bunny", in contrast, they tend to conclude that "to gorp" is transitive, since it now has a direct object (Naigles & Kako, 1993). Thus, a rather rudimentary syntactic analysis (such as counting the number of noun phrases and analyzing their positions) can constrain the interpretation of novel verbs.

Of course, semantic and syntactic bootstrapping accounts are not mutually exclusive, and infants might well use both routes in complementary ways. Both syntax and the lexicon might initially develop in parallel and cross-fertilize each other. This possibility is particularly important for the issue studied here, relating to how function words are acquired and used during language acquisition. From a syntactic bootstrapping perspective, the syntactic information carried by function words would be clearly helpful for learning new (content) words, as function words indicate syntactic roles and syntactic categories. For example, in a language like English, a word following a determiner is likely to be a noun, while a word following an auxiliary is likely to be a verb. Therefore, when hearing a novel word that is accompanied by a function word, infants might interpret it as referring to a novel object if the function word marks it as a noun (Brown, 1957), as referring to a novel action if the function word marks it as a verb (Bernal, Litz, Millote & Christophe, 2007; Brown, 1957), and as referring to a property when the function word marks it as an adjective (Waxman & Booth, 2001). This capacity seems to be present early in life, as infants as young as 14-month-old start using the syntactic information provided by function words to interpret new content words (Waxman &

Booth, 2003).

While function words might facilitate the acquisition of content words by providing syntactic cues, they are more problematic from a vocabulary acquisition perspective, as they have no clear referents. As a result, unless infants can identify function words *as* function words, these words should impair vocabulary acquisition – because infants might consider them as meaningless "noise". In order to take advantage of the syntactic information provided by function words, infants thus need to identify them early on. In the next section, we will discuss a number of cues that might allow them to solve this problem.

3.1.4 Cues to identify function words

To identify function words, and to distinguish them from content words, infants might rely on two types of surface cues: phonological properties (Shi, Werker & Morgan, 1999; Shi, Morgan & Allopenna, 1998; Cutler, 1993), and distributional cues (Gervain, Nespor, Mazuka, Horie & Mehler, 2008; Shi, Morgan & Allopenna, 1998).

Function and content words tend to have different phonological properties. Compared to content words, function words are often shorter, simpler and unstressed. In English, these differences are salient enough for neonates to notice them (Shi et al., 1999). However, even though such results demonstrate that very young infants are sensitive to these phonological differences, it remains unclear whether infants actually use them for language acquisition. In fact, the phonological differences between content and function words vary from one language to another. For example, English function words have reduced vowels (the [ðə], of [əv], etc.) and tend to start with certain consonants (th-, wh-) that are not commonly used in content words. In contrast, Hungarian function words do not have reduced vowels, and their initial consonants occur in many content words. In that language, function and content words differ mainly in the number of syllables (Gervain et al., 2008). In French, some function words are even homonymous with content words. For example, the sound [vo] can be both a determiner (as in "vos" – your) and a noun (as in "veau" – veal). Therefore, infants can use phonological cues to find function words only after having learned enough about their native language to identify the relevant cues[2].

Distributional cues, in contrast, seem to be relatively consistent across languages. Function words tend to occur at the edges of prosodic units, and therefore at utterance boundaries (Christophe, Millote, Bernal & Lidz, 2008; Gervain et al., 2008; Shi, Morgan & Allopenna, 1998). Moreover, as all languages contain only a limited number of function words, their frequency of occurrence is much higher than that of content words. Indeed, except for proper names and nicknames such as 'Mummy' and 'Daddy', or non referential interjection such as 'look!', the 20 or so most frequent words in child-directed speech are function words. This cue appears to be consistent across languages (see

[2] Some prosodic cues might seem to be universal cues to function words. Specifically, function words are systematically less stressed than the content words they occur with (e.g., Nespor & Vogel, 1986). Note, however, that there are also some content words that receive systematically less stress than other content words; for example, in head-complement languages, pre-nominal adjectives tend to be less stressed than the nouns they occur with. It thus seems that watching out for "less stressed" words for identifying function words is not totally reliable.

Gervain et al., 2008, for Italian and Japanese; Shi, Morgan & Allopenna, 1998, for Mandarin Chinese and Turkish).

Gervain and colleagues (2008) showed that 7-month-old infants are sensitive to these distributional cues. In an artificial grammar learning experiment, they found that Italian infants preferred frequent syllables to occur at the beginning of a unit (i.e., a bisyllabic word), whereas Japanese infants preferred the frequent syllables to occur at the end of a unit. These preferences correlate with the word order of the participants' native language. Indeed, (frequent) function words tend to occur at the beginning of units in Italian, especially in utterance-initial positions; in Japanese, in contrast, function words tend to occur at the end of units, especially in utterance-final positions. These results suggest that 7-month-old infants are sensitive to variations of frequency of occurrence, and can use this cue to organize their input.

While these results suggest that distributional and, to a lesser extent, phonological cues to function words are available in infant–directed speech, and that, to some extent, infants seem to be able to process them, there is another question that has never been addressed: can infant learners actually use these cues to identify function words? This question is important, because there are numerous demonstrations showing that a perceptual sensitivity (such as the infants' sensitivity to word frequency) is not necessarily used in all circumstances where it might be useful. For example, rats are sensitive to light flashes, as they can associate them with electroshocks; however, they cannot use this sensitivity to associate light flashes with visceral sickness. Conversely, they are sensitive to tastes, as they can associate them to visceral sickness; however, they cannot use this sensitivity to associate tastes with electroshocks (e.g., Garcia, Hankins, &

Rusiniak, 1974). Hence, although rats are sensitive to both tastes and electroshocks, they cannot use these sensitivities for all kinds of associations.

The distinction between being sensitive to a cue and being able to use it is especially important for a cue such as frequency of occurrence. As many if not most animals are sensitive to this cue, this sensitivity might have evolved for non-linguistic reasons, raising the question of whether it can be used for aspects of syntax acquisition as well. For example, pigeons can categorize events by frequency (Keen & Machado; 1999; Machado & Cevik, 1997). Although they share the sensitivity to frequency of occurrence with human infants, they clearly cannot use it to acquire function words – because they do not acquire language. Mutatis mutandi, human infants might well be sensitive to acoustic or distributional differences between function words and content words – without using these potentially useful differences to discover function word candidates. Here, we start addressing this issue, asking whether a language-invariant cue to function words – their high frequency of occurrence – allows infants to attribute different properties to potential function words and to potential content words.

3.1.5 Our study

Different cues have been proposed to be useful for identifying function words (Christophe et al., 2008; Gervain et al., 2008; Shi et al., 1998). However, it has never been shown whether infants can actually use them. Here, we start addressing this issue. We present infants with a word-learning situation, and ask whether they are more likely to attribute content-word-like properties to infrequent items than to frequent items. Based on the hypothesis that it should be easier to associate objects with content words than with function word, we asked whether infants would be more likely to associate a visual object with a determiner or rather with a noun when listening to an unknown language.

In Experiment 7, we simply confirmed that the paradigm used in Experiments 8-11 and stimuli used for Experiments 8a and 8b allow infants to learn an association between a bisyllabic label and a visual object.

In Experiment 8a, we exposed Italian 17-month old infants to short, naturally recorded sentences in a foreign language (i.e. in French). All sentences contained two frequent French determiners, "ce" ([sə]; "this") and "vos" ([vo]; "your"), and several relatively less frequent content words. Note that these words were less frequent than the determiners not only in French in general, but, crucially, also in the language sample to which the infants were exposed. This familiarization phase was followed by a teaching phase in which an object-label association was taught. Specifically, a visual object was presented together with a bisyllabic phrase consisting of a determiner and a noun (e.g., "ce chat", [səʃa]; "this cat"), both taken from the familiarization corpus. Following this, we assessed which of the two words (e.g., the determiner "ce", "this", or the noun "chat", "cat") was more strongly associated with the object. If infants consider frequent items as function word candidates and the infrequent items as content word candidates, the object should be more strongly associated with the less frequent items, as content words are more likely to have observable referents. Hence, we would expect the object to be associated more strongly with the noun than with the determiner.

To assess this, infants saw the object from the teaching phase and a novel object, both presented side-by-side on a computer screen. At the same time, they heard a label that was derived from that used during the teaching phase. Specifically, compared to the original label, we changed either the determiner (e.g., "*vos* chats"; [voʃa], "*your* cats", derived from "*ce* chat"; [səʃa], "*this* cat") or the noun (e.g., "ce *met*"; [səmɛ], "this *dish*", again derived from "ce *chat*"; [səʃa], "this *cat*").

To assess whether the object was more strongly associated with the determiner or with the noun, we measured how likely infants were to orient first towards the familiar object from the teaching phase as opposed to towards the new object. If it is easier to associate content words with objects, and if infants consider frequent items as function word candidates and the infrequent items as content word candidates, they should be more likely to orient towards the familiar object when the derived label had a new determiner than when the derived label had a new noun.

In addition to frequency, natural speech such as the stimuli used in Experiment 8a might present other, especially acoustic and phonological, cues to function words. In Experiment 8b, we asked whether these cues alone would be sufficient to explain the results of Experiment 8a. Specifically, Experiment 8b was identical to Experiment 8a, except that infants were not exposed to French sentences at the beginning of the experiment. While the stimuli used in Experiment 8b had the same acoustic and phonological properties as those employed in Experiment 8a, word-frequency was no longer available as a cue to function words. If the distributional properties of the familiarization of Experiment 8a contribute to the results, we would expect different results in Experiment 8b.

In Experiments 9, 10 and 11, we ask whether frequency alone, is a sufficient cue for infants to understand what part of a label is the referential part, and what part is not.

We ask whether infants are more likely to associate an object with a frequent or an infrequent syllable, when these do not differ in pitch, duration or intensity. In each experiment, we first exposed infants to an artificial speech stream where one of two frequent syllables alternate with one of eighteen relatively infrequent syllables. In a second phase, infants saw an object on the screen and heard the repetition of a bisyllabic label consisting in the association of one frequent and one infrequent syllables. When now presented with both the previous and now familiar object and a novel object, we ask whether infants would be more likely to orient towards the familiar object when hearing a label with a *new frequent* and the *previous infrequent* syllables, or when hearing a label with a *new infrequent* and the *previous frequent* syllables, thus asking whether they formed a stronger association between the object and the frequent or the infrequent syllable. All three experiments differed only in the initial exposure phase. Experiments 9 and 10 were identical, except for which syllables were frequent in the familiarization. In Experiment 9, the syllables then used as label-initial syllables were very frequent. In Experiment 10, the syllables then used as label-final syllables were very frequent. Experiment 11 was identical to Experiment 10 except that the familiarization stream was segmented.

3.2 Pairing objects and labels – Experiment 7

The goal of this Experiment 7 was to ensure that the parameters used in the teaching and test phases of Experiments 8-11 would allow infants to learn the association between objects and labels. The experiment was identical to Experiment 8b, except that the test labels were either identical to those used in the teaching phase, or fully different from those used in the teaching phase. That is, in the same-label condition, the test labels comprised both the same determiner and the same noun as the label from the teaching phase; in the novel-label condition, in contrast, both the determiner and the noun were different.

3.2.1 Materials and Methods

3.2.1.1 Participants

Twenty-five Italian 17-month old infants were tested. Nine were excluded for fussiness (7) or equipment failure (2). The remaining 16 infants (6 males, 10 females, age range: 17 months and 02 days – 17 months and 27 days) were included in the final analysis. Importantly, these infants acquired Italian and had no experience with French (the language we used for our stimuli).

3.2.1.2 Stimuli

Objects

The objects used in the teaching and test phases were simple three-dimensional shapes generated as 3D animations in Maya 6.0 (Autodesk, Inc., San Rafael, CA), using a frame rate of 25 fps, the H.264 codec and the mov container format. Figure 3.1 shows the two objects we used. One was a blue three-dimensional cross. The other was a green pile of rings (similar to the belly of the emblem of the Michelin tire brand). Both objects were symmetrical, had similar perceived volumes and perimeters, and were found to be similarly attractive for 17-month old infants in a pilot study. However, the two objects had clearly different shapes and colors. During the teaching phase, one of the objects was presented in an animated movie of 33 s. During this movie, the object moved from one side to the other while rotating around its axes.

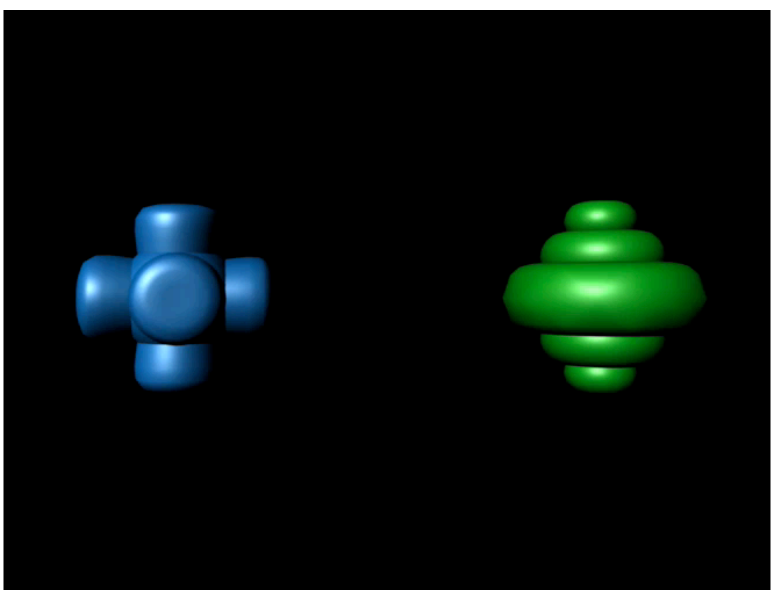

Figure 3.1 - Objects used in all experiments. During the test phase, the two objects were presented statically as shown in the figure. During the teaching phase, in contrast, they slowly moved on the screen.

Labels

Four tokens of the phrases “ce chat” ([səʃa], “this cat”), “vos mets” ([vomɛ], “your dishes”), “ce met” ([səmɛ], “this dish”) and “vos chats” ([voʃa], “your cats”) were recorded from the same native speaker of French who also produced the familiarization sentences. (Note that in French, even though the spelling varies, the words “chat/chats” and “met/mets” are pronounced identically in the singular and plural forms.) The labels were recorded and stored in the same way as the French sentences described above. Labels had an average duration of 500 ms.

Apparatus

In this and all other experiments reported here, infants were tested individually. They sat on a parent’s lap 80 cm from a 17-inch LCD screen in a dimly lit, sound-attenuated cubicle. Parents wore dark sunglasses throughout the experiment to avoid all parental influence on the infants’ behavior.

3.2.1.3 Procedure

Teaching Phase

In the teaching phase, infants saw a visual object on the screen and simultaneously heard a speech label. We used the two shapes described above and two labels composed of a determiner and a noun (“ce chat” - ([səʃa], “this cat”; “vos mets” - [vomɛ], “your dishes”). All four words had occurred during the familiarization phase. Moreover, during the familiarization phase, the nouns selected for the teaching phase had occurred only in the determiner-noun combinations that were also used in the teaching phase, but never with the other determiner. The four stimulus combinations resulting from two objects and two labels were counterbalanced across infants. Each infant learned only one object-label combination.

During the teaching phase, the object moved from one side to the other on an LCD screen while rotating around its axes. Simultaneously, a label was repeated 23 times. Two repetitions of the label were separated by about 900 ms of silence, yielding a total duration of 33 s.

The presentation movie could be interrupted if the infant looked away for more than 2 s. However, all the 16 infants included in the analysis looked at the entire movie without interruption.

Test Phase

Following the familiarization phase and the teaching phase, infants completed four test trials. Two successive trials were separated by the presentation of the central fixation attractor (i.e. a white cross moving back and forth). The experimenter started each trial by pressing a key when the infant was looking at the central fixation attractor.

Once the experimenter started a trial, the fixation attractor disappeared, and the two objects presented in Figure 3.1 appeared on the computer screen. The objects then remained static and visible on the screen for the entire trial duration. Two seconds after the trial started, the central fixation attractor appeared again for 3.32 s. A test label (see below) was first pronounced while the central fixation attractor was still visible on the screen; the offset of the label was synchronized with the attractor's disappearance.

Following this, the test label was repeated five more times, two consecutive repetitions being separated by 1s of silence.

Each of the objects on the screen occupied a surface of 8.5 cm x 9 cm. The centers of the two objects were separated by about 19 cm. For each infant, the position of objects was counterbalanced across trials. We also counterbalanced the position in which the blue object appears in the first trial across infants.

There were two test conditions: the same-label condition and the novel-label condition. The two conditions were identical except for the test label used. As mentioned above, in the same-label condition, the label was identical to the one used in the teaching phase. In the novel-label condition, the label was composed of a different determiner *and* a different noun. Specifically, infants who heard the label “ce chat” (“this cat”) during the teaching phase heard “ce chat” (“this cat”) in the same-label condition, and “vos mets” (“your dishes”) in the novel-label condition. Infant who heard the label “vos mets” (“your dishes”) in the teaching phase, heard “vos mets” (“your dishes”) in the same-label condition, and “ce chat” (“this cat”) in the novel-label condition. Infants completed two trials in the same-label condition, and two trials in the novel-label condition.

3.2.1.4 Analysis

We defined two windows of interest for the analysis of the infant eye gaze. Each was a square of 11 cm x 13 cm, centered on one of the objects. Infants’ looking behavior was monitored using a Tobii 1720 Eye-tracker system and the Clearview 2.5.1 software package. Only infants for whom the eye tracker data for each trial contained at least one fixation of at least 100 ms in one of the two windows of interest were included in the analysis. Infants not meeting these criteria were excluded for insufficient eye tracker data. Considering the four test trials, infants for whom more than 70% of the total fixation time was spent in only one of the windows of interest were considered to exhibit a side bias and were rejected from further analysis.

As dependent variable for the main analysis, we considered the first look to one of the objects following the first presentation of the test label. The first look was defined as the first uninterrupted fixation of at least 100 ms in one of the two windows of interest described above.

In each test trial, infants were coded either as first looking at the familiar object from the teaching phase, or as first looking at the novel object. Infants who did not look at any object in at least one of the four test trials were excluded from the analysis. For each infant and each condition, we computed the proportion of trials in which they first oriented towards the familiar object from the teaching phase. We then assessed whether these proportions differed between the conditions using a two-tailed Wilcoxon signed rank test.

3.2.2 Results

There was no trial in which infants already fixated an object (rather than the central attractor) before the offset of the label. Figure 3.2 shows how often infants first looked towards the familiar object from the teaching phase as a function of the experimental condition. Infants were significantly more likely to first look at the object from the teaching phase in the same-label condition (*M* = 1.19, *SD* = .66, *Mdn* = 1) than in the novel-label condition (*M* = .62, *SD* = .62, *Mdn* = 1). Wilcoxon signed rank test, $W = 50$, $p = 0.015$, $CI_{.95,\text{Mdn difference}} = 0, 1$.

These results were further confirmed by the distribution analysis. Nine infants out of 16 directed their first fixation more often towards the object from the teaching phase in the same-label condition than in the novel-label condition, one infant directed the first fixation more towards that object in the novel-label condition than in the same-label condition, and six infants directed their first fixation equally often towards either object in both conditions. This distribution was significantly different from chance, $p = 0.025$ (exact multinomial test).

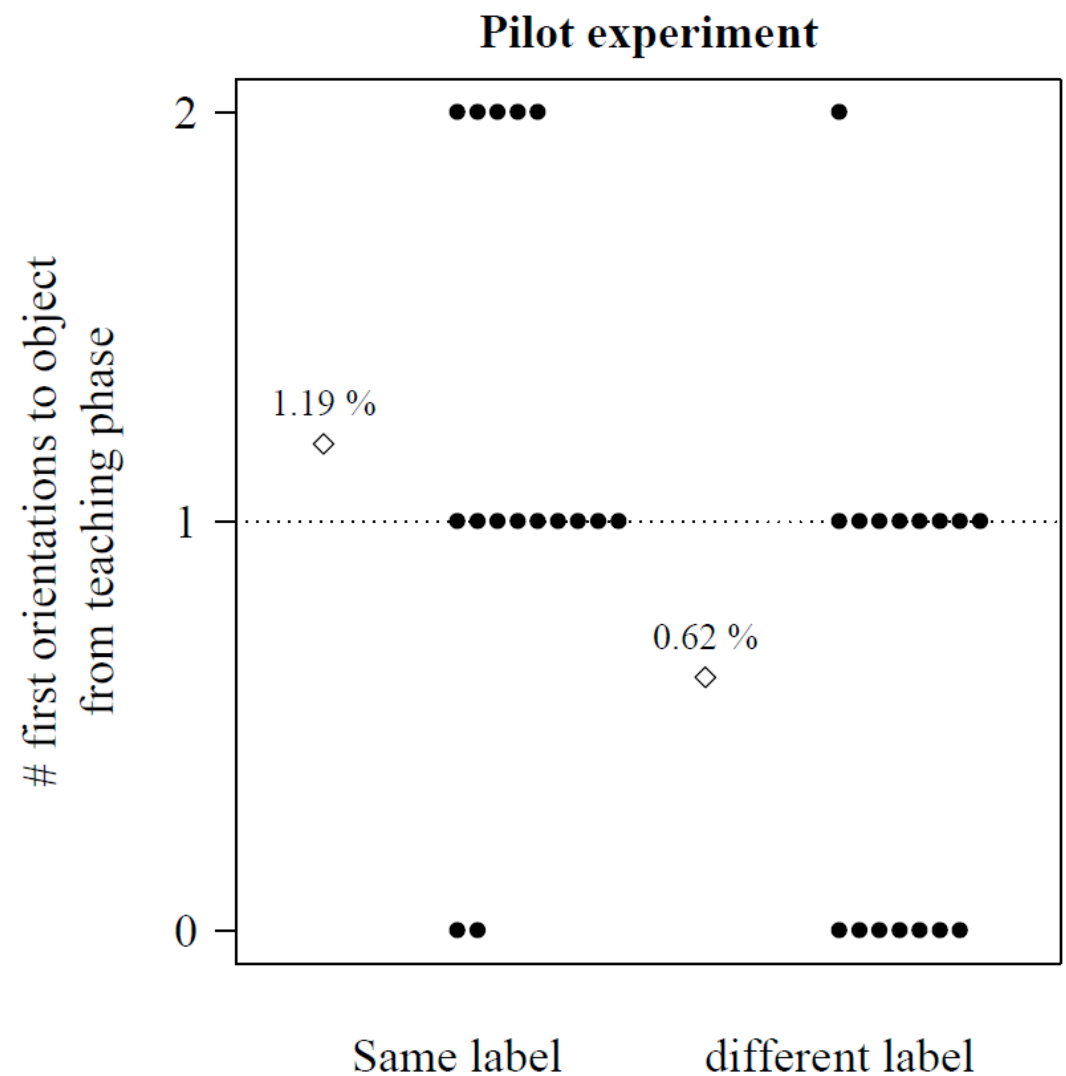

Figure 3.2 - Results of the pilot experiment, Experiment 7. Proportion of first looks directed towards the familiar object from the teaching phase. Dots represent the means of individual participants, diamonds sample averages, and the dotted line the chance level of 50%. When the label used during the test phase was identical to that presented during the teaching phase, infants were more likely to fixate first the familiar object from the teaching phase compared to the condition where the label used during the test phase differed from that presented during the teaching phase.

3.3 Identifying function words in a foreign language – Experiment 8

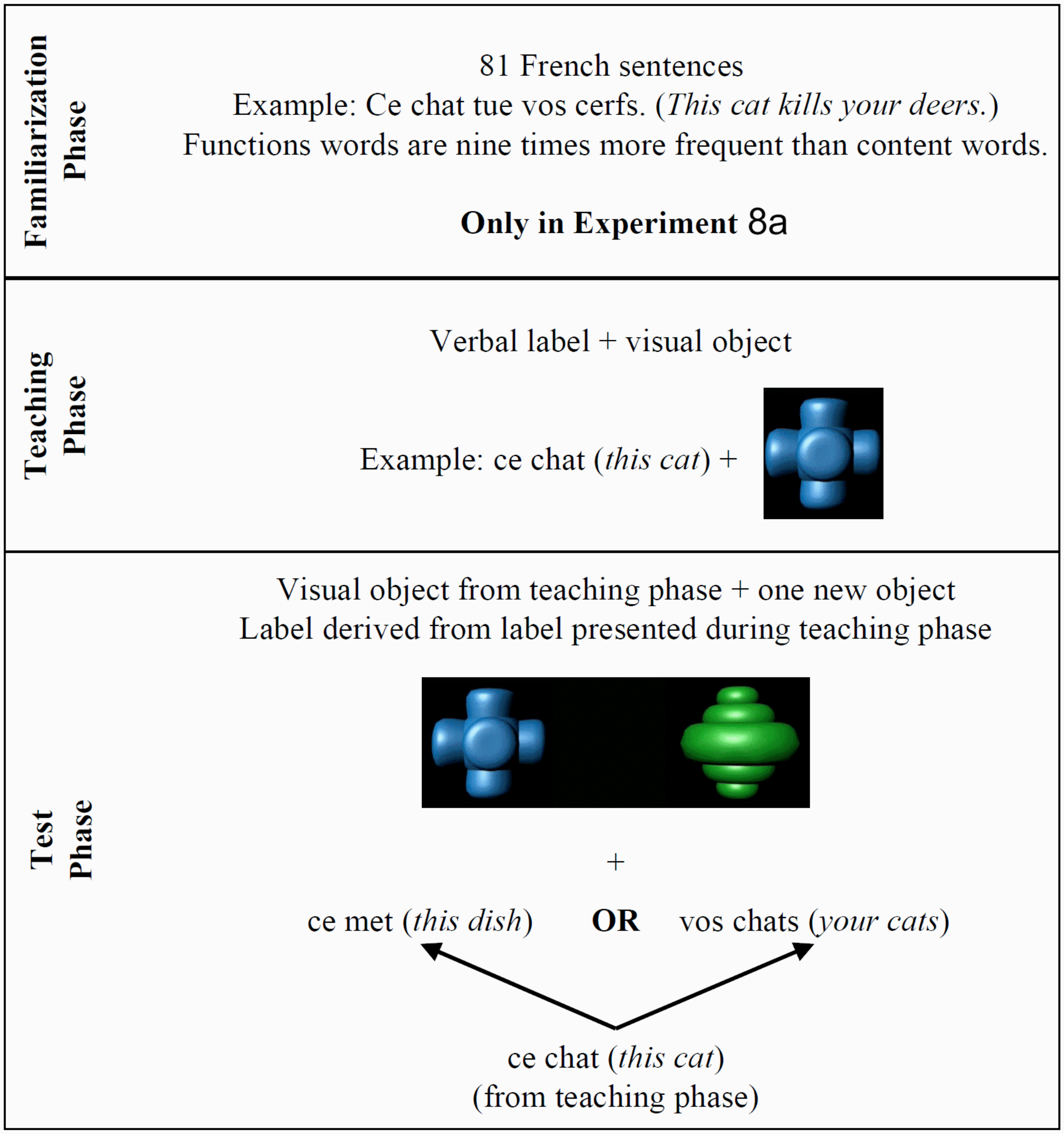

Figure 3.3 - Schematic representation of the design of Experiments 8a and 8b. (Top) Italian infants were first familiarized to a sequence of simple French sentences in which determiners were much more frequent than nouns. This familiarization phase was administered only in Experiment 8a but not in Experiment 8b. (Middle) Following the familiarization phase (in Experiment 8a), or the start of the experiment (in Experiment 8b), infants saw a visual object on a computer screen and simultaneously heard a label composed of a determiner and a noun. We reasoned that they would associate the label with the object. (Bottom) Following the teaching phase, infants took part in the test phase. They saw two visual objects on the screen, presented side-by-side. One was the familiar object from the teaching phase, while the other one was novel. Simultaneously, infants heard two types of labels, both derived from the label used during the teaching phase. The label had either the same noun but a different determiner compared to the label played in the teaching phase, or the same determiner and a different noun. We measured which of the two objects infants would fixate first, as a function of the type of label they heard, using a Tobii eye-tracker.

3.3.1 Experiment 8a

3.3.1.1 Materials and Methods

Figure 3.3 presents the experimental paradigm for Experiments 8a and 8b. Experiment 8a consisted of three phases: the familiarization phase, the teaching phase, and the test phase. Two successive phases were separated by a visual fixation attractor, that is, a white cross presented centrally on the screen and moving back and forth. The experimenter started each phase by pressing a key when the infant was looking at the central fixation attractor.

Participants

Twenty-eight Italian 17-month old infants were tested. Twelve were excluded for fussiness (5), equipment failure (insufficient eye tracker data, 3), the mother not following experimental instructions (1) or side bias (3). The remaining 16 infants (4 males, 12 females, age range: 17 months and 5 days - 17 months and 30 days) were included in the final analysis. Importantly, all infants acquired Italian and had no experience with French (the language used for our stimuli).

Stimuli

French Sentences

In the familiarization phase, we presented infants with French sentences that each contained (in addition to a verb) two highly frequent determiners and two less frequent nouns. The list of words is presented in Appendix B. Eighty-one sentences were created (see Appendix C). Each sentence was five words long and conformed to the following pattern: determiner-noun-verb-determiner-noun (e.g. “ce chat tue vos cerfs”; “this cat kills your deers”). The two determiners were always “ce” ([sə]; “this”) and “vos” ([vo]; “your”). Each determiner could be followed by nine different nouns, yielding to a total of 18 different determiner-noun combinations. Nine different verbs were used. Appendix B shows the phonetic transcriptions of all words, their number of occurrences in sentence-initial and sentence-final positions (that is, before or after the verb), respectively, and their English translations. All words were monosyllabic and had either a consonant-vowel or a consonant-vowel-consonant syllable structure. As a result, the determiners “ce” and “vos” were the most frequent words (and syllables) during the familiarization.

Forty-one sentences started with the determiner “ce” and forty started with the determiner “vos”. The sentences were recorded from a female native speaker of French. The sentences, and all speech stimuli used in the present study, were recorded in a sound-attenuated chamber using a Sony ECM-S959C microphone connected to an M-Audio pre-amplifier and stored in the aiff file format (sample rate 44.1 kHz, sample size 16 Bit). Files were normalized to a mean intensity of 60 dB using PRAAT (Version 5.0.25) sound processing software (Boersma & Weenik, 2008). Sentences lasted 1.4 s on average.

Objects

We used the same objects and same visual stimuli as in Experiment 7.

Labels

Four tokens of the phrases “ce chat” ([səʃa], “this cat”), “vos mets” ([vomɛ], “your dishes”), “ce met” ([səmɛ], “this dish”) and “vos chats” ([voʃa], “your cats”) were recorded from the same native speaker of French who also produced the familiarization sentences. (Note that in French, even though the spelling varies, the words “chat/chats” and “met/mets” are pronounced identically in the singular and plural forms.) The labels were recorded and stored in the same way as the French sentences described above. Labels had an average duration of 500 ms.

Two of these labels were used in the familiarization phase: “ce chat” or “vos mets”. The other two labels were used in the test phase: “ce met” and “vos chats”.

Procedure

Familiarization

The 81 French sentences were played from a loudspeaker located behind the screen. They were presented in random order with a silence of 1 s between two sentences. While the sentences were played, the computer screen showed a silent movie. The movie consisted of a rotating chessboard changing its rotation direction every 30 s. The movie was chosen to be attractive enough to keep infants attentive, while containing no three-dimensional shape that could be interpreted as an object. The entire familiarization lasted 3 min 15 s.

Teaching Phase

In the teaching phase, infants saw a visual object on the screen and simultaneously heard a speech label. We used the two shapes described above and two labels composed of a determiner and a noun ("ce chat" - ([səʃa], "this cat"; "vos mets" - [vomɛ], "your dishes"). All four words had occurred during the familiarization phase. Moreover, during the familiarization phase, the nouns selected for the teaching phase had occurred only in the determiner-noun combinations that were also used in the teaching phase, but never with the other determiner. The four stimulus combinations resulting from two objects and two labels were counterbalanced across infants. Each infant learned only one object-label combination.

During the teaching phase, the object moved from one side to the other on an LCD screen while rotating around its axes. Simultaneously, a label was repeated 23 times. Two repetitions of the label were separated by about 900 ms of silence, yielding a total duration of 33 s.

The presentation movie could be interrupted if the infant looked away for more than 2 s. However, all the 16 infants included in the analysis looked at the entire movie without interruption.

Test Phase

Following the familiarization phase and the teaching phase, infants completed four test trials. Two successive trials were separated by the presentation of the central fixation attractor (i.e. a white cross moving back and forth). The experimenter started each trial by pressing a key when the infant was looking at the central fixation attractor.

Once the experimenter started a trial, the fixation attractor disappeared, and the two objects presented in Figure 3.1 appeared on the computer screen. The objects then remained static and visible on the screen for the entire trial duration. Two seconds after the trial started, the central fixation attractor appeared again for 3.32 s. A test label (see below) was first pronounced while the central fixation attractor was still visible on the screen; the offset of the label was synchronized with the attractor's disappearance. Following this, the test label was repeated five more times, two consecutive repetitions being separated by 1s of silence.

Each of the objects on the screen occupied a surface of 8.5 cm x 9 cm. The centers of the two objects were separated by about 19 cm. For each infant, the position of objects was counterbalanced across trials. We also counterbalanced the position in which the blue object appears in the first trial across infants.

