Scuola Internazionale Superiore di Studi Avanzati
International School for Advanced Studies

# Investigating Two Domain-General Processes in Early Infancy: Disjunctive Inference and Reorientation of Attention

Candidate:

Milad Ekramnia

Supervisor:

Jacques Mehler

Thesis submitted for the degree of Doctor of Philosophy in Cognitive Neuroscience
Trieste, 2016

# Abstract

A characteristic of the more evolved nervous systems is the ability to process information in an abstract amodal domain. The existence of this capability, necessitates the presence of mental processes that are amodal and therefore, can act on a broad range of internal and external stimuli. Investigating the early development of the interaction between the amodal mental processes and their domain of action on mental representations, can shed light on the extents of the computations that can be accommodated by these processes.
In this thesis through a series of eye-tracking studies in pre-verbal infants, we attempted to investigate the early development of some of these interactions from two different domains.

In one domain, we addressed if logical operators, as a subset of the mental processes, are available to pre-verbal infants; so they can be utilized in combining and assessing the several mental images involved in an inference process. For this purpose, we introduced a face-voice association paradigm, in which infants could potentially use disjunctive inference to disambiguate the context and make the right face-voice pairings. We showed that the performance of the 10-month-old infants suggests that they might be able to perform this association through the process of disjunctive inference based on the elimination of the incorrect alternative. We furthermore, used the pupillometry data and results from an adult control group to suggest a time-frame for the steps of this process.

In another domain, we studied the integration of abstract visual icons with attentional shift. In one hand we showed that arrows can trigger an attentional shift in the 4-month-old infants but not 8-

month-olds. We further showed that this reorientation of attention might be due to the triangular area of the icon. These striking results, although should await further confirmations, suggest an early sensitivity to the features of these icons, which can trigger a top-down reorientation of attention (as we tried to eliminate the possibility of a bottom-up process). A sensitivity that possibly disappears later in the development.

On the other hand, we showed that 8-month-olds and not 4-month-olds can assign an attentional shift to an arbitrary icon in a very few number of trials. These results together suggest a mixed picture for attribution of attentional shift to the icons; however indicating that a volitional attribution of attention to arbitrary icons can be carried out by infants as young as 8 months of age.

# Acknowledgments

When Jacques kindly accepted me as a member of his laboratory, with a background in Physics and Leech Electrophysiology, the only thing I brought with myself to the lab, was merely a curiosity in the field of developmental psychology. Now after three years, thanks to Jacques and the wonderful members of the lab, I'm finishing this thesis with a deeper and wider passion in infant studies. I awe the greatest thanks to Jacques for introducing me to the field, growing my curiosity, and helping me to persuade my interests even though, they were out of the focus of the lab; and more importantly sharing with me precious insights on how to conduct a proper study and how to learn from the failures. His precious remarks would always stay with me in my career. I also awe great thanks to Marina Nespor, for all her precious support and detailed comments during all these years.

I'm specially grateful to Bahia Guellai, a former member of the lab and a wonderful friend; the paradigm used in chapter two, stemmed directly from the long conversations we had together. I'd like to warmly thank Hanna Marno, Alissa Ferry, Amanda Saksida and Erno Teglas, for their continuous support, insightful comments and all the fruitful conversations. Ana Flo, Yliana Rodriguez, Alan Angus, Perrine Brusini, Yamil Vidal Dos Santos, and Michela Bonfieni for their patience in helping me with elaborating the ideas and correcting my several ignorances, and moreover, for all the fun we shared together out of the lab. Furthermore, these studies could not have been done without the continuous help of Francesca Gandolfo and Marijana Sjekloća, for arranging the appointments with the parents and the very smooth organization in the laboratory.

I'm greatly thankful to other professors of our institute, John Nicholls, Alessandro Treves, Daniele Amati, Mathew Diamond and my ex-supervisors Vincent Torre and Mojtaba Alae, for the numerous insightful conversations we had together. I also awe special thanks to three of my close friends, Arash Fassihi, Nader Nikbakht and Sina Tafazzoli, for all the interesting insights they shared with me from Systems Neuroscience.

Moreover, I'd like to thank my wonderful friends, Brinduzha Chituleanu, Fahimeh Bafti Zadeh, Sophi Rosay, Diletta Pozzi, RaeAnn Wiseman, Georgette Argiris, Julia Franzoi, Caterina Marino, Ben von Harling, Mahdi Eshaghi, along with numerous others who kept my esteem high and turned these years into a joyful and memorable journey.

Last but not least, I'd like to thank my parents and my sister, Mina Ekramnia, for their continuous support and motivation in all these years.

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (*FP7/2007-2013)* / ERC *grant agreement* n° 269502 (*PASCAL*).

# Chapter 1. General Motivation

## 1.1. Representation of Concepts

Concepts are considered to be the building blocks or basic elements of thought. In cognitive psychology, the word *concept* is frequently used to refer both to concrete entities (for instance the concept of dog) and to a subset of mental processes (such as the concept of logical negation).

Stepping aside from the debates on what entities can be included in the category of concepts, a valid and related question can be how the concepts are represented in the brain and how they interact with each other. Philosophers and cognitive psychologists have proposed several frameworks in this regard; here we briefly explain two of these frameworks that have been addressed in the past decades (see Margolis & Laurence, 2014; and Margolis, 1998, for a discussion).

### 1. Concepts as Mental Representations

Pioneered by Locke (1690/1975) and matured by Fodor (1975), concepts can be considered as a subset of mental representations that are based on the mental states toward propositions (also referred to as *propositional attitudes*; for instance *believing* in the validity of a proposition is one attitude or a *mental state* toward that proposition). This view received a widespread support from cognitive psychologists (Fodor, 1987) and constitutes the core of *Language of Thought hypothesis*, a theory that considers mental representations as tokens and describes mental processes as operations on these tokens (Fodor, 1975).

## 2. Concepts as Abilities

The theory of concepts as abilities, holds that concepts are in fact the ability to discriminate the referents of a concept from the referents of the adjacent concepts. For instance the concept of dog, is considered to be the ability to discriminate dogs from non-dog entities. This ability further accommodates the species to draw certain inferences about the referents of that concept, based on the common features of all the known referents; for instance the concept of dog may accommodate an inference such that all the animals that are included in this category are having four legs (Kenny, 2010).

Although more empirical evidence is needed to assess these two perspectives, but one can speculate that these two views are not mutually exclusive and both can contribute to the formation and maintenance of a concept. Moreover, one can consider an interaction between these two aspects of concepts, for instance a mental representation holding a concept may facilitate the abilities of the species in discriminating between the referents and non-referents of that concept; on the other hand, the abilities that accommodate a concept may further fine-tune the mental representation that corresponds to that concept.

## 1.2. Abstract Entities

Regardless of the framework we choose to describe concept representation, the concepts can be classified into concrete and abstract. According to a classical categorization that was first pioneered by Plato in his proposal in distinguishing between “Forms and Sensibles”, concrete concepts have a reference to a physical or a concrete item in the world but the abstract concepts lack such an existence outside of the mind. However, there is not a unanimously agreed boundary for defining abstract concepts. Here instead of using the term *abstract concept,* more generally we can use the term *abstract entities* to avoid presupposing that they are necessarily concepts.

The term *abstract* in the literature is used to refer to several different entities. It may refer to the amodal representation of a concrete entity such as an amodal representation of a character (Quiroga, 2012), domain general rules and functions that guide the species in their thought processes and inferences (*Wallis & Miller, 2001*), lexical items and abstract icons that can be mapped to any mental representation (Waxman & Gelman, 2009), etc.

Electrophyisological studies provide a strong support for abstract entities in a neuronal level. There is a battery of evidence mainly from non-human animals that various representations that are referred to as abstract can be observed in the response patterns of a single neuron. For instance it has been reported that single neurons represent the concept of nest in rats (Lin, Chen, Kuang, Wang & Tsien, 2007) or the concept of a specific actress in humans (Quiroga, 2012), or newly learned rules by chimpanzees (Wallis et al. 2001, Freedman, Riesenhuber, Poggio & Miller, 2001) or abstract representation of numbers in crowns (Ditz and Neider, 2015).

Regarding humans, there have been several studies to explore the early development of abstract processes/representations in infancy; to name a few, infants within the first year of life, show a preferential look toward any face-like cartoon (Goren, Sarty & Wu, 1975), show a differentiation between solid and non-solid substances (Hespos, Ferry & Rips, 2009), show multi-modal understanding of numerosity (Izard, Sann, Spelke & Streri, 2009), show an ability to discriminate between patterns and associating rules to different patterns (Marcus, Vijayan, Rao & Vishton, 1999; Kovacs & Mehler, 2009) and so on (see Carey, 2009, for a comprehensive discussion).

These studies suggest that infants from the first year of life, enjoy a rich repertoire of abstract processes and furthermore, they have the ability to acquire new rules through abstraction of the instances.

In this thesis, we attempted to further explore our understanding on the domain of action of the abstract processes, by investigating their interaction with abstract and non-abstract mental representations. This understanding can play a key role in explaining the complex capabilities that infants exhibit, such as conceptualization processes and deductive inferences.

## 1.3. Main Framework

In a series of studies we sought to focus on abstract functions as a subset of abstract entities through exploring their early integration with the mental representations (here by *function* we are referring to a kit of mental processes that are in service of carrying a well-defined action). Abstract functions are particularly interesting as the domain of action of a single function can extend to several modalities, with inputs ranging from perceptual to conceptual mental representations.