Each infants completed two test trials in the "same-determiner" condition, and two test trials in the "same-noun" condition. The two conditions were identical except for the test label used. In both conditions, the label differed only in one word from the label heard during the teaching phase. In the same-determiner condition, the label had the same determiner as during the teaching phase, but a different noun. In the same-noun condition, the label had the same noun as during the teaching phase, but a different determiner. Specifically, infants who had heard the label "ce chat" ("this cat") during the teaching phase heard "ce met" ("this dish") in the same-determiner condition, and "vos chats" ("your cats") in the same-noun condition. Infant who had heard the label "vos mets" ("your dishes") in the teaching phase, heard "vos chats" ("your cats") in the same-determiner condition and "ce met" ("this dish") in the same-noun condition. Importantly, while all words had occurred during the familiarization phase, the specific label phrases used during the test phase were all new, and had not occurred during the familiarization phase.

Analysis

We defined two windows of interest for the analysis of the infant eye gaze. Each was a square of 11 cm x 13 cm, centered on one of the objects. Infants' looking behavior was monitored using a Tobii 1720 Eye-tracker system and the Clearview 2.5.1 software package. Only infants for whom the eye tracker data for each trial contained at least one fixation of at least 100 ms in one of the two windows of interest were included in the analysis. Infants not meeting these criteria were excluded for insufficient eye tracker data. Considering the four test trials, infants for whom more than 70% of the total fixation time was spent in only one of the windows of interest were considered to exhibit a side bias and were rejected from further analysis.

As dependent variable for the main analysis, we considered the first look to one of the objects following the first presentation of the test label. The first look was defined as the first uninterrupted fixation of at least 100 ms in one of the two windows of interest described above.

In each test trial, infants were coded either as first looking at the familiar object from the teaching phase, or as first looking at the novel object. Infants who did not look at any object in at least one of the four test trials were excluded from the analysis. For each infant and each condition, we computed the proportion of trials in which they first oriented towards the familiar object from the teaching phase. We then assessed whether these proportions differed between the conditions using a two-tailed Wilcoxon signed rank test.

We further analyzed the individual choices of each infant using a "distribution analysis". Infants were categorized into three groups: (i) those associating the familiar object from the teaching phase more with the determiner than with the noun ("*determiner associators*"), (ii) those associating the familiar object more with the noun than with the determiner ("*noun associators*"), and (iii) those associating the familiar object equally with the noun and with the determiner ("*neutral associators*"). Determiner associators were infants who were more likely to look first towards the object from the teaching phase in the same-determiner condition than in the same-noun condition. Noun associators were infants who were more likely to look first towards the object from the teaching phase in the same-noun condition than in the same-determiner condition. Finally, neutral associators were infants who were equally likely to look first towards the object from the teaching phase in the same-noun condition and in the same-determiner condition.

We determined the expected distribution of noun associators, determiner associators and neutral associators, respectively, in the following way. In each trial, infants scored 1 if they first looked at the familiar object, and 0 if they first looked at the novel object. With two trials per condition, each infant could obtain a total score of 0, 1 or 2 in each condition (Note that all infants included in the analysis completed all four trials.) If infants oriented randomly towards the two objects in the test phase, the probabilities to orient first towards the familiar object on 0, 1 and 2 occasions were .25, .5 and .25, respectively for each condition (assuming that the conditions are statistically independent). To be a determiner associator, infants had to fall into one of the following three cases: (i) one initial orientation towards the familiar object in the same-determiner condition, and no such orientation in the same-noun condition; (ii) two initial orientations towards the familiar object in the same-determiner condition, and one such orientation in the same-noun condition; (iii) two initial orientations towards the familiar object in the same-determiner condition, and no such orientations in the same-noun condition. Summing over the individual probabilities of these cases, the probability of being a determiner associatior was .5*.25+.25*.5+.25*.25 = .3125. Symmetrically, the probability of being a noun associator was also .3125. To be a neutral associator, infants had to orient to the familiar object as often in both conditions. These infants could orient twice to the familiar object for each condition, once for each condition, or zero times for each condition. The probability of being a neutral associator was, therefore, .25*.25+.5*.5+.25*.25 = .375. Given these probabilities, we used an exact multinomial test to assess whether the observed distribution of determiner associators, noun associators and neutral associators was expected by chance.

3.3.1.2 Results

Figure 3.4 shows how often infants first looked towards the familiar object from the teaching phase as a function of the experimental condition. Infants were significantly more likely to look first to the familiar object in the same-noun condition (M = 1.38, SD = .72, Mdn = 1.5) than in the same-determiner condition (M = .69, SD = .70, Mdn = 1), Wilcoxon signed rank test, $W = 12$, $p = 0.016$, $CI_{.95,\text{Mdn difference}} = 0, 1$.

These results were further confirmed by the distribution analysis. Eleven infants out of sixteen were noun associators, two were determiner associators, and three were neutral associators. This distribution was significantly different from the distribution expected by chance, $p = 0.0077$ (exact multinomial test).

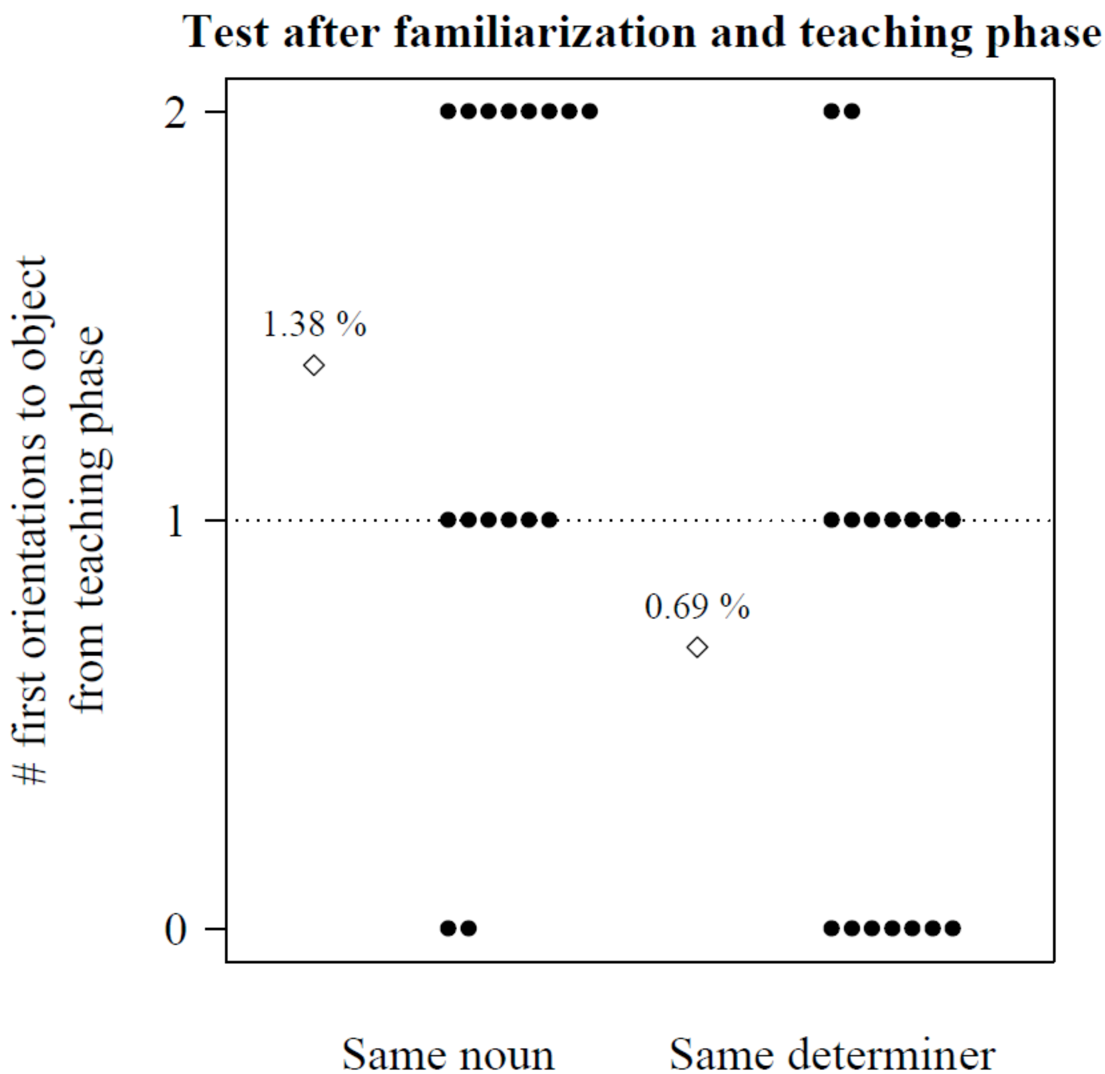

Figure 3.4 - Results of Experiment 8a. Proportion of first looks directed towards the familiar object from the teaching phase. Dots represent the means of individual participants, diamonds sample averages, and the dotted line the chance level of 50%. When the label used during the test phase had the same noun but a different determiner as the label presented during the teaching phase, infants were more likely to fixate first the familiar object from the teaching phase compared to the condition where the label used during the test phase had the same determiner but a different noun as the label presented during the teaching phase.

To make sure that each infant's fixation pattern was not based on what they happened to fixate *before* the offset of the label, we also performed separate analyses for trials where infants already fixated an object (rather than the central attractor) before the offset of the label. In total, there were 14 such trials (21.9 % of the trials). Among these 14 trials, 8 occurred in the same-determiner condition and 6 in the same-noun condition. After removing these 14 trials, our central pattern of results remained unchanged: infants were significantly more likely to look first to the familiar object in the same-noun condition ($M = 1.14$, $SD = .77$) than in the same-determiner condition ($M = .5$, $SD = .65$), Wilcoxon signed rank test, $W = 5$, $p = 0.020$.

3.3.1.3 Discussion

Experiment 8a asked whether infants could use cues provided by the speech signal to identify possible function words and possible content words, even when exposed to an unfamiliar language. Italian 17-month old infants were first exposed to a series of French sentences. Following this, they saw a visual object and heard simultaneously a bisyllabic label consisting of a highly frequent determiner and a less frequent noun (e.g., "ce chat", "this cat"). We hypothesized that potential content words should be more likely to be associated with a referent than potential function words. If this is the case, and if Italian infants can use cues provided by French speech to identify function words, then they should associate the visual object more strongly with the noun from the label than with the determiner from the label.

To test this hypothesis, infants then saw two objects presented side-by-side. One was the visual object they had just seen, while the other one was novel. Simultaneously, infants heard a label derived from that used during the teaching phase. The label was derived either by changing the determiner and keeping the noun the same, or by changing the noun and keeping the determiner the same. Infants were more likely to first orient towards the familiar object when the derived label had the same noun as the label from the teaching phase, compared to the condition where the derived label had the same determiner as the label from the teaching phase. These results suggest that infants formed a stronger association between the object and the noun than between the object and the determiner.

While the results of Experiment 8a suggest that infants treated words with a relatively low frequency as more content-word like than words with a relatively high frequency, it is less clear which cues infants used to identify potential function words. Infants might have used the frequency of the determiners to identify them as potential function words, as the determiners were much more frequent than any other content word in the familiarization phase of Experiment 8a. However, they might also have used a different cue. Specifically, in French, as in most languages of the world, content words are usually more salient than function words because they bear the main prominence (Nespor & Vogel, 1986). If infants pay more attention to more salient syllables than to less salient syllables, they might have associated the object more strongly with nouns than with determiners for this reason, and not because of their sensitivity to distributional information.

Experiment 8b was designed to control for this possibility. Experiment 8b was identical to Experiment 8a except that infants did not undergo the familiarization phase, but only the teaching and test phases. Crucially, these two phases used exactly the same stimuli as in Experiment 8a. As the stimuli used in these phases were naturally recorded, the nouns were still more salient than the determiners. Hence, if acoustical properties such as stress, or any other property specific to the stimuli of the teaching and test phases, are sufficient to account for the infants' preferential association between nouns and the visual objects, we would expect similar results as in Experiment 8a. Conversely, if the distributional properties such as the frequency of occurrence determined infants' behavior in Experiment 8a, we would expect different results in Experiment 8b.

3.3.2 Experiment 8b

3.3.2.1 Materials and Methods

Experiment 8b was identical to Experiment 8a, except that there was no familiarization phase prior to the teaching and test phases.

Twenty-one new Italian 17-month old infants were tested. Five were excluded for fussiness (1), equipment failure (1) or side bias (3). The remaining 16 infants (5 males, 11 females, age range: 17 months and 2 days - 17 months and 22 days) were included into the final analysis. As in Experiment 8a, these infants acquired Italian and had no experience with French (the language we used for our stimuli).

3.3.2.2 Results

Figure 3.5 shows how often infants first looked towards the familiar object from the teaching phase as a function of the experimental condition. Infants started looking towards the familiar object from the teaching phase as often in the same-noun condition (M = 1.13 SD = .62, Mdn = 1) as in the same-determiner condition (M = 1.06, SD = .68, Mdn = 1), Wilcoxon signed rank test, $W = 25$, $p = 0.82$, $CI_{.95,\text{Mdn difference}}$ = -.5, .5.

We further analyzed the results of Experiment 8b using the distribution analysis. Six infants out of 16 were noun associators, five were determiner associators, and five were neutral associators. This distribution was not significantly different from chance, as assessed by an exact multinomial test, $p > 0.9$.

Figure 3.6 shows the comparison of Experiments 8a and 8b. We calculated for each infant the difference in proportions of first orienting towards the familiar object in the same-noun condition and in the same-determiner condition, respectively. This difference in Experiment 8a was marginally different from that in Experiment 8b, Wilcoxon rank sum test, $U = 313$, $p = 0.054$, $CI_{.95,\text{Mdn difference}}$ = 0, 1.

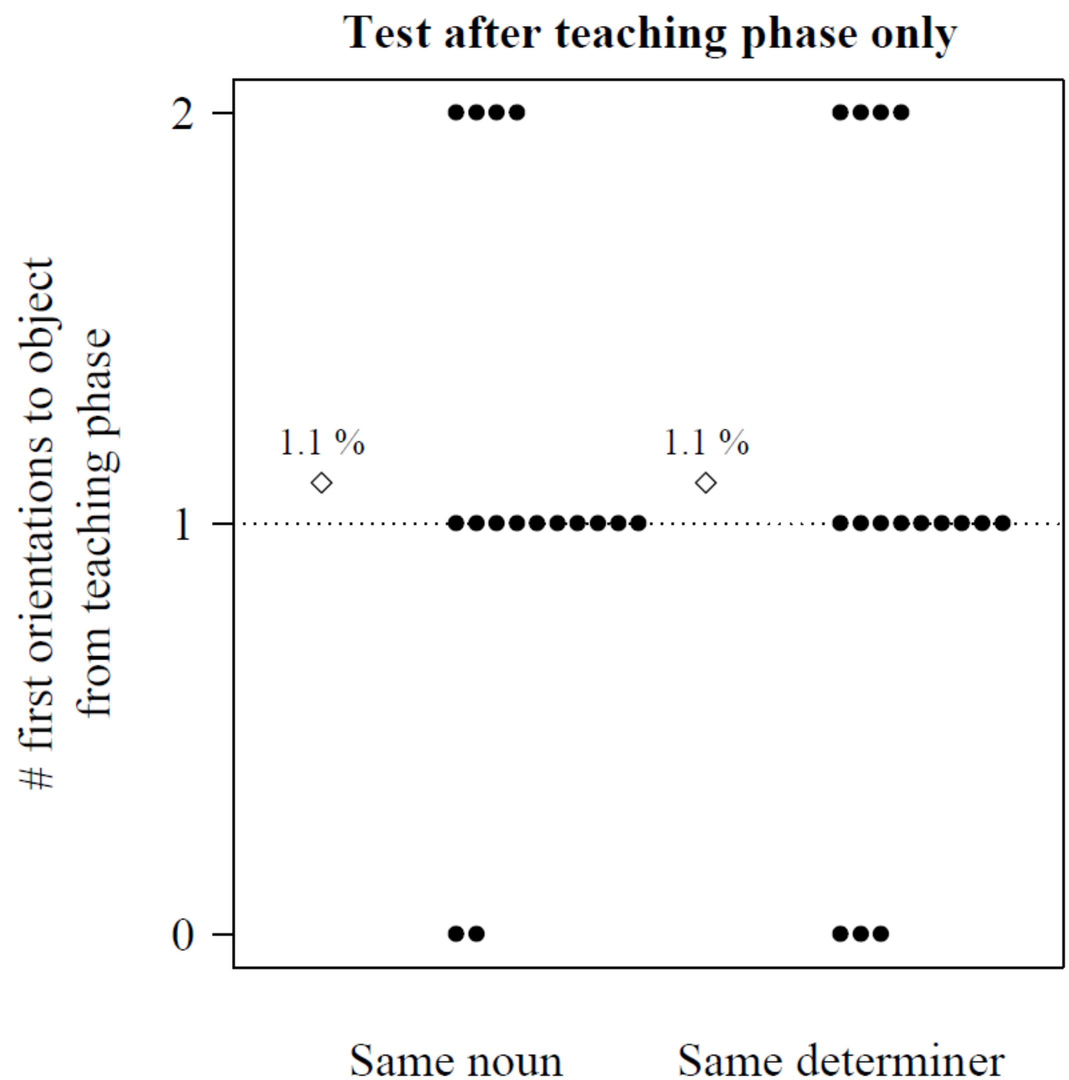

Figure 3.5 - Results of Experiment 8b. Proportion of first looks directed towards the familiar object from the teaching phase. Dots represent the means of individual participants, diamonds sample averages, and the dotted line the chance level of 50%. During the test phase, infants were as likely to fixate first the familiar object from the teaching phase as to fixate first a new object, both when the label used during the test phase had the same noun but a different determiner as the label presented during the teaching phase, and when the label used during the test phase had the same determiner but a different noun as the label presented during the teaching phase.

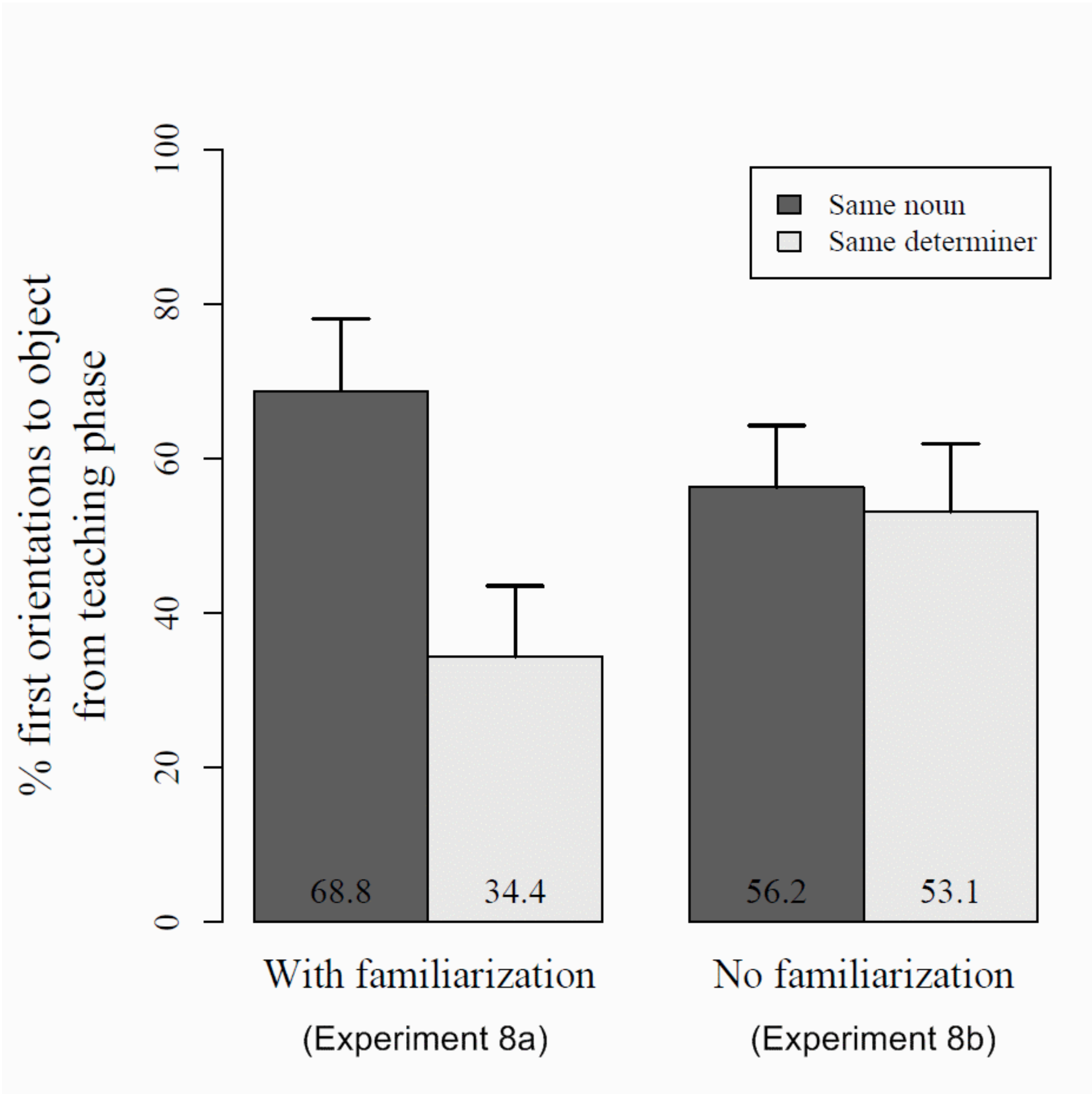

Figure 3.6 - Results of Experiments 8a and 8b. Proportion of first looks directed towards the familiar object from the teaching phase. Bars show the sample averages in the different conditions and errors bars the standard errors from the average.

To make sure that each infant's fixation pattern was not based on what they happened to fixate *before* the offset of the label, we performed separate analyses for trials where infants already fixated an object (rather than the central attractor) before the offset of the label. Infants were not looking at the central fixation when the test label was pronounced in 7 trials (10.94% of all trials). Among these 7 trials, 4 occurred in the same-determiner condition and 3 in the same-noun condition. After removing these 7 trials, our central pattern of results remained unchanged: infants started looking towards the familiar object from the teaching phase as often in the same-noun condition (M = 1.063, SD = .57) as in the same-determiner condition (M = .878, SD = .72), Wilcoxon signed rank test, $W = 12$, $p = 0.562$. In that analysis, the difference between Experiment 8a and 9b only showed a trend, Wilcoxon rank sum test, $U = 254$, $p = 0.104$.

3.3.2.3 Discussion

In Experiment 8a, infants preferentially associated visual objects with nouns rather than with determiners. A possible conclusion from this experiment is that infants might have used the very high frequency of the determiners to identify them as function words, and thus to disprefer them as carriers of meaning. However, infants might also have relied on acoustic or phonological properties of the words used in the teaching and test phases, especially on the greater saliency of nouns compared to determiners.

Experiment 8b controlled for this possibility by replicating Experiment 8a without the initial familiarization. As a result, Experiment 8b provided infants with all acoustic and phonological cues present in Experiment 8a, but determiners were no longer more frequent than nouns.

If the results of Experiment 8a had been carried exclusively by the greater acoustic saliency of nouns relative to determiners, one would expect similar results in Experiment 8a and 9b. In contrast, if distributional information, and especially the high frequency of determiners, influenced the infants' behavior in Experiment 8a, the results in Experiment 8b should differ from those in Experiment 8a.

In contrast to Experiment 8a, where infants' first fixations tended to be directed towards the familiar object in the same-noun condition, and towards the novel object in the same-determiner condition, infants' first fixations in Experiment 8b were directed equally often to either object. It thus seems safe to conclude that, in Experiment 8b, infants did not preferentially associate visual objects with nouns compared to determiners. Hence, the high frequency of determiners in the familiarization phase of Experiment 8a seems to play a crucial role for establishing the preferential association between visual objects and nouns.

This interpretation is strengthened when comparing the results of Experiment 8b to those of Experiment 7. Experiment 7 was similar to Experiment 8b, except that, during the test phase, infants were presented either with a label that was identical to that used during the teaching phase, or with an entirely novel label with both a different determiner and a different noun (as opposed to Experiments 8a and 8b, where the test labels changed either the determiner or the noun, but not both). Compared to the condition with the old label from the teaching phase, the infants' first look was directed more often towards the novel object when presented with the entirely novel label. These results contrast with those of Experiment 8b, where infants were equally likely to orient to either object in both conditions. A tentative conclusion is that infants recognized that parts of the test labels in Experiment 8b were identical to the labels used during the teaching phase; in fact, if they had considered a label with a changed noun or with a changed determiner as a novel label, they should orient towards the novel object as in Experiment 7.

Hence, although infants in Experiment 8b seemed to recognize that parts of the test labels were identical to the label used during the teaching phase, and although all acoustic and phonological cues provided in Experiment 8a were also present in Experiment 8b, we did not observe the preferential association between visual objects and nouns as in Experiment 8a. Together, these results thus suggest that this preferential association cannot be explained based exclusively on acoustic or phonological properties of the labels. Rather, the combined results of Experiment 8a and 9b suggest that the distributional cues implemented in the familiarization phase of Experiment 8a influenced the infants' behavior in the successive phases. Specifically, we suggest that infants prefer to associate a novel object with an infrequent word rather than with a frequent word, because infrequent words are more likely to be content words and frequent words are more likely to be function words.

We may however consider an alternative account of our data. Indeed, utterances in Italian (the language of our participants) often start with function words (Gervain et al., 2008), which are not directly associated to referents. By 17-months of age, infants may thus have learned that the first word of an utterance is not referential. Even though the results of Experiment 8b suggest that position alone cannot account for the results of Experiment 8a, infants might require the familiarization phase of Experiment 8a, not to detect frequency distributions, but to understand that labels are composed of two words – rather than of a single, bisyllabic word. Specifically, they might notice that the transitional probabilities between determiners and nouns are low; according to many authors (e.g., Aslin, Saffran, & Newport, 1998; Saffran, Aslin, & Newport, 1996), they might postulate word boundaries where transitional probabilities are low. As a result, infants might consider the first part of the labels as non-referential in Experiment 8a (where it would correspond to a word), but not in Experiment 8b (where it would correspond to the first syllable of a bisyllabic word). We will reconsider this account after presenting Experiments 9-11.

3.3.3 Appendix – List of words

Table 3.1

Words used in the sentences of the familiarization phase of Experiment 8a

Word	**Pronunciation**	**Occurrences in sentences beginning with "ce"**	**Occurrences in sentences beginning with "vos"**	**English Translation**
		Nouns following the determiner "ce"		
Rat	[ʁa]	5	4	rat
chat	[ʃa]	4	5	cat
bas	[ba]	5	4	stocking
pot	[po]	5	4	jar
show	[ʃo]	5	4	show
corps	[kɔʁ]	4	4	body
pet	[pɛ]	4	5	fart
fait	[fɛ]	4	5	fact
verre	[vɛʁ]	5	5	glass

Nouns following the determiner "vos"

cas	[ka]	5	5	cases
pas	[pa]	5	4	steps
mas	[ma]	4	4	Provencal houses
mots	[mo]	4	5	words
seaux	[so]	5	4	buckets
bords	[bɔʁ]	5	5	sides
mets	[mɛ]	5	4	dishes
quais	[kɛ]	4	4	platforms
cerfs	[sɛʁ]	4	5	stags

Verbs

lie/lient	[li]	3	5	links/link
scie/scient	[si]	4	5	saws/saw
nie/nient	[ni]	5	5	denies/deny
noue/nouent	[nu]	5	3	knot/knot
joue/jouent	[ʒu]	4	5	plays/play
loue/louent	[lu]	5	4	praise(s)
voit/voient	[vwa]	5	5	sees/see
noit/noient	[nwa]	4	4	drowns /drown
tue/tuent	[ty]	6	4	kills/kill

3.3.4 Appendix – List of sentences

Table 3.2

List of sentences used in the familiarization phase of Experiment 8a

Sentences beginning with "ce"	Sentences beginning with "vos"
Ce rat tue vos cerfs.	Vos mas lient ce rat.
Ce pet joue vos quais.	Vos quais voient ce rat.
Ce show noie vos bords.	Vos seaux nouent ce show.
Ce verre tue vos seaux.	Vos cas tuent ce rat.
Ce fait lie vos seaux.	Vos bords scient ce chat.
Ce pot scie vos pas.	Vos mas louent ce corps.
Ce pot voit vos bords.	Vos pas voient ce fait.
Ce pet noie vos mas.	Vos mots noient ce pet.
Ce bas tue vos cas.	Vos cas noient ce bas.
Ce rat voit vos mets.	Vos cas scient ce corps.
Ce fait loue vos quais.	Vos pas louent ce verre.
Ce corps tue vos bords.	Vos mets scient ce fait.
Ce pet nie vos seaux.	Vos quais nient ce corps.
Ce corps nie vos mets.	Vos mots lient ce bas.
Ce corps voit vos mots.	Vos mets nient ce verre.
Ce corps loue vos pas.	Vos mas jouent ce show.
Ce bas nie vos pas.	Vos bords jouent ce bas.
Ce bas noue vos mets.	Vos cerfs scient ce verre.
Ce show noue vos mots.	Vos mets jouent ce chat.

Ce chat noue vos quais.
Ce pot nie vos cerfs.
Ce corp scie vos cerfs.
Ce fait voit vos cas.
Ce show lie vos cerfs.
Ce pet noue vos cas.
Ce bas joue vos seaux.
Ce chat nie vos mas.
Ce chat noie vos cas.
Ce chat scie vos seaux.
Ce show joue vos pas.
Ce fait tue vos mas.
Ce verre scie vos quais.
Ce show tue vos mets.
Ce rat noue vos bords.
Ce pot noie vos mets.
Ce pot loue vos mots.
Ce rat lie vos pas.
Ce verre voit vos mas.
Ce bas loue vos bords.
Ce rat joue vos mots.
Ce verre loue vos cas.

Vos bords nient ce pet.
Vos mots tuent ce fait.
Vos quais noient ce pot.
Vos bords tuent ce verre.
Vos mas scient ce pot.
Vos cerfs voient ce bas.
Vos bords lient ce fait.
Vos mots voient ce verre.
Vos pas noient ce chat.
Vos cerfs louent ce fait.
Vos mots nient ce chat.
Vos seaux louent ce pot.
Vos cas nient ce pot.
Vos quais tuent ce show.
Vos cerfs nouent ce chat.
Vos mets lient ce pet.
Vos cas lient ce show.
Vos seaux voient ce corps.
Vos seaux jouent ce rat.
Vos cerfs jouent ce pet.
Vos pas nouent ce pet.

3.4 Frequency as a cue to identify function words – Experiment 9

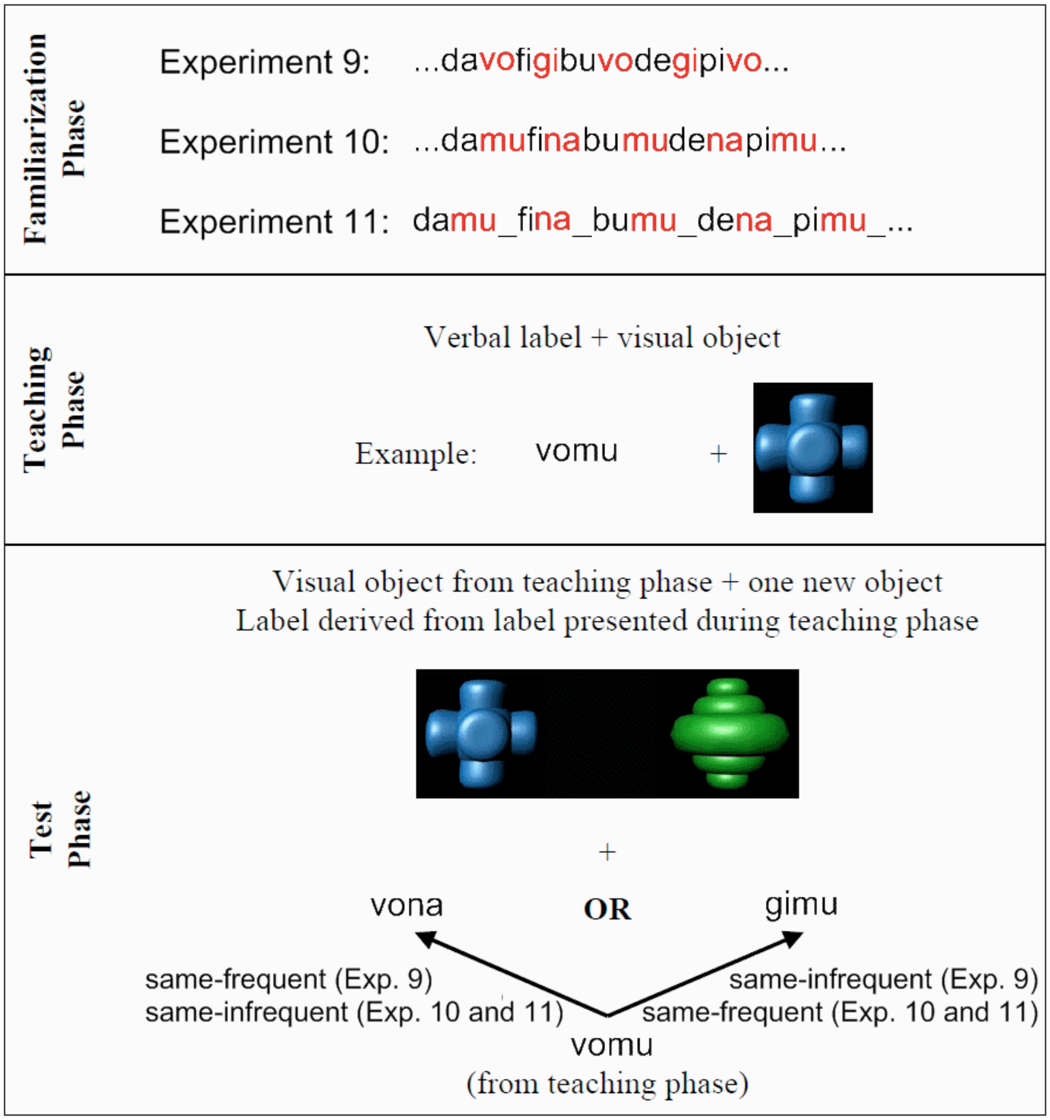

Figure 3.7 - Schematic representation of the design of Experiments 9, 10 and 11. (Top) Infants were first familiarized to an artificial speech stream, where two syllables were more frequent than others. The frequent syllables were *gi* and *vo* in Experiment 9, *na* and *mu* in Experiments 10 and 11. The stream was continuous and ramped at the beginning and at the end for Experiments 9 and 10. In contrast, in Experiment 11, the stream was segmented through an insertion of a silent pause after each frequent syllable. (Middle) Following the familiarization phase, infants saw a visual object on a computer screen and simultaneously heard a label composed of a frequent and an infrequent syllables. We reasoned that they would associate the label with the object. (Bottom) Following the teaching phase, infants took part in the test phase. They saw two visual objects on the screen, presented side-by-side. One was the familiar object from the teaching phase, while the other one was novel. Simultaneously, infants heard two types of labels, both derived from the label used during the teaching phase. The label had either the same infrequent but a different frequent syllable compared to the label played in the teaching phase, or the frequent and a different infrequent syllable. We measured which of the two objects infants would fixate first, as a function of the type of label they heard, using a Tobii Eye-Tracker.

3.4.1 Materials and Methods

Figure 3.7 presents the experimental paradigm for Experiments 9, 10 and 11. Experiment 9 consisted of three phases: a familiarization phase, a teaching phase, and a test phase. Two successive phases were separated by a visual fixation attractor, that is, a white cross presented centrally on the screen and moving back and forth. The experimenter started each phase by pressing a key when the infant was looking at the central fixation attractor.

3.4.1.1 Participants

Forty-three Italian 17-month old infants were tested. Twenty were excluded for fussiness (7), equipment failure (insufficient eye tracker data, 8), the mother not following experimental instructions (2) or side bias (3). The remaining 23 infants (17 males, 7 females, age range: 17 month and 4 days - 18 month and 6 days) were included in the final analysis.

3.4.1.2 Stimuli

<u>Familiarization Stream</u>

In the familiarization phase, we presented infants with an artificial speech stream. The stream alternated between one of two very frequent syllables (*gi* or *vo*) and one of eighteen infrequent syllables (*va, da, ga, na, ta, pi, si, ni, fi, ko, bo, go, ku, bu, pu, mu, fe, de*). The two frequent syllables were therefore nine times more frequent than each of the infrequent syllables.

The stream was generated by the speech synthesizer MBROLA (fr4) with a phoneme duration of 120 ms and a monotonous pitch of 200 Hz. Frequent and infrequent syllables did *not* differ in terms of pitch, intensity or duration. The whole familiarization stream lasted 2 minutes.

<u>Objects</u>

We used the same objects and same visual stimuli as in Experiment 8.

<u>Labels</u>

Four bisyllabic labels were generated, associating the two frequent syllables (*gi* and *vo*) and two infrequent syllables of the familiarization (*mu* and *na*), thus generating the labels *vomu*, *gina*, *vona* and *gimu*.

Two of these labels were used in the familiarization phase: *vomu* or *gina*. The other two labels were used in the test phase: *gimu* and *vona*. The labels were generated by the speech synthesizer MBROLA (fr4) with a phoneme duration of 120 ms and a monotonous pitch of 200 Hz.

3.4.1.3 Procedure

<u>Familiarization</u>

The familiarization stream was played from a loudspeaker located behind the screen. While it was played, the computer screen showed a silent movie. The movie consisted of a rotating chessboard changing its rotation direction every 30 s. The movie was chosen to be attractive enough to keep infants attentive, while containing no three-dimensional shape that could be interpreted as an object. The entire familiarization lasted 2 min XX.

Teaching Phase

In the teaching phase, infants saw a visual object on the screen and simultaneously heard a speech label. We used the two shapes described above and two labels composed of a frequent and an infrequent syllable from the familiarization stream. The four stimulus combinations resulting from two objects and two labels were counterbalanced across infants. Each infant learned only one object-label combination.

During the teaching phase, the object moved from one side to the other on an LCD screen while rotating around its axes. Simultaneously, a label was repeated 24 times. Two repetitions of the label were separated by 900 ms of silence, yielding a total duration of 32 s.

The presentation movie could be interrupted if the infant looked away for more than 2 s. However, all the 24 infants included in the analysis looked at the entire movie without interruption.

Test Phase

Following the familiarization phase and the teaching phase, infants completed four test trials. Two successive trials were separated by the presentation of the central fixation attractor (i.e. a white cross moving back and forth). The experimenter started each trial by pressing a key when the infant was looking at the central fixation attractor.

Once the experimenter started a trial, the fixation attractor disappeared, and the two objects presented in Figure 3.1 appeared on the computer screen. The objects then remained static and visible on the screen for the entire trial duration. Two seconds after the trial started, the central fixation attractor appeared again for 3.32 s. A test label (see below) was first pronounced while the central fixation attractor was still visible on the screen; the offset of the label was synchronized with the attractor's disappearance. Following this, the test label was repeated five more times, two consecutive repetitions being separated by 900 ms of silence.

Each of the objects on the screen occupied a surface of 8.5 cm x 9 cm. The centers of the two objects were separated by about 19 cm. For each infant, the position of objects was counterbalanced across trials. We also counterbalanced the position in which the blue object appears in the first trial across infants.

Each infant completed two test trials in the "same-frequent" condition, and two test trials in the "same-infrequent" condition. The two conditions were identical except for the test label used. In both conditions, the label differed only in one syllable from the label heard during the teaching phase. In the same-frequent condition, the label had the same frequent syllable as during the teaching phase, but a different infrequent syllable. In the same-infrequent condition, the label had the same infrequent as during the teaching phase, but a different frequent (see Table 1).

3.4.1.4 Analysis

We defined two windows of interest for the analysis of the infant eye gaze. Each was a square of 11 cm x 13 cm, centered on one of the objects. Infants' looking behavior was monitored using a Tobii 1720 Eye-tracker system and the Clearview 2.5.1 software package. Only infants for whom the eye tracker data for each trial contained at least one fixation of at least 100 ms in one of the two windows of interest were included in the analysis. Infants not meeting these criteria were excluded for insufficient eye tracker data. Considering the four test trials, infants for whom more than 70% of the total fixation time was spent in only one of the windows of interest were considered to exhibit a side bias and were rejected from further analysis.

As dependent variable for the main analysis, we considered the first look to one of the objects following the first presentation of the test label. The first look was defined as the first uninterrupted fixation of at least 100 ms in one of the two windows of interest described above.

In each test trial, infants were coded either as first looking at the familiar object from the teaching phase, or as first looking at the novel object. Infants who did not look at any object in at least one of the four test trials were excluded from the analysis. For each infant and each condition, we computed the proportion of trials in which they first oriented towards the familiar object from the teaching phase. We then assessed whether these proportions differed between the conditions using a two-tailed Wilcoxon signed rank test.

We further analyzed the individual choices of each infant using a "distribution analysis". Infants were categorized into three groups: (i) "frequent-associators", (ii) "infrequent-associators" and (iii) "neutral associators". "Frequent-associators" are infants who are more likely to look first towards the object from the teaching phase in the same-frequent condition than in the same-infrequent condition. "Infrequent-associators" are infants who are more likely to look first towards the object from the teaching phase in the same-infrequent condition than in the same-frequent condition. Finally, "neutral associators" are infants who are equally likely to look first towards the object from the teaching phase in the same-infrequent condition and in the same-frequent condition.

We determined the expected distribution of infrequent-associators, frequent-associators and neutral associators, respectively, in the following way. In each trial, infants scored 1 if they first looked at the familiar object and 0 if they first looked at the novel object. With two trials per condition, each infant could obtain a total score of 0, 1 or 2 in each condition (Note that all infants included in the analysis completed all four trials.) If infants orient randomly towards the two objects in the test phase, the probabilities to orient first towards the familiar object on 0, 1 and 2 occasions are .25, .5 and .25, respectively for each condition (assuming that the conditions are statistically independent). To be a frequent-associator, infants have to fall into one of the following three cases: (i) one initial orientation towards the familiar object in the same-frequent condition, and no such orientation in the same-infrequent condition; (ii) two initial orientations towards the familiar object in the same-frequent condition, and one such orientation in the same-infrequent condition; (iii) two initial orientations towards the familiar object in the same-frequent condition, and no such orientations in the same-infrequent condition. Summing over the individual probabilities of these cases, the probability of being a frequent-associator is .5*.25+.25*.5+.25*.25 = .3125. Symmetrically, the probability of being an infrequent-associator is also .3125. The probability to be a neutral associator is therefore 1-.3125-.3125 = .375. Given these probabilities, we used an exact multinomial test to assess whether the observed distribution of frequent-associators, infrequent-associators and neutral associators was expected by chance.

3.4.2 Results

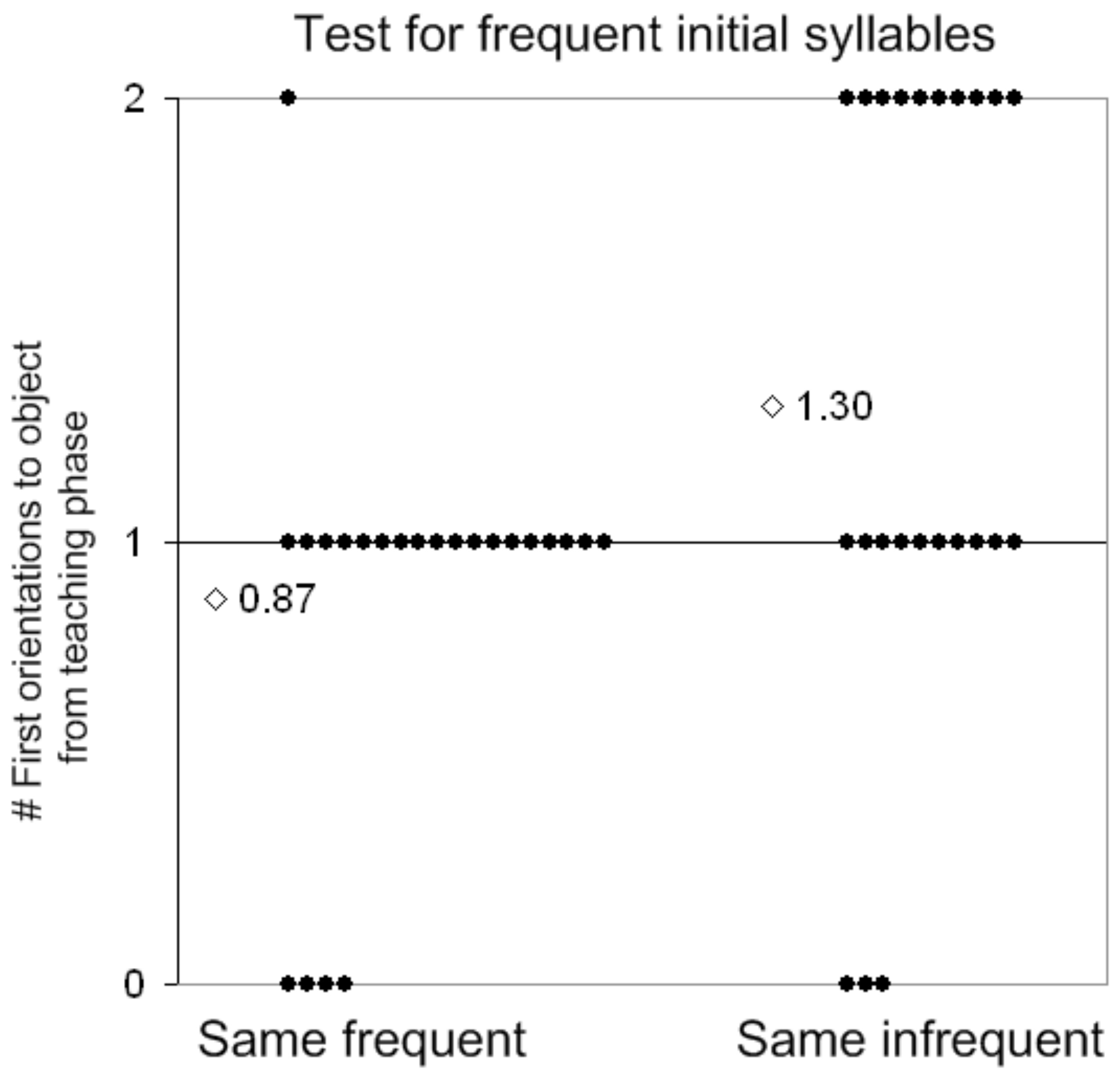

Figure 3.8 - Results of Experiment 9. Mean number of first looks directed towards the familiar object from the teaching phase. Error bars represent the standard error from the average. When the label used during the test phase had the same infrequent but not the same frequent syllable as the label used during the teaching phase, infants were more likely to first orient towards the object from the teaching phase than in the condition where the label used during the test phase had the same frequent but not the same infrequent syllable as the label used during the teaching phase.

Figure 3.8 shows how often infants first looked towards the familiar object from the teaching phase as a function of the experimental condition. Infants were faster to orient towards one or the other object in the same-frequent (RT = .89 s) than in the same-infrequent (RT = 1.02 s) condition. However this difference was not significant, $t(22) = .69$, $p = .50$. Infants were significantly more likely to look first to the familiar object in the same-infrequent condition (M = 1.30, SD = .70) than in the same-frequent condition (M = .87, SD = .46), Wilcoxon signed rank test, W = 19.5, p = .037.

These results were further confirmed by the distribution analysis. Eleven infants out of twenty-three were infrequent-associators, three were frequent-associators, and nine were neutral associators. However, this distribution was not significantly different from the distribution expected by chance, p = 0.097 (exact multinomial test).

To make sure that each infant's fixation pattern was not based on what they happened to fixate *before* the offset of the label, we also performed separate analyzes for trials where infants already fixated an object (rather than the central attractor) before the offset of the label. In total, there were 17 such trials (18.5% of the trials). Among these 17 trials, eight occurred in the same-infrequent condition and nine in the same-frequent condition. After removing these 17 trials, our central pattern of results remained unchanged: infants were significantly more likely to look first to the familiar object in the same-infrequent condition (M = 1.09, SD = .73) than in the same-frequent condition (M = .57, SD = .51), Wilcoxon signed rank test, W = 19.5, p = 0.018.

3.4.3 Discussion

The results of Experiment 9 replicate and extend those previously reported in Experiment 8a. After listening to a speech stream alternating between frequent and infrequent syllables, infants paired more strongly the infrequent than the frequent syllable of a label to a novel object. However, in contrast to our previous work, the familiarization speech stream was not produced by a natural speaker of a foreign language, but was generated by an artificial speech synthesizer. This allowed us to ensure that all syllables all along the experiment had the same pitch, intensity and duration. In consequence, our present results exclude the influence of any acoustic factor on the effect reported in Experiment 8a. This strengthens our view that infants consider very frequent syllables less likely than infrequent syllables to be associated to referents.

However, in addition to frequency, a second distributional cue may play a role in the effect that we observed. Indeed, in Experiments 8a and 9, frequent syllables always appeared in initial position. This position is also congruent with the position of determiners and more generally function words in Italian, the language being acquired by our participants. If the use of frequent syllables in Experiment 9 reflects a property that infants learned from exposure to their language, the effect might be restricted to initial frequent syllables. In contrast, if the non-referential use of frequent syllables reflects a mechanism supporting language acquisition, it should be generalizable to final frequent syllables.

In Experiment 10, we ask whether the effect observed in Experiment 9 for initial frequent syllables can be obtained for final frequent syllables. Furthermore, Experiment

10 also serves as a control for an alternative explanation of our results: infants may simply rely on final rather than initial syllables when learning a novel object label. If the later were true, frequency would not play any role in the results of Experiment 9, and we should expect the same results in Experiment 10.

3.5 The role of position in function word identification – Experiments 10 and 11

3.5.1 Experiment 10

3.5.1.1 Materials and Methods

Experiment 10 was identical to Experiment 9, except for the familiarization stream. Instead of *gi* and *vo*, which constitute the initial syllables of the object labels, the frequent syllables were *na* and *mu*, which constitute the final syllables of the object labels. The infrequent syllables were *va, da, ga, ta, gi, pi, si, ni, fi, ko, bo, go, vo, ku, bu, pu, fe* and *de*.

Participants

Fourty-one Italian 17-month old infants were tested. Seventeen were excluded for fussiness (8), equipment failure (insufficient eye tracker data, 8) or the mother not following experimental instructions (1). The remaining 24 infants (17 males, 7 females, age range: 17 month and 5 days - 18 month and 7 days) were included in the final analysis.

3.5.1.2 Results

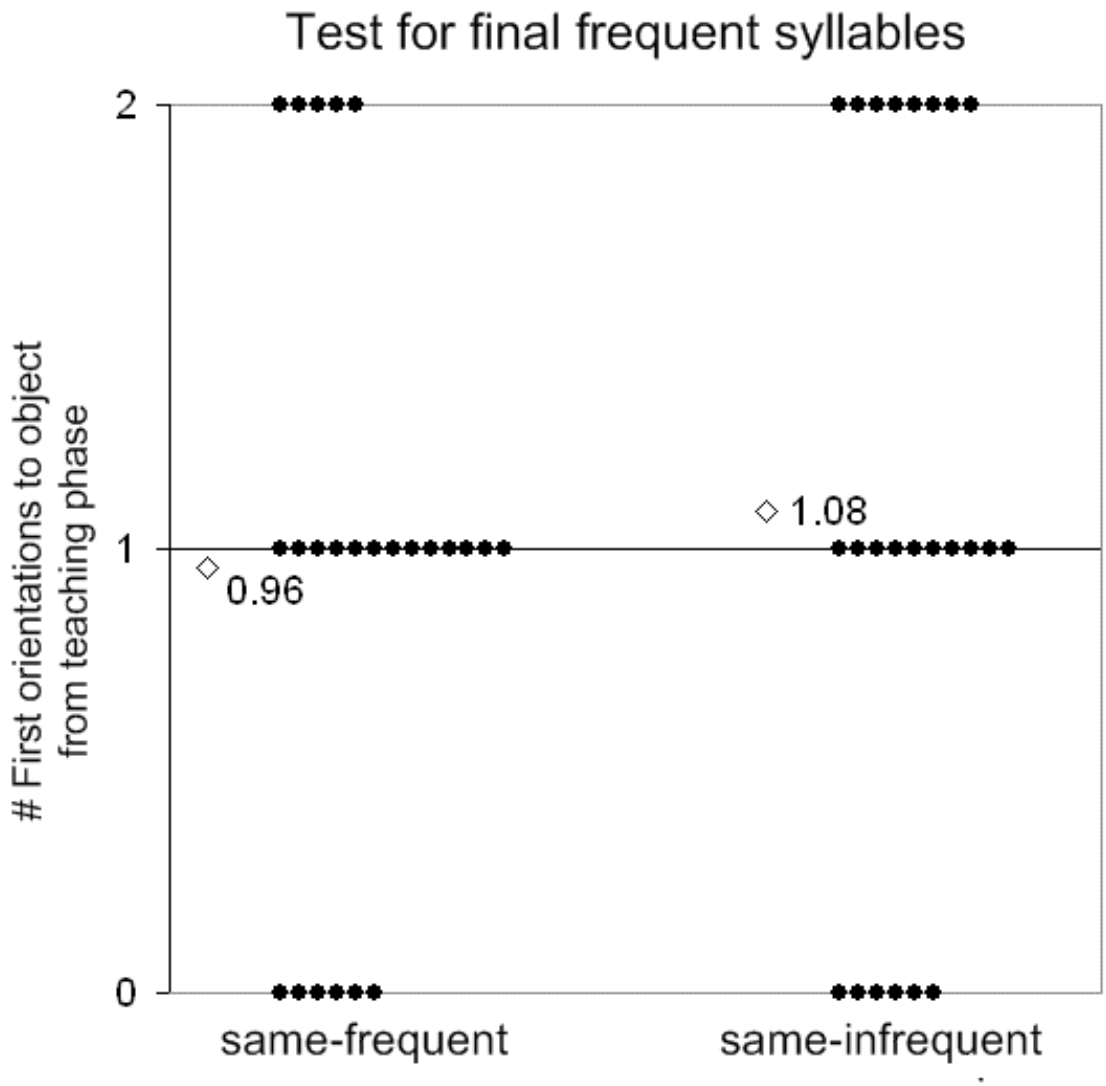

Figure 3.9 - Results of Experiment 10. Mean number of first looks directed towards the familiar object from the teaching phase. Error bars represent the standard error from the average. When the label used during the test phase had the same infrequent but not the same frequent syllable as the label used during the teaching phase, infants were not more likely to first orient towards the object from the teaching phase than in the condition where the label used during the test phase had the same frequent but not the same infrequent syllable as the label used during the teaching phase.

Figure 3.9 shows how often infants first looked towards the familiar object from the teaching phase as a function of the experimental condition. Infants were not faster to orient to one or the other object in the same-frequent (*RT* = 1.02 s) or in the same-infrequent (*RT* = 1.11 s) condition, t(23) = .60, p = .55. Infants started looking towards the familiar object from the teaching phase as often in the same-infrequent condition (*M* = 1.08 *SD* = .78) as in the same-frequent condition (*M* = .96, *SD* = .69), Wilcoxon signed rank test, $W = 81$, $p = 0.54$.

We further analyzed the results of Experiment 10 using the distribution analysis. Ten infants out of twenty-four were infrequent-associators, nine were frequent-associators, and five were neutral associators. This distribution was not significantly different from chance, as assessed by an exact multinomial test, $p > 0.22$.

To make sure that each infant's fixation pattern was not based on what they happened to fixate *before* the offset of the label, we also performed separate analyzes for trials where infants already fixated an object (rather than the central attractor) before the offset of the label. In total, there were 33 such trials (34% of the trials). Among these 33 trials, fifteen occurred in the same-infrequent condition and eighteen in the same-frequent condition. After removing these 33 trials, our central pattern of results remained unchanged: infants started looking towards the familiar object from the teaching phase as often in the same-infrequent condition (*M* = .75 *SD* = .74) as in the same-frequent condition (*M* = .58, *SD* = .65), Wilcoxon signed rank test, $W = 46.5$, $p = 0.51$.

3.5.1.3 Discussion

Experiment 10 differed from Experiment 9 in that the final syllables of object labels, rather than the initial syllables, occurred frequently in the previous familiarization stream. Contrary to Experiment 9, infants did not appear to form a stronger association with either the frequent or the infrequent syllable of a label. In Experiments 8a-8b, we had already shown that infants who did not undergo the familiarization phase and were directly asked to learn a novel object-label pair did not rely more on the first or last syllable of the label. Here, Experiment 10 further shows that frequency has no influence on the pairing of the label if the frequent syllable appears in final position.

Thus, two implications follow from the results of Experiment 10. First, the non-referential use of initial frequent syllables in Experiment 9 was a consequence of frequency, and not solely of the initial position. Second, the non-referential use of frequent syllables appears restricted to syllables occurring in initial position of a label. However, the latter implication may not be fully warranted yet. Indeed, even though no segmentation cue was provided in the familiarization of Experiments 9 and 10, both young infants (Gervain, Nespor, Mazuka, Horie & Mehler, 2008) and adults (Gervain et al., submitted) tend to segment continuous streams alternating between frequent and infrequent syllables, with respect to the distribution of frequent elements in their native language. In particular, Italian-learning 7-month-olds tend to perceive an artificial speech stream very similar to our familiarization stream, as a series of short sequences starting with frequent syllables. Thus, if we generalize Gervain et al.'s results to our situation, frequent elements might be perceived as initial in the familiarization of Experiment 10, whereas they appear in final position in the successive teaching and test phases. This incongruence in position may be responsible for the absence of effect. The frequent syllables extracted from the familiarization phase may be associated to the utterance-initial position, and infants may not recognize or be unable to use them in final position.

In Experiment 11, we ask whether the absence of a significant effect in Experiment 10 was due to a mismatch in the perceived position of frequent items in the familiarization and successive phases, or whether it meant that the effect is restricted to initial position. To this end, Experiment 11 was identical to Experiment 10, but frequent-final segmentation was forced by inserting silent pauses after each frequent syllable in the familiarization. The position of frequent syllables was therefore final in every phase of the experiment.

3.5.2 Experiment 11

3.5.2.1 Materials and Methods

Experiment 11 was identical to Experiment 10, except for the familiarization stream. The familiarization stream was segmented so that each frequent syllable was followed by a phoneme-long (120 ms) silent pause. In consequence, infants should perceive the familiarization stream as a series of sequences ending with a frequent syllable.

Participants

Forty-six Italian 17-month-old infants were tested. Twenty-two were excluded for fussiness (11), equipment failure (insufficient eye tracker data, 8), the mother not following experimental instructions (1) or side bias (2). The remaining 24 infants (17 males, 7 females, age range: 17 month and 05 days - 18 month and 01 day) were included in the final analysis.

3.5.2.2 Results

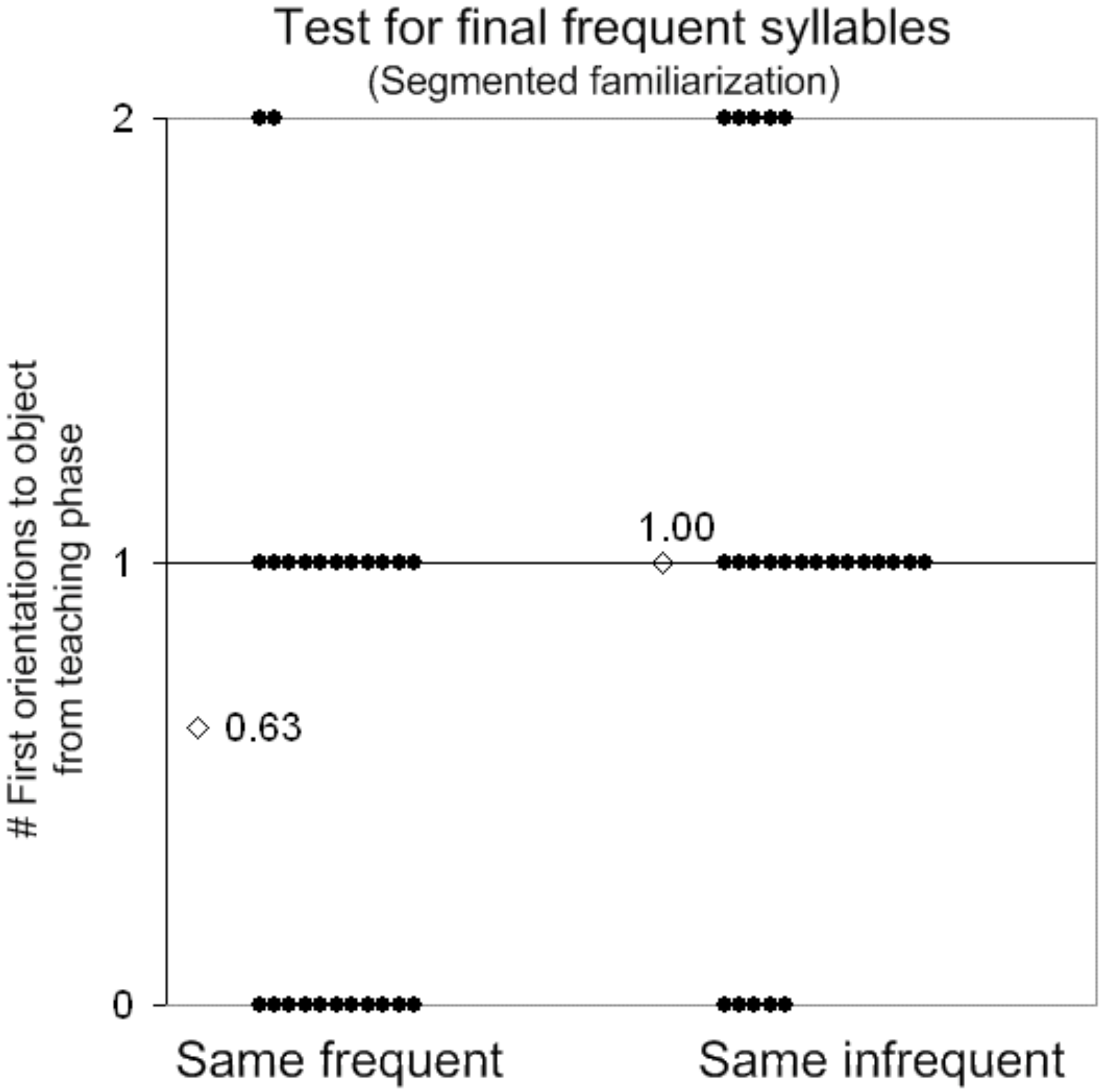

Figure 3.10 - Results of Experiment 11. Mean number of first looks directed towards the familiar object from the teaching phase. Error bars represent the standard error from the average. When the label used during the test phase had the same infrequent but not the same frequent syllable as the label used during the teaching phase, infants were more likely to first orient towards the object from the teaching phase than in the condition where the label used during the test phase had the same frequent but not the same infrequent syllable as the label used during the teaching phase.

Figure 3.10 shows how often infants first looked towards the familiar object from the teaching phase as a function of the experimental condition. Infants were not faster to orient to one or the other object in the same-frequent (*RT* = .93 s) or in the same-infrequent (*RT* = .90 s) condition, $t(23) = .18$, $p = .86$. Infants were significantly more likely to look first to the familiar object in the same-infrequent condition (*M* = 1, *SD* = .66) than in the same-frequent condition (*M* = .625, *SD* = .65), Wilcoxon signed rank test, $W = 32$, $p = .039$

These results were further confirmed by the distribution analysis. Twelve infants out of twenty-four were infrequent-associators, four were frequent-associators, and eight were neutral associators. However, this distribution was not significantly different from the distribution expected by chance, $p = 0.14$ (exact multinomial test).

To make sure that each infant's fixation pattern was not based on what they happened to fixate *before* the offset of the label, we also performed separate analyzes for trials where infants already fixated an object (rather than the central attractor) before the offset of the label. In total, there were 30 such trials (31% of the trials). Among these 30 trials, 13 occurred in the same-infrequent condition and 17 in the same-frequent condition. After removing these 30 trials, our central pattern of results remained unchanged: infants were significantly more likely to look first to the familiar object in the same-infrequent condition (*M* = .75, *SD* = .68) than in the same-frequent condition (*M* = .375, *SD* = .49), Wilcoxon signed rank test, $W = 12$, $p = 0.033$.

3.5.2.3 Discussion

In Experiment 11, contrary to Experiment 10, infants associated a novel object with the infrequent rather than with the frequent syllable of a label. Thus, we replicate the effect obtained in Experiments 8a and 9, using final frequent syllables. These results contrast with the absence of preference in Experiment 10. In Experiment 11, contrary to Experiments 9 and 10, the familiarization stream was segmented, so that frequent syllables would be perceived in final position of all bigrams constituting the stream. This was done to ensure that the frequent syllables would be perceived in the final position all along the experiment. Thus, we interpret that the failure in Experiment 10 was not necessarily due to an inability to use frequent items as non-referential in final position. Rather, this result can be due to a mismatch between the perceived position of the frequent items in the familiarization and subsequent phases of the experiment.

3.6 Identifying function words on the basis of frequency information

In the course of language acquisition, infants are confronted with the hard task of associating chunks of speech with semantic referents. The Gavagai problem, as stated by Quine (1960), illustrates the difficulty to select the appropriate referents in the world. Here, however, we consider the difficulty of identifying the appropriate speech chunk to associate to a given referent. Assuming one identified that a rabbit is the intended referent of the utterance “gavagai!”, what part of the utterance does actually mean “rabbit”? In particular, “gavagai” may well mean “*the* rabbit”, and thus should be stripped of the part that means “the”. To solve this problem, we propose that very frequent words are spontaneously used in a non-referential way.

In fact, the most frequent words in a language (i.e. determiners, prepositions, pronouns and conjunctions) do not refer to any particular perceivable object, action or property. Rather, they are mainly in the service of syntax. In line with their largely syntactic role, these words are called function words, as opposed to content words such as nouns, verbs, adjectives and adverbs. Function words are characterized by various distributional and phonological properties. However, it is unknown whether young infants can actually use such properties to identify potential function words. In the present study, we started addressing this issue, asking whether 17-month old infants can identify potential function words, based on a language-universal cue, namely their high frequency of occurrence.

3.6.1 Summary of our results

In Experiment 8a, we presented infants with short sentences in a language unknown to them. These sentences contained frequent and infrequent words, corresponding to actual function and content words, respectively. We asked whether infants were more likely to associate the relatively infrequent words with external referents as compared to the relatively frequent words. Our results suggest that infants associated the object more strongly with the infrequent noun than with the frequent determiner.

When the initial familiarization to French sentences was removed (Experiment 8b), however, infants failed to preferentially associate the object with the noun, although the results of Experiment 7 shows that infants can associate labels with objects in the absence of familiarization. Hence, the combined results of Experiment 8b and Experiment 7 suggest that the preferential association between content words and objects observed in Experiment 8a cannot be explained solely due to acoustic or phonological properties of the function words and content words. Rather, our results suggest that infants can use the distributional properties of words to infer certain of their properties.