In these studies we decided to avoid assessing functions at their intersect with the domain of language. This will help us to investigate the basic dynamics and domain of activity of functions before they integrate with language-related processes.

To this end, in one hand we investigated if infants can use logical operators in combining and assessing propositions (in this thesis we refer to propositions as falsifiable mental images that construct the possible outcomes of an inference process, furthermore, we do not try to draw a line between operators and functions and we use the two terms interchangeably). On the other hand, we focused on reorientation of attention, as a fundamental mental process, and studied its integration with abstract icons. Here we briefly describe the two approaches:

**Approach 1.** In this approach, we chose to investigate the capability of infants in using logical operators. In this regard, we focused on disjunctive inference (for instance, 'Sue is either at home or at school, Sue is not at home, so she is at school'). Disjunctive inference is a favorable process, as it combines two propositions and a cue, and validates one of the propositions through eliminating a negated alternative (the cue).

To assess if this inference is available to the pre-verbal infants or if they can integrate it with propositions, we introduced an ambiguity to the infants in a face-voice association context (i.e. either face 1 is female or face 2 is female; given that face 1 is male, so it can not be female; therefore face 2 is female), and we asked if infants can apply negation and disjunctive inference to disambiguate the situation. The ambiguity implied two possible alternatives as two propositions that can be translated to two mental images. In this example the logical operator embedded in disjunctive inference can be viewed as a set of mental processes that inputs the two mental images and correspondingly assigns a validity to them based on the cue.

**Approach 2.** Furthermore, to better understand how mental processes can integrate with abstract tokens, we investigated the capability of pre-verbal infants in associating a simple function, i.e. reorientation of attention, with abstract visual icons. In this case the icons would be tokens that can signal a reorientation of attention. Such an icon can then go through a process of conceptualization and turn into a symbol of attentional shift.

We chose to explore this integration by two classes of icons:

**Feature-Relevant Icons.** We first decided to characterize to what extent an attentional shift integrates with an icon that has features that can facilitate this association. In our studies, we used arrow for this purpose; as its features (a rectangle with a triangular head) may inherently play a role in the integration of the icon with an attentional shift, and may consequently facilitate the process of conceptualization of the icon as a symbol for attentional cuing.

**Feature-Irrelevant Icons.** Infants naturally go through the process of assigning functions to abstract items in the context of word learning. In this study we tried to characterize if the mechanism of assigning functions to abstract tokens is available to the young infants in a visual domain. For this purpose we studied the learning trends of associating an attentional shift to arbitrary visual icons in two age groups of 4 months and 8 months.

These two lines of studies, although they address two very different functions, they can potentially extend our understanding on the extents that the pre-verbal infants can naturally combine mental processes with various mental representations, in the lack of the facilitations that are provided by language related processes and ostention.

The first approach will be addressed in chapter 2 and the second approach will be addressed in chapter 3. At the introduction of each chapter I will give a more comprehensive overview on the literature and provide the motivation behind the studies.

# Chapter 2. Disjunctive Inference in a Face-Voice Context

## 2.1. Introduction

### 2.1.1. Negation and Disjunctive Syllogism

To begin with, we first define some terms in syllogism that we will frequently use in the text.

#### 2.1.1.1. Defining the Terms

**Negation** is generally referred to as being the complement or the opposite of an affirmative statement. Moreover, negation is considered to be less informative than an affirmative statement. E**xample:** 'Paris isn't the capital of Spain' is less informative than 'Paris is the capital of France'.

**Modus Ponens** refer to propositions such as:

If *A* then *B*, given *A*, then one would infer *B*.

**Example:** If it's cloudy, it rains. It's cloudy, therefore it rains.

**Modus Tollens** refer to the inference rules in the following form:

If *A* then *B*, not *B*, then not *A*.

**Example:** If it's cloudy, it rains. It's not raining, so it's not cloudy.

**Disjunction** is a compound of two propositions that are connected by *OR*. The *OR* connective can be inclusive or exclusive. In inclusive O*R* both propositions can be valid at the same time, however, in exclusive *OR* only one of the two propositions can be valid.

**Example:** Maria may have tea *OR* may have coffee.

In this example, if *OR* is inclusive, both of the propositions can be valid, Maria can have either tea or coffee or both; however, if *OR* is exclusive, only one of the two propositions can be valid (either Maria has tee or she has coffee).

Here the two propositions can potentially be represented as two possible mental representations.

**Disjunctive Inference,** also referred to as the process of elimination, refers to inferences in which disjunction is the first premise and negation is the second premise.

The general form of Disjunctive inference can be written as:

If $A$ *OR* $B \rightarrow$ True

$\sim (B \rightarrow$ True)

Then $A \rightarrow$ True

**Example:** Maria went to restaurant or she went home;

Maria did not go to restaurant.

So Maria went home.

In this example, the outcome of the inference is the same for both inclusive and exclusive *OR*.

Therefore in order to follow a disjunctive inference we need to have an understanding of disjunction, so constructing the two mental representations for the two propositions that the disjunction is based on (Namely *A* and *B* in the example above). Furthermore we need to be able to follow the negation to exclude one of the two mental representations and finally infer that the other one is the correct alternative.

#### 2.1.1.2. Context Dependency of Deductive Inferences

Development of human deductive reasoning has been a subject of extensive research since Jean Piaget proposed his influential framework (Piaget, Cook & Norton, 1954; Piaget, 1953). Piaget suggested that deductive inference can be observed in early childhood at its earliest. He proposed children as young as 5 to 7 years old can make deductive inferences based on concrete concepts; and deductive inferences based on abstract concepts appear at a later age around 11-12 years. He referred to these two stages as familiar and abstract stages of deductive reasoning and termed them as concrete operational stage and formal operational stage respectively. The inferences in familiar contexts are also referred to as pragmatic inferences.

According to Piaget's theory, adults should be able to perform equally good in both concrete and abstract contexts; however, Wason (1968) in his famous card selection task showed that the majority of the adult participants failed in applying Modus Tollens when the presented context was symbolic (thematic), but in a pragmatic context the majority of participants succeeded (Wason & Shapiro, 1971; Thompson, 2000).

Similar results have been reported in children as young as 2.5 years of age (Bowerman, 1986; Kuhn, 1977; Cheng & Holyoak, 1985), especially the authors reported a higher performance in the inferences that were presented in a pragmatic framework based on valid knowledge-based arguments (also referred to as pragmatic schema), for instance, '*If you are older than 18 years old, you can drink*', '*can you drink?*'. These results suggest that logical operators and inference processes in general, may integrate better with mental representations that are referring to concrete events rather than to arbitrary abstract tokens.

Regarding the early integration of logical operators with mental representations, we will focus on disjunctive inference and negation as an integral part of it in the concrete contexts. These two inference processes are at the core of formal logic; however, their early development in pre-verbal infants is very poorly understood. In the following two sections, we provide a brief review on the literature of development of negation and disjunctive inference.

#### 2.1.1.3. Development of Negation

In the context of negation there have been several studies trying to track the earliest production of negative sentences by toddlers in their natural language.

In a classic work in 1970, Bloom registered the conversations of 3 children starting from 19 months of age, this led her to introduce three types of negations in English language. A*bsence* (there is no pen), *rejection* (I don't want anymore) and *inhibition* (no swimming). She further proposed that the production time-line of these three classes of negation in children is as follows: *absence*, *rejection* and *inhibition* (Bloom, 1970; Bloom, 1993; Cameron-Faulkner, Lieven & Theakston, 2007). Choi (1987) later fine-tuned this classification by including 6 additional classes of negation therefor 9 classes in total: non-existence, prohibition, rejection, failure, denial, inability, epistemic negation, normative and inferential negation. She found all the 9 categories in the utterances of English-speaking, French-speaking and Korean-speaking children, so considered them as the universal sub-types of negation (Choi, 1987).

The battery of studies on negation has focused mainly on the production of negation words. It's been shown that the first instance of negation is with utterance of *No* and *Not* that occurs as early as 12 months of age, mainly to refer to non existence and inhibition; and denial emerges later around 19-23 months of age (Pea 1978); however, elaboration of the production continues until 4 years of age. Based on this general pattern of development Klima and Bellugi (1966) introduced 3 stages for the development of negation; in stage 1, children are able to meaningfully utter *No* and *Not*; however, the use of these negators would produce wrong phrases when toddlers want to signal denial; in stage 2, use of *Don't* and *Can't* would become part of the negation utterances;

however, these negators are still single units since *Do* or *Can* are still not attested at this stage, and in the 3rd stage, infants are able to readily make negation with *'nt* units at the end of the verbs and using *Didn't*, *Isn't*, *Won't*, etc. In the literature on negation, denial is considered to be the most cognitively demanding type of negation compared to absence or rejection, and appears to be the last to emerge among the three main subtypes. An explanation for this observation is that negation as absence or negation as rejection refers to the violation of an expectation for an immediate object or an immediate desire (Bloom, 1970; Choi,1988; Pea, 1978; Tam & Stokes, 2001). However, in case of negation as denial, we need to construct a parallel mental representation of the affirmative clause of the given proposition, before actually negating it. For example to comprehend a proposition like '*there is no eagle in the sky'*, we first construct the counter-factual form of it, that would be *'there is an eagle in the sky*', and then negating it (Kaup, 2007).