Experiments 9 to 11 further confirmed these results. In these experiments, infants were first familiarized to an artificial speech stream alternating between frequent and infrequent syllables. A rotating chessboard was the only visual stimulus presented during this phase. Subsequently, infants learned to pair an object with a label consisting in the association of one frequent and one infrequent syllable. When infants had been familiarized with a continuous familiarization stream, we found that infants associated the object more strongly with the infrequent syllable than with the frequent syllable, only if that syllable was in initial position (Experiment 9) but not if it appeared in final position (Experiment 10). If, however, infants had been familiarized with a segmented familiarization stream, so that frequent syllables were perceived as sequence final during the familiarization, we found that infants associated the object more strongly with the infrequent initial syllable than with the frequent final syllable (Experiment 11).

Altogether, our results suggest that infants are less likely to associate objects with very frequent words (or syllables) than with relatively less frequent words (or syllables), raising the possibility that they might consider these words as function word candidates[3].

---

[3] Another way to explain our results is the possibility that infants might consider relatively infrequent words as better candidate for content words. However, infants participating in Experiment 1 are more likely to have extracted the frequent, rather than the infrequent words from the familiarization phase. Therefore, we consider that if they attributed a property to some words, they must have attributed a property to the

Below, we will discuss possible mechanisms that might account for these results, and that might also contribute to the identification of function words in natural language acquisition.

3.6.2 Word frequency influence on word-object associations

Why are 17-month old infants more likely to associate an object with an infrequent than with a frequent word? Below, we will consider four possible reasons and conclude that, despite being at first sight plausible accounts of our data, none is fully satisfactory. We then suggest that, just as children use various biases and heuristics for learning content words (Markman, 1990; Hirsh-Pasek, Golinkoff, & Hollich, 2000), they might also have universal biases for identifying function words, especially a bias to consider highly frequent words as function word candidates.

3.6.2.1 The role of position

In the discussion of Experiments 8a and 8b, it was proposed that the frequency in the familiarization only served to individuate the frequent syllables as segmented words, but was not directly responsible for their non-referential use. Infants may have associated more strongly the object with the second rather than the initial word, because they learned from seventeen months of exposure to Italian that the initial words of utterances

---

frequent words.

are usually not referential (i.e. because those are often function words; Gervain et al., 2007). The absence of preference for the noun in Experiment 8b would then be due to their perceiving the label as one bisyllabic word, instead of two monosyllabic words.

This view should be rejected in the light of Experiments 9-11. Indeed, it would have predicted identical results in Experiments 9 and 10. Specifically, it would have predicted that, in both experiments, the frequent syllables should be perceived as segmented words. Infants' hypothesized knowledge that utterance initial words are not referential should lead them in both Experiments 9 and 10 to rely on the final syllable, that is the infrequent syllable in Experiment 9 and the frequent syllable in Experiment 10. In contrast, infants showed no preference in Experiment 10.

Moreover, the familiarization stream of Experiment 11 was segmented in bisyllabic sequences, so that infants may consider the frequent syllable as constituents of segmented bisyllabic words, rather than as individual monosyllabic words. This would predict that infants should not rely more on the frequent or the infrequent syllable when pairing the label with the referent object. In contrast, we show in Experiment 11 that infants favored the initial infrequent syllable.

Frequency distributions, rather than position, therefore appears to be directly responsible for the preference of infants to pair one or the other syllable of a label to an object in our experiments.

3.6.2.2 Can "mutual exclusivity" account for the preferential association of objects with infrequent words?

The mutual exclusivity assumption (Markman, 1990) might account for our results. When learning words, infants assume that one object has only one label (Halberda, 2003; Markman, 1990). Liittschwager and Markman (1994), for instance, showed that 16-month-olds find it more difficult to learn a new label for an object they can already name, than for an unfamiliar object. Consequently, when infants have to find the referent of a novel word, they tend to choose an object for which they have no previous name, rather than an already labeled object (Markman, Wasow & Hansen, 2003). Potentially, the results of Experiments 8a, 9 and 11 might be due to a very similar principle: in addition to assuming that one type of object has only one name, they might also assume that one label refers only to one type of object. If so, infants might choose a label that is not yet associated with a referent when choosing among possible labels for a novel object.

As frequent words occurred by definition much more often than infrequent words during the familiarization phase, infants had many more opportunities to associate them with some referent. Once they had learned a referent for a frequent word, infants might not be willing to associate this word to a second referent. Therefore, when a novel object was accompanied by a label composed of a frequent and an infrequent words, and if the frequent word was already associated to some other referent, infants should associate the novel object with the infrequent word.

While plausible, the explanation can hold only if infants had already assigned a meaning to the determiners *before* the teaching phase of Experiments 8a, 9 and 11. This, however, is unlikely for several reasons. First, our participants had never heard French or the artificial language before taking part in the experiment, and none of the words we used have a meaning in their native language (i.e. Italian). Therefore, the infants could have learned a meaning for one or both of the frequent words only during the familiarization phase. During the familiarization phase, however, the infants' attention was attracted mainly by a rotating chessboard-like display presented on the screen. This bi-dimensional display, however, was chosen because it is, we believe, unlikely to be perceived as an object (see also Stager & Werker, 1997). Given that the display was unlikely to be perceived as an object, it is also unlikely that the infants associated it with one of the frequent words.

Second, the determiners were never presented in isolation; rather they were embedded in sentences and surrounded by other French words in Experiment 8a and by meaningless syllables in Experiments 9 and 11. Each of the frequent words represented 20% of the words heard during the familiarization in Experiment 8a, and 25% in Experiments 9 and 11. To the best of our knowledge, it has never been tested whether 17-month-old infants can associate a word and a visual referent, when the target word is embedded in meaningless speech.

Finally, the familiarizations in Experiments 9 and 10 differed only for which syllables were more frequent. Thus, if participants in Experiment 9 were able to associate one of the frequent words to a referent, participants in Experiment 10 should have done the same. Thus, an account of Experiment 9 based on the mutual exclusivity principle would predict that, also in Experiment 10, infants should associate the referent object with infrequent rather than frequent words. However, this was not supported by our data, and an account based on the mutual exclusivity if at all should play a minor role.

3.6.2.3 Can novelty account for the preferential association of objects with infrequent words?

A second account of the preferred association of novel object with infrequent elements is based on the relative familiarity or novelty of the different stimuli. Specifically, infants may tend to preferentially associate novel objects to less familiar words than to very familiar words (Golinkoff, Mervis & Hirsh-Pasek, 1994; Mervis & Bertrand, 1994). Frequent words being more familiar than infrequent words, this view predicts that infants would focus on the infrequent syllables when labeling the novel object from the teaching phase. However, this view does not make any prediction about the role of position, and would expect similar results in Experiments 9, 10 and 11. The absence of preference for the infrequent syllable in Experiment 10 therefore argues against a novelty account of our results.

3.6.2.4 Can "cross-situation statistics" account for the preferential association of objects with infrequent words?

A third possible explanation of our results is related to what has been called cross-situational statistical learning. For learning the meaning of words, children might track the frequency with which a word and an object co-occur (Smith & Yu, 2007;

Vouloumanos, 2008; Yu & Smith, 2007). In the most commonly assumed form of cross-situational statistics, learners keep track of the number of times a given word and a given object occur together. According to the axiom that what fires together wires together, the associative link between the word and the object representations is strengthened each time they occur together. The meaning of a word is then determined by selecting the referent with the strongest associative strength in a "winner-takes-all" fashion.

This simple view of cross-situational statistics does not account for our data. Indeed, the teaching phase is the only phase where infants were likely to form associations between objects and their labels. In that phase, the object occurred as often with the frequent as with the infrequent words. According to the version of the cross-situational statistics outlined above, the associations between the frequent word and the object should thus be as strong as between the infrequent word and the object. Therefore this view does not provide an account for our data.

There is an alternative version of cross-situational statistics, however, that may at first glance seem more consistent with our results. In fact, the aforementioned version of cross-situational statistics does not take into account how often the object appears without any word being pronounced, and how often a word is uttered in the absence of any object. Alternatively, the associative strength between an object and its label might be determined by counting how often the object and the label co-occur, and by comparing this number to the number of times the object and the label are encountered in total. (Technically, such a measure of associative strength could be expressed in terms of mutual information or related statistics.) The label of an object might then still be determined by selecting the label with the strongest associative link.

This version of cross-situational statistics seems to be consistent with our data. Indeed, recall that each frequent word occurred nine times as often as each infrequent word during the familiarization phase. In contrast, during the teaching phase, only one frequent and one infrequent words occurred, and they occurred equally often. Moreover, each frequent word appeared many more times in the familiarization than in the teaching phase. In Experiment 8a, the frequent word used in the word teaching phase appeared 81 times in the familiarization, and only 23 times paired with the referent object. The infrequent word, on the other hand appeared only 9 times in the familiarization, and 23 times with the paired object. In Experiments 9-11, the frequent syllable used in the word teaching phase appeared 126 times in the familiarization, and only 24 times paired with the referent object. The infrequent syllable appeared 14 times in the familiarization, and 24 times paired with the referent object. Accordingly, the mutual information between the frequent word and the object is lower than that between the infrequent word and the object.

While this account is in agreement with the results of Experiments 8a, 9 and 11, it cannot account for the whole of our data. In fact, the mutual information between each of the two syllables constituting the label and the object is the same in Experiments 9 and 10; it is higher between the object and the infrequent syllable, than between the object and the frequent syllable. The mutual information explanation would have then predicted reliance on the infrequent syllable in both Experiments 9 and 10, which was not verified in Experiment 10.

Moreover, this account is inconsistent with previous data. Specifically, this account predicts that frequently hearing a word in absence of its referent should *always* be detrimental for acquiring its meaning, a prediction that seems problematic both based on every day experience, and on previous experiments. For example, Swingley (2007) familiarized 18-month-old infants with a novel word form, but, crucially, without presenting its referent. Following this familiarization, he administered a word learning task where infants had the opportunity to associate a word with a visually presented object. Results showed that infants were better at learning a referent for the word they had been familiarized with compared to learning a referent for a novel word. In contrast to the predictions of the cross-situational statistics model, hearing a word in the absence of its referent is thus not necessarily detrimental for associating this word with a meaning, as infants were *better* at learning the meanings for the words that they had heard frequently.

3.6.3 Prior expectations guiding function word acquisition

We therefore favor a fourth explanation, that infants are equipped with a bias that treats frequent words as less referential than infrequent words. Moreover, the results obtained in Experiments 9-11 suggest that infants do not solely extract frequent syllables from the familiarization stream, but assign a position to these frequent items. In Experiments 9 and 10, no segmentation cue was provided in the familiarization stream, so that the way infants perceive and organize the continuous stream is likely to depend on the language of their environment (i.e., Italian). In fact, Gervain et al. (2008) showed that Italian 7-month-old infants tend to perceive similar streams to our own as a series of syllable sequences starting with the frequent syllables. If participants indeed perceived frequent syllable as initial in the familiarization stream, the position of frequent syllables in the object labels was congruent with that of the familiarization for Experiment 9 but not for Experiment 10. In contrast, segmentation cues were provided in Experiment 11, so that frequent syllables occupied final sequential positions both in the familiarization stream and in the object labels. We observe that frequent syllables were relatively dispreferred compared to infrequent syllables in label-object associations only when their position was coherent all along the experiment (Experiment 9 and 12), but not when it varied (Experiment 10).

Interestingly, there is nothing in the familiarization stream that explicitly says that the frequent words should be used in a non-referential way, only if they are found in the same position where they frequently appeared. In fact, our results suggest that infants do not solely extract the frequent words from the familiarization streams, but rather a frequent frame where a slot must be filled on the side of the frequent syllables, either to the right or to the left. This property is characteristic of function words (especially determiners), which tend to occupy a specific position relative to their lexical complement (Chan, 2008). For example, a determiner such as “the” may be represented as “theN”, where the symbol N can be replaced by any noun.

Moreover, infants know that these frequent frames are not referential, and, in contrast, associate a referent object to the lexical complement filling the frame. In fact, when hearing a novel word that is accompanied by a function word, infants are more likely to interpret it as referring to a novel object if the function word marks it as a noun (Brown, 1957) and as referring to a novel action if the function word marks it as a verb (Bernal, Litz, Millote & Christophe, 2007; Brown, 1957). Interestingly, in these experiments, infants show their understanding of the word on new utterances where it is pronounced in isolation or with other function words. This suggests that the frame and the novel word are used differently. While the frame orients the possible meanings of a novel word by assigning it to a syntactic class, only the novel word, not the frame, is associated to the referent.

3.6.4 Conclusion

Natural human languages are characterized by the existence of two broad classes of elements: content elements that are linked to semantic referents, and function elements that mainly serve syntax. Understanding what elements may refer to objects, properties or actions in the world, and what elements may not, may constitute one of the first steps of language acquisition. Here, we presented three experiments suggesting that infants are equipped with a bias that treats very frequent elements as non referential. This may allow infants to solve the second Gavagai problem based on a distributional analysis. Moreover, we propose that the connection infants make between high frequency and non-referential use support the existence of core linguistic knowledge that bridge distributional and linguistic properties. Specifically, we propose that infants are endowed with the correspondence between the distributional class of frequent words, and a linguistic class regrouping function word candidates, defined by their use (e.g. non-referential use). Such prior expectation would allow infants to identify early on function and content words, and use them appropriately, focusing on content words when building the lexicon and on function words when learning syntax.

On pourrait comparer la raison à une matière infiniment combustible, mais qui néanmoins ne s'embrase d'elle-même. Il faut qu'une étincelle soit jetée dans l'âme.

De l'étymologie en général

A. W. Schlegel, 1846

CHAPTER 4

General Discussion

4.1 Perceptual trigger to learning mechanisms

In the first part of this thesis, we showed that two mechanisms participating in language acquisition, the encoding of word-like sequences and the generalization of relational structures are triggered by different speech categories. Human speech differs from other animals' vocalic productions by an alternation between two speech sound categories, consonants and vowels. We showed that these two categories have different roles in language acquisition at least from the age of 12-months. By this time, word memory relies more on consonants, whereas the generalization of structural patterns is best when these are implemented by vowels. Thus we propose that the mechanisms underlying the acquisition of the lexicon and of syntax privilege different computational units: consonants are privileged to carry categorical distinctions for the lexicon, and vowels signal structural relations that may inform syntax.

Our results adhere with the proposition that language acquisition relies at least in part on the identification of specific computational units, such as consonants and vowels, and on constraints imposed on their functions. For instance, computing all kinds of statistical information can help infants to segment words from the speech input and facilitate the word to world mapping (Graf Estes, Evans, Alibali & Saffran, 2007). However, statistics can in principle be computed on anything that infants represent at some level, or are sensitive to (e.g. transition probabilities between syllables, consonants, vowels, articulatory features, liquids, open vowels, etc). Yang (2004) proposed that, for statistical computations to be useful, they must be restricted to pre-defined informative units such as syllables and phonological categories. In fact, the experimental work of Bonatti, Peña, Nespor and Mehler (Bonatti et al., 2005; Mehler et al., 2006) showed that adults compute better statistics such as transition probabilities over consonants than over vowels. Here, we extend this view and show that also the encoding in memory and the extraction of structural relations are constrained by 12-months of age.

Are these constraints to speech processing innate, meaning that they are not acquired? We show that the constraints on the processing of vowels and consonants have not yet emerged by 6-months of age. Moreover, these infants may not discriminate the broad categories of consonants and vowels yet. Indeed, 6-month-olds exhibited difficulty in discriminating between vowel-repetition and consonant-repetition exemplars

(Experiments 4 and 5). At that age, infants may still represent syllables in a holistic way, thus forming representations that are mainly occupied by the acoustic substrate of vowels (see Experiment 2). Brain maturation or further exposure to the speech input may be necessary for infants to form the categories of consonants and vowels. We have proposed that the category of vowels may first be identified as the ensemble of atomic syllables, those syllabic sounds that enter in the constitution of other syllables, and which themselves cannot be divided in a syllable and a residual. The distinction between consonants and vowels may nevertheless be acquired shortly after 6-months, as the production of canonical consonant-vowel syllables around 7-months of age (Oller, 1980) requires motor programs to correctly articulate. Such motor program may rely on some representations of consonants and vowels.

Once the consonant and vowel categories are represented and used, it is not cleat whether their respective lexical and structural functions are automatically set or whether these must be acquired. The specific roles of consonants and vowels may constitute the output of earlier computations that operate on unspecialized speech categories. In particular, we have proposed that the lexical role of consonants may first require the formation of language-specific phonemic categories between the ages of 10- and 12-months.

In sum, speech processing is constrained in a way that may help language acquisition by organizing a division of labor between consonants and vowels. These constraints are in place by the end of the first year of life, when infants just start building the lexicon. The perceptual or distributional properties of consonants and vowels appear indeed to trigger different learning mechanisms. Stimulus-specific constraints suggest the existence of domain-specific computations. Considering that humans mainly use speech to convey linguistic content, speech-specific computations may yield linguistic content such as word representations and information about syntactic organization.

4.2 Prior representations guiding the acquisition of function words

In the second part of this thesis, we showed that infants infer certain functional properties from distributional properties. In fact, our results suggest that frequency determines the use of words, showing that infants act as if they established a category of words defined by their high frequency of occurrence, and avoid using these words referentially, associating an object referent to the infrequent rather than to the frequent syllables of a speech label (Experiments 8-12).

The non-referential use of frequent items is characteristic of the use of function words in language. Indeed, these words are not associated to specific and concrete referents such as objects or actions, but can be used to mark and recognize syntactic categories by adults and children as young as 14-month-olds (Bernal, Lidz, Millote, & Christophe, 2007; Waxman & Booth, 2001, 2003). We considered and excluded several alternative accounts for our results. In particular, we show that our results cannot be explained by cross-situation statistics, and conclude that infants may have assigned functional characteristics typical of function words to the class of very frequent words.

Thus, our results suggest a bootstrapping mechanism, allowing infants to infer linguistic knowledge from distributional information. Precisely, infants' cognitive system can make use of the correspondence between the distributional class of very frequent words and the syntactic class of function words to identify potential function words. The statistician abilities of infants meet with prior expectations incorporated in infants' cognition system to make language acquisition possible. Such expectations must include the definition of an abstract class of function words that is defined both by its linguistic use (i.e., no referent) and its distributional properties (i.e. high frequency of occurrence).

Still, it is not fully demonstrated by our experiments, that the expectations infants bring to the task were not acquired from exposure to linguistic input. A scenario where infants acquire the non-referential property of frequent syllables should resemble the following picture. Infants learn very early on the most frequent words of their language (by 8-months, see Shi & Lepage, 2008). Further discovering that these very frequent words do not refer to specific objects, object properties or actions in the world, they eventually extend this characteristic to the rest of very frequent words. Gervain et al. (2008) showed that 7-month-olds generalize a distributional property (i.e., initial or final position of utterances) learned from the linguistic input to frequent syllables presented in an artificial speech stream, thus suggesting that they represent something like the class of frequent items and can assign properties to that class. However, no experimental evidence shows that they could similarly generalize to all frequent items a functional property such as constraints on the association of words and referents.

Furthermore, it is unclear how infants could learn that frequent words are not referential, if they are not biased to look for non-referential words. The associationist assumption when considering word learning is that infants are forming word-object associations on the basis of repeated co-occurrences of labels and objects, a phenomenon modeled by cross-situational statistics. We have shown that cross-situational statistics cannot account for our data, especially for the position effect observed in Experiments 10-12. However, in more naturalistic conditions, if one supposes those statistics are computed over prosodically defined words rather than syllables, these may lead infants to initially learn labels that incorporate the frequent function words, thus believing that a rabbit is called *therabbit*, but also *arabbit*. From such evidence, they may infer that a rabbit is actually called *rabbit*, thus deleting from the initial redundant labels the frequent particles *the* and *a.* Still, there is no reason why they should generalize that property to all frequent syllables, as they do in our experiments, unless they have a prior expectation that there exists a *class* of non-referential words.

Thus, in the absence of evidence that infants could learn from the input that frequent syllables are non-referential, our results are better explained by the hypothesis that infants are equipped with core knowledge of the existence of a class of non-referential words. Moreover, infants appear to know that members of this class of non-referential words are signaled by their high frequency of occurrence.

This distributional bootstrapping phenomenon may initiate a cascade of bootstrapping mechanisms, allowing infants to bootstrap into the acquisition of syntax and the lexicon. Precisely, the acquisition of a few words that refer to objects and actions could allow infants to first split the function word category into those words that mark nouns, verbs or adjectives (lexical bootstrapping). For example, infants will notice that the function word “the” precedes the word “cookie” that refers to a physical object, whereas “is” precedes “eating” that refers to an action. Assuming, like Pinker (1984), that infants are equipped with knowledge about the syntactic classes of nouns and verbs, they may infer from this that “the” marks nouns and “is” marks verbs. Consequently, novel words can be assigned to their corresponding syntactic category according to the function words they are associated with, and syntactic bootstrapping can constrain the interpretation of novel words, nouns being initially more likely to refer to objects, adjectives to object properties and verbs to actions (Bernal et al., 2009; Brown, 1957; Waxman & Booth, 2003).

4.3 Early language acquisition

Both series of experiments that were presented and discussed above point at mechanisms allowing infants to identify and distinguish the sources of information relevant for the lexicon and for syntax, respectively. By the end of the first year of life, infants know that consonants carry more lexical information, whereas vowels are informative for structural relations. Before eighteen months, they further know that relatively infrequent words are good candidates for new lexical entries, whereas most frequent words may carry a different type of information (e.g. indicating syntactic classes or syntactic roles).

These mechanisms appear crucial to language acquisition for two reasons. They constraint the learning mechanisms infants employ to acquire language, and they allow learners to develop the lexicon and syntax in parallel.

4.3.1 Language-specific constraints to learning mechanisms

4.3.1.1 Constraints on memory

Potentially, infants may encode all dimensions of a stimulus in memory. When they hear their mother saying "What a beautiful baby!", they may encode the voice of their mother, the different syllables, their order, the different segments, their order, the absolute variations of pitch, intensity and duration, the relative variations of pitch, intensity and duration, etc.. Are all these dimensions of the stimulus equally relevant to the task of learning language? Encoding all dimensions may in fact be counter-productive, impeding infants to recognize the same word pronounced by different speakers, or by the same speaker with different intonations, e.g. *baby!*, *baby?*. To develop their lexicon, infants must thus learn or know what are the relevant dimensions to encode.

The relevance of certain dimensions varies cross-linguistically. Infants must thus learn some of the relevant dimensions in their language. For example, French-learners loose the ability to encode lexical stress, because that information is redundant in French (Dupoux, Pallier, Sebastian & Mehler, 1997; Dupoux, Peperkamp & Sebastian-Galles, 2001, 2010; Dupoux, Sebastian-Galles, Navarete & Peperkamp, 2008; Peperkamp & Dupoux, 2002). Infants also learn language-dependent phonemic categories, thus ignoring the within-category variations that are irrelevant to their language (Mehler & Dupoux, 1994; Werker & Tees, 1984).

There may be also universal constraints on word encoding, that dimensions of the input, which are universally more informative for the lexicon. Cross-linguistically, consonants are more informative than vowels for word recognition, because they are more numerous and because they are perceived categorically (see Chapter 2). Therefore, infants may be endowed with constraints on memory, focusing on consonants when learning words. Similar memory constraints have been observed in other domains of cognition. For example, object memory appears to rely on shape, rather than color (see Appendix A; Tremoullet et al., 2002).

4.3.1.2 Constraints on generalizations

As we discussed in Chapter 1, the induction problem (Goodman, 1955) makes constraints on generalization a necessity. We discussed constraints on the shape that generalizations can take. However, constraints on the input of these mechanisms can also be beneficial. Indeed, when aiming at drawing generalizations over limited input, learners may potentially generalize the patterns observed in any dimension of the input. This may quickly overload infants' capacity.

In particular, different dimensions of speech stimuli are often correlated. For example, the syllabic structure and stress position in a word are related, heavy syllables (CVC or CVV) being much more likely to carry the stress than light syllables. Therefore, generalizing the patterns of both syllable structure distribution, and stress distribution would be redundant.

Concerning consonants and vowels, some authors have observed that consonants and vowels are not freely associated, but front vowels tend to come with labial consonants (Stoel-Gammon, 1983), whereas velar consonants tend to come with back vowels (Tyler & Langsdale, 1996), because those combinations are easier to produce. Consequently, regularities implemented over vowels may also be reflected over consonants. Generalizing the two implementations of a same regularity would represent a useless load to the learner's cognitive capacities.

Constraints on the input that is considered for generalizing structural patterns and relations should thus facilitate learning. They allow infants to quickly identify relevant rather than casual and uninformative structures, and avoid generalizing redundant structures on different sources of information.

4.3.2 Learning in parallel

Separating the sources of information for syntax and the lexicon allows infants to work on both tasks in parallel. Generalizing an abstract pattern and encoding specific tokens in memory are sometimes conflicting tasks. For example, Gomez (2002) showed that 18-month-olds generalize a non adjacent-dependency between the first and last words of three-word sequences, if there were 24 possible words occurring in the middle position, but not if there were 3 or 12. Gomez (2002) interprets these results, proposing that infants notice non-adjacent dependencies, only when adjacent dependencies become unpredictable. Another interpretation, however, is that when the set of exemplar sequences is limited (3 and 12), infants try to remember each of them. Only when the set of exemplars largely exceeds the limits of learners' memory capacity, do they generalize the pattern. Thus, accurately memorizing exemplars may delay generalizations.

In return, generalization may damage memories. Generalization indeed usually implies a transition from one level of representation where exemplary tokens are represented, to a more abstract level of representation. This transition may cause a lost of sensitivity to some dimensions that are represented in the first, but not in the second level of representation. For example, if one generalized the repetition pattern after hearing exemplars such as *folo* and *dulu*, he now represents something like $c_1V_1c_2V_1$. In that type of representations, the actual vowels he heard are not represented but replaced by a more abstract and general representation: vowel. The learner may thus not be able to say whether he heard the word *dulu* or not. Another example of such a phenomenon is categorical perception, when a stimulus is represented as the class to which it belongs. If lower levels of representation, where the precise features of the stimulus are encoded, are not conserved, the system will loose the ability to distinguish different tokens of the same class. This happens for example with phoneme perception (Werker & Tees, 1984) and with face perception of unfamiliar phenotypes (i.e. the "other race effect"; Kelly, Quinn, Slater, Lee, Ge & Pascalis, 2007; Meissner & Brigham, 2001).

Thus, separating the sources of information relevant for memorizing tokens and building the lexicon on one hand, and structural generalizations and categorization on the other hand, may avoid conflicts between these processes and thus facilitate language acquisition.

4.4 Core linguistic representations

Language acquisition is always described as a hard task, even though infants all around the globe achieve it continuously. Thus, language acquisition is *not* a hard task for who has the adequate learning mechanisms and prior knowledge. None of the learning mechanisms that have been considered in this thesis or elsewhere appears to be language-specific. However, our experimental work suggests that young infants are equipped with language specific knowledge. In particular, we showed that they know what sources of information are relevant for different general learning mechanisms such as encoding in memory, generalizations and associations.

Such core linguistic knowledge may guide the first steps of language acquisition, allowing infants to bootstrap into language. It fills the gap between a distributional and/or perceptual description of the input, which should be available to young infants, and more abstract core linguistic representations. In Chapter 3, we have already discussed at length one example of a core linguistic representation. We showed that infants form a category of words based on a distributional property, high frequency. Moreover, infants have expectations about the use of these words: they do not use them referentially. Thus, we propose that this category of frequent words may constitute a core linguistic representation. Of course, further learning will be necessary to attribute each of these words a more refined classification and a syntactic function, yielding proper function words. Core linguistic representations thus constitute the foundations of linguistic knowledge, and should be enriched by further learning and experience.

We now consider two linguistic representations, which are observed cross-linguistically, triggered by perceptual features, and play a role in the linguistic system. These representations, we propose may thus constitute core linguistic representations.

4.4.1 The syllable

Many authors have argued for the syllable as a natural unit in speech perception and representation. A classical argument in favor of this thesis states that young children and illiterates handle better syllables than segments in meta-linguistic tasks (Bertelson, de Gelder, Tfouni & Morais, 1989; Liberman, Shankenweiler, Fisher & Carter, 1974; Morais, Bertelson, Cary & Alegria, 1986; Morais, Cary, Alegria & Bertelson, 1979; Treiman & Breaux, 1982). For instance, Liberman, Shankenweiler, Fisher and Carter (1974) showed that preschool, kindergarten and first-grade children are better at counting syllables than phonemes in speech utterances.

Mehler and colleagues (1981) showed that French adults are faster to detect speech sequences in words where they correspond to a syllable (e.g. *pal* in *pal-mier* or *pa* in *pa-lace*) than in words where they do not correspond to a syllable (e.g. *pa* in *pal-mier* or *pal* in *pa-lace*). These results, however, were invalid for English speakers (Cutler, Mehler, Norris & Segui, 1986; Cutler, Butterfield & Williams, 1987). This variety of results across languages suggests that speakers of different languages may process differently different speech units. The syllable may not be the favored unit of speech perception for adult English speakers, but this does not mean that they do not represent syllables.

Other experiments, in fact suggest that English speakers do represent syllables. Wood and Day (1975), for instance, found that participants were unable to ignore the vowel following a consonant in a classification task. Participants were asked to classify syllables according to their first consonants. Participants were slower when the vowel of the syllables varies (e.g. *ba, bi, bu* vs. *da, di, du*) than when it was constant (i.e. *ba* vs. *da*). These results remained true, even if participants were explicitly instructed to focus on the consonants and ignore the vowels (see also Day & Wood, 1972; Eimas, Tartter & Miller, 1981; Eimas, Tartter, Miller & Keuthen, 1978; Tomiak, Mullenix & Sawusch, 1987). These results thus suggest that participants automatically integrated the consonant and vowel into one representation, i.e. the syllable.

As discussed in section 2.8, the syllable appears to be a unit that young infants and neonates use when processing speech sequences (Bertoncini, Floccia, Nazzi & Mehler, 1995; Bijeljac-Babic, Bertoncini & Mehler, 1993; Bertoncini & Mehler, 1981; van Ooyen, Bertoncini, Sansavini, & Mehler, 1997). Syllable representations are likely to be triggered by the wave envelope, i.e. the spectral power in low frequencies (Bertoncini & Mehler, 1981; Stevens, 2002). As is shown on Figure 4.1, the sequence *tsp*, which does not correspond to a well-formed syllable, and the sequences *tap*, *tlp* and *tlup*, which are possible well-formed syllables in more than one language (e.g., Slavic languages), have very different spectrograms. Infants may thus detect the maxima of spectral power in low frequencies, and build syllables around these maxima.

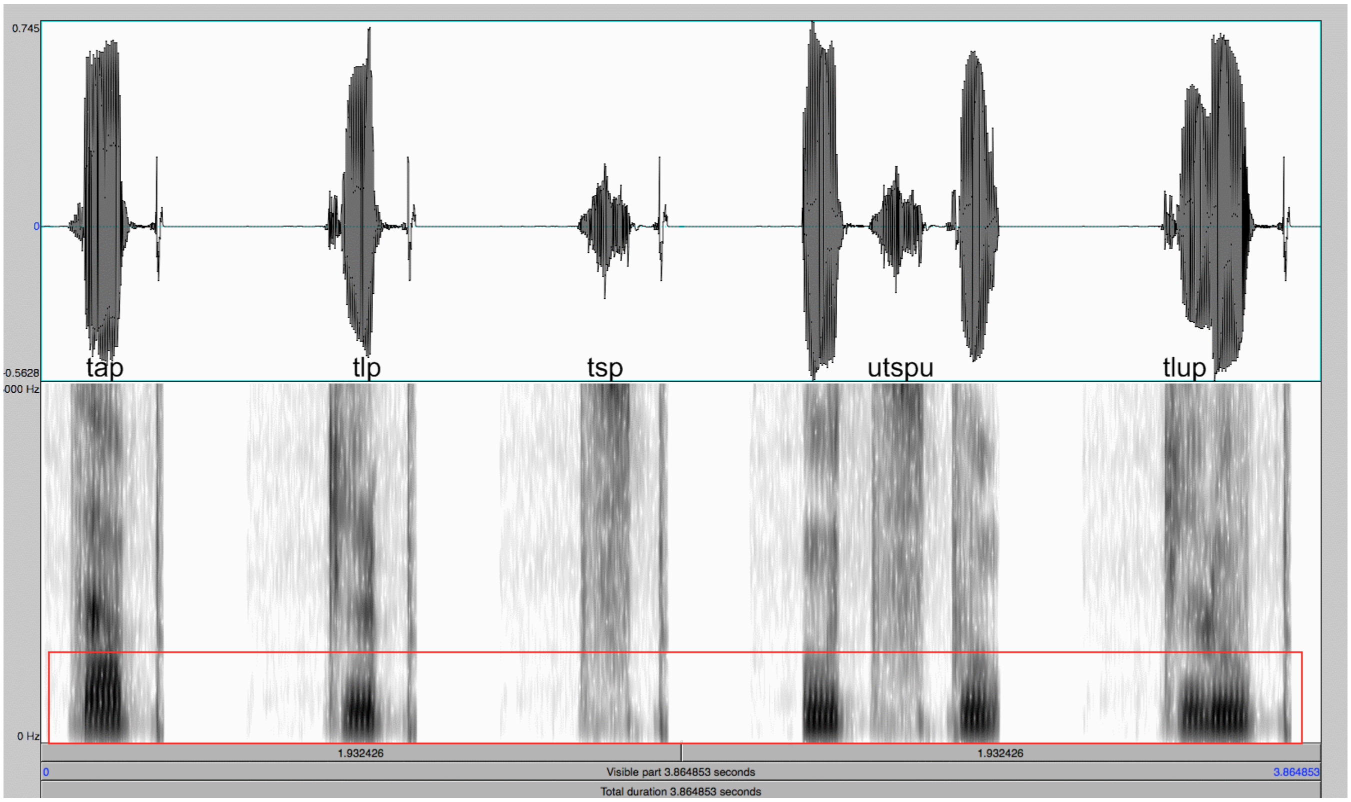

Figure 4.1 – Sound wave and spectrogram of *tap*, a well-formed syllable, *tlp*, a possible syllable in some languages, *tsp*, an ill-formed syllable, *utspu*, a bisyllabic speech sequence and *tlup*, a well-formed syllable. Well-formed syllables can easily be identified by a burst of energy in the low frequencies (red box), which is not observed for the sequence *tsp*.

What evidence do we have that syllables are relevant to the linguistic system? The syllable represents a fundamental unit to explain many linguistic regularities such as phonotactic constraints (Goldsmith, 1990) and stress assignment (Halle & Vergnaud, 1988). Nespor and Vogel (1986/2008) defined the syllable as a phonological domain, the smallest constituent of the prosodic hierarchy. They motivate their claim by a number of phonological rules that apply within the domain of syllable. One example consists in the Schwa Insertion in certain varieties of Dutch (Booij, 1981; Trommelen, 1983; van der Hulst, 1984). According to this rule, a schwa can be inserted between a liquid and a noncoronal obstruent, only if these belong to the same syllable. Thus, Schwa Insertion applies to the examples in (2) but not to those in (3):

(2)

a. park:	$[\text{park}]_\sigma$	→	par[ə]k	'park'
b. helpster:	$[\text{help}]_\sigma$ $[\text{ster}]_\sigma$	→	hel[ə]pster	'helper'
c. melkauto:	$[\text{melk}]_\sigma$ $[\text{au}]_\sigma$ $[\text{to}]_\sigma$	→	mel[ə]kauto	'milk van'

(3)

a. parkiet:	$[\text{par}]_\sigma$ $[\text{kiet}]_\sigma$	→	*par[ə]kiet	'parakeet'
b. pulpig	$[\text{pul}]_\sigma$ $[\text{pig}]_\sigma$	→	*pul[ə]pig	'pulpy'
c. melkerij:	$[\text{mel}]_\sigma$ $[\text{ke}]_\sigma$ $[\text{rij}]_\sigma$	→	*mel[ə]kerij	'milk farm'

(from Nespor & Vogel, 1986/2008)

Syllable also represents the unit to form the higher constituents in the prosodic hierarchy, i.e. foot (Nespor & Vogel, 1986/2008) and possibly of word forms. Syllables are also proposed to play a role in speech production (Levelt & Wheeldon, 1994). Thus the universality, early use, and linguistic relevance of the syllable suggest it may constitute a core linguistic representation. As we discussed in section 2.8, this representation may be enriched when infants discover the categories of consonants and vowels. However, it keeps playing a role in adult speech perception, production and representation.

4.4.2 The word form

Word form is another linguistic representation that is found universally across languages, and thus a promising candidate for constituting a core linguistic representation. Here, we talk about word forms, in order to distinguish that type of representations from words, which include lexical and semantic information. Word forms are defined as sequence of speech that will be associated to semantic referents and receive syntactic specification, in order to form proper words. In this sense, word forms are thus clearly relevant to the linguistic system, as a fundamental constituent of words. Below, we discuss two mechanisms that have been proposed to trigger word form representations: TP computations and edge detection.

Two series of experiments aimed at testing Saffran et al.'s (1996) original claim that TPs were used to extract words in a continuous speech stream, thus yielding word representations. Graf Estes, Evans, Alibali and Saffran (2007) tested this issue with 17-month-old infants. Their participants first undertake a classical segmentation task, where they listened to a continuous speech stream consisting in a series of "words" defined by high TPs within words and low TPs between words. Following this, they were presented with a novel object that was associated to a label consisting in one of the statistically defined words, a non-word that had not appeared in the previous stream, or a part-word that occurred as often as the words in the stream but had low internal TPs. Only those infants for whom the label was statistically defined in the familiarization stream learned to pair the label and the object. This experiment suggests that statistical computations feed word representations.

However, Endress and Mehler (2009) argued against this interpretation. In a series of experiments, they exposed adult participants to continuous speech streams, containing statistically defined *words*, and subsequently asked them to recognize sequences of syllables that constituted the stream. They showed that a sequence such as *tazeRu*, which never occurred in the stream but had high internal adjacent TPs between *ta* and *ze*, and between *ze* and *Ru*, and high non-adjacent TP between *ta* and *Ru*, was accepted in the same manner as sequences that actually occurred (i.e., *tazepi*, *mizeRu*, *tanoRu*) 600 times in the stream. Observing that a correlation between frequency of occurrence and familiarity is a classical footprint of word processing, Endress and Mehler (2009) thus argue that high internal TPs alone do not trigger the creation of word representations. Rather, they suggest that the presence of prosodically defined edges, due for example to silent pauses or final syllable lengthening, is the most important cue signaling word boundaries, and triggering the formation of a word representation.

How can we combine the two results? On one hand, Graf Estes and colleagues (2007) show that high internal statistical dependencies facilitate the association of a label with an object. On the other hand, Endress and Mehler bring convincing arguments that such statistical structure is insufficient to trigger the formation of a word representation, which necessitates the presence of edges. Endress and Mehler (2009) propose that participants in Graf Estes et al.'s experiments do in fact form word representations, but only when the labels are presented in isolation in the pairing task. High internal TPs thus facilitate the pairing of a word with an object, but are not directly responsible for the formation of word representations.

Mechanisms triggering word representations may therefore be found in the study of prosodically marked boundaries. Neonates are already sensible to certain prosodic boundaries, such as those separating two phonological phrases (Christophe, Dupoux, Bertoncini, & Mehler, 1994; Christophe, Mehler, & Sebastian-Galles, 2001). They can indeed discriminate between two bisyllabic items containing the same phonemic information, but one corresponds to the internal part of a word, i.e. *mati* in *mathé**mati**cien*, and the other item straddles a prosodic boundary, i.e. *mati* in *pyja**ma ti**gré*. Interestingly, neonates appear as sensitive to prosodic boundaries of the language of their environment, as to those of a foreign language (Christophe, Mehler, & Sebastian-Galles, 2001). Gout, Christophe and Morgan (2004) showed that infants can use these cues by the age of 10 months to recognize familiar words in fluent speech. Moreover, Bion, Benavides-Varela and Nespor (in press) showed that 7-month-olds can at least use pitch variations to extract and remember a word. In their experiment, infants first listened to a continuous speech stream alternating between two syllables, a high pitch syllable, e.g. *mi*, and a low pitch syllable, e.g. *ta*. In a test phase, infants show different interest for the words that corresponded to the high-low pattern, i.e. *mita*, and that corresponding to the low-high pattern, i.e. *tami*, suggesting that infants segmented the stream according to pitch variations. Still, more experimental work is necessary to understand whether prosodic boundaries trigger word representations in young infants.

In an Event-Related Potential experiment presented in Annex B, we showed that the memory of 3-month-old infants exhibits and edge effect similar to the traditional U-curve observed in adults' memory experiments (Henson, 2001). In a change detection paradigm adapted from Dehaene-Lambertz and Dehaene (1994), each trial corresponded to four 5-syllabic continuous speech sequences. The first three sequences were identical and correspond to the habituation sequence. The fourth sequence (e.g. *fominegadu*), the target, may differ from the habituation sequence in the first (habituation: *Shuminegadu*), third (habituation: *fomisogadu*) or final syllable (habituation: *fominegali*). An ERP analysis revealed that infants detected syllable changes in the first and final syllables, but not in the internal syllable. These results suggest that, just like for adults, memory is triggered by perceptual edges.

Word form representations are thus triggered by perceptual cues. As we stated above, the linguistic role of word forms is to form associations with referents. Interestingly, even before infants routinely and efficiently associate word forms with referents (around 14- or 17-months of age, Stager & Werker, 1997), they already *expect* word forms to have referents. There is ample evidence that common labels help 6-month-old and older infants to form object categories (Balaban & Waxman, 1997; Fulkerson & Waxman, 2007; Waxman & Braun, 2005). Apparently, hearing a common label, but not a common tone sequence is an invitation for infants to discover what those entities have in common. Xu (2002) presented 9-month-old infants with a scene where two objects successively came out from behind an occluder. One different word was pronounced when each object came out. When the occluder was removed, and one or two objects were revealed, infants' looking times suggest that they expected two objects behind the occluder. Interestingly, they still expected two objects even if these were physically identical. These results were not obtained, however, with other acoustic stimuli such as two different tones, environmental noises or emotional expressions. In another study, an experimenter looked into a box and pronounced two words. The content of the box was then revealed to 12-month-old infants. Their looking behavior suggests that they expected two identical objects if they had heard twice the same label, e.g., *Look! A blicket! Look! A blicket!*, but two different objects if they had heard two different labels, e.g. *Look! A blicket! Look! A stad!* (Dewar and Xu, 2007). Altogether, these results strongly suggest that infants expect word forms to refer to object kinds.

Perceptual triggers such as the prosodic edges thus yield representations for which infants have strong unlearned expectations. These two features exemplify our conception of core linguistic representations.

4.4.3 What is new in the 'core linguistic representations' approach?

Discussing rationalist approaches of the mind, Chomsky (1966) noticed: "*The strong assumptions about innate mental structures made by the rationalistic psychology and philosophy of mind eliminated the necessity for any sharp distinction between a theory of perception and a theory of learning. In both cases, essentially the same processes are at work; a store of latent principles is brought to the interpretation of the data of sense.*" (Chomsky, 1966, *Cartesian Linguistics*)

By proposing the notion of core linguistic representations, we change the focus from the mechanisms to the representations involved in language acquisition. In recent years, two types of mechanisms were proposed to account for language acquisition: statistic computations, and symbolic computations. Statistic computations, however, are insufficient because they are inherently unable to yield the hierarchical phrase structures required by natural languages (Chomsky, 1957, Marcus, 1998). On the other hand, the use of symbolic rules by pre-lexical infants has not been fully demonstrated. Evidence in favor of such computations is limited to the generalization of repetition-based structures (Marcus et al., 1999; Kovacs & Mehler, 2009b), which may in fact rely on a perceptual primitive sensitive to identity relations (Endress, Nespor & Mehler, 2009).

Thus, to leave the impasse where we find the search for dedicated language acquisition mechanisms, we propose to focus instead on the representations involved. Studies in the framework of the core cognition hypothesis have shown that infants are born with a series of conceptual representations, which are triggered by definite perceptual properties. In Chapter 1, we briefly discussed the notion of core knowledge of objects (Carey, 2009; Spelke, 1990; Spelke et al., 1994). Spelke (1990) showed that infants divide the world into object entities according to the principles of cohesion, boundedness, rigididty, and no action at distance. When a given entity in the world obeys each of these principles, it automatically triggers an object representation in infants' mind. No learning mechanism is required here. Learning mechanisms only play a role to enrich initial representations.

Similarly, the notion of core linguistic representations allows us to bypass the search for mechanisms that initially bridge distributional and/or perceptual features and linguistic knowledge, stating that infants are equipped with input analyzers that *recognize* the corresponding entities in the perceptual input. The burden consequently falls on identifying the perceptual and distributional properties that trigger core linguistic representations. High energy in low frequencies of the spectrum may trigger syllabic representations, and prosodic edges may trigger word representations. In this thesis, we have proposed that high frequency of occurrence triggers function word representations.

An account of language acquisition should thus state what core representations are involved, how these are triggered by the perceptual input, and how they may be enriched in the course of life.

4.5 Conclusion

Asking whether language acquisition relies in part on dedicated learning mechanisms and/or language-specific representations, we conducted two series of experiments, showing that learning mechanisms such as encoding in memory, associations and structural generalization do not apply blindly to all stimuli, but are triggered or impeded by specific perceptual or distributional properties. Specifically, we showed, that 12-month-old infants rely on consonants when encoding words in memory, and on vowels when generalizing structural relations. Such specialization, however, emerges only after 6-months of age. We further proposed that younger infants might initially represent speech as a sequence of syllables, and only later make use the consonant-vowel distinction.

Furthermore, we showed that seventeen-month-old infants form a category of very frequent words, and are reluctant to associate referents to these words. We propose that this phenomenon may constitute the first step in the acquisition of function words, thus arguing that infants are endowed with the knowledge of a class of words, defined by distributional (i.e., frequency and position) and functional properties. Our experimental work suggests that infants are endowed with core linguistic knowledge, allowing them to focus on appropriate sources of information to develop syntax and the lexicon in parallel.

These observations lead us to formulate the hypothesis that language acquisition relies on core linguistic representations, which are triggered by perceptual or distributional properties of the input. Such representations can then enter several learning mechanisms, allowing infants to bootstrap into language acquisition. We proposed that the syllable and the word form are prototypical examples of core linguistic representations. Furthermore, our view incorporates the prosodic and distributional bootstrapping approaches of syntax acquisition, which already hypothesized that certain syntactic knowledge (e.g. linguistic parameters) was triggered by perceptual or distributional properties.

The notion of core linguistic representation moves the focus of our investigations from the mechanisms to the representations involved in language acquisition. We thus propose that the specificity of language acquisition may not reside in dedicated mechanisms, but in specific representations on which general or dedicated mechanisms operate. Our research program should thus focus on identifying the core linguistic representations, how these are triggered by the input, and how they can be enriched in the course of life. Beyond our results, we thus hope to have provided a novel framework for further research on the acquisition of a human-unique ability, the language faculty.

CHAPTER 5

Appendix A - Shape and Color

In Chapter 2, we have shown that 12-month-old infants privilege consonants when encoding words in memory, or when accessing to these words' representations, and extract with more ease structural regularities over the vocalic tier than over the consonantal tier. This suggests that both memory and the mechanisms sustaining structural generalizations are constrained in the speech domain. In this series of experiments, we ask whether homolog constraints can be observed in the domain of object cognition.

5.1 The shape bias in object representations

Several experiments reported that 12-month-old infants privilege shape when creating object representations. Tremoulet, Leslie and Hall (2000) reported that infants could identify hidden objects according to their shape, but not with respect to their color. Participants in their experiments saw two objects hidden behind an occluder. When the occluder was removed, they were surprised if one of the object had changed in shape, but not if one had changed in color. Xu, Carey and Quint (2004) confirmed the superiority of shape to trigger the formation of an object representation, showing that 12-month-olds failed to use color, size, pattern and within-basic-level-kind shape differences to individuate objects, but could use cross-basic-level-kind shape differences. Xu et al., (2004) considered two possible accounts for the above data. Their favored account is that shape is more informative for kind differences (Booth & Waxman, 2002), which support object individuation. In agreement with this hypothesis, Bonatti, Frot, Zangl and Mehler (2002) found that even 10-month-olds could use conceptual knowledge available to them to individuate objects. Precisely, they could individuate a car and a face, profiting from their knowledge of the human-object conceptual difference, but not a car and a cup. This suggests that what underlies the formation of separate object representations is primarily the conceptual recognition of objects.

Moreover, Xu and colleagues showed how linguistic labels, which usually refer to object kinds rather than object individuals (Booth & Waxman, 2002; Markman, 1994; Waxman, 1999; Waxman, Senghas & Benveniste, 1997), guide young infants' individuation of objects. Xu (2002) showed that 12-month-olds would search for two objects hidden in a box if they heard two different labels, but would be satisfied with retrieving only one object when they heard the same label repeated twice. Dewar and Xu (2007) repeatedly opened a box to show 9-month-old infants two different objects, or two identical objects. In a test phase, before the box was opened, the two objects were described with two distinct labels ("I seea wug!", "I see a dak!"), or with one repeated label ("I see a zav!", "I see a zav!"). Infants would expect the two different objects if these differed in shape but not if they only differed in color.

Altogether these experiments strongly support the bond between shape and kind membership, and the role of conceptual knowledge in guiding infants' individuation and recognition of objects. However, an alternative account is that certain shape differences may be more salient and consequently better encoded in object representations. This attentional and/or perceptual account could explain both why shape is better associated with linguistic labels, and why it better serves object individuation. In fact, the attention hypothesis predicts a shape bias in any task involving objects processing. Shape would therefore be privileged both when encoding objects in memory and when generalizing structural relations.

5.2 Our study

In Experiment A1, we first seek new evidence for a shape bias in object representations, asking whether infants would base their prediction of the location of a puppet's appearance on the shape or color of objects. As in Experiments 1-7, infants sat in front of a black screen showing two windows, where a toy could appear. A toy appeared in one window after the presentation of a green disk in the center, and another toy appeared on the other window after the presentation of an orange triangle. We asked where infants would expect a toy to appear when now presented with an orange disk or a green triangle, thus basing their expectations onto shape or color.

Furthermore, Experiments A2 asked whether infants would find it easier to generalize an identity relation based on shape or color. If the shape bias observed in tasks involving object representations is purely due to shape being perceptually more salient or attracting more attention, an advantage for shape should be observed in Experiment A2 as in Experiment A1.

5.3 Experimental evidence for the shape bias – Experiment A1

In Experiment A1, we seek new evidence that 12-month-old infants rely more on shape than color when encoding and/or accessing to object representations in memory.

The paradigm of Experiment A1 is presented in Figure 5.1 and was adapted from Experiment 1. The experimental design was identical to that of Experiment 1, except that colorful shapes instead of words were used as cues to the location of the toys' appearances. Data was analyzed like that of Experiment 1.

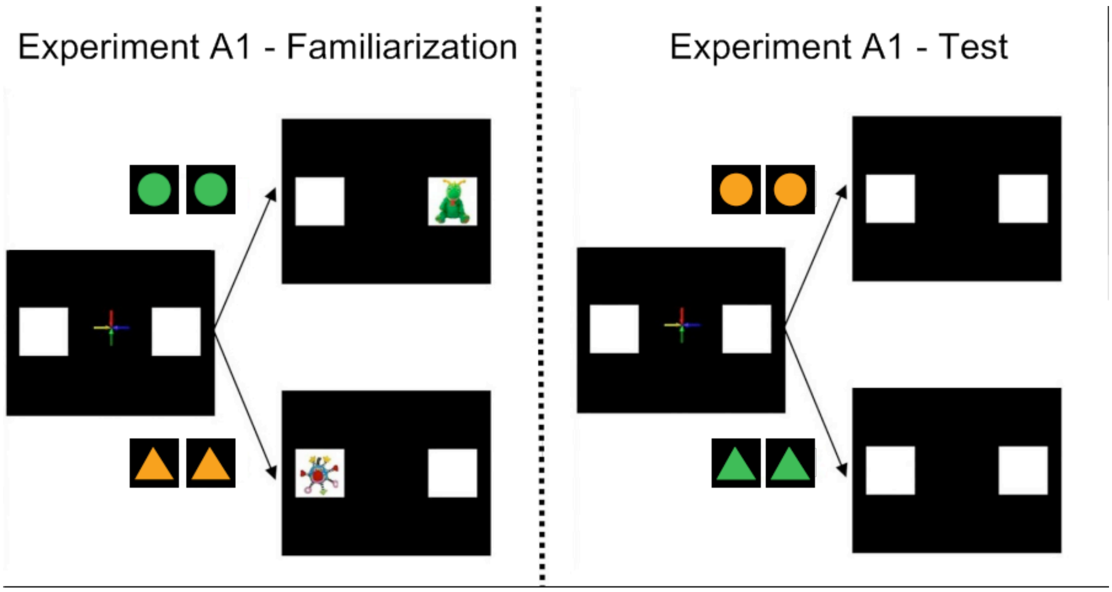

Figure 5.1 Paradigm of Experiments A1. Participants took 32 familiarization trials and 8 test trials.

5.3.1 Materials and methods

5.3.1.1 Participants

Twenty-four infants were included in the analysis; age range 11 month 15 days to 12 month 15 days. Eight other infants participated in the study but were excluded due to fussiness (5), equipment failure (2) or the mother not following instructions (1).

5.3.1.2 Stimuli

Visual cues to the location of toys' appearances consisted in the repeated presentation of colorful shapes. In Experiment A1, two shapes (disk and triangle) and two colors (green and orange) were used. Thus four visual stimuli were created, a green disk, a green triangle, an orange disk and an orange triangle. Each occupied a square with a side length of 4 cm. Two objects sharing neither shape nor color were used in the Familiarization (i.e., the green disk and the orange triangle). The two remaining visual cues were used in the Test. Colored geometrical shapes were synthesized with Adobe Photoshop CS version 8.0 on a computer running Mac OS X, version 10.5.7.

5.3.1.3 Procedure

The Familiarization phase consisted of 32 Familiarization trials. Familiarization trials started with a display of two white squares on the sides and a central attention-grabber. When the infant looked at it, the attention-grabber disappeared, and one colorful shape was presented twice for 800 ms, with separation of 600 ms between the two presentations. Half the infants saw either a green disk or an orange triangle, and half saw either an orange disk or a green triangle. A meaningless attractive sound lasting 500 ms was played in synchrony with the appearance of each colorful shape. One second after the visual cue disappeared, a toy appeared in one of the squares, contingent on the visual cue: one colorful shape predicted the toy's appearance in one of the squares, while the other colorful shape predicted the toy's appearance in the other square. The pairing of the visual cues with toy-locations was counterbalanced across participants.

During test, infants were exposed to 8 trials in a pseudo-random order. Test trials were similar to the familiarization trials, except that infants saw visual cues consisting in the shape of one of the familiarization visual cues, and the color of the other. For instance, if the Familiarization visual cues were a green disk and an orange triangle, the Test visual cues were a green triangle and an orange disk. No toy ever appeared in the test trials. Two seconds after the visual cue disappeared, the next trial started.

5.3.2 Results

The first fixation data for familiarization trials is presented in Figure 5.2. Infants anticipated to one or the other side in 42% of the trials. We computed the proportion of infants showing a correct anticipatory look for each trial. A linear regression analysis showed that the rate of correct anticipations increased during the familiarization ß=.007, R^2=.13, t(30)=2.07, P = .047.

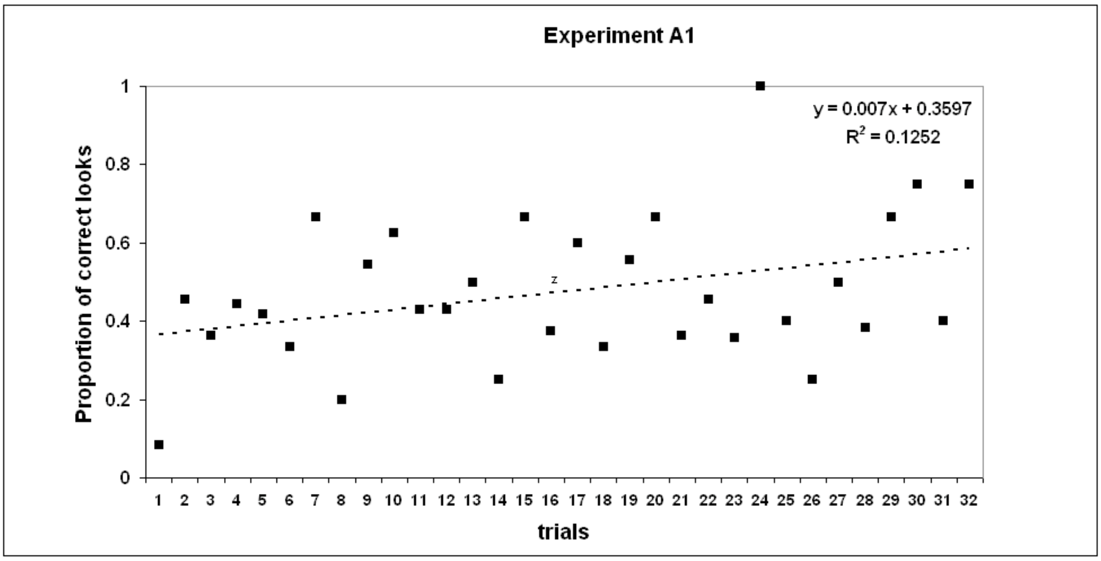

Figure 5.2 - Proportion of correct anticipatory looks for each familiarization trial in Experiment A1. The dotted line depicts the corresponding linear regression.

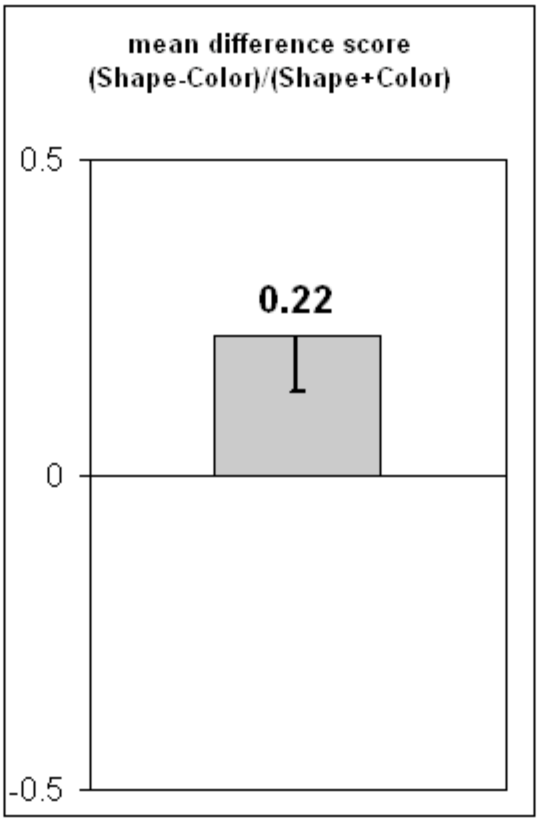

Figure 5.3 - Mean difference score for Experiment A1 considering the first fixations. Infants looked more at the side predicted by shape. Error bars represent standard error.

In the test phase, infants looked to the left or the right in 63% of the trials. Considering first fixations (see Figure 5.3), infants' mean difference score was .22, which was significantly greater than 0, $t(23) = 2.64$; $P = .015$; d’ = .54. Fifteen infants obtained a positive difference score, six infants a negative difference score, and three infants a null difference score. The binomial test assessing whether statistically more infants obtained a positive difference score than a negative difference score was not significant, $P = .078$.

These results were confirmed by the analysis of the overall accuracy, infants' mean difference score was .22, which was significantly greater than 0, $t(23) = 2.47$; $P = .022$; d’ = .50. Fifteen infants obtained a positive difference score, seven infants a negative difference score, and two infants a null difference score. The binomial test assessing whether statistically more infants obtained a positive difference score than a negative difference score was not significant, $P = .13$.

Thus, altogether, infants privileged the prediction made by shape rather than that made by color.

5.3.3 Discussion

In this experiment, we first taught infants that one colored geometrical shape, e.g. a green disk, predicted a toy’s appearance in one location of the screen, and that another colorful geometrical shape, e.g. an orange triangle, predicted a toy’s appearance in another location of the screen. We then asked whether infants would base their expectation on the location of a toy’s appearance on the shape or the color of an ambiguous stimulus, i.e. an orange disk or a green triangle. Our results show that 12-month-old infants based their expectation onto shape, rather than color. At least three accounts of this phenomenon can be proposed.

First, when learning the association between visual cues and locations in the Familiarization phase, infants may encode better shape than color. That is, the object representation they need to form to learn to predict the toys' appearances may consist mainly in the shape of objects. They would learn that a disk, whatever its color, predicts the right side of the screen. In this view, shape is privileged in the formation of object representations.

Second, infants may well encode shape and color, as well as other perceptual dimensions, during Familiarization, thus learning that a *green* disk predicts the right side of the screen. But they would rely more on shape than color when comparing the object representation they have formed to the novel test stimulus. That is, they will find that an orange disk is more similar to a green disk than to an orange triangle. In this view, shape is privileged in retrieval object representation.

Both the previous views state that the shape bias is a structural constraint of the mechanisms underlying the formation or access to object representations in memory. A third account, however, states that shape is perceptually more salient than color. In consequence, shape should be privileged both in the formation of and access to object representations. In fact, the perceptual account of the shape bias predicts that shape should be privileged over color in any task involving objects. This hypothesis will be tested in Experiments A2.

5.4 Generalizing identity relations over shape and color – Experiments A2-A3-A4

5.4.1 Same-shape vs. same-color – Experiment A2

The paradigm of Experiment A2 is presented in Figure 5.4. Experiments A2 requires infants to process colored geometrical shapes as those of Experiment A1. However, in that task, infants need not form object representation in long-term memory, but solely learn and generalize the same-different relationship between two successively presented stimuli. The perceptual account predicts that the shape bias should be verified in this task as in Experiment A1. If, however, the shape bias is specific to the processes of formation or retrieval of object representations, no shape bias should be observed.

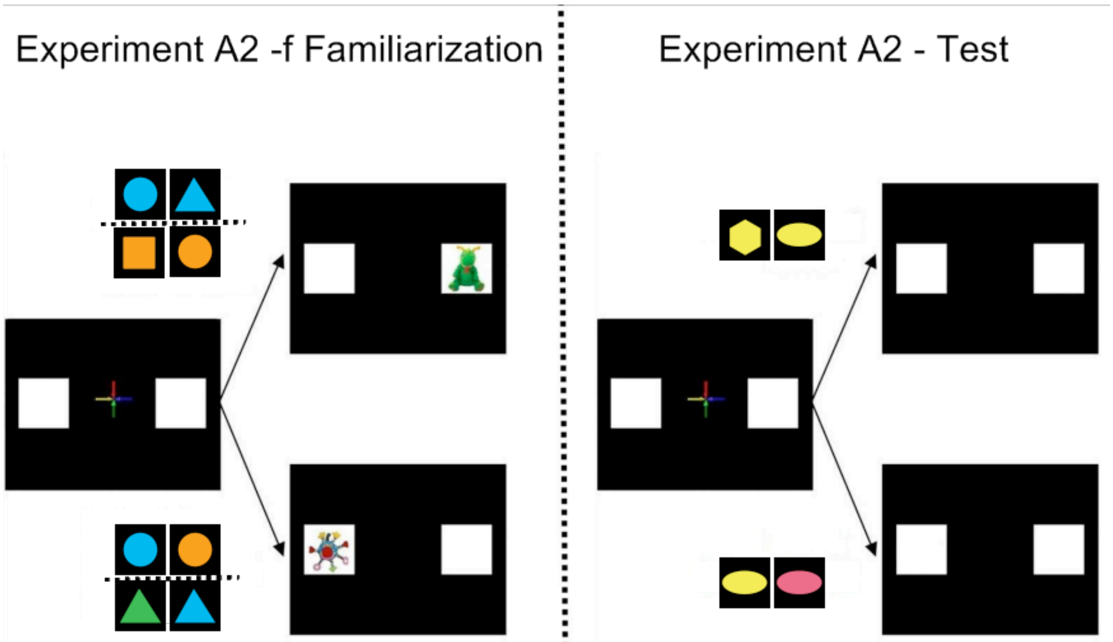

Figure 5.4 – Paradigm of Experiments A2. Participants took 32 familiarization trials and 8 test trials.

5.4.1.1 Materials and Methods

<u>Participants</u>

Twenty-six infants were included in the analysis; age range 11 month 15 days to 12 month 15 days. Six other infants participated in the study but were excluded due to fussiness (5), or the mother not following instructions (1).

<u>Stimuli</u>

Visual cues to the location of toys' appearances consisted in the repeated presentation of colorful shapes. In Experiment A2, three shapes (disk, triangle and square) and three colors (green, blue and orange) were used in the familiarization. Two different shapes (ellipse and hexagon) and two different colors (pink and yellow) were used in the test phase. Each visual stimuli occupied a square with a side length of 4 cm.. Twelve pairs of colorful shapes were created for the familiarization. Six pairs consisted in two colorful shapes sharing the same color (green disk – green square; green square – green triangle; orange square – orange triangle; orange triangle – orange disk; blue disk – blue square; blue triangle – blue disk), and six pairs consisted in two colorful shapes sharing the same shape (green disk – blue disk; blue disk – orange disk; blue triangle – orange triangle; orange triangle – green triangle; green square – blue square; orange square – green square). Four pairs were created for the test. Two pairs consisted in two colorful shapes sharing the same color (yellow ellipse – yellow hexagon; pink hexagon – pink ellipse), and two pairs consisted in two colorful shapes sharing the same shape (yellow ellipse – pink ellipse; pink hexagon – yellow hexagon). Colored geometrical shapes were synthesized with Adobe Photoshop CS version 8.0 on a computer running Mac OS X, version 10.5.7.

Procedure

The Familiarization phase consisted of 32 Familiarization trials. Familiarization trials started with a display of two white squares on the sides and a central attention-grabber. When the infant looked at it, the attention-grabber disappeared, and two colorful shapes were presented sequentially for 800 ms each, with a separation of 600 ms between the two pictures. The two colorful shapes could share either the same color, or the same shape. A meaningless attractive sound lasting 500 ms was played in synchrony with the appearance of each colorful shape. One second after the visual cue disappeared, a toy appeared in one of the squares, contingent on the visual cue: the same-color relation predicted the toy's appearance in one of the squares, while the same-shape relation predicted the toy's appearance in the other square. The pairing of the visual cues with toy-locations was counterbalanced across participants.