Regarding the comprehension of negation words, Nordmeyer and Frank (2013) addressed the saliency of the sentences that are set to be negated, as a possible factor that undermines the performance of toddlers in looking at the alternative choice, when the alternative choice is less salient or less attractive. For instance, when toddlers at the age of 2 and 3 years, were asked to look at the boy with no apples, they performed better when the image of the boy with apples was contrasted with an image of a boy with toys, rather than an image of a boy with empty hands. This effect however, dropped significantly in the 3-year-olds.

In a recent study, Austin, Theakston, Lieven, and Tomasello (2014), addressed the comprehension of negation as denial, when it is conveyed either as a gesture, a word or a sentence. They showed that in a two bucket paradigm, when the experimenter was signaling the participant that the toy is not in this bucket (one of the buckets out of two), infants were not able to follow the denial word or denial sentence until their 2nd birthday. But by 2 years and 4 months they showed an ability to also comprehend a denial gesture. Hummer, Wimmer, and Antes (1993), proposed approximately the same age for comprehension of denial, to be at 1.8 to 2.1 years, by asking yes/no questions from the toddlers. For instance by asking '*is this a dog?'* while pointing to photo of a cat.

These studies are based on verbal comprehension of negation. The observed failures in infants, however, does not necessarily point to the lack of the concept of negation, but it can also refer to the lack of an appropriate mapping between the words related to negation such as *no* and *not* and the more abstract mental operation that support negation; or even if the mapping is already established, the saliency of the negated mental representation may outcompete the saliency of the alternative mental representation; both these factors can contribute to the failure of toddlers, even if the concepts of different types of negation are potentially available to them.

#### 2.1.1.4. Development of Disjunctive Inference

To Address the availability of the negation operator in the computational resources of the infants, the alternative explanations for the failure of toddlers that we mentioned in the previous section, signify the importance of designing tasks that do not rely on verbal comprehension of negation. Alternatively we can indirectly prompt infants to utilize negation in an ambiguous context, where they have a natural tendency to disambiguate the situation. Some examples of these contexts are finding a puppet in a bucket, word learning, or as we will further address, the context of face-voice association.

One of the inference processes that is based on negation, is the disjunctive syllogism. In this process, negation or as more generally referred to, the process of elimination, is a crucial part of the inference. Disjunctive syllogism, has been considered as a favorable inference process that can assess negation in the non-verbal contexts. Here we briefly review the studies on this process in early infancy.

The studies on disjunctive inference in a non-verbal context have mainly been focused on the invisible displacement tasks; tasks in which the participants know that a puppet is hidden in one of the several possible locations (several buckets for instance). Participants can exclude the possibilities one by one by checking each location; or similarly use the provided cue that a certain possibility is eliminated, in order to limit their possible choices (Premack & Premack, 1994; Hill, Collier-Baker & Suddendorf, 2012; Watson & Gergely, 2001, Call & Carpenter, 2001). These studies have shown that children as young as 3-4 years of ages consistently search and exclude the possibilities until they reach the hidden object. A critique to the experiments based on this type of paradigm can be if exclusion of a possible solution is performed in the context of disjunctive inference, there should be an increased level of confidence toward the available possibilities, by eliminating each one. Subsequently this higher confidence can translate to a lower reaction time in the available alternative space. Although Watson & Gergely (2001), used this argument to show that dogs could not follow disjunctive syllogism to perform the task, but on the other hand, relying on the same measure, the children did not increase their searching speed.

These results can potentially account for an alternative scenario in which the participants exhaust all the options until they find the desired object independent of refining the space of possible alternatives after each failure.

In another line of studies, disjunctive inference has been suggested as an alternative strategy for word learning in infants. In this context it has been shown that when infants as young as 17 months are provided with a familiar and a novel object, they prefer to associate the novel object to a novel label rather than a familiar label. This behavior has been observed in several studies (Bion, Borovsky & Fernald, 2013; Golinkoff, Hirsh-Pasek, Bailey & Wenger, 1992; Halberda, 2003; Horst & Samuelson, 2008; Markman & Wachtel, 1988; Carey & Bartlett, 1978).

To explain the mechanism behind associating the novel object to the novel label, by analyzing the looking pattern of infants while hearing a novel label and looking at a familiar and a novel object, Halberda (2006) suggested that infants reject the possible alternatives (in this context, the possibility that the familiar object can have a novel label) until they are left with the only possible choice (that the novel label can only be associated with the novel object). Halberda concluded that infants are using disjunctive syllogism to perform this novel to novel association. The main evidence to support his argument was the fact that infants before fixating on the novel object consistently fixated on the familiar object (a process dubbed as *double checking*). As Halberda concluded, the double checking of the wrong item arise from the process of elimination or rejection of the incongruent alternative.

In a study by Mather and Plunkett (2012), the authors in addition of providing to the infants, a familiar labeled object and a novel object without label, they also provided a familiar unlabeled object. Twenty-two-month-olds showed a preferential look only toward the novel object while they were hearing a novel label, although the familiar unlabeled object could also be a potential referent to the novel label. The observed novelty preference that is referred to as Novel-Name Nameless-Category principle (Mervis & Bertrand, 1994; Golinkoff et al. 1992) is an alternative mechanism suggested for the process of word learning, that contrary to the mechanisms that are based on disjunctive syllogism, this mechanism does not necessitate the process of elimination. Accordingly, Mather and Plunkett suggested that infants may not be using only one single mechanism for word learning (mechanisms based on disjunctive syllogism), but several processes might be involved in this context.

### 2.1.2. An Alternative Context to Assess Disjunctive Syllogism

The studies on disjunctive inference have mainly focused on the second year of life and mainly in the context of word learning or invisible displacement tasks. Here we propose a new context to explore disjunctive syllogism in which pre-verbal infants are introduced to two possible choices in an ambiguous situation. The framework we introduced has a high resemblance to an everyday situation.

Suppose you are passing behind two men; they are standing next to each other. Regarding age and other outfit features they are similar and you can't see their faces. One of them starts to talk, but as you cannot see them, there is no way of understanding which one is talking. Then one of them turns and talks with you with a voice different from the voice you just heard. So you understand that the first voice could not be from the person who is facing you but it is from the other person. So if the person who turned to talk with you, faces back again like before, and you hear the first voice again, this time you can be confident that this voice is not from the person who turned back but from the other person.

In this example, at the beginning there was an ambiguous situation, because you do not have enough evidence to assume which person was talking. When one of the men turns and talks with a different voice, you can disambiguate the situation by excluding the possibility that the first voice was from this person and consequently you infer that it could only be from the other man. This inference can be disjunctive inference. At the end, when both of the men are facing back as in the beginning, and you hear the first voice another time, based on the previous cue, you can be sure that the voice is being uttered by the other man who did not turn. Figure 2.1. shows the framework of this inference based on the context of formal logic.

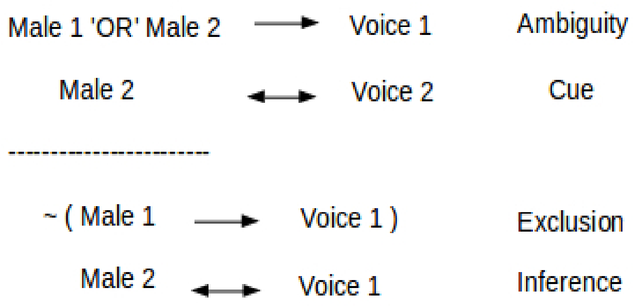

**Fig. 2.1.** The different steps of the disjunctive inference in a face-voice context. In the first premise, *'OR'* is exclusive, meaning that it cannot be the case that voice 1 could be associated with both male 1 and male 2, but it can be associated only with one of them. Furthermore it is important to note that from the moment of receiving the second premise, the *'Cue'*, it is possible to disambiguate the situation using a disjunctive inference.

In the context described above, the exclusion can take place either at the time of receiving the cue or later when the ambiguous situation is presented again.

The exclusion process can be based on a negation operation, as one out of the two possibilities is being negated or being rejected in this step. However, since we are not certain about the computational mechanism behind the exclusion step, and to avoid presupposing that the negation operation is the only possible mechanism for exclusion, we decided to use the more general term *exclusion* instead of *negation* in the rest of the chapter, and leave it for further discussions if this process should be referred to as negation.

In order to explore the disjunctive inference in pre-verbal infants, we developed a context similar to the described example, with the following constraints:

Our main goal was to design a paradigm in which pre-verbal infants could potentially use disjunctive inference to exhibit a preferential look to one side of the screen, so their preferential look could reliably be measured as an index of their performance. We furthermore wanted to avoid a familiarization phase, so we could directly assess the inference process independent of the efficiency of a familiarization block; and finally we sought to design a paradigm in which the pupil profile could reliably be registered to track the possible components of the pupil dilation that could be relevant to the different stages of inference process. To take these concerns into account, we decided to base our paradigm on the preferential look at a talking face. In the following sections, we first briefly review the literature on the face-voice association in early infancy to explain our motivation for the used stimuli.

### 2.1.3. Face Preferences

We chose to base our paradigm on the preferential look at a talking face, because there is a load of evidence suggesting infants from the first minutes of life show a preference for looking at face-like stimuli and in particular, talking faces. Here we briefly review the literature on face-voice association in infancy.