During test, infants were exposed to 8 trials in a pseudo-random order. Test trials were similar to the familiarization trials, except that infants saw novel visual cues, which consisted in pairs of novel objects sharing either the same color or the same shape. No toy ever appeared in the test trials. Two seconds after the visual cue disappeared, the next trial started.

5.4.1.2 Results

The first fixation data for familiarization trials is presented in Figure 5.5. Infants anticipated to one or the other side in 39% of the trials. We computed the proportion of infants showing a correct anticipatory look for each trial. A linear regression analysis yield non significant results for the same-color condition pairs, ß=-.005, R^2=.022, t(14)=-.56, *P* = .59; and for the same-shape pairs, ß=-.001, R^2=.001, t(14)=-.135, *P* = .89.

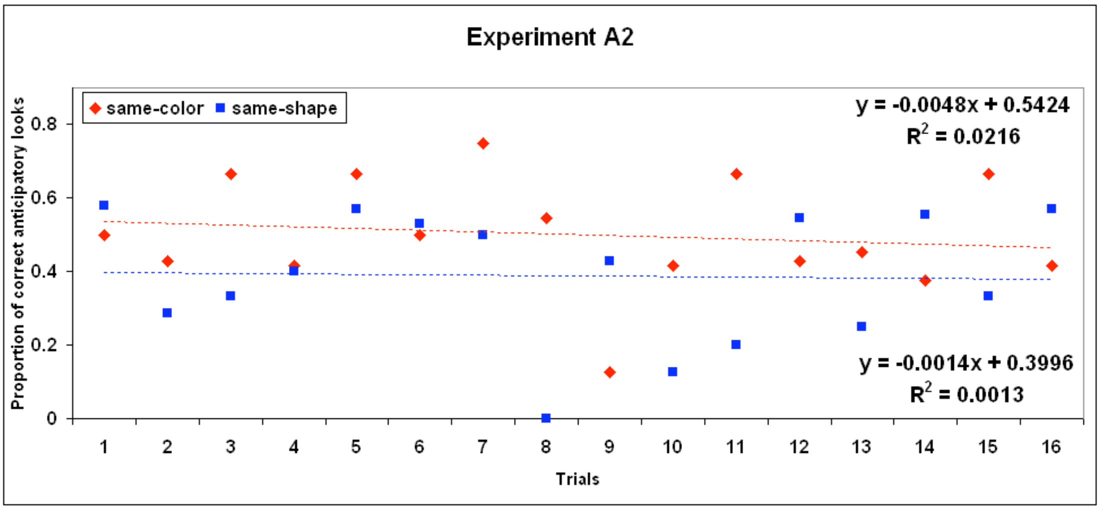

Figure 5.5 - Proportion of correct anticipatory looks for each familiarization trial in Experiment A2. The dotted lines depict the linear regression for same-color (red) and for same-shape (blue) pairs, respectively.

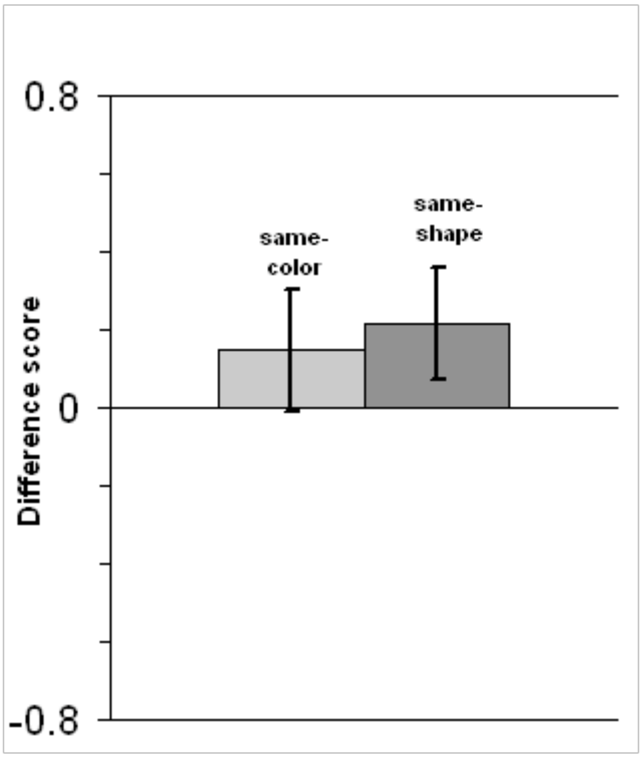

Figure 5.6 - Mean difference scores for Experiment A2 considering the first fixations. Error bars represent standard errors.

In the test phase, infants looked to the left or the right in 57% of the same-color pairs, and in 66% of the same-shape pairs. Considering first fixations (see Figure 5.6), infants' mean difference score was .15 for the same-color pairs, which was not significantly greater than 0, $t(24) = .93$; $P = .36$; $d' = .19$, and .22 for the same-shape pairs, which was not significantly greater than 0, $t(23) = 1.49$; $P = .15$; $d' = .30$. Regarding the same-color pairs, twelve infants obtained a positive difference score, seven a negative difference score and six a null difference score. Regarding the same-shape pairs, eleven infants obtained a positive difference score, eight a negative difference score and five a null difference score. Binomial tests showed that for neither type of pairs, significantly more infants obtained positive than negative difference scores, *Ps* > .35.

Three participants did not provide data for one or the other condition. Considering the 23 remaining participants, we averaged their difference score for both condition. The mean average difference score was .17, with was marginally greater than 0; $t(22) = 1.86$; $P = .076$; $d' = .39$. Fifteen infants obtained a positive average difference score, five a negative average difference score, and three a null average difference score. A binomial test showed that significantly more infants obtained a positive than a negative average difference score, $P = .04$.

Qualitatively, twenty out of twenty six infants obtained a positive difference score for one or the other type of pairs. Interestingly, if they obtained a negative difference score for the other type of pairs, the absolute value of the negative score was lower than that of the positive score (see Figure 5.7). Such a pattern suggests that some infants may have learned to predict the toy appearance for one side, but they did not all learn for the same side.

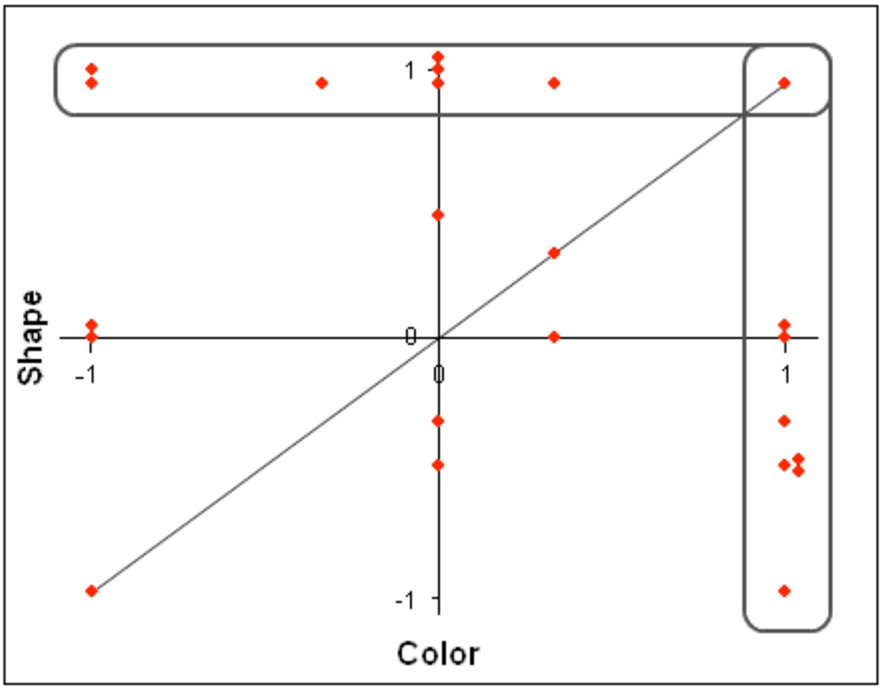

Figure 5.7 – The x-axis represents the difference scores considering first fixations for same-color pairs, and the y-axis represents the difference score for same-shape pairs. Each red dot represents one participant in Experiment A2. Most infants obtained a high positive difference score for at least one of the types of pairs.

Similar results were obtained analyzing the overall accuracy. Infants' mean difference score was .14 for the same-color pairs, which was not significantly greater than 0, t(24) = .85; *P* = .40; d' = .17, and .17 for the same-shape pairs, which was not significantly greater than 0, t(23) = 1.10; *P* = .28; d' = .22. Regarding the same-color pairs, thirteen infants obtained a positive difference score, seven a negative difference score and five a null difference score. Regarding the same-shape pairs, eleven infants obtained a positive difference score, nine a negative difference score and four a null difference score. Binomial tests showed that for neither type of pairs, significantly more infants obtained positive than negative difference scores, *Ps* > .26.

Three participants did not provide data for one or the other condition. Considering the 23 remaining participants, we averaged their difference score for both condition. The mean average difference score was .14, with was not significantly greater than 0; t(22) = 1.36; *P* = .19; d' = .28 . Fourteen infants obtained a positive average difference score, five a negative average difference score, and four a null average difference score. A binomial test showed that the number of infants who obtained a positive average score was marginally greater than the number of infants who obtained a negative average score, *P* = .06.

Qualitatively, twenty out of twenty six infants obtained a positive difference score for one or the other type of pairs. Interestingly, if they obtained a negative difference score for the other type of pairs, the absolute value of the negative score was lower than that of the positive score. Such a pattern suggests that some infants may have learned to predict the toy appearance for one side, but they did not all learn for the same side.

5.4.1.3 Discussion

In this experiment, we did not find evidence that infants could generalize the "same" relation over shape or color. However, as always with negative results, it is not yet possible to conclude. We consider two possible accounts for our data.

First, infants may not be able to generalize a "same" relation in the object domain in these conditions. Particularly, infants may not be able to detect and generalize a "same" relation, implemented on one dimension only of the objects, i.e. shape or color.

An alternative account, however, suggested by the qualitative observation of the behavior of individuals, is that 12-month-olds can generalize the "same" relation over shape and over color; but they are not better for one or the other relation.

Therefore, some infants learn the same-shape relation, and others learn the same-color relation. As often in our paradigm (see Chapter 2; Kovacs, 2008; Kovacs & Mehler, 2009b), if there is no strong preference for one of the two patterns presented to infants, we obtain null group results.

To assess which of these two accounts is correct, Experiment A3 asks whether 12-month-olds can generalize the same-shape relation, when contrasted by pairs of colorful shapes differing both in shape and color.

5.4.2 Same-shape vs. different – Experiment A3

The paradigm of Experiment A3 is presented in Figure 5.8. Experiment A3 is identical to Experiment A2, except for the visual stimuli. As in Experiment A2, the appearance of a toy in one location of the screen was associated to pairs of colorful shapes sharing the same shape. In contrast to Experiment A2, however, the appearance of the toy in the other location was associated to pairs of colorful shapes deferring both in shape and color.

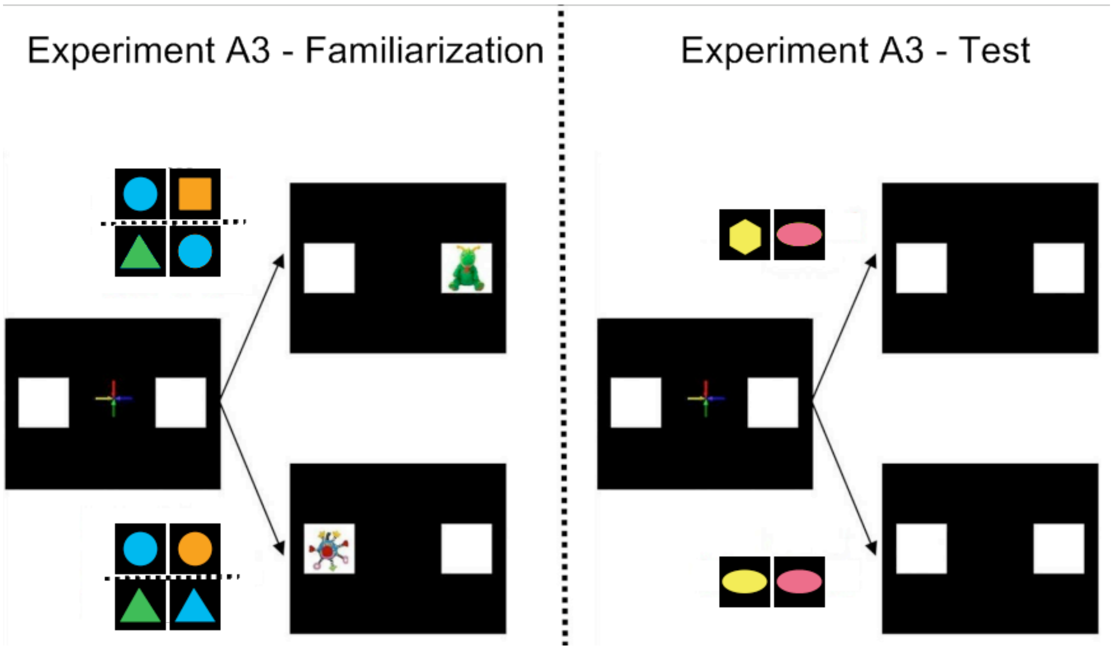

Figure 5.8 – Paradigm of Experiments A3. Participants took 32 familiarization trials and 8 test trials.

5.4.2.1 Material and Methods

<u>Participants</u>

Twenty-five infants were included in the analysis; age range 11 month 15 days to 12 month 15 days. Nine other infants participated in the study but were excluded due to fussiness (8), or equipment failure (1).

<u>Stimuli</u>

The same colorful shapes as in Experiment A2 were used in this experiment. For familiarization trials, six same-shape pairs consisted in two colorful shapes sharing the same shape (green disk – blue disk; blue disk – orange disk; blue triangle – orange triangle; orange triangle – green triangle; green square – blue square; orange square – green square), and six different pairs consisted in two colorful shapes differing both in shape and color (blue disk – green triangle; blue square – orange triangle; orange square – green disk; orange triangle – blue disk; green triangle – orange square; green disk – blue square). For test trials, two pairs consisted in two colorful shapes sharing the same shape (yellow ellipse – pink ellipse; pink hexagon – yellow hexagon), and two pairs consisted in two colorful shapes differing both in shape and color (pink ellipse – yellow hexagon; yellow hexagon – pink ellipse).

5.4.2.2 Results

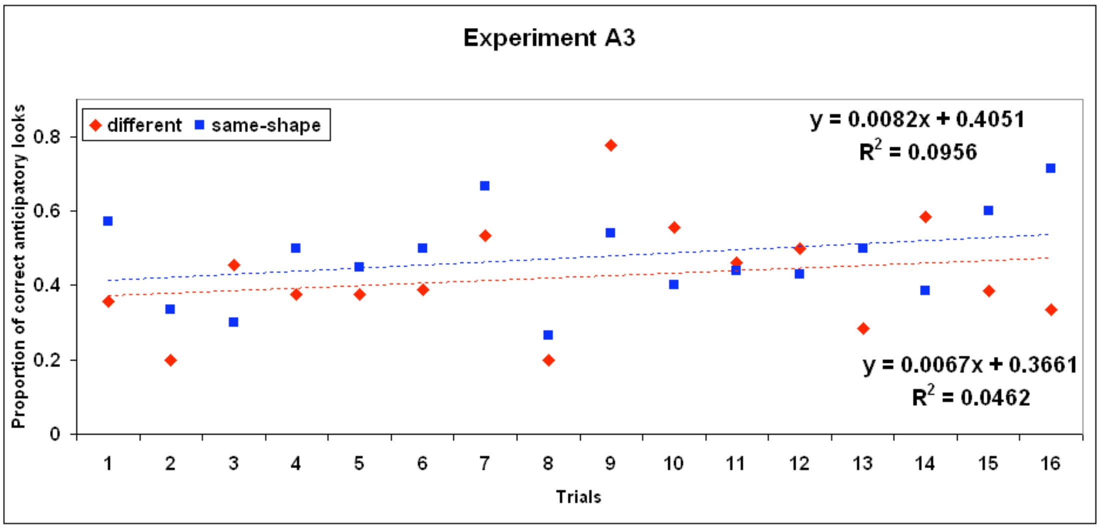

Figure 5.9 - Proportion of correct anticipatory looks for each familiarization trial in Experiment A3. The dotted lines depict the linear regression for same-shape (blue) and for different (red) pairs, respectively.

The first fixation data for familiarization trials is presented in Figure 5.9. Infants anticipated to one or the other side in 47% of the trials. We computed the proportion of infants showing a correct anticipatory look for each trial. A linear regression analysis showed non significant increases in correct anticipations for the same-shape pairs, ß=-.008; R^2=.096; t(14)=1.22; *P* = .24; and for the different pairs, ß=.007; R^2=.046; t(14)=.82; *P* = .42.

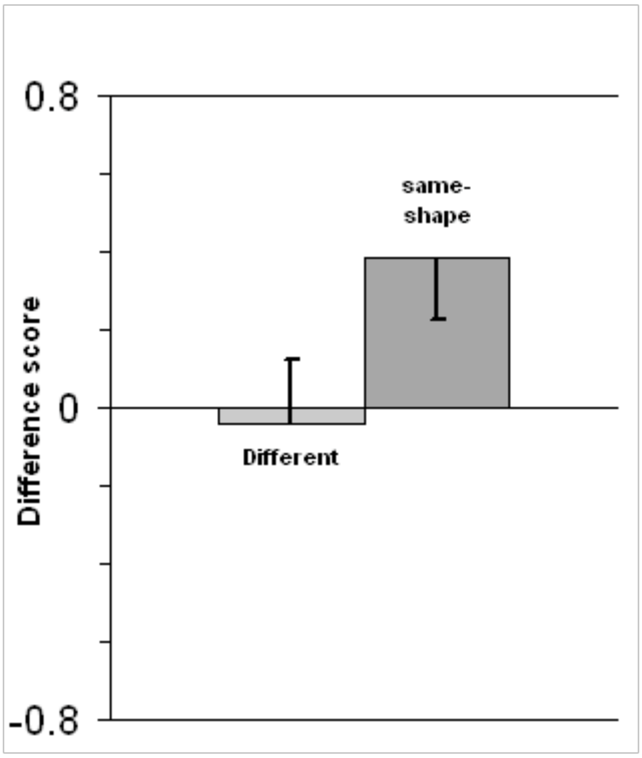

Figure 5.10 - Mean difference scores for Experiment A3 considering the first fixations. Infants correctly generalized the prediction made by the same-shape pairs, but not that made by the different pairs. Error bars represent standard errors.

In the test phase, infants looked to the left or the right in 61% of the same-shape trials, and in 50% of the different trials. Two infants did not provide data for the same-shape trials, and one infant did not provide data for the different trials, so that 23 infants were included in the analysis of the same-shape trials, and 24 in the analysis of the different trials. Considering first fixations (see Figure 5.10), infants' mean difference score was .38 for the same-shape pairs, which was significantly greater than 0, $t(22) = 2.40$; $P = .025$; d' = .50, and .-.04 for the different pairs, which did not significantly differ from 0, $t(23) = -.25$; $P = .80$; d' = .05. Regarding the same-shape pairs, sixteen infants obtained a positive difference score, seven a negative difference score. This difference was however not significant, as assessed by a binomial test, *P* = .09. Regarding the different pairs, nine infants obtained a positive difference score, ten a negative difference score and five a null difference score.

Considering overall accuracy, infants' mean difference score was .20 for the same-shape pairs, which was not significantly greater than 0, t(22) = 1.18; *P* = .25; d' = .24, and -.27 for the different pairs, which did not significantly differ from 0, t(23) = -1.60; *P* = .12; d' = .33. Regarding the same-shape pairs, thirteen infants obtained a positive difference score, eight a negative difference score and two a null difference score. Regarding the different pairs, eight infants obtained a positive difference score, thirteen a negative difference score and three a null difference score. Binomial tests showed that for neither type of pairs, significantly more infants obtained positive than negative difference scores, *Ps* = .38.

5.4.2.3 Discussion

Analyzing the direction of first fixations, the results of Experiment A3 show that 12-month-old infants are able to generalize the "same" relation implemented over shape. These results however were not confirmed statistically when analyzing overall accuracy, even though the data showed qualitatively the same pattern of results. This discrepancy may mean that infants' predictions are somewhat weak in this experiment.

Still, in the light of these results, we can now interpret the results of Experiment A2. The failure of Experiment A2 is not due to infants' inability to generalize the relation of identity over shape or color, as they could do so over shape in Experiment A3. Rather, we can conclude, as we suggested, that infants are able to generalize the identity relation both on shape and color, but neither of these two relations is easier than the other. To be able to conclude, we need however to assess whether 12-month-olds can generalize the same-color relation.

5.4.3 Same-color vs. different – Experiment A4

The paradigm of Experiment A4 is presented in Figure 5.11. Experiment A4 is identical to Experiment A2 and A3, except for the visual stimuli. As in Experiment A2, the appearance of a toy in one location of the screen was associated to pairs of colorful shapes sharing the same color. In contrast to Experiment A2, however, the appearance of the toy in the other location was associated to pairs of colorful shapes deferring both in shape and color.

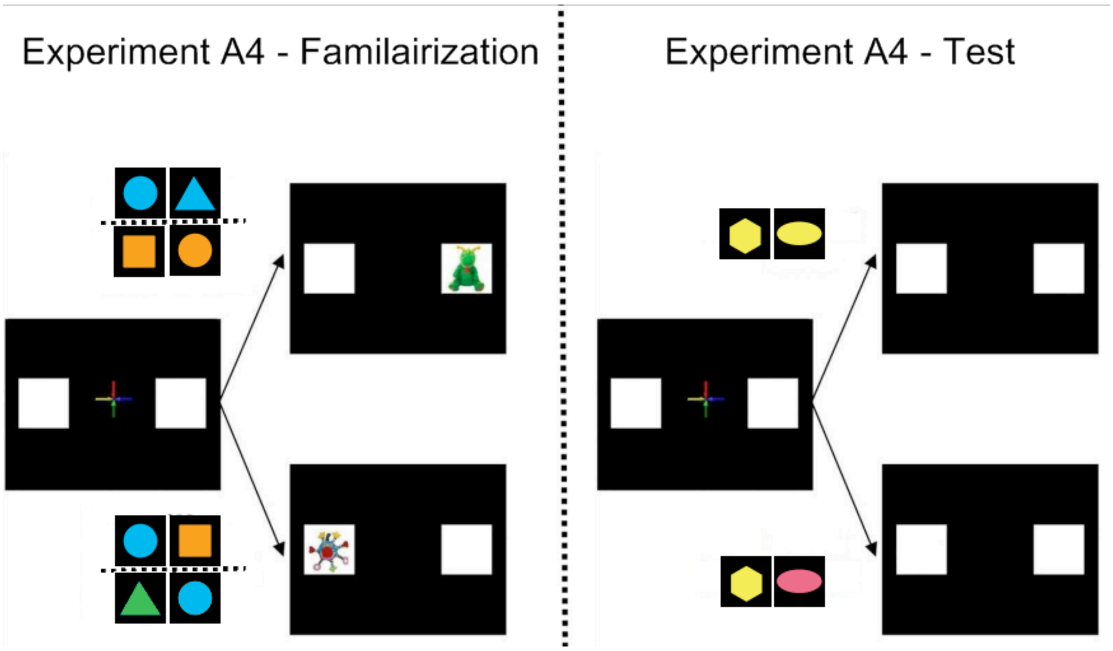

Figure 5.11 – Paradigm of Experiments A4. Participants took 32 familiarization trials and 8 test trials.

5.4.2.1 Material and Methods

Participants

Twenty-one infants were included in the analysis; age range 11 month 15 days to 12 month 15 days. Four other infants participated in the study but were excluded due to fussiness (2), equipment failure (1), or the mother not following instructions (1).

Stimuli

The same colorful shapes as in Experiment A2 were used in this experiment. For familiarization trials, six same-color pairs consisted in two colorful shapes sharing the same color (green disk – green square; green square – green triangle; orange square – orange triangle; orange triangle – orange disk; blue disk – blue square; blue triangle – blue disk), and six different pairs consisted in two colorful shapes differing both in shape and color (blue disk – green triangle; blue square – orange triangle; orange square – green disk; orange triangle – blue disk; green triangle – orange square; green disk – blue square). For test trials, two pairs consisted in two colorful shapes sharing the same color (yellow ellipse – yellow hexagon; pink hexagon – pink ellipse), and two pairs consisted in two colorful shapes differing both in shape and color (pink ellipse – yellow hexagon; yellow hexagon – pink ellipse).

5.4.2.2 Results

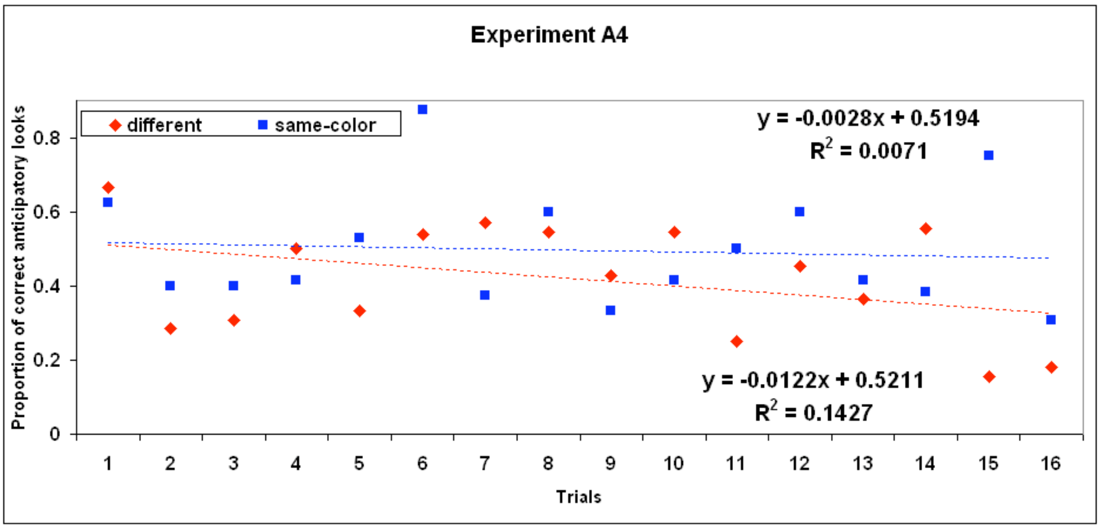

Figure 5.12 - Proportion of correct anticipatory looks for each familiarization trial in Experiment A4. The dotted lines depict the linear regression for same-color (blue) and for different (red) pairs, respectively.

The first fixation data for familiarization trials is presented in Figure 5.12. Infants anticipated to one or the other side in 56% of the trials. We computed the proportion of infants showing a correct anticipatory look for each trial. A linear regression analysis showed a non significant decrease in correct anticipations for the same-color pairs, ß=-.002; R^2=.003; t(14)=-.20; P = .85; and a larger but statistically non-significant decrease for the different pairs, ß=-.011; R^2=.13; t(14)=-1.46; P = .17.

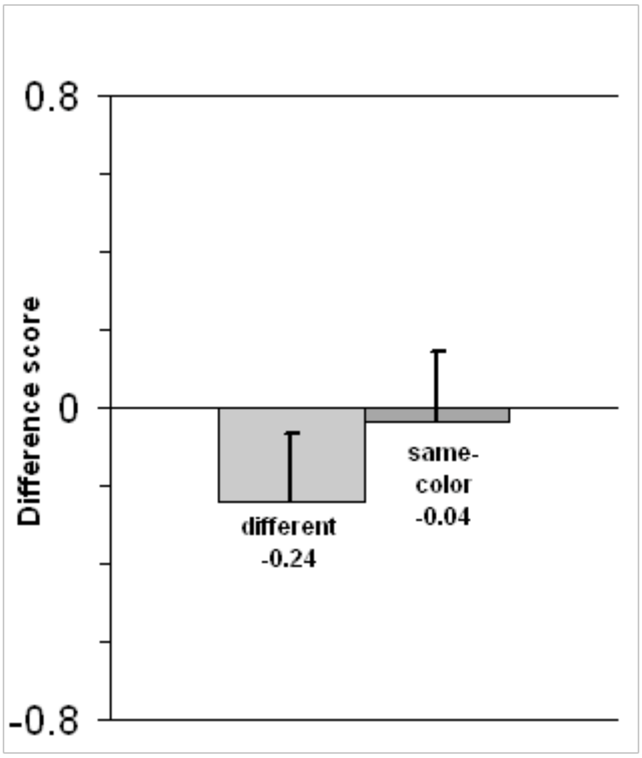

Figure 5.13 - Mean difference scores for Experiment A4 considering the first fixations. Error bars represent standard errors.

In the test phase, infants looked to the left or the right in 59% of the same-color trials, and in 61% of the different trials. Two infants did not provide data for the same-color trials, and one infant did not provide data for the different trials, so that 19 infants were included in the analysis of the same-color trials, and 20 in the analysis of the different trials. Considering first fixations (see Figure 5.13), infants' mean difference score was -.035 for the same-color pairs, which statistically did not differ from 0, t(18) = -.19; *P* = .85; d' = .045, and .-.24 for the different pairs, which did not significantly differ from 0, t(19) = -1.37; *P* = .186; d' = .31. Regarding the same-color pairs, seven infants obtained a positive difference score, eight a negative difference score, and four a null difference score. Regarding the different pairs, seven infants obtained a positive difference score, ten a negative difference score and three a null difference score. Binomial tests showed that for neither type of pairs, significantly more infants obtained positive than negative difference scores, *Ps* > .62.

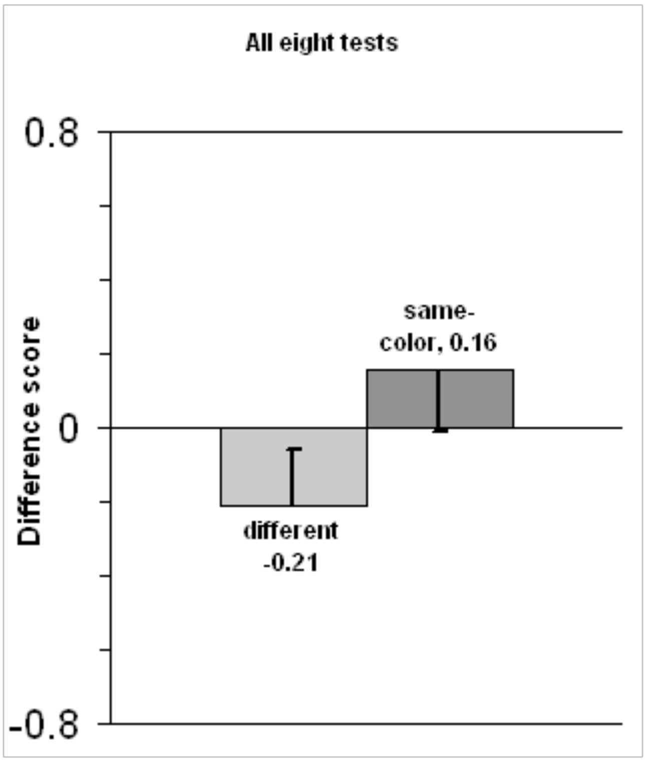

Figure 5.14 - Mean difference scores for Experiment A4 considering the overall accuracy. Error bars represent standard errors.

Considering overall accuracy (see Figure 5.14), infants' mean difference score was .16 for the same-color pairs, which was not significantly greater than 0, t(18) = 1.96; *P* = .35; d' = .22, and -.21 for the different pairs, which did not significantly differ from 0, t(19) = -1.21; *P* = .24; d' = .27. Regarding the same-color pairs, eight infants obtained a positive difference score, five a negative difference score and six a null difference score. Regarding the different pairs, seven infants obtained a positive difference score, ten a negative difference score and three a null difference score. Binomial tests showed that for neither type of pairs, significantly more infants obtained positive than negative difference scores, *Ps* > .58.

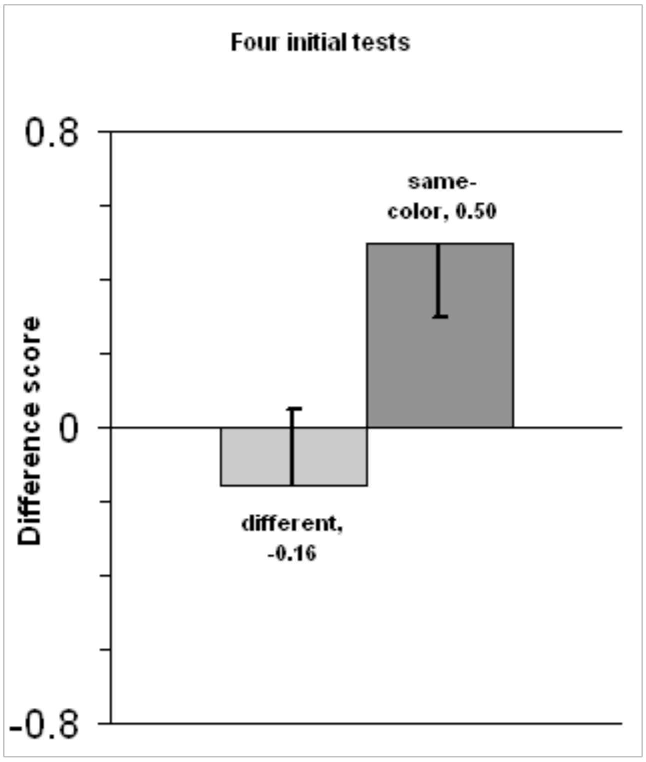

Figure 5.15 - Mean difference scores for Experiment A4 considering overall accuracy. Infants correctly generalized the prediction made by the same-color pairs, but not that made by the different pairs. Error bars represent standard errors.