Infants since birth are exposed to a rich environment of faces. Johnson and Goren, showed that newborns not only follow faces but also they follow face-like patterns, more than scrambled faces, blank faces, or faces with non-facial features (Goren, Sarty & Wu, 1975). They showed that this preference however, declines after the first month of life. Johnson further discussed that this early preference could be due to seeking the adult con-specifics. He further argued that however, later this mere following of facial cues, evolves into a facial processing capability that conveys an identity to the observed face (Johnson, Dziurawiec, Ellis & Morton, 1991). In another study, Frank, Vul, and Johnson (2009) used animation cartoons, to analyze the looking pattern of infants at three, six and nine months of age toward the faces. In this study they showed a significant increase over age in attending the faces rather than other salient features in the scenes.

In another line of research, there is a long standing debate in explaining the preference of infants to novel or familiar faces. Since Piaget there is an understanding that infants have a general preference for familiar faces, a view that is supported by the preference of newborns and older infants toward their mothers' face (Walton & Bower, 1992; Bushneil, Sai & Mullin, 1989), but on the other hand in the habituation paradigms generally a preference for the novel face has been reported in contrast to a face that is habituated in the familiarization blocks (Tighe & Leaton, 1976; Pascalis & Schonen, 1994).

Another question in the development of face recognition in infancy is regarding gender preference. The majority of studies show that infants have a preference for the female face over a male face, a pattern that is more salient if the primary caregiver is a female rather than a male (Ramsey-Rennels & Lanlois, 2006; Quinn et al., 2008; Quinn et al., 2002). These observations are mainly explained on the bases of higher exposure of infants to female faces rather than male faces in the first months of life.

### 2.1.4. Voice Preferences

Regarding voice preferences in young infants, it has been shown that at the early months of life infants show a preference for hearing the voice of their mothers and this preference then later toward eight months, generalizes to any female voice (Standley & Madsen, 1990). However, this generalization does not apply to male voices (Decasper & Prescott, 1984).

In a simple framework it can be noted that infants' preference over different voices can be framed as the following: Mother's voice > Female voice > Male voice (Werker & McLeod, 1989). It has also been shown that there is a preference for infant-directed speech over normal speech at various ages in the first year of life (Werker & McLeod, 1989; Friedlander, 1968; Glenn & Cunningham, 1983).

### 2.1.5. Face-Voice Association

Integration of faces and voices have been explored in several studies in early infancy. It's been reported that 4- to 5-month-old infants raise expectations regarding the visual features of the auditory stimuli they are hearing (Spelke, 1976; Vouloumanos, Druhen, Hauser & Huizink, 2009). Spelke and Owsley (1979) showed that 3.5-month-olds show a preferential look at the face of their parents whose voice is being presented. In a habituation paradigm, Bahrick, Hernandez-Reif, and Flom (2005) showed that 4-month-olds can detect the violation of an arbitrary pairing of face and voice, similar results were shown by Brookes and colleagues (2001) at the age of three months. However, the ability to remember the association occurs between 4 to 6 months of age.

In another study Jordan and Brannon (2006), for the purpose of investigating the abstract number system in infancy, showed that 7-month-olds can match the number of human faces they see with the number of voices they hear at the same time, suggesting a one to one match between a face and a voice.

Furthermore in three experiments on young children at the age of four and five years, Moher, Feigenson and Halberda (2010), compared the face-voice association with word learning contexts. The authors suggested that mechanisms that are involved in face-voice association are similar to the mechanisms involved in word learning. These mechanisms include one to one mapping bias (similar to a mutual exclusivity bias) and fast mapping of faces to voices. They showed that participants could use an association established by these mechanisms as a reliable evidence, based on which they could continue inferring further face-voice associations.

Based on the mentioned literature, we decided to provide a context to the infants, at the age of 10 months, in which they find an ambiguity in locating the face that represents a female voice. We hypothesized that infants try to find the female face and to achieve this, they can apply disjunctive inference.

# Exp.1. Inference in an Ambiguous Situation

## 2.2.1. Experimental Design

We developed a context in which the participants were shown two static faces, namely face one and face two. During presentation of the two faces, we presented a female voice. Observing two faces with one voice, introduced an ambiguity to the participants in associating the voice to one of the two faces.

Next we presented one of the two faces alone, face two for instance, accompanied by a male voice. Here infants could make an association between this face and the male voice.

In the end, we presented the two faces again, accompanied with the same female voice. Here we expected the participants to infer that since face two was associated earlier with the male voice, only the other face, face one, could represent the ongoing voice (the female voice).

In the following we will provide a detailed description of the different phases of this paradigm.

### 2.2.1.1 Ambiguity Phase

We provided an ambiguous situation by presenting two static cartoon faces on the screen. The cartoon faces were different from each other and they lacked any gender related cue.

During the presentation of the two faces side by side, we presented a female voice. The participants at this phase did not have any cue which could have enabled them to assume which face was the talking face. Therefore in this phase, the participants could make two mental images of the two possible alternatives, which later would be validated based on the cue.

#### 2.2.1.2. Cue Phase

In the cue phase, we presented one of the two faces accompanied with a male voice. We will refer to this face as the cue face. We proposed that the participants would understand that the presented single face in this phase is a male, and since a face can either be a male or a female, the possibility that it could be associated with the first female voice is subsequently rejected.

#### 2.2.1.3. Test Phase

To test whether participants could apply disjunctive inference to the ambiguity phase based on the cue; at test phase we presented the same two faces again with the same female voice, identical to that of the first phase.

According to disjunctive syllogism, while participants hear the female voice, we expected them to show a preferential look not at the cue face, since that face was inferred to be a male, but at the opposite face. We will refer to this condition as Female-Male-Female condition, or FMF for short.

In addition to the FMF condition, we introduced a control condition, in which during the presentation of the cue face, instead of presenting a male voice we presented the same female voice as of the first and the third phase. We hypothesized that in the test phase, while infants were watching both faces and hearing the same female voice again, they show a preference for looking at the cue face. We refer to this condition as Female-Female-Female condition or FFF for short. We tested this condition to assure that infants could make an association between the voices and the static cartoon faces on the screen, within the time frame of the phase.

### Summarizing the Hypotheses:

In the test phase of both the FMF and FFF conditions, we expected to observe a preferential look at the female face since participants were always listening to a female voice at this phase.

In the FFF condition, we expected the preferential look to be toward the face that was presented as the cue, since the same voice was presented as in the cue phase; while in the FMF condition since the voices were different from the cue phase to the test phase, we predicted to observe a preferential look at the face other than the cue face.

### 2.2.2. Stimuli

#### 2.2.2.1. Visual Stimuli

Each participant was presented with nine trials. Therefore we designed nine pairs of faces (see Figure 2.2).

The faces were static cartoons, designed especially to satisfy three main criteria:

First, not having any gender specific features to avoid biasing the decision of participants, since the goal of the decision task was to specify which face is male and which is female solely based on the voice cues and not based on facial features.

Second, the faces had to be ambiguous in whether they were talking or silent. Therefore the facial features were designed with a mouth half open to allow this ambiguity.

Third, to try to minimize the relative attractiveness of one face over the other, as well as for the further analysis of the pupil data, the facial features and the colors were set to be identical in every set, and the only varying feature was the outer shape of the faces.

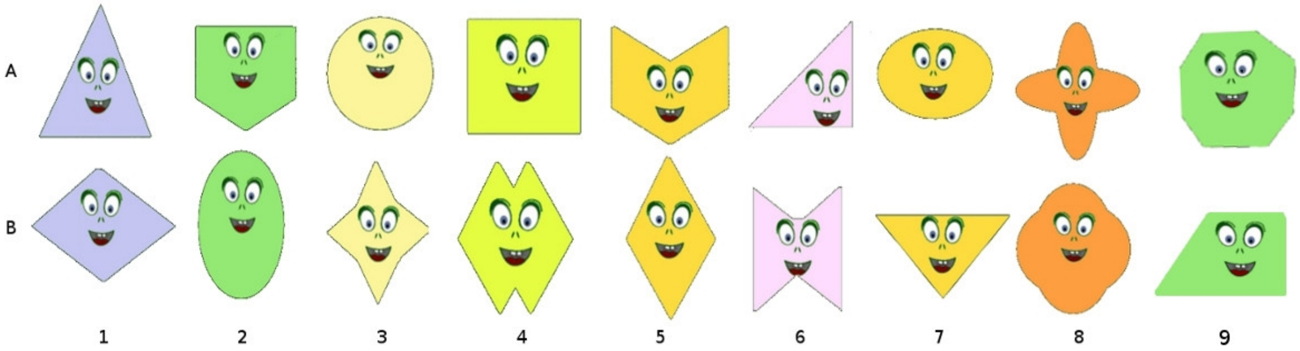

**Fig. 2.2.** The nine sets of faces used as the visual stimuli. In each trial, a different pair of faces was presented. All the faces had an identical set of facial features and they were only differing in their geometrical shape. Having mouths half open was to convey an ambiguity if the faces were silent or talking.

The faces were created by GNU Image Manipulator (GIMP) software. v.2.8.1 and they were confined to a region of 320 x 280 pixels on a screen resolution of 1280 by 1024 pixels.

In the entire time of the experiment, the background color was set to black, and the faces were centered at the mid-line of the screen. When two faces were presented, they were placed on the two sides of the screen, approximately 75 pixels from the left and right edges. When only one face was on the screen, it was centered.

In each of the three phases of a trial, the faces were presented on the screen for seven seconds, with a delay of approximately 150 ms at the transitions in between the phases.

### 2.2.2.2. Audio Stimuli

The voices that were accompanying the faces, were uttering one sentence pronounced by adult native Italian speakers. Each speaker read the same sentence: “Lungo il fiume su andiamo, tutto insieme saltelliamo”. The voices were all infant directed (Werker & McLeod, 1989) and each voice was presented only once during the whole session of the experiment to avoid any effect due to memorization of the face correspondence of a voice. Depending on how fast the voices were read by the narrator, the duration varied between 3.7 to 4.7 seconds.

The recorded voices were normalized in their amplitude to 65 db using Praat software (v. 5.0.42).

### 2.2.3. Participants

For each participant we tested only one condition. Therefore, participants were divided into two groups, the FMF and the FFF groups.

#### 2.2.3.1. FMF Condition

Thirty-eight healthy monolingual Italian infants were tested in this condition, from 41 to 47 weeks, with a mean of 44.5 weeks, SD = 1.65 weeks.

Five infants were excluded due to lack of attention and not providing sufficient number of trials.

#### 2.2.3.2. FFF Condition

Thirty-seven healthy monolingual Italian infants were tested in this condition, from 40 to 48 weeks, with a mean of 43.5 weeks, SD = 1.5 weeks.

Five subjects were excluded due to fuzziness and therefore not providing sufficient number of trials. Two more subjects were excluded due to crying at the beginning of the experiment.

All the infants had an APGAR number over 7 out of 10 and they were declared by their parents to be full term. The participants were recruited from Trieste, Italy, through sending invitation letters to a random selection of parents, whose babies fit the age range of the study. The parents were all reimbursed for attending the experiment and the infants received a certificate of attendance.

### 2.2.4. Apparatus

The visual stimuli were presented on a 15” monitor, equipped with a Tobii eye-tracker T120, that recorded the location of gaze at 60 Hz.

The participants sat on a fixed chair, in the lap of their parents, approximately 70 cm from the monitor, with their eye-sight aligned with the center of the monitor.

Parents were asked to wear opaque sunglasses, so they could not interfere with the performance of the infant and we could further be sure that the eye tracker was only registering the eyes of the infant.

The experimental room was completely darkened, so that the only light was emitted by the screen. The audio was presented in stereo from two loud speakers on the left and the right sides of the monitor. The audio was always identical from both of the loud speakers.

### 2.2.5. Procedure

Before the experiment, the subjects passed a five point calibration protocol, included in the Tobii studio. The calibration consisted of fixating on five points, four at the corners and one at the center of the screen. A yellow duckling attractor was shown on the screen to attract attention if it was needed.

After the calibration, infants watched a short animation for a duration of six seconds, we will refer to it as the task initial animation. This animation meant to orient the attention of infants on the screen and allow a sufficient time for the pupil adaptation to the light condition of the room.

The animation was a bell, shaking at the center of the screen on a black background.

After this task initial animation, the trials started. Nine trials were presented in total. Each trial began with a 2250 ms of a shaking bell animation, accompanied by the three phases. The 2250 ms animation at the beginning of each trial was used to let pupil reach its baseline (see Figure 2.3).

**Fig. 2.3.** The time line of the task initial animation followed by the first trial. Other than the task initial animation, a 2250 ms animation was the starting point of each trial, after this short animation at the beginning of each trial, the three phases were proceeding, with approx 150 ms delay in between.

In the first phase, two faces were presented on the sides of the screen. A female voice started after two seconds of silence.

In the second phase only one of the two faces were presented at the center of the screen. After 1500 ms of silence, infants heard either a male voice (in the FMF condition) or the same female voice (in the FFF condition).

In the third phase, the same two faces as in phase one were shown again, with the same female voice, starting after two seconds of silence. This phase was similar to phase one, with the only difference that the two faces were always counterbalanced in side to avoid possible effects due to side bias (see Figure 2.4).

The total time of a trial was 23.250 seconds. The periods of silence before the voice onset, were used to measure the base lines for the pupillometry analysis.

Furthermore if the participants were distracted from attending the screen, at the transition between two trials a few seconds of a Pixar Animation *'For the Birds'* was shown to them, to redirect their attention to the screen.

In order to have enough trials in each condition per subject, we presented only one condition to each participant.

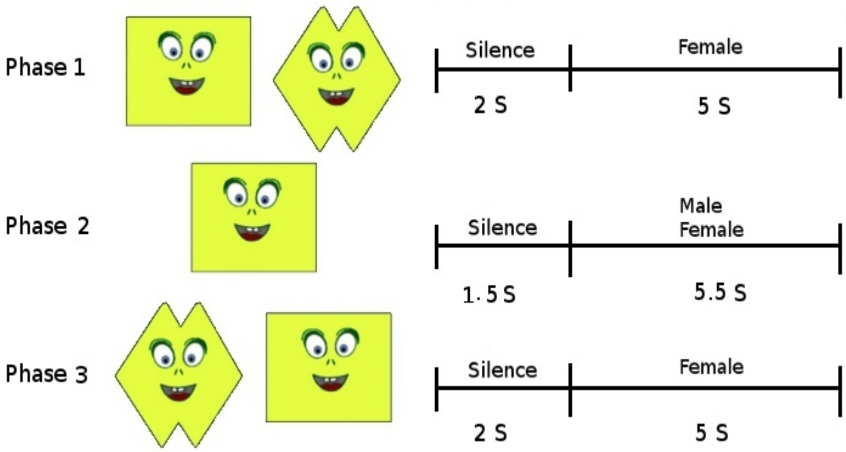

**Fig. 2.4.** Time-line of the three phases of a single trial. The voices and the faces presented in each trial occurred only once during the experiment. In both of the conditions, the voices in the first phase and the third phase were an identical female voice; in the second phase, the same female voice was presented in the FFF condition, and a male voice was presented in the FMF condition. All the voices started after a period of silence, which was meant for measuring the baselines for pupillometry purposes.

Based on our hypothesis, in the test phase (phase three) of the FFF condition, we predicted that we would observe a tendency for looking at the cue face, since the voices at the cue phase and at the test phase were the same. However, in the FMF condition, we expected to find a tendency for looking at the face other than the cue, since the two voices in the cue phase and the test phase were different.

### 2.2.6. Scoring

The coordinates of the gaze, the gaze quality and the pupil diameter of the two eyes at each gaze, were collected for further analyses.

All the analyses were done using the Mathworks Matlab software (v.2015b), Statistics Toolbox™ and Machine Learning Toolbox™ of Matlab.

One trial out of nine in the FMF condition, was excluded from the analysis, because the content of the female voice was not matching the rest of the voices. So we decided to avoid any possible perturbation induced in the pupil dilation analysis. However, exclusion of this trial did not change the behavioral data. In the FFF group, to have the same number of trials as in the FMF group, we similarly eliminated the corresponding trial.

#### 2.2.6.1. Inclusion Criteria

##### 2.2.6.1.1. RoI Settings

We divided the screen into three Regions of Interest (RoI) to tag the location of the gaze points. The stimuli always appeared on the mid-line of the vertical span of the screen. We further set an upper and lower threshold for the area of the stimuli, to eliminate the gaze points that were not vertically adjacent to the faces. We considered the gazes with a $y$ coordinate, within the upper or the lower 20% of the vertical axis as invalid (see Figure 2.5).

In the phases one and three, there were two faces on the screen; therefore, the left and right RoIs were considered to be valid, and in phase two that contained only one face at the center, only the central RoI was considered to be valid.

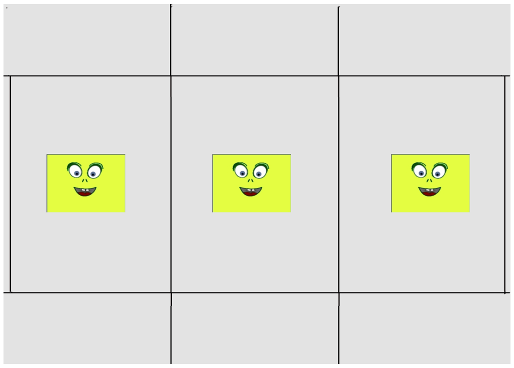

**Fig. 2.5.** The borders of the three regions of interest (RoI), with a sample cartoon face to demonstrate the relative scales. In the phases one and three of each trial, the left and the right RoIs were considered to be valid, and in phase two, only the central RoI was considered to be valid. The background color in this figure is set to gray just for demonstrative purposes, the background color in the experiments was black.

#### 2.2.6.1.2. Valid Look

We defined Valid Look as the ratio of time the participant looked inside the valid RoI(s).

Valid Look = Time looked in a valid RoI / Total duration of the phase

Since the sampling rate of the eye-tracker was at 60 Hz, we measured time in the units of **time bin,** that is equivalent to 16.67 ms, representing a single gaze point registered by the eye tracker. A gaze point was regarded as valid, if it was inside an RoI and having a quality score registered by the eye-tracker between 0-1 out of 4 (4 is equivalent to poor quality). The quality score was assigned to each point automatically by the eye-tracker if the data of both eyes was available and consistent.

#### 2.2.6.1.3. Side Bias

It was necessary for the subjects to consider the two faces presented in the first phase, as two possible alternatives, and looking at the two faces at the first phase, was considered as an essential part of the paradigm. Hence to assure that infants looked at both of the faces in the first phase, we set a threshold of 15% of total look in the valid RoIs, for the minimum ratio of look in the RoI of each face in phase one.

We furthermore measured the side bias over the course of trials, by defining the left bias as following:

Left Bias = Overall look at the left RoI / (Overall look at the Left RoI + Overall look at the Right RoI)

We set a threshold of 15% for the bias threshold. Therefore if the overall Left Bias was more than 85% or less than 15%, the corresponding participant was excluded.