Thus, considering all eight test trials, infants show no evidence of having learned and generalized the same-color relation. However, qualitatively, the pattern of results observed for overall accuracy matched our predictions. In the absence of reinforcement due to the absence of toys in the test phase, extinction effects may be observed. Thus, we also ran our analyses considering only the first 4 test trials. Three infants did not provide data for the same-color test trials, and one infant did not provide data for the different pairs test trials. Thus, sixteen infants were included in the analysis of the same-color test trials, and nineteen in the analysis of the different test trials. This did not provide significant results for the pattern of first fixations; *Ps* > .20. Considering overall accuracy (see Figure 5.15), however, infants' mean difference score was .5 for the same-color pairs, which was significantly greater than

0, t(15) = 2.45; *P* = .023; d' = .61, and -.16 for the different pairs, which did not significantly differ from 0, t(18) = -.77; *P* = .45; d' = .176. Regarding the same-color pairs, eleven infants obtained a positive difference score, three a negative difference score and two a null difference score. A binomial test comparing the number of infants obtaining a positive score and the number of infants obtaining a negative score yield marginally significant results; *P* = .057. Regarding the different pairs, six infants obtained a positive difference score, nine a negative difference score and four a null difference score A binomial test comparing the number of infants obtaining a positive score and the number of infants obtaining a negative score showed that this difference was not statistically significant; *P* > .60.

5.4.2.3 Discussion

In Experiment A4, as for Experiment A3, we did not find congruent results considering first fixations and overall accuracy. This suggests that infants' predictions in these experiments are weak. Still, we found evidence suggesting that 12-month-old infants are able to generalize the same-color relation. These results therefore confirm our interpretation of Experiment A2. Infants are able to generalize both the same-shape and the same-color relations, and do not handle better one or the other.

5.5 Conclusion

In Experiments A2-A4, we assessed infants' ability to learn and generalize identity relations. We showed that infants could learn and generalize the "same" relation over shape and over color. Furthermore, we tested whether it was easier for

12-month-olds to learn and generalize the same-shape or the same-color relation. No preference was observed. In fact, we observed that about the same number of infants learning the relation over color, as the number of infants learning the relation over shape.

Thus, the advantage for shape observed in Experiment A1 was not observed in a task that did not require storing object representations in long-term memory. We thus propose that the well-known shape advantage for object individuation (Tremoullet et al., 200; Xu et al., 2004) reflects a structural bias of memory, rather than a perceptual bias.

CHAPTER 6

Appendix B – Edge effect in 3-month-olds' word memory

6.1 Introduction

An important issue to progress in our understanding of language acquisition is to clarify what information the infants can process. Numerous adult experiments showed an advantage for elements positioned at perceptual edges for memory and generalization processes. For example adults remember better the first and last elements of a list to remember (Henson, 2001). Endress, Scholl and Mehler (2005) showed that adults can generalize a repetition-based structure in the speech domain, only if the repetition occurs in an edge position; i.e. they generalize the structure ABCDEFF, but not the structure ABCDDEF. Pena and colleagues (2002) showed that the insertion of subliminal edges (i.e. 25 ms pauses) in a continuous speech stream triggers the generalization of a non-adjacent dependency after only two minutes of exposure, whereas that dependency was not otherwise generalized even after 30 minutes of exposure. Perceptual edges appear therefore to play a crucial role for triggering generalizations (see also Endress & Bonati, 2007).

In Experiment B1, we asked whether infants process differently syllables that are positioned at perceptual edges, and syllables that are not. Precesiely, we ask whether they would detect a syllable change if it occurred in the first, third or final position of a repeated five-syllabic sequence. Our paradigm was adapted from Dehaene-Lambertz and Dehaene (1994). We recorded infants' encephalogrammes while they listened to repeated five-syllabic sequences. Each trial consisted in four five-syllabic continuous speech sequences. The first three sequences were identical and correspond to the habituation sequence. The fourth sequence (e.g. *fominegadu*), the target, could differ from the habituation sequence in the first (habituation: *Shuminegadu*), third (habituation: *fomisogadu*) or final syllable (habituation: *fominegali*).

6.2 Materials and methods

6.2.1 Participants

Twenty-two infants were included in the analysis. Eight were excluded for yielding less than 30 artifact-free trials. The remaining 14 infants (age range 3 months 07 days – 3 months 24 days, mean age 3 months 18 days) provided on average 42 artifact-free trials. Twenty-one other infants participated to the experiment but were excluded without being analyzed for listening to less than 40 trials (10 per condition).

6.2.2 Stimuli

Twelve set of four five-syllabic sequences were created with the following eight syllables: *fo, du, ne, shu, so, li, mi, ga*. The syllables *mi* and *ga* always occurred in the position 2 or 4 of the sequences. For each set, one sequence constituted the target (e.g., *fomidugane*), and the three other sequences could differ from the target in the initial syllable (e.g., *shumidugane*), the third syllable (e.g., *fomisogane*) or the final syllable (e.g., *fomidugali*). Three syllable changes could occur: *fo* to *shu* (*sh* being pronounced as in *shoes*), *du* to *so*, and *ne* to *li*. Each change consisted in the change of two consonantal features and one vocal feature (i.e. close vs. mid-close vowels). The different sets were constructed so that all specific changes occurred in all locations across sets (i.e. the *fo* to *shu* change occurred in initial position for one set, in third position for another set, and in final position for a third set). The target for each set could be *fomidugane, fominegadu, duminegafo, dumifogane, nemifogadu, nemidugafo, fogadumine, foganemidu, duganemifo, dugafomine, negafomidu* or *negadumifo*. Five-syllabic pseudo-words were synthesized with MBROLA (fr4) with a phoneme duration of 120 ms and a monotonous pitch of 200 Hz. There was no silent pause between two syllables within a word.

6.2.3 Procedure

Each participant was randomly assigned a set of four five-syllabic sequences. Infants sat on the lap of their mother or father, wearing a129 electrodes HydroCel Geodesic Sensor Net, adapted to the circumference of its head. They faced a screen showing a short silent movie, which repeated from the beginning at the beginning of each trial. Each trial consisted in the presentation of four five-syllabic speech sequences. Each sequence lasted 1200 ms, and two sequences were separated by 600 ms of silence. The first three sequences constituted the habituation and were identical. The fourth sequence constituted the target, and was the same across all trials for a given participant. It could be identical to the habituation sequence (no-change condition), or differ from the habituation sequence in the first (initial syllable change condition), in the third (third syllable change condition), or in the final syllable (final syllable change condition). As the target was the same in all conditions, a difference of evoked potential in the target between conditions was necessarily due to the encoding of the sequence repeated in habituation.

Between each trial, one of five different sounds lasting between 2 and 4 s were played to maintain infants' attention and avoid interferences of sequence memory between trials. Table 6.1 summarizes the design of the experiment.

Table 6.1 Paradigm of Experiment B1

Condition	Habituation	Target
No change	fomidugane x3	fomidugane
Initial syllable change	**SHU**midugane x3	**FO**midugane
Third syllable change	fomi**SO**gane x3	fomi**DU**gane
Final syllable change	fomiduga**LI** x3	fomiduga**NE**

6.2.4 Data analysis

EEG was recorded from 129 electrodes HydroCel Geodesic Sensor Nets (EGI recording system). Scalp voltages were amplified, low-passes filtered at 100Hz and digitized at 250Hz. The signal was then digitally filtered between 0.3 and 20 Hz. Epochs starting 220ms before the onset of the first, third and final syllables, respectively, and ending 460 ms after it were extracted from the continuous signal. The epochs automatically edited to reject bad channels for each trials (voltage exceeding threshold of 80 μV) and trials contaminated by body movement (at least 50% bad channels). Infants with less than 30 artifact-free trials were rejected from the analysis. The artifact-free trials were averaged for each subject in two conditions, change and no-change, for the first, the third and the final syllable separately. The epochs for each condition were then re-referenced to an average reference corrected for polar average reference effect (PARE correction), and corrected to a 22 ms baseline.

Our analysis aims at understanding whether infants detect the syllable change in each of the three studied serial positions. Thus, to select the sites and time-windows of interest, we computed the main effect map, by averaging together the epochs for the change condition for the first, third and final syllable on one hand, and the epochs for the no-change condition for the first, third and final syllable on the other hand. A t-test was computed for each electrode and each data-point. We selected the sites that regrouped at least three adjacent electrodes with significant main effect ($P < .01$) in a time window lasting at least 20 ms. Two groups of six electrodes were selected in the anterior region, symmetrically disposed to the left and the right of Fz. Two time windows were selected: 140-200 ms and 284-392 ms. In addition to being the time windows were the main effect appeared the most significant, thes two time windows also correspond to the two peaks elicited by syllables and reported by Dehaene-Lamberts & Dehaene (1994).

For each participant, we computed the average scalp voltage for each time window and each group of electrode. Among infants included in the final analysis, the set of stimuli 1 was assigned to 5 infants, each of the sets 3, 5, 6, 8 and 12 were assigned to 1 infant, and the sets 7 and 9 were assigned to 2 infants each. Altogether, except for the predominance of the set 1, the number of infants who were assigned each set of words was thus balanced. Because there were few infants for each set of stimuli, we did not include that factor in our analysis. We run a repeated measure analysis of variance, with syllable position (3), change (2), lateralization (2) and time (2) as within-subject factors. We are particularly interested in the interactions with change and syllable position.

6.3 Results

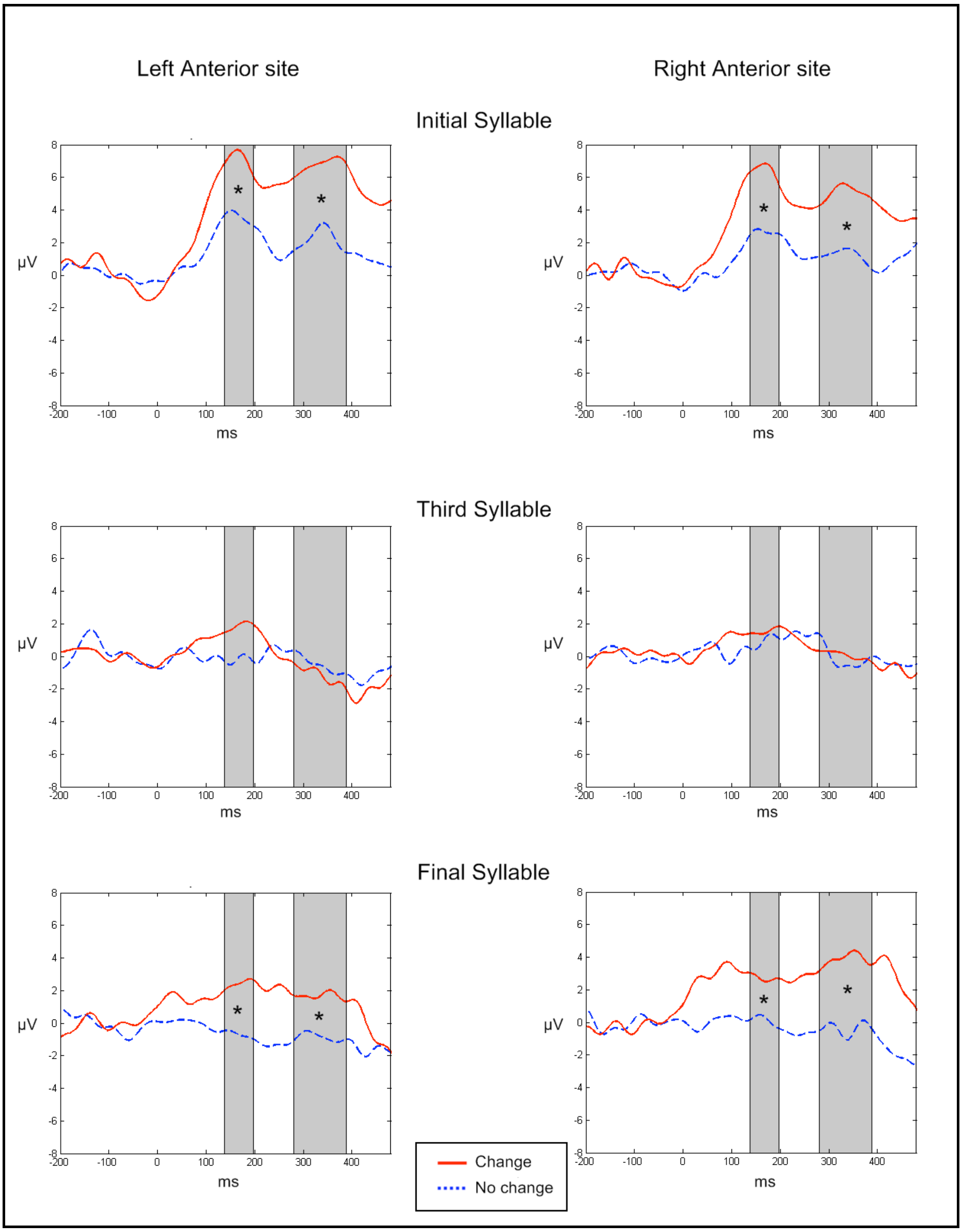

Figure 6.1 – Event related potentials for the first, third and final syllables in anterior left and right regions. The red full lines and dotted blue lines represent the ERPs to the change and tno-change condition, respectively. The time windows of interest are depicted in gray.

In addition to the main effect of change; F(1,13) = 18..93, $P < .0008$; for which electrodes and time-windows were selected, the 3x2x2x2 repeated-measure ANOVA yield a significant effect syllable position; F(2, 26) = 5.97, $P = .007$. Post-hoc analysis showed that the initial syllable elicited higher response than the third ($P = .009$, Bonferroni corrected) and final syllable ($P = .045$). This effect may reflect the fact that the baseline was taken on silence for the initial syllable, whereas it was taken on a syllable for the third and final syllables. The lateralization x syllable position interaction was significant; F(2, 26) = 4.71; $P = .018$, reflecting stronger effect of syllable position in the left site ($Ps < .000001$) than in the right site ($Ps < .00001$). Most importantly, the syllable position by change by lateralization by time interaction was significant; F(2, 26) = 3.90, $P = .033$. It reflects that the effect of change was significant for the initial syllable in both sites and both time windows ($Ps < .00002$), for the final syllable in both sites and both time windows ($Ps < .005$), but only showed a marginal effect ($P = .056$) for the left site in the first time window for the third syllable (all other $Ps > .99$).

6.4 Discussion

In Experiment B1, we found strong evidence that infants detected the change of a syllable when it occurred in an edge position, either initial or final. In contrast, we found little evidence that they could detect a syllable change when it occurred in the middle of the sequence, in the third syllable.

Each syllable elicited two peaks of activation in anterior regions, as previously described in 3-month-olds for isolated syllables (Dehaene-Lambertz & Dehaene, 1994; Dehaene-Lambertz & Baillet, 1998). Our peak 1 reached a maximum about 170 ms after the syllable onset, and peak 2 reached a maximum about 390 ms after the syllable onset. These timings occurred slightly earlier than what has previously been reported, which may be due to the faster speech rate used in our stimuli (syllables lasted 240 ms in our experiment; 289 ms in Dehaene-Lambertz & Dehaene, 1994; and 275 ms in Dehaene-Lambertz & Baillet, 1998). In previous reports, an effect of syllable change was only observed for the second peak, and was stronger for a phonological than for an acoustic change. In these experiments, the stimuli consisted in syllables with the same vowel, and varying only in one consonantal feature (i.e. place of articulation).

In contrast to previous reports, we found strong effects of syllable change detection both for peak 1 and 2, for the edge syllables, and a marginal effect of syllable change in the third position for the first peak only. Our syllabic pairs varied in two consonantal features and one vocalic feature, and were thus phonologically and acoustically more different that the stimuli used by Dehaene-Lambertz and colleagues. A tentative interpretation of our pattern of results is that peak 1 is only sensitive to low-level acoustic information, whereas peak 2 is sensitive to phonological information. Thus, we suggest that edge syllables are encoded both at an acoustic and at a phonological level, whereas the third syllable may be weakly encoded at an acoustic level only.

In conclusion, we showed that edge prevalence in speech memory is present in 3-month-olds. Edges appear to be better processed and encoded at a higher level than non-edgy syllables. Future experiments should now investigate what learning mechanisms can apply to edges, and what type of information can be extracted.

References

Akhtar, N. & Tomasello, M. (1996). Two-year-olds learn words for absent objects and actions. *British Journal of Developmental Psychology*, 14, 79 – 93.

Altmann, G. T. (2002). Learning and development in neural networks–the importance of prior experience. *Cognition*, 85(2), B43-50.

Baker, M. (2001). *The atoms of language*. New York, NY: Basic Books.

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? Journal of Experimental Child Psychology, 64, 3–26.

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development,* 62, 875-890.

Baldwin, D. A. (1993). Infants' ability to consult the speaker for clues to word meaning. *Journal of Child Language*, 20, 395-418.

Banks, M., Aslin, R. & Letson, R. (1975). Sensitive period for the development of human binocular vision. *Science*, 190, 675-7.

Benavides-Varela, S., Gomez, D. M., Bion, R. A. H., Macagno, F., & Mehler, J. (in prep.) Memory for words in neonates: a brain imaging study.

Benavides-Varela, S., Hochmann, J-R., Macagno, F., Nespor, M., & Mehler, J. (in prep.) Newborn's brain memorizes vowels.

Berent, I., & Perfetti, C. A. (1995). A rose is a reez: The two cycles model of phonology assembly in reading English. *Pshychological Review*, 102, 146-184

Berent, I., Vaknin, V., & Marcus, G. F. (2007). Roots, stems, and the universality of lexical representations: Evidence from Hebrew. *Cognition,* 104, 254-286.

Bernal, S., Lidz, J., Millote, S., & Christophe, A. (2007). Syntax constrains the acquisition of verb meaning. Language Learning and Development, 3, 325–341.

Bertelson, P., de Gelder, B., Tfouni, L. V., & Morais, J. (1989). Metaphonological abilities of adult illiterates: New evidence of Heterogeneity. *European Journal of Cognitive Psychology*, *1*, 239–250.

Bertoncini, J., Bijeljac-Babic, R., Jusczyk, P. W., Kennedy, L., Mehler, J. (1988). An Investigation of Young Infants' Perceptual Representations of Speech Sounds, *Journal of Experimental Psychology : General*, *117*, 21-33.

Bertoncini, J., Floccia, C., Nazzi, T., Mehler, J. (1995) Morae and Syllables: Rythmical Basis of Speech Representations in Neonates. *Language and Speech, 38*, 311-329

Bijeljac-Babic, R., Bertoncini, J., Mehler, J. (1993) How Do Four-Day-Old Infants Categorize Multisyllabic Utterances? *Developmental Psychology, 29*, pp.711-723

Bion, R.H, Benavides Varela, S., & Nespor, M. (in press) Acoustic markers of prominence influence infants' and adults' memory for speech sequences. Language and Speech.

Blackwell, H. R., & Schlosberg, H. (1943). Octave generalization, pitch discrimination, and loudness thresholds in the white rat. *Journal of Experimental Psychology*, 33, 5.

Bloom, P. (2000). How Children Learn the Meanings of Words. Cambridge, MA: MIT Press.

Bloom, P., & Markson, L. (1998). Capacities underlying word learning. Trends in Cognitive Sciences, 2, 67-73.

Boersma, P., & Weenink, D. (2008). Praat: doing phonetics by computer (Version 5.0.25) [Computer program]. Retrieved from http://www.praat.org/

Bonatti, L., Frot, E., Zangl, R., Mehler, J. (2002). The Human First Hypothesis: Identification of conspecifics and individuation of objects in the young infant*.* *Cognitive Psychology*, 44, 388-426.

Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations: The role of consonants and vowels in continuous speech processing. Psychological Science, 16, 451-459.

Bonatti, L. L., Peña, M., Nespor, M., & Mehler, J. (2007). On consonants, vowels, chickens, and eggs. Psychological Science, 18, 924-925.

Booij, G. (1981). *Generatieve Fonologie van het Nederlands*. Utrecht: Het Spectrum.

Booth, A. E., & Waxman, S. R. (2002) Word learning is ‘smar’: evidence that conceptual information affects preschoolers’ extension of novel words. *Cognition*, 84, B11-B22.

Brainard, M. S., & Doupe, A. J. (2002). What songbirds teach us about learning. *Nature*, 417, 351-8.

Brent, M.R., & Siskind, J.M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition*, 81, B33–B44.

Broca, P. (1861) Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie. *Bulletin de la Société Anatomique* tome XXXVI: 330-357

Broadbent, D. E., & Gregory, M. (1964). Accuracy of recognition for speech presented to the right and left ears. *Quarterly Journal of Experimental Psychology,* 16, 359–360.

Brown, R. (1957). Linguistic determinism and parts of speech. Journal of Abnormal and Social Psychology, 55, 1-5.

Bushnell, I. W. R. (1979). Modification of the externality effect in young infants. *Journal of Experimental Child Psychology*, 28, 211-229.

Caramazza, A., Chialant, D., Capasso, R., & Miceli, G. (2000). Separable processing of consonants and vowels. Nature, 403, 428-430.

Carey, S. (2009). *The Origin of Concepts*. New York: Oxford University Press.

Carey, S. & Bartlett, E. (1978). Acquiring a single new word. Papers and Reports on Child Language Development, 15, 17-29.

Carey, S. & Spelke, E. (1996). Science and core knowledge. *Philosophy of Science, 63(4),* 515-533.

Carreiras, M., Gillon-Dowens, M., Vergara, M., & Perea, M. (2009). Are vowels and consonants processed differently? Event-related potential evidence with a delayed letter paradigm. Journal of Cognitive Neuroscience, 21, 275-288

Caselli, M. C., Bates, E., Casadio, P., Fenson, L., Fenson, J., Sanderl, L. & Weir, J. (1995). A cross-linguistic study of early lexical development. Cognitive Development, 10, 159-99.

Chan, B. H-S. (2008). Code-switching, word order and the lexical/functional category distinction. *Lingua*, 118, 777-809.

Chomsky, N. (1957). Syntactic structures. The Hague: Mouton.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press

Chomsky, N. (1966). Cartesian Linguistics: Acquisition and use of language. In *Cartesian Linguistics*. New York: Harper and Row.

Chomsky, N. (1970). *Current Issues in Linguistic Theory*. Berlin and New York: Mouton de Gruyter.

Chomsky, N. (1980). On cognitive structures and their development: a reply to Piaget. In M. Piatelli-Palmarini (Ed.), Language and Learning: The Debate between Jean Piaget and Noam Chomsky (pp. 35–54). Cambridge, MA: Harvard University Press.

Chomsky, N. (1981). *Lectures on Government and Binding: The Pisa Lectures*. Holland: Foris Publications

Chomsky, N. (2000). *New horizons in the study of language and mind*. Cambridge, MA: Cambridge University Press.

Chomsky, N., & Lasnik, H. (1993). The theory of principles and parameters. In J. Jacobs et al. (eds.) *Syntax: An International Handbook of Contemporary Research*, Vol. 1. Walter de Gruyter, pp. 506-569

Christiansen, M.H., Allen, J., & Seidenberg, M.S. (1998). Learning to segment speech using multiple cues: a connectionist model. *Language and Cognitive Processes*, 13, 221–268.

Christiansen, M. & Curtin, S. (1999). Transfer of learning : rule acquisition or statistical learning ? *Trends in Cognitive Sciences*, 3, 289-290.

Christophe, A., Dupoux, E., Bertoncini, J., Mehler, J. (1994) Do Infants Perceive Word Boundaries? An Empirical Approach to the Bootstrapping Problem for Lexical Acquisition.*Journal of the Acoustical Society of America, 95,*1570-1580.

Christophe, A., Mehler, J., & Sebastián-Gallés, N. (2001), Infants' Sensitivity for Prosodic Boundary Cues, Infancy, 2, 385-94.

Christophe, A., Millotte, S., Bernal, S., & Lidz, J. (2008). Bootstrapping lexical and syntactic acquisition. Language and Speech, 51, 61–75

Christophe, A., M. Nespor, E. Dupoux, M.-T. Guasti and B. v. Ooyen (2003) Reflexions on prosodic bootstrapping: its role for lexical and syntactic acquisition. Developmental Science. 6.2. 213-222.

Corballis, M. C. (2009) Do rats learn rules? *Animal Behavior*, 78, e1–e2

Cutler, A. (1993). Phonological cues to open- and closed-class words in the processing of spoken sentences. Journal of Psycholinguistic Research, 22, 109-131.

Cutler, A., Butterfield, S., & Williams, J. N. (1987). The perceptual integrity of syllabic onsets, *Journal of Memory and Language*, *26*, 406-418.

Cutler, A., Dahan, D., & Van Donselaar, W.A. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40*, 141-202

Cutler, A., Mehler, J., Norris, D., Segui, J. (1986). The Syllable's Differing Role in the Segmentation of French and English, *Journal of Memory and Language*, *25*, 385-400

Cutler, A., Sebastián-Gallés, N., Soler-Vilageliu, O., & van Ooijen, B. (2000). Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons. Memory and Cognition, 28, 746-755.

Crain, S., & Nakayama, M. (1987). Structure dependence in grammar formation. *Language*, 63, 522–543.

Csibra, G. (2003). Teleological and referential understanding of action in infancy. *Philosophical Transactions of the Royal Society, London B, 358*, 447-458

Dabrowska, E. (2001). Learning a morphological system without a default: the Polish genitive. Journal of Child Language, 28, 545-574.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. Behavioral Research Methods, Instruments, & Computers, 28, 125-127.

Day, R. S., & Wood, C. C. (1972). Mutual interference between two linguistic dimensions of the same stimuli, *Journal of the Acoustical Society of America*, *52*.

Dehaene-Lambertzs, G., & Baillet, S. (1998) A phonological representation in the infant brain. *NeuroReport*, 9:1885—1888.

Dehaene-Lambertz, G., & Dehaene, S. (1994). Speed and cerebral correlates of syllable discrimination in infants. *Nature*, 370, 292-295.

Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science*, 298, 2013-2015

Demany, L., & Armand, F. (1984). The perceptual reality of tone chroma in early infancy. *Journal of the Acoustical Society of America*, 76, 57-66.

Dent, M.L., Brittan-Powell, F., Dooling, R.J., & Pierce, A. (1997). Perception of synthetic /ba/-/wa/ speech continuum by budgerigars (Melopsittacus undulatus). *Journal of the Acoustical Society of America, 102*, 1891-1897.

Deutsch, D. (1972). Octave generalization and tune recognition. *Perception & Psychophysics*, 11, 411-412.

Dewar, K. & Xu, F. (2007) Do 9-month-old infants expect distinct words to refer to kinds? *Developmental Psychology, 43*, 1227-1238.

Donegan, P., & Stampe, D. (1983). Rhythm and the holistic organization of language structure. Papers from the Annual Regional Meeting, Chicago Linguistic Society: Parasession on the Interplay of Phonology, Morphology and Syntax, 337–353.

Dupoux, E., Pallier, C., Sebastian, N., & Mehler, J. (1997). A destressing "deafness" in French? *Journal of Memory and Language*, 36, 406-421

Dupoux, E., Peperkamp, S., & Sebastian (2001). A robust method to study stress 'deafness'. *Journal of the Acoustical Society of America*, 110, 1606-1618

Dupoux, E., Peperkamp, S., & Sebastian-Galles, N. (2010). Limits on bilingualism revisited: Stress 'deafness' simultaneous French-Spanish bilinguals. *Cognition*, 114, 266-275.

Dupoux, E., Sebastian-Galles, N. Navarete, E., & Peperkamp, S. (2008). Persistent stress `deafness': the case of French learners of Spanish. *Cognition*, 106,682-706

Dyer, F. C., & Dickinson, J. A. (1994). Development of sun compensation by honeybees: How partially experienced bees estimate the sun's course. *Proceedings of the National Academy of Sciences, USA, 91*, 4471-4474.

Ehret, G., & Haack, B. (1981). Categorical perception of mouse pup ultrasounds by lactating females. *Naturwissenshaften, 68*, 208.

Elbert, S. H., & Pukui, M. K. (1979), Hawaiian Grammar, Honolulu: The University Press of Hawaii

Eimas, P. D., Tartter, V. C., & Miller, J. L. (1981). Dependency relations during the processing of speech, In P. D. Eimas & J. L. Miller (Eds.), *Perspectives on the Study of Speech*. Hillsdale, NJ: Erlbaum.

Eimas, P. D., Tartter, V. C., Miller, J. L., & Keuthen, N. J. (1978). Asymetric dependencies in processing phonetic features, *Perception & Psychophysics*, *23*, 12–20.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179–211.

Elman, J. L. (1993). Learning and development in neural networks: the importance of starting small. *Cognition*, 48, 71-99.

Emlen, S. (1975). The stellar orientation system of a migratory bird. *Scientific American*, 233, 102-111.

Endress, A.D. & Bonatti, L.L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition*, 105, 247-299.

Endress, A. D., Dehaene-Lambertz, G., & Mehler, J. (2007) Perceptual Constraints and the Learnability of Simple Grammars. Cognition, 105, 577-614

Endress, A. D., & Mehler, J. (2009). The surprising power of statistical learning: When fragment knowledge leads to false memories od unheard words. Journal of Memory and Language, 60, 351-367.

Endress, A. D., Nespor, M., & Mehler, J. (2009) Perceptual and Memory Constraints on Language Acquisition. Trends in Cognitive Sciences, 13, 348-353

Endress, A. D., Scholl, B. J., & Mehler, J. (2005). The role of salience in the extraction of algebraic rules. Journal of Experimental Psychology. General, 134, 406-419.

Faber, A., 1992. Phonemic segmentation as epiphenomenon: evidence from the history of alphabetic writing. In: Downing, P., Lima, S., Noonan, M. (Eds.), *The Linguistics of Literacy*. John Benjamins, Amsterdam, pp. 111–134.

Fenk-Oczlon, G., & Fenk, A. (2005). Crosslinguistic correlations between size of syllables, number of cases, and adposition order. In Gertraud Fenk-Oczlon & Christian Winkler (eds.), Sprache und Natürlichkeit: Gedenkband für Willi Mayerthaler, 75–86. Tübingen: Gunther Narr.

Fenson, L., Dale, P., Reznick, J., Bates, E., Thal, D., Pethick, J. (1994). Variability in early communication development. Monographs of the Society for Research in Child Development, 59.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, **20**, 104–113.

Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development*, **64**, 637–656.

Fischer, J. (1998). Barbary macaques categorize shrill barks into two call types. *Animal Behaviour, 55*, 799-807.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. Proceedings of the National Academy of Sciences of the United States of America, 99, 15822-15826.

Fitch, W.T., & Hauser, M.D. (2004). Computational Constraints on Syntactic Processing in a Nonhuman Primate. Science, 303, 377

Fitch, W.T., Hauser, M.D., & Chomsky, N., (2005). The evolution of the language faculty: Clarifications and implications. Cognition, 97, 179-210

Fulkerson, A. L., & Waxman, S. R. (2007). Words (but not tones) facilitate object categorization: evidence from 6- and 12-month-olds. *Cognition*, 105, 218-228.

Gallistel, C. R. (2000) The replacement of general-purpose learning models with adaptively specialized learning modules, in M.S. Gazzaniga, Ed. *The Cognitive Neurosciences*. 2d ed. (1179-1191) Cambridge, MA. MIT Press.

Gallistel, R., Brown, A., Carey, S., Gelman, R., & Keil, F. (1991). Lessons from animal learning for the study of cognitive development. In S. Carey & R. Gelman (eds.), *The Epigenesis of Mind: Essays in Biology and Cognition.* Hillsdale, NJ: Erlbaum, 3-36.

Garcia, J., Hankins, W.G., & Rusiniak, K.W. (1974). Behavioral regulation of the milieu interne in man and rat. Science, 185, 824-831.

Gardner, R.A., Gardner, B.T., & Van Cantfort, E. (1989). Teaching sign language to chimpanzees. Albany: State University of New York Press.

Gardner, T. J., Naef, F., & Nottebohm, F. (2005). Freedom and rules: the acquisition and reprogramming of a bird's learned song. *Science* 308(5724):1046–49

Gerken, L. A., Jusczyk, P. W., & Mandel, D. R. (1994). When prosody fails to cue syntactic structure: nine-month-olds' sensitivity to phonological vs. syntactic phrases. Cognition, 51, 237-265.

Gervain, J., Nespor, M., Mazuka, R., Horie, R., & Mehler, J. (2008). Bootstrapping word order in prelexical infants: a Japanese-Italian cross-linguistic study, Cognitive Psychology, 57, 56-74.

Gervain, J., Sebastián-Gallés, N., Díaz, B., Laka, I., Mazuka, R., Yamane, N., Nespor, M., & Mehler, J. (under review). Word Frequency Bootstraps Word Order: Cross-Linguistic Evidence. *Cognitive Science*

Geschwind, N., & Levitsky, W. (1968). Human brain: Left-right asymmetries in temporal speech region, *Science*, 161, 186-187.

Ghim, H-R., & Eimas, P. D. (1988). Global and local processing by 3- and 4-month-old infants. *Perception & Psychophysics*, 43, 165-171.

Gil, D. (1986). A prosodic typology of language. Folia Linguistica, 20, 165–231.

Gillette, J., Gleitman, H., Gleitman, L.R., & Lederer, A. (1999). Human simulations of vocabulary learning. Cognition, 73, 135-176

Gleitman, L.R. (1994). Words, words, words… Philosophical Transactions of the Royal Society of London. Series B, 346, 71-77.