No infant from the FMF group was excluded due to side bias, and from the FFF group one infant was excluded.

### 2.2.6.2. Target Look

To measure the performance, we defined the Target Look as the ratio of the valid looking time that the infant looked at the correct face. We can formulate this as following:

Target Look = Number of time bins within the correct RoI / Total number of valid time bins

A time bin, was considered valid if its corresponding gaze coordinates were falling inside the left or the right RoI. The correct choice in the FMF condition was the face other than the cue face, and in the FFF condition was the cue face.

#### 2.2.6.3. Trial Validation and Subject Validation

We set the following criteria to validate a trial:

1. The participant had to pass the threshold we considered for Valid Look in all the three phases.
2. The threshold of Valid Look, for the phases 1 and 3 was considered 50% and for phase two, 80%. The choice of 80% was to assure a full attention to the cue phase (phase 2), in order to avoid artifacts in the pupil analysis.

In supplementary material 1, we show that the measured performance did not depend on the values of the Valid Look in any of the phases of a trial.

A subject was eliminated from the analysis if it provided less than three valid trials.

#### 2.2.6.4. Trimming of the Time Span

The duration of the third phase in all the trials was seven seconds, and the female voices started after two seconds of silence. However, the durations of the female voices were different from each other, lasting from 3.7 to 4.7 seconds. Therefore, for instance in a trial where the female voice had a duration of 3.7 seconds, after the voice offset, there were 1.3 seconds of silence to the end of the phase. This short span of silence after end of the voice, could induce a disengagement of the participants. Therefore for the analysis, the time span in all the trials, was trimmed to 3.7 seconds, to limit the scanning duration to a period in which in all the trials there was an ongoing female voice.

As a results, the duration to scan the Target Look was from 2 to 5.7 seconds from the phase onset

(onset of the faces), that corresponds to the voice onset till end of the shortest voice.

### 2.2.7. Results

After applying the thresholds on the data, 34 participants were included for the FMF condition, providing 135 trials and 29 infants were included for the FFF condition, providing 120 trials.

#### 2.2.7.1. Overall Results

We calculated the Target Look averaged over subjects. In the FMF group, the analysis resulted in, Mean $_{FMF}$ = 0.59 (chance level at 0.5), SE = 0.03, and in the FFF group, Mean $_{FFF}$ = 0.54, SE = 0.03.

To compare the measures of Target Look across conditions, we first tested if the obtained Target Looks followed a normal distribution. The Lilli test for normality (a test to examine if the data is following a normal distribution), failed to reject the null hypothesis of normality of the populations of the Target Look, with p-values of 0.5 and 0.5 for the FMF and the FFF groups (0.5 is the largest tabulated number the test can provide).

We then compared the ratio of looking at the cue face across the two conditions, (note that the ratio of look at the cue face in the FFF condition is equivalent to the Target Look and in the FMF condition is the incorrect face and equivalent to (1 – Target Look), see Figure 2.6). A two-way *t*test comparing this measure across the two conditions resulted in t(61) = 2.67, p = 0.0095.

Moreover, we compared the Target Look of the two groups versus the chance level (0.5), a one-way *t*test resulted in t(33) = 2.82, p = 0.0082 in the FMF group, and t(28) = 0.7,  p = 0.48 in the FFF group.

The significant trend of looking at the Target Face in the FMF condition, was in line with our hypothesis, suggesting that infants potentially made an association between the female voice and the Target Face, by successfully excluding the cue face as an invalid alternative, as it was previously associated with the male voice.

However, the ratio of looking at the cue in the FFF group, contrary to our hypothesis, did not show any overall significance.

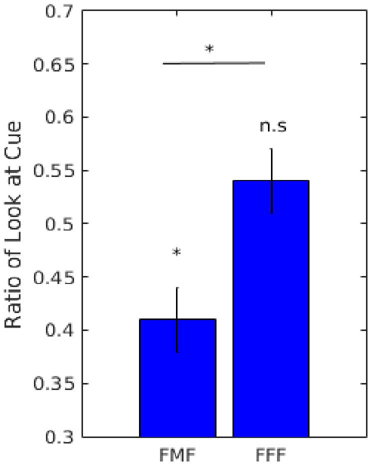

**Fig. 2.6.** Comparison of the FMF versus FFF group. The figure shows the overall ratio of looking at the cue face, averaged across subjects. Chance level is at 0.5. The errorbars show one standard error from the mean. In the FMF group, the ratio of looking at the cue face (opposite to the target face), was significantly below chance, t(33) = -2.82, p = 0.0082; while in the FFF group the ratio of looking was above chance, however, not significant, t(28) = 0.7, p = 0.48.

### 2.2.7.2. Dynamics of Target Look

The duration of the test phase was 5.7 seconds. This span of time was long enough to let us to explore the dynamics of the Target Look as a function of time. The temporal dynamics of the Target Look could potentially provide us information on the timing of the inference process involved in the test phase (Nordmeyer & Frank, 2013, Halberda, 2006).

For investigating the dynamics of looking pattern as a function of time, the time span was divided into bins of 200 ms, and the gaze distribution within each time bin was averaged over valid trials across subjects. The time onset was set to 300 ms from the onset of the face stimuli at the third phase (see Figure 2.7). In the FMF group, the dynamics of the Target Look suggests that within the first two seconds of the phase, there is a preferential look at the target face (Mean = 0.57, SE = 0.03), a trend that is amplified after the voice onset and then briefly reaches the chance level. However, as can be seen in Fig. 2.7, in the FFF group, looking at the correct face did not last for the whole duration of the phase.

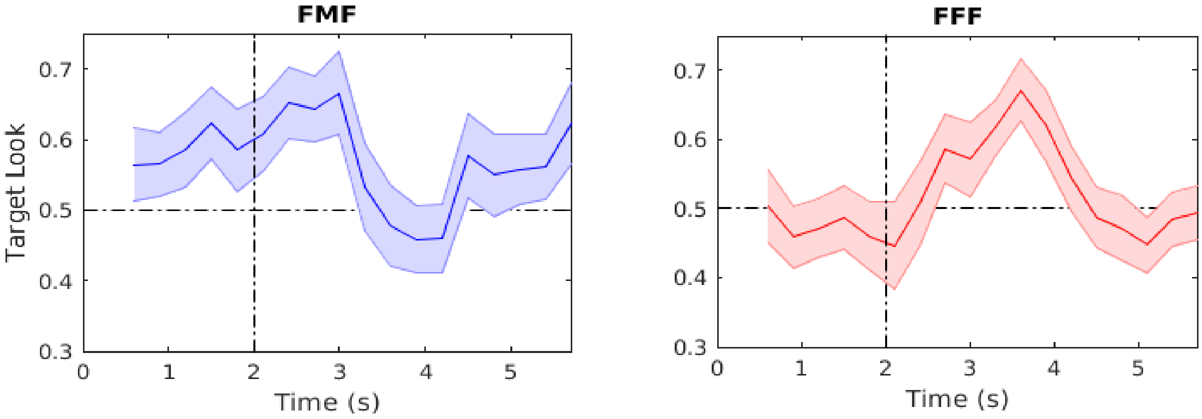

**Fig. 2.7.** Target Look over time across both conditions. The figures depict the Target Look over time, averaged over subjects in 200 ms bins. Onset of analyses was at 300 ms from face onset. The vertical dashed line shows the voice onset at second two. The horizontal line shows the chance level at 0.5. The shaded area is one standard error from the mean at each time bin.

In the FFF condition, a more detailed analysis, revealed that the mean Target Look was significantly above chance from 3.2 seconds after the face onset to the end of the trial (see Figure 2.7). Mean = 0.58, SE = 0.034. We ran a one-way *t*test comparing the Target Look average over subjects within this period versus the chance level, the test resulted in t(28) = 2.38, p = 0.024. However, from voice onset up to 3.2 seconds (1.2 seconds from the face-onset), there was not a significant tendency to look at any direction, Mean = 0.49, SE = 0.045, t(28) = -0.038, p = 0.96.

Since we applied our statistical test twice in this analysis, we used Bonferroni multiple comparison correction to determine the significance of the tests. The correction resulted in declaring the Target Look from 3.2 s to the end of the test phase (5.7 s) being significant and the Target Look from 2 s to 3.2 s not being significant.

Therefore, although in the FFF condition, the overall Target Look was not significantly different from the chance level, the Target Look from 1.2 seconds from the voice onset to the end of the phase, was proven to be significant toward the correct face.

#### 2.2.7.3. Performance Based on Initial Target Look

The temporal dynamics of Target Look in the FMF condition in Fig. 2.7, suggested that within the first two seconds of the test phase, there was a tendency to look at the Target face ($\text{Mean}_{\text{FMF - Initial target Look}} = 0.57$, $SE = 0.03$). We asked if the temporal dynamics observed after the voice onset of the FMF condition was only depending on the initial look within the first 2 seconds, or alternatively even if the infants were looking at the incorrect face prior to the voice, at the voice-onset, they changed their looking profile toward the correct face.

To address this question, we performed a demonstrative analysis, taking all the trials that had at least 3 valid time bins prior to the voice onset (the first 2 seconds, for this consideration some of the trials that were included in the overall analyses were not included in this analysis, since they did not have enough initial gaze points).

We then divided the trials based on their average Target Look within the first 2 seconds, and distributed them into two categories. One category with an average initial Target Look below 50% (46 trials) and a category of trials with an average initial Target Look over 50% (50 trials). The Mean Target Look of the two categories of trials, averaged over subjects was as follow: $\text{Mean}_{\text{FMF – Higher}} = 0.56$, $SE = 0.05$ and $\text{Mean}_{\text{FMF – Lower}} = 0.58$, $SE = 0.02$.

We applied the same analysis for the FFF group as well (63 trials were included in the below 50% category and 41 trials were included in the higher 50% category, with an overall mean of initial

Target Look equal to Mean $_{\text{FFF - Initial target Look}}$ = 0.47, SE = 0.025. The mean Target Look of the two categories averaged over subjects were: Mean $_{\text{FFF – Higher}}$ = 0.55, SE = 0.063 and Mean $_{\text{FFF – Lower}}$ = 0.55, SE = 0.04).

Figure 2.8, shows the temporal dynamics of Target Look in these two categories, across subjects.

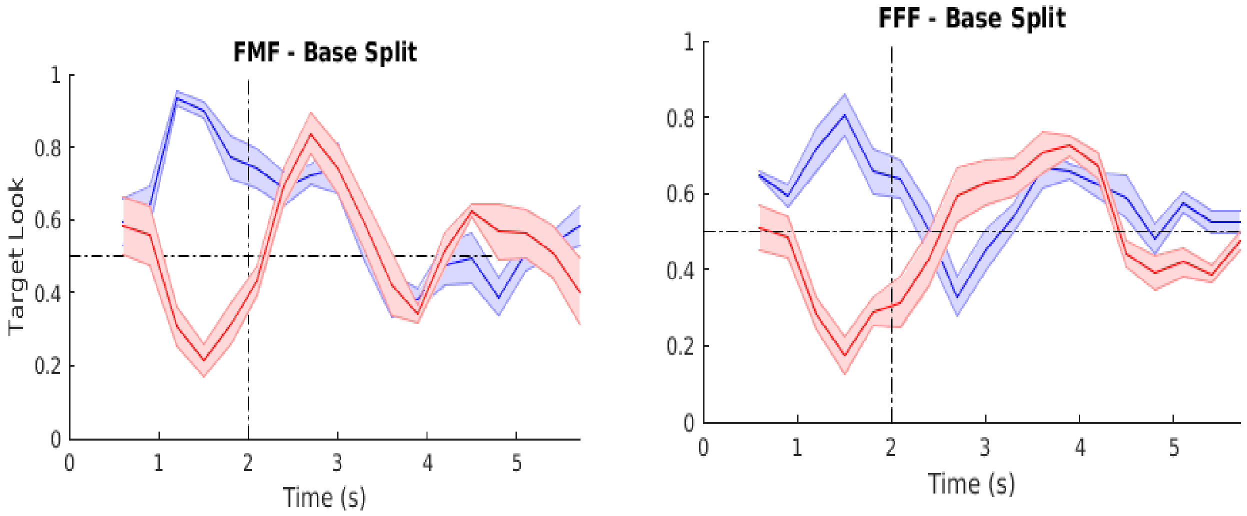

**Fig. 2.8.** The across subject temporal dynamics of the trials with an average initial Target Look below 50% (Red) and trials with an average initial Target Look over 50% (Blue), for the two groups of FMF and FFF. The errorbars are one standard error from the mean. The vertical dashed line indicates the time of the voice onset and the horizontal dashed line represents the chance level at 0.5.

As can be seen in Figure 2.8, also the trials with an incorrect initial Target Look, showed a shift toward the target face by the onset of the voice, a pattern similar to the trials with a correct initial Target Look. A two-way *t*test analysis comparing the population of mean Target Look of the trials form the two categories, resulted in: t (93) = -0.79, p = 0.43 for the FMF condition and  t (101) = 0.061, p = 0.95 for the FFF condition, suggesting that the Target Looks did not depend on the initial pre-voice Target Look.

### 2.2.8. Interim Conclusion

The results from the two groups of 10-month-old infants suggest a significant difference in looking patterns in the FMF condition against the chance level and against the FFF condition. This pattern of performance in the FMF condition, in line with our hypothesis, could potentially be due to a disjunctive inference, in the following steps: associating the cue face to the male voice in the cue phase, and subsequently eliminating face from the possible referents of the female voice, and finally inferring that the female voice can only be associated with the other face (the target face).

Moreover, in line with our hypothesis, we observed a preferential look toward the correct face in the FFF condition. However, this tendency was only significant from 1.2 seconds after the voice onset. This can be due to an initial, brief tendency for looking at the face other than the cue, before fixating on the cue face.

In two control experiments, we will try to address the possible explanations for the observed preferential look in the FMF condition. In experiment two, we address the recency effect of the cue face in directing the infants' gaze toward the target face, and in experiment 3, we will address the role of the ambiguity phase (first phase) in the inference process.

## Exp.2. Role of the Cue

### 2.3.1. Introduction

In the FMF condition of experiment one, an explanation alternative to disjunctive inference could be based on an avoidance in looking at the recent face (cue face) due to a partial habituation to the cue face, (Tighe & Leaton, 1976).

In the cue phase of the FMF condition, infants observed one single face; therefore in the test phase, they could had been more inclined to look at the face other than the cue, merely due to the fact that it was *'relatively more novel'* compared to the cue face (with a difference in exposure time of 7 seconds equivalent to the duration of the cue phase).

If participants had a preference for looking at the target merely due to its partial novelty, they did not need to rely on the voice cues to make an inference but a partial novelty preference could account to the observed Target Look in the FMF condition.

Moreover, in the FFF condition, one could reason that the tendency for looking at the cue face due to the female voice, overcame the preference for looking at the partially novel face, and as a result we could observe an overall preferential look at the cue face.

Therefore in order to assess if the behavioral result in the FMF condition, was not due to the recency effect of the cue face, we tested another condition in which we omitted the male voice in the cue phase, and hence there was only a period of silence in this phase. We refer to this new condition as Female-Silence-Female or FSF for short.

Our hypothesis was that if infants in the test phase of the FMF condition, were looking at the target, merely due to a recency effect; in the FSF condition we predicted to observe a similar pattern of Target Look as in the FMF condition.

### 2.3.2. Stimuli

The stimuli were identical to the stimuli used in experiment one (See section 2.2.2).

### 2.3.3. Participants

Twenty-five healthy monolingual Italian infants were tested for this condition, from 39 to 46 weeks, with a mean of 42 weeks. SD = 1.6 weeks,

Three subjects were excluded due to lack of attention, and one more subject due to crying at the beginning of experiment.

### 2.3.4. Apparatus

The apparatus was identical to that of experiment one (See section 2.2.4).

### 2.3.5. Procedure

The procedure used in this experiment was identical to experiment one (See section 2.2.5), with the only difference of omitting the male voice from the cue phase of the trials.

### 2.3.6. Scoring

For measuring the performance, similar to the FMF condition, we considered the ratio of looking at the face other than the cue, to define the Target Look. Considering that in the FSF condition there was no voice in the cue phase, therefore there was not a unique correct choice in this condition. However, since we predicted a partial novelty preference as an alternative hypothesis to the disjunctive inference, we set the correct face as in the FMF condition.

All other scoring details were identical to those of experiment one (see section 2.2.6).

One subject was excluded due to side bias.

### 2.3.7. Results

After applying the threshold criteria, 87 trials from 21 participants were included in the analyses. To calculate the Target Look we first checked the normality of the population of Target Looks across subjects. We used Lilli test of normality for this purpose. The distribution of the Target Looks averaged over subjects, followed a normal distribution with a p-value of 0.5, in failing to reject the null hypothesis of normality.

We then applied a *t*test, to compare the average Target Look of subjects against the chance level (0.5), the test failed to confirm a significant Target Look. Mean $_{FSF}$ = 0.49, SE = 0.02, t(20) = -0.08, P = 0.92.

We further calculated the Target Look as a function of time, similar to experiment one, in order to characterize if at any period of time, the Target Look diverged from the chance level. The analysis showed an initial tendency to look at the opposite face (Mean = 0.59, SE = 0.03); However, after the voice onset, the tendency declined to the chance level for the rest of the phase (see Figure 2.9).

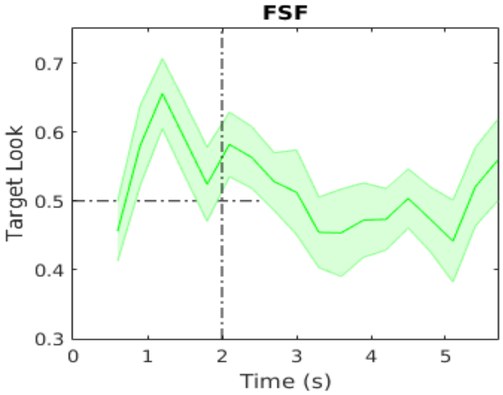

**Fig. 2.9.** Temporal dynamics of Target Look averaged over subjects in the FSF group. The overall Target Look did not pass the significance test. Mean: 0.49, p-value: 0.92, the vertical dashed line indicates the voice onset, the horizontal dashed line indicates the chance level at 0.5. The errorbar shows one standard error from the mean.

### 2.3.8. Interim Conclusion

The results of the FSF condition suggest that unlike the FMF condition, infants failed to show an overall tendency for looking at the target face. Consequently, these results negate the alternative hypothesis that in the FMF condition the Target Look resulted from low level factors, such as the recency effect of the cue face.