Gleitman, L.R., & Gleitman, H. (1992). A picture is worth a thousand words, but that's the problem: the role of syntax in vocabulary acquisition. Current Directions in Psychological Science, 1, 31-35.

Gleitman, L. R., & Wanner, E. (1982). The state of the state of the art. in Language acquisition: The state of the art, eds Wanner E, Gleitman LR (Cambridge University Press, Cambridge, UK), pp 3-48.

Goldsmith, J. A. (1990). *Autosegmental and metrical phonology*. Cambridge, MA: Basil Blackwell.

Golinkoff, R. M., Church Jacquet, R., Hirsh-Pasek, K., & Nandakumar, R. (1996). Lexical principles may underlie the learning of verbs. Child Development, 67, 3101-3119.

Golinkoff, R., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language*, 21, 125–155.

Gomez, R. L. (2002) Variability and Detection of Invariant Structure. *Psychological Science*, 13, 431-436.

Gómez, R. L., & Gerken, L. A. (1999). Artificial grammar learning by one-year-olds leads to specific and abstract knowledge. Cognition, 70, 109–135

Goodman, N. (1955). *Fact, fiction and forecast*. Cambridge : Harvard University Press.

Gout, A., Christophe, A. & Morgan, J. (2004). Phonological phrase boundaries constrain lexical access: II. Infant data. *Journal of Memory and Language, 51,* 547-567

Gouzoules, S., Gouzoules, H., & Marler, P. (1984). Rhesus monkey (*Macaca mulatta*) screams: representational signaling in the recruitment of agonistic aid. *Animal Behaviour, 32*, 182-193.

Graf Estes, K., Evans, J. L., Alibali, M., & Saffran, J. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. Psychological Science, 18, 254-260.

Greenberg, J. (1963) *Universals of Language*. Cambridge: MIT Press.

Halberda, J. (2003). The development of a word-learning strategy. Cognition, 87, B23–B34.

Halle, M., & Vergnaud, J. R. (1988). *An Essay on Stress*. Cambridge, Mass.: MIT Press.

Hauser, M.D., Chomsky, N., & Fitch W.T. (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? Science, 298, 1569

Hauser, M.D., Newport, E.L., & Aslin, R.N. (2001). Segmentation of the speech stream in a nonhuman primate: Statistical learning in cotton top tamarins. *Cognition, 78,* B53-B64.

Havy, M. (2009). Spécificité phonétique et acquisition lexicale en situation de développement normal et atypique. Thèse de Doctorat en Psychologie Cognitive, Université Paris Descartes, soutenance publique: 21 décembre 2009.

Havy, M., Bertoncini, J., & Nazzi, T. (submitted) Phonetic processing during acquisition of new words in 3-to-6-year-old French speaking deaf children, equipped with cochlear implants

Havy, M., & Nazzi, T. (2009). Better processing of consonantal over vocalic information in word learning at 16 months of age. Infancy, 14, 439-456.

Hayes, B. (1995). *Metrical Stress Theory: Principles and Case Studies*.Chicago: The University of Chicago Press.

Hebb, D. O. (1949). *The organization of behaviour*. New York : JohnWiley & Sons.

Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. Child Development, 58, 1021-1034.

Henson, R.N. (2001) Serial order in short-term memory. *Psychologist*, 14, 70–73

Hirsh-Pasek, K., Golinkoff, R. M., & Hollich, G. (2000). An emergentist coalition model for word learning: Mapping words to objects is a product of the interaction of multiple cues. In R. M. Golinkoff, K. Hirsh-Pasek, L. Bloom, L. B. Smith, A. L. Woodard, N. Akhtar, M. Tomasello, & G. Hollich (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 179 – 186). New York: Oxford University Press.

Hochmann, J-R., Azadpour, M., & Mehler, J. (2008). Do humans really learn A^nB^n artificial grammars from exemplars?, *Cognitive Science*, 32, 1021-1036

Hochmann, J-R., Endress, A. D., & Mehler J. (2010). Word frequency as a cue for identifying function words in infancy. Cognition, 115, 444-457.

Hofsten, C., Dahlstrom, E., & Fredrikson, Y. (2005). 12-month-old infants' perception of attention direction in static video images. Infancy, 8, 217–231.

Hubel, D., & Wiesel, T. (1970). The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *Journal of Physiology*, 206, 419-36.

Hulst, H. van der (1984). *Syllable Structure and Stress in Dutch*. Dordrecht: Foris.

Humboldt, W. von. (1836/1999), *On Language*. Translated by Peter Heath and edited by Michael Losonsky. (Cambridge: Cambridge University Press).

Hume, D. (1739/2003). *A treatise of human nature*. Project Gutenberg.

Itô, J. and A. Mester (1986), The Phonology of Voicing in Japanese, Linguistic Inquiry, 17, 49-73.

Jackendoff, R., & Pinker, S. (2005). The nature of the language faculty and its implications for evolution of language (Reply to Fitch, Hauser, and Chomsky). Cognition, 97, 211-225.

Jakobson R (1960). Why mama and papa? In Perspectives in Psychological Theory, Essays in Honor of Heinz Werner, New York, 124-134.

Johnson, S. P., Fernandes, K. J., Frank, M. C., Kirkham, N. Z., Marcus, G. F., Rabagliati, H., & Slemmer, J. A. (2009). Abstract rule learning for visual sequences in 8- and 11-month-olds. Infancy, 14, 2-18.

Jusczyk, P. W. (1997). The discovery of spoken language. Cambridge, MA: MIT Press.

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. Child Development, 64, 675-687.

Kaminski, J., Call, J. & Fischer, J. (2004) Word learning in a domestic dog: evidence for fast mapping. *Science* 304, 1682-1683.

Keidel, J. S., Jenison, R. L., Kluender, K. R., & Seidenberg, M. S. (2007). Does grammar constrain statistical learning? Commentary on Bonatti et al. (2005). Psychological Science, 18, 922-923.

Kelly, Quinn, Slater, Lee, Ge, & Pascalis (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing, *Psychological Science*, 18, 1084-1089.

Kemler Nelson, D. G., Hirsh-Pasek, K., Jusczyk, P. W., & Wright-Cassidy, K. (1989). How prosodic cues in motherese might assis language learning. Journal of Child Language, 16, 55-68.

Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology,* 15**,** 166–171.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: evidence for a domain general learning mechanism, Cognition 83, B35–B42

Kiparsky, P. (1993). Paninian Linguistics. Encyclopedia of Languages and Linguistics.

Kluender, K.R., Diehl, R.L., & Killeen, P.R. (1987). Japanese quail can learn phonetic categories. *Science, 237*, 1195-1197.

Knobel, M., & Caramazza, A. (2007). Evaluating computational models in cognitive neuropsychology: The case from the consonant/vowel distinction. Brain and Language, 100, 95-100.

Kouider, S., Halberda, J., Wood, J., & Carey, S. (2006). Acquisition of English number marking: the singular-plural distinction. Language Learning and Development, 2, 1-25.

Kovács, Á. M. (2008). Learning two languages simultaneously: mechanisms recruited for dealing with a mixed linguistic input. SISSA, Trieste. PhD. dissertation.

Kovács, Á. M., & Mehler, J. (2009a). Cognitive gains in 7-month-old bilingual infants. Proceedings of the National Academy of Sciences, 106(2), 6556-6560.

Kovács, Á. M., & Mehler, J. (2009b). Flexible learning of multiple speech structures in bilingual infants. Science, 325, 611-612.

Kuhl, P. (1989). On babies, birds, modules, and mechanisms: A comparative approach to the acquisition of vocal communication. In R.J. Dooling & S.H. Hulse (Eds.), *The comparative psychology of audition.* (pp. 379-422). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. Perception and Psychophysics, 50, 93-107.

Kuhl, P.K., & Miller, J.D. (1975). Speech perception by the chinchilla: Voicedvoiceless distinction in alveolar plosive consonants. *Science, 190*, 69-72.

Kuhl, P.K., & Padden, D.M. (1982). Enhanced discriminability at the phonetic boundaries for the voicing feature in macaques. *Perception and Psychophysics, 32*, 542-550.

Kuhl, P. K., Williams, K. A. , Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. Science, 255, 606-608.

Landau, B., & Gleitman, L.R. (1985). Language and experience: Evidence from the blind child. Cambridge, MA: Harvard University Press.

Langus, A., Marchetto, E., Bion, A.H., & Nespor, M. (under review). Can prosody be used to discover hierarchical structure in continuous speech?

Lee, H. W., Rayner, K., & Pollastek, A. (2001). The relative contribution of consonants and vowels to word identification during reading. Journal of Memory and Language, 44, 189-205

Lehiste, I. (1970). Suprasegmentals. MIT Press, Cambridge, Mass.

Lenneberg, E. H. (1967). *Biological foundations of language*. New York : John Wiley and Sons.

Levelt, W. J. M., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary, *Cognition*, *50*, 239–269.

Liberman, A.M. (1982) On finding that speech is special. *American* Psychologist, 37, 148-167.

Liberman, A.M. (1996). *Speech: A special code*. Cambridge, MA: MIT Press.

Liberman, I. Y., Shankenweiler, D., Fisher, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child, *Journal of Experimental Child Psychology*, *18*, 201–212.

Liberman, M. Y., & Streeter, L. A. (1978). Use of nonsense-syllable mimicry in the study of prosodic phenomena, *Journal of the Acoustical Society of America*, *63*, 231–233.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised, *Cognition*, 21, 1-36.

Lidz, J., Gleitman, H., & Gleitman L.R. (2003). Understanding how input matters: verb learning and the footprint of universal grammar. Cognition, 87, 151-178

Lidz, J., Waxman, S., & Freedman, J. (2003) What Infants Know about Syntax but Couldn't have Learned: Experimental Evidence for Syntactic Structure at 18-months, *Cognition* 89, B65-B73.

Liittschwager, J.C., & Markman, E.M. (1994). Sixteen and 24-month-olds' use of mutual exclusivity as a default assumption in second label learning. Developmental Psychology, 30, 955–968.

Lindauer, M. (1957). Sonnenorientierung der Bienen unter der Aequatorsonne und zur Nachtzeit. *Naturwissenschaften,* 44, 1-6.

Linell, P., 2005. *The Written Language Bias in Linguistics: Its Nature, Origins and Transformations.* Routledge, London.

Lock, A. (1980). The guided reinvention of language. London: Academic Press.

Locke, J. (1690/2004). An essay concerning humane understanding. Project Gutenberg.

Maddieson (2008). Consonant-Vowel ratio. In: Haspelmath, Martin & Dryer, Matthew S. & Gil, David & Comrie, Bernard (eds.) The World Atlas of Language Structures Online. Munich: Max Planck Digital Library, chapter 3. Available online at http://wals.info/feature/3

Marcus, G. F. (1998). Rethinking eliminative connectionism. Cognitive Psychology, 37, 243-282.

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule-learning in seven-month-old infants. Science, 283, 77-80.

Markman, E. M. (1990). Constraints children place on word meanings. Cognitive Science 14, 57-77.

Markman, E. M. (1994). Constraints on word meaning in early language acquisition, *Lingua*, 92, 199-227.

Markman, E. M., & Abelev, M. (2004). Word learning in dogs? *Trends in Cognitive Sciences,* 8, 479-481.

Markman, E. M., & Hutchinson J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic vs. thematic relations. *Cognitive Psychology*, 16, 1-27.

Markman, E. M., Wasow, J. L., & Hansen M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology*, 47, 241-275.

Marler, P. (1991). The instinct to learn. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind* (pp. 37-66). Hillsdale, NJ: Lawrence Erlbaum Associates.

Marler, P., & Tamura, M. (1962). Song "dialects" in three populations of white-crowned sparrows, *The Condor*, 64, 368-377.

Massaro, D. W. (1974). Perceptual units in Speech Recognition. *Journal of Experimental Psychology*, *102*, 199–208.

Massaro, D. W. (1987). Categorical partition: A fuzzy-logical model of categorization behavior. In S. Harnad (Ed.), *Categorical Perception: The grounwork of cognition*, 254–283. Cambridge, UK: Cambridge UniversityPress.

May, B., Moody, D.B., & Stebbins, W.C. (1989). Categorical perception of conspecific communication sounds by Japanese macaques, Macaca fuscata. *Journal of the Acoustical Society of America, 85*, 837-847.

Maye, J., Weiss, D., & Aslin, R. N. (2008). Statistical phonetic learning in infants: facilitation and feature generalization. *Developmental Science*, 11, 122-134.

Maye, J., Werker, J. F., & Gerken, L. A. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82, B101–B111

Mazuka, R. (1996). How can a grammatical parameter be set before the first word? In J.L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 313–330). Mahwah, NJ: Lawrence Erlbaum Associates.

McCarthy, J. (1991), L'infixation réduplicative dans les langages sécrets, Langages, 101, 11-29.

McMurray, B. and Aslin, R. N. (2004). Anticipatory eye movements reveal infants' auditory and visual categories. *Infancy, 6,* 203-229.

McShane, J. (1979). The development of naming. *Linguistics*, 17, 879-905.

Mehler, J. (1981). The Role of Syllables in Speech Processing: Infant and Adult Data, *Philosophical Transactions of the Royal Society*, *B 295*, 333-352.

Mehler, J., Dommergues, J. Y., Frauenfelder, & U., Segui, J.(1981). The Syllable's Role in Speech Segmentation, *Journal of Verbal Learning and Verbal Behavior*, *20*, 298-305.

Mehler, J., & Dupoux, E. (1994). *What Infants Know*. Blackwell.

Mehler, J., Peña, M., Nespor, M., & Bonatti, L. L. (2006). The "soul" of language does not use statistics: Reflections on vowels and consonants. *Cortex*, 42, 846-854.

Meissner C. A., & Brigham J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, 7, 3–35.

Mervis, G. B. (1987). Child-basic object categories and early lexical development. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 201-233). Cambridge: Cambridge University Press.

Mervis, C. B., & Bertrand, J. (1994). Acquisition of the novel name–nameless category (N3C) principle. *Child Development*, 65, 1646–1662.

Milewski, A. E. (1976). Infants' discrimination of internal and external pattern elements. *Journal of Experimental Child Psychology*, 22, 229-246.

Miller, G. A. (1996). The Science of Words (2nd edition), W. H. Freeman.

Mintz, T. H. (2002). Category induction from distributional cues in an artificial language. *Memory & Cognition , 30*, 678-686.

Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition, 90,* 91-117.

Mintz, T. H. (2006). Finding the verbs: distributional cues to categories available to young learners. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action Meets Word: How Children Learn Verbs,* p. 31-63. New York: Oxford University Press.

Mintz, T.H., Newport, E.L. & Bever, T. G. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science, 26*, 393-424.

Monaghan, P., Chater, N., & Christiansen, M. H. (2005). The differential contribution of phonological and distributional cues in grammatical categorization. *Cognition, 96*, 143–182.

Monaghan P., Christiansen, M. H., & Chater, N. (2007). The Phonological-Distributional Coherence Hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive Psychology,* 55, 259–305.

Mondragon E, Murphy RA, Murphy VA (2009) Rats do learn XYX rules. *Animal Behavior*, 78, e3–e4

Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy Training and Speech Segmentation, *Cognition*, *24*, 45–64.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously ?, *Cognition*, *7*, 323–331.

Morgan, J. L., & Demuth, K. (Eds) (1996). Signal to syntax: Bootstrapping from speech to grammar in early acquisition. Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Mueller, J.L., Bahlmann, J. & Friederici, A.D**.** (2010). Learnability of embedded syntactic structures depends on prosodic cues. *Cognitive Science, 34,* 338-349.

Murdock, G. P. (1959). Cross language parallels in parental kin terms. Anthropological Linguistics, 1, 1-5.

Murphy, R. A., Mondragon, E., & Murphy, V. A. (2008). Rule learning by rats, *Science*, 319, 1849-1851.

Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Büchel, C., & Weiller, C. (2003). *Nature Neuroscience,* 6, 775

Naigles, L.G., & Kako, E. T. (1993). First contact in verb acquisition: Defining a role for syntax. Child Development, 64, 1665-1687.

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. Cognition, 98, 13-30.

Nazzi, T., & Bertoncini, J. (2009). Consonant specificity in onset and coda positions in early lexical acquisition. Language and Speech 52.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: towards an understanding of the role of rhythm. Journal of Experimental Psychology. Human Perception and Performance, 24, 756-766.

Nazzi, T., Floccia, C., Moquet, B., & Butler, J. (2009). Bias for consonantal over vocalic information in French- and English-learning 30-month-olds: crosslinguistic evidence in early word learning. Journal of Experimental Child Psychology, 102, 522–537.

Nazzi, T., & New, B. (2007). Beyond stop consonants: Consonantal specificity in early lexical acquisition. Cognitive Development, 22, 271-279.

Nelson, D. A., & Marler, P. (1989). Categorical perception of a natural stimulus continuum: Birdsong, *Science*, 244, 976-978

Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. Lingue & Linguaggio, 2, 201-227.

Nespor, M., Shukla, M., & Mehler, J. (in press) Stress-timed vs. syllable timed languages. Companion to Phonology.

Nespor, M., Shukla, M., van de Vijver, R., Avesani, C., Schraudolf, H., & Donati, C. (2008). Different phrasal prominence realization in VO and OV languages. Lingue e Linguaggio. 7, 1-28.

Nespor, M., & Vogel, I. (1986). Prosodic Phonology. Berlin: Mouton de Gruyter. 2ndt edition, 2008. Dordrecht. Foris.

New, B., Araujo, V., & Nazzi, T. (2008). Differential processing of consonants and vowels in lexical access through reading. Psychological Science, 19, 1223-1227

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. Cognitive Psychology, 48, 127-162.

Nottebohm F. 1993. The search for neural mechanisms that define the sensitive period for song learning in birds. *Netherlands Journal of Zoology*, 43, 193–234.

Oller, D. K. (1980). The emergence of the sounds of speech in infancy. In G. H. Yeni-Komshian, J. F. Kavanagh, & C. A. Ferguson (Eds.), *Child phonology, volume 1: Production* (pp. 93-112). New York: Academic Press.

Pallier, C. (1994). Rôle de la syllable dans la perception de la parole: Etudes attentionnelles. Thèse de Doctorat de Sciences Cognitives de l'EHESS nouveau régime soutenue le 12 décembre 1994. http://www.pallier.org/papers/Pallier_phdthesis.pdf

Pascalis, O., de Schonen, S., Morton, J., Deruelle, C., & Fabre-Grenet, M. (1995). Mother's face recognition by neonates: a replication and an extension. *Infant Behavior and Development*, 18, 79-85.

Pavlov, I. P. (1927). Conditioned reflexes : An investigation of the physiological activity of the cerebral cortex (G. V. Anrep, Ed.). London : Oxford University Press. (http ://psychclassics.yorku.ca/Pavlov/)

Pearl, L. (2007). Necessary bias in natural language learning, Ph.D. Dissertation, University of Maryland, College Park, http://hdl.handle.net/1903/6984

Pearl, L. (2009). Learning English Metrical Phonology: When Probability Distributions Are Not Enough, *Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America (GALANA 2008)*, ed. Jean Crawford et al., 200-211. Somerville, MA: Cascadilla Proceedings Project.

Pearl, L., & Lidz, J. (2009). When domain general learning fails and when it succeeds: Identifying the contribution of domain specificity*, Language Learning and Development, 5*, 235-265.

Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-Driven Computations in Speech Processing, Science, 298, 604-607.

Peña, M., Maki, A., Kovačić, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., Mehler, J. (2003) Sounds and Silence: An optical topography study of language recognition at birth. *Proceedings of the National Academy of Science,* 100**,** 11702-5

Peperkamp, S. & Dupoux, E. (2002). A typological study of stress 'deafness'. In: C. Gussenhoven & N. Warner (eds.) *Laboratory Phonology 7.* Berlin: Mouton de Gruyter

Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. F., Dupoux, E., Fazio, F., Mehler, J. (1996) Brain processing of native and foreign languages. *NeuroReport*, 7, 2439-2444

Pinker, S. (1984). Language Learnability and Language Development. Cambridge, MA: MIT Press

Pinker, S. (1994). The language instinct. New York: William Morrow.

Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? Cognition, 95, 201–236.

Pisoni, D. B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics*, 13, 253-260.

Plato (360 BCE). Sophist, translated by F. M. Cornford. in *Plato: the collected dialogues*, eds Hamilton E, Cairns H (Princeton University Press, Princeton, N.J.) , p 998.

Pons, F., & Toro, J.M. (2010). Structural generalizations over consonants and vowels in 11-month-old infants. *Cognition,* 116*,* 361-367.

Port, R., 2006. The graphical basis of phones and phonemes. In: Munro, M., Bohn, O.-S. (Eds.), *Second Language Speech Learning: The Role of Language Experience in Speech Production and Perception*. John Benjamins, Amsterdam, pp. 349–365.

Port R (2007) How are words stored in memory? Beyond phones and phonemes. *New Ideas in Psychology* 25,143–170.

Port R (2008) All is prosody: Phones and phonemes are the ghosts of letters. Proceedings of the 4th International Conference on Speech Prosody, 7–16.

Port, R.F., & Leary, A., (2005). Against formal phonology. *Language* 81, 927–964.

Premack, D. (1976). *Intelligence in Ape and Man*. Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey.

Prunet, J. F., Beland, R., & Adrissi, A. (2000). The mental representation of Semitic words, *Linguistic Inquiry*, 31.

Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

Redington, M. & Chater, N. (1997). Probabilistic and distributional approaches to language acquisition. Trends in Cognitive Sciences, 1, 273-281.

Regier, T. & Gahl, S. (2004). Learning the unlearnable: The role of missing evidence. *Cognition 93,*147-155.

Ramus, F., Hauser, M. D., Miller, C. T., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and cotton-top tamarin monkeys. *Science*, 288, 349-351.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73, 265-292.

Rodriguez, P. (2001). Simple recurrent networks learn context-free and context-sensitive languages by counting. *Neural Computation*, 13, 2093-118.

Rodriguez, P., Wiles, J. & Elman, J. L. (1999). A recurrent neural networkthat learns to count. *Connection Science*, 11(1), 5–40.

Rohde, D. & Plaut, D. (1999). Language acquisition in the absence of explicit negative evidence : how important is starting small ? *Cognition*, 72, 67-109**.**

Rosenblatt, F. (1957). The perceptron : A perceiving and recognizing automaton (project para) (Tech. Rep. No. 85-460-1). Ithaca, New York : Cornell Aeronautical Laboratory.

Rosenblatt, F. (1958). The perceptron : a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65, 386-408.

Rumbaugh, D. M. (Ed.). (1977). *Language learning by a chimpanzee: The LAN A project.* New York: Academic Press.

Rumelhart, D. E., Hinton, G. E. & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323, 533–536.

Rumelhart, D. E. & McClelland, J. L. (1986). On learning the past tenses of english verbs. In J. L. McClelland, D. E. Rumelhart & The PDP Research Group (Eds.), *Parallel distributed processing : Explorations in the microstructure of cognition* (Vol. 2 : Psychological and Biological Models, pp. 216–71). Cambridge, MA : MIT Press.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926-1928.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by adults and infants. *Cognition,* 70, 27-52.

Saffran, J. R., Pollak, S. D., Seibel, R. L., & Shkolnik A. (2007). Dog is a dog is a dog: infant rule learning is not specific to language. *Cognition*, 105:669–680.

Saffran, J. R., & Wilson, D. P. (2003). From syllables to syntax: Multi-level statistical learning by 12-month-old infants. Infancy, 4, 273-284.

Saussure, F. de (1916). *Cours de linguistique générale*. Paris : Payot (ed. by Charles Bally & Albert Sechehaye in collaboration with Albert Riedlinger).

Savage, C., Lieven, E., Theakston, A., & Tomasello, M. (2003). Testing the abstractness of young children's linguistic representations: Lexical and structural priming of syntactic constructions? *Developmental Science*, *6-5*, 557-567.

Segui, J. (1984). The syllable: A basic perceptual unit in speech perception? In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and Performance X*, 165–181. Hillsdale, NJ: Erlbaum

Segui, J., Dupoux, E., & Mehler, J. (1990). The role of the syllable in speech segmentation,phoneme identification, and lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing: psycholinguistic and computational perspectives*, chap. 12, pp. 263–280. Cambridge, Mass.: MIT Press.

Seidenberg, M. S. & McClelland, J. L. (1989). A Distributed, Developmental Model of Word Recognition and Naming. *Psychological Review, 96,* 523-568

Selkirk, E.O. (1984) *Phonology and Syntax. The Relation between Sound and Structure.* Cambridge. MA. MIT Press.

Seyfarth, R.M., Cheney, D.L., & Marler, P. (1980). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science, 210*, 801-803.

Shafer, V. L., Shucard, D. W., Shucard, J. L., & Gerken, L. (1998). An electrophysiological study of infants' sensitivity to the sound patterns of English speech. *Journal of Speech, Language and Hearing Research*, *41*, 874–886.

Shankweiler, D., & Studdert-Kennedy, M. (1967). Identification of consonants and vowels presented to left and right ears. *Quarterly Journal of Experimental Psychology,* 19**,** 59–63.

Shi, R., & Lepage, M. (2008). The effect of functional morphemes on word segmentation in preverbal infants.*Developmental Science*. *11* (3), 407-413

Shi, R., Morgan, J., & Allopenna, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: a cross-linguistic perspective. *Journal of Child Language*, 25, 169-201.

Shi, R., Werker, J., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy*, *10*, 187–198.

Shi, R., Werker, J., & Morgan, J.L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, *72*, B11–B21.

Shukla, M., & Nespor, M. (2010). Rhythmic patterns cue word order. In N. Erteschik-Shir (ed.) *The Sound Patterns of Syntax*. Oxford. Oxford University Press. pp. 174-188.

Shukla, M., Nespor, M., & Mehler, J. (2007). An interaction between prosody and statistics in the segmentation of fluent speech. *Cognitive Psychology*, 54, 1-32.

Simion, F., Regolin, L., Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the National Academy of Sciences*, 105, 809-813

Skinner, B. F (1938). *The behavior of organisms: An experimental analysis*. New York: Appleton-Century-Crofts.

Skinner, B. F (1957) *Verbal Behavior*. Appleton-Century-Crofts, New York.

Smith, L., & Yu, C. (2007). Infants rapidly learn words from noisy data via cross-situational statistics. In D. S. McNamara & J. G. Trafton (Eds.) *Proceedings of the 29th Annual Cognitive Science Society* (pp. 653-658). Austin, TX: Cognitive Science Society.

Snowdon, C.T. (1987). A naturalistic view of categorical perception. In S. Harnad (Ed.), *Categorical Perception.* (pp. 332-354). Cambridge, UK: Cambridge University Press.

Soja, N., Carey, S., & Spelke, E. (1985). Constraints on word learning. Paper presented at the 1985 Biennial Convention of the Society for Research in Child Development, Toronto, Canada.

Spelke, E. S. (1990). Principles of object perception. *Cognitive Science, 14*, 29-56.

Spelke, E. S. (1994). Initial knowledge: Six suggestions. *Cognition, 50*, 431-445.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review, 99*, 605-632.

Spelke, E. S., Katz, G., Purcell, S. E., Ehrlich, S. M., & Breinlinger, K. (1994). Early knowledge of object motion: Continuity and inertia. *Cognition, 51*, 131-176.

Sperber, D., & Wilson, D. (2004). Relevance Theory. in G. Ward and L. Horn (eds) *Handbook of Pragmatics*. Oxford: Blackwell, 607-632.

Spinoza, B. (1677). Compendium Grammatices Linguae Hebraeae French Translation Abrégé de Grammaire Hébraïque. Paris: Librairie Philosophique J. Vrin, 1968.

Staal, J. F. (1962). A method of linguistic description: the order of consonants according to Pānini. *Language,* 38, 1-10.

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word learning tasks. *Nature*, *388*, 381-382.

Stevens, K. N. (2002). Toward a model for lexical access based on acoustic landmarks and distinctive features. *Journal of the Acoustical Society of America,* 111, 1872–1891.

Stoel-Gammon, C. (1983). Constraints on consonant-vowel sequences in early words, *Journal of Child Language*, 10, 455-457.

Swingley, D. (2005). Statistical clustering and the contents of the infant vocabulary. *Cognitive Psychology*, 50, 86-132.

Swingley, D. (2007). Lexical exposure and word-form encoding in 1.5-year-olds. *Developmental Psychology*, *43*, 454–464.

Terrace, H.S. (1979). *Nim.* New York: Knopf.

Thothathiri, M., & Snedeker, M. (2008). Syntactic priming during language comprehension in three- and four-year-old children. *Journal of Memory and Language*, *58*, 188-213.

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, *74*, 209-253.

Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.

Tomiak, G. R., Mullenix, J. W., & Sawusch, J. R. (1987). Integral perception of phonemes: Evidence for a phonetic mode of perception, *Journal of the Acoustical Society of America*, *81*, 755–764.

Toro, J. M., Nespor, M., Mehler, J., & Bonatti, L. L. (2008a). Finding words and rules in a speech stream: Functional differences between vowels and consonants. *Psychological Science*, 19, 137-144.

Toro, J. M., Shukla, M., Nespor, M., & Endress, A. D. (2008b). The quest for generalizations over consonants: Asymmetries between consonants and vowels are not the by-product of acoustic differences. *Perception and Psychophysics*, 70, 1515-1525

Toro, J.M., & Trobalón, J.B. (2005). Statistical computations over a speech stream in a rodent. *Perception & Psychophysics,* 67, 867-875.

Treiman, R., & Breaux, A. M. (1982). Common phoneme and overall similarity relations among spoken syllabes: their use by children and adults, *Journal of Psycholinguistic Research*, *11*, 581–610.

Tremoulet, P. D., Leslie, A. M., & Hall, D. G. (2000). Infant individuation and identification of objects. *Cognitive Development*, 15, 499-522.

Trommelen, M. (1983). *The Syllable in Dutch.* Dodrecht: Foris.

Tyrrell, D., Stauffer, L., & Snowman, L.(1991). Perception of abstract identity/difference relationships by infants. *Infant Behavior and Development*, 14, 125-129.

Tyler, A. A., & Langsdale, T. E. (1996). Consonant-Vowel interactions in early phonological development. *First Language*, 16, 159-191.

Tyrrell, D., Zingaro, M., & Minard, K.(1993). Learning and transfer of identity – difference relationships by infants. *Infant behavior and Development*, 16, 43-52.

van Ooyen, B., Bertoncini, J., Sansavini, A., Mehler, J. (1997). Do weak syllables count for newborns? *Journal of the Acoustical Society of America, 102,* 3735-3741

Vickery, T. J., & Jiang, Y. V. (2009). Associative grouping: perceptual grouping of shapes by association. *Attention, Perception & Psychophysics*, 71, 896-909.

Vouloumanos, A. (2008). Fine-grained sensitivity to statistical information in adult word learning. *Cognition*, *107*, 729–742.

Vouloumanos, A., Hauser, M. D., Werker, J. F., & Martin, A. (2010). The tuning of human neonates' preference for speech. *Child Development*, *81*, 517-527.

Vouloumanos, A., & Werker, J. F. (2004). Tuned to the signal: The privileged status of speech for young infants. *Developmental Science, 7,* 270-276

Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental Science, 10,* 159-164.

Waxman, S. R. (1999). Specifying the scope of 13-month-olds' expectations for novel words. *Cognition*, 70, B35–B50.

Waxman, S.R., & Booth, A.E. (2001). Seeing pink elephants: Fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology*, 43*,* 217-242.

Waxman, S.R., & Booth, A.E. (2003). The origins and evolution of links between word learning and conceptual organization: new evidence from 11-month-olds. *Developmental Science*, *6*, 128–135

Waxman, S. R., & Braun, I. E. (2005). Consistent (but not variable) names as invitations to form object categories: new evidence from 12-month-old infants. *Cognition*, 95, B59–B68.

Waxman, S. R., Senghas, A., & Benveniste, S. (1997). A cross-linguistic examination of the noun-category bias: its existence and specificity in French- and Spanish-speaking preschool- aged children, *Cognitive Psychology*, 32, 183-218.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49-63.

Wittgenstein, L. (1953). *Philosophical investigations*. Oxford : Blackwell.

Wood, C. C., & Day, R. S. (1975). Failure of selective attention to phonetic segments in consonant-vowel syllables, *Perception & Psychophysics*, *17*, 346–350.

Wright, A., Rivera, J., Hulse, S., Shyan, M., & Neiworth, J. (2000). Music perception and octave generalization in rhesus monkeys. *Journal of Experimental Psychology: General, 129,* 291–307.

Wyttenbach, R.A., May, M.L., & Hoy, R.R. (1996). Categorical perception of sound frequencies by crickets. *Science, 273*, 1542-1544.

Xu, F. (2002). The role of language in acquiring object kind concepts in infancy. *Cognition, 85*, 223-250

Xu, F., Carey, S., & Quint, N. (2004). The emergence of kind-based object individuation in infancy. *Cognitive Psychology*, 49, 155-190.

Yang, C. D. (2002). *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press.

Yang, C. D. (2004). Universal Grammar, statistics or both? *Trends in Cognitive Sciences*, 8, 451-456.

Yoshida, K. A., Pons, F., Cady, J. C., & Werker, J. F., (2006). Distributional Learning and Attention in Phonological Development, *Paper presented at the annual meeting of the XVth Biennial International Conference on Infant Studies, Westin Miyako, Kyoto, Japan*

Yu, C., & Smith, L. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*, 414-420

Zuberbuhler, K., Noe, R., & Seyfarth, R.M. (1997). Diana monkey long-distance calls: messages for conspecifics and predators. *Animal Behaviour, 53*, 589-604.