We can conclude that in the FMF condition, the process behind looking at the target face was essentially entangled with the male voice presented in the cue phase, and the omission of the male voice nullified the Target Look as was observed in the FSF condition.

# Exp.3. Role of the Ambiguity

## 2.4.1. Introduction

In the FMF condition in experiment one, we first introduced two alternative choices in the form of an ambiguity. We then presented a cue that participants could use to exclude one choice and exhibit a preferential look at the other one. The presented ambiguity in phase one that constitutes the disjunction premise, is an integral part of a disjunctive syllogism. However, it is possible that infants could make the inference without considering the ambiguity phase.

The potential process alternative to disjunctive syllogism, that infants could utilize to look at the target face, was novel to novel association, similar to the Novel-Name-Nameless-Category principle (N3C), (Mervis & Bertrand, 1994, Golinkoff et al., 1992). This principle that was proposed as an alternative mechanism in the context of lexical acquisition does not necessitate the exclusion of each option during the inference process, however, the subject can make a direct mapping between the novel label and the un-named (or novel) object.

So we postulated that in our paradigm it was possible that infants were not attending the first phase of the trials, or they were forgetting or ignoring the first phase due to constraints in their short term memory. In this case, the context of the paradigm turns into a novel to novel association task, and novel to novel mapping could potentially be *the* mechanism that infants could use in the FMF condition.

If this process was utilized by infants in the FMF condition, eliminating the ambiguity stage should not make a difference in the performance at the test phase. Moreover, the removal of the ambiguity stage would make the target face and the female voice presented in the test phase completely novel for the participants, since they would not be exposed to these stimuli in the first phase.

To address if this alternative mechanism could be the mechanism to explain the Target Look in the FMF condition, we tested a control condition in which the ambiguity was not provided as part of the trial, and we asked if participants still exhibit a preferential look at the Target Face.

Therefore we hypothesized that if by elimination of the first phase, we observe a similar Target Look as in the FMF condition, it would provide a support for the assumption that in the FMF condition, infants could have used novel to novel mapping to associate the female voice to the target face.

On the other hand, if eliminating the first phase does not result in a significant Target Look, we could infer that novel to novel mapping could not be an alternative mechanism in the FMF condition. By eliminating the first phase, our paradigm consisted of the second and the third phases only. We will refer to this condition as the Male-Female or MF condition for short.

Aside from this condition, we tested a control condition, analogue to FFF condition, but again with elimination of the first phase. We refer to this control condition as Female-Female, or FF condition for short. We hypothesized that participants in the this condition would have a preference for looking at the cue face, and not at the novel face. Through this control condition, we wanted to clarify the role of ambiguity in the performance of the FFF condition and also contrast it with the performance of the MF condition.

### 2.4.2. Stimuli

For the two conditions of MF and FF, as we eliminated the first phase, we could increase the number of trials of the experiment. We chose to test 12 trials in each condition, 3 trials more compared to the 9 trials of experiment 1. For this purpose we added 3 extra sets of faces and 3 pairs of male and female voices. The criteria to design the faces are explained in section 2.3.1 and the details of preparing the voices are explained in section 2.3.2.

### 2.4.2. Participants

Similar to experiment one, in order to have enough number of trials per subject, we tested only one condition in each experiment; therefore we distributed the participants into two groups of MF and FF conditions.

#### 2.4.2.1. MF Condition

Thirty-three healthy monolingual Italians were tested for this condition, from 38 to *44 weeks* and with a mean of 41 weeks, SD = 1.8 weeks.

Four subjects were excluded due to fuzziness and not passing enough number of trials, two more subjects were eliminated due to crying at the beginning of the experiment.

#### 2.4.2.2. FF Condition

Twenty-seven healthy monolingual Italians were tested for this condition with from *40 to 45 weeks* and a mean of 42, SD = 1.4 weeks.

Three subjects were excluded due to lack of attention and not passing enough number of trials.

All the infants were full term and had an APGAR number over 7 out of 10. The participants were recruited from Trieste, Italy, through sending invitation letters to a random selection of parents, whose babies fit the age range of the study. The parents were all reimbursed for attending the experiment and the infants received a certificate of attendance.

### 2.4.3. Apparatus

The eye-tracker setup and the protocol used to calibrate were identical to the one explained in Experiment one (see section 2.2.4).

### 2.4.4. Procedure

The stimuli and the experimental design at the MF and FF conditions were similar as described previously in experiment one (section 2.6), except for the following differences:

1. The phase one was removed from the trials, therefore the trials were starting at phase two.
2. The number of trials were increased to 12, by addition of new sets of faces and voices.
3. In the cue phase, the voice onset started at 4.5 s from the face onset, compared to 1.5 second in experiment 1. Subsequently the duration of the cue phase, increased from 7 seconds to 10 seconds.

The reason to delay the voice onset for an extra three seconds in the cue phase, was for considerations for the pupil analyses. Due to the fact that by omission of the first phase, the cue face was novel for the participants, therefore we wanted to let enough time for the pupil diameter to reach a stable diameter before onset of the voice; due to novelty of the face; so we could avoid having overlapping effects in the dilation profile due to the novelty of the face and the inference process.

### 2.4.5. Scoring

The methods used to define the RoIs and the inclusion criteria were the same as previously explained in experiment one (see section 2.2.6).

Four subjects from the MF condition and one subject from the FF condition were excluded due to side bias.

### 2.4.6. Results

In the first step of the analyses, to have the same number of trials as in experiment one, first we describe the Target Look of the first eight trials of the MF and FF conditions.

130 trials from 23 subjects in the MF condition, and 146 trials from 24 subjects in the FF condition passed the criteria for the analyses.

In the MF condition, contrary to our hypothesis, we did not observe a significant tendency toward any of the faces at the test phase. $\text{Mean}_{\text{MF}} = 0.50$, SE = 0.019. We applied a one-way *t*test comparing the Target Look average across subjects versus the chance level (0.5), the results were $t(22) = -0.05$, $p = 0.96$. The overall performance in the FF condition was also at chance, $\text{Mean}_{\text{FF}} = 0.51$, SE = 0.012, a one way *t*test comparing the overall Target Look versus chance resulted in $t(23) = 0.36$, $p = 0.72$.

To further investigate the dynamics of the inference processes, we calculated the Target Look as a function of time, identical to the analyses described in experiment one (see Figure 2.10).

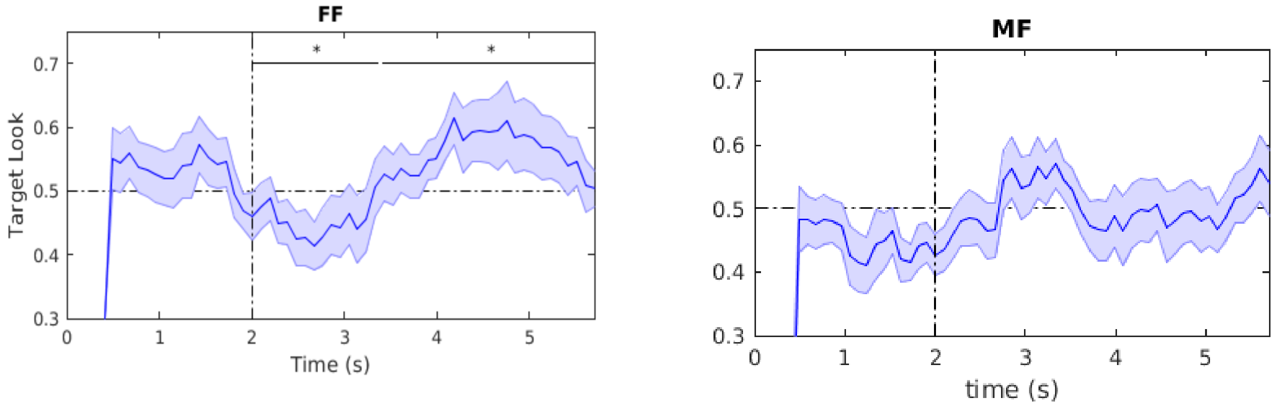

**Fig. 2.10.** Temporal dynamics of Target Look in the FF condition **(Left)** and the MF condition **(Right)**. The vertical dashed line shows the onset of voice. Only in the FF condition, the dynamics of the Target Look significantly deviated from the chance level (corrected for multiple comparison by Bonferroni-Holm method), exhibiting first a look at the novel face and then a look at the cue face. The errorbars are one standard error from the mean.

In the FF condition, based on the temporal dynamics, we observed a period of looking at the novel (incorrect) face, lasting between 2 to 3.5 seconds from the face onset. The Target Look average over subjects in this period was $\text{Mean}_{\text{FF - 2s : 3.5s}} = 0.42$, SE = 0.03, The mean Target Look in this period was different from chance resulting from a one way *t*test, $t(22) = -2.38$, $p = 0.028$.

After 3.5 seconds from the face onset up to the end of the test phase, the Target Look was significantly above chance; $\text{Mean}_{\text{FF - 3.5s : 5.7s}} = 0.57$, SE = 0.022, a one way *t*test resulted in $t(22) = 3.04$, $p = 0.006$. The analyses on these two segments consisted of 125 and 143 valid trials respectively.

The tests to compare the Target Look versus chance in both of the above time-spans in the FF condition, passed the significancy limit after correcting for multiple comparison problem with the

Bonferroni-Holm method, however, based on the more stringent Bonferroni correction, only the second time span from 3.5 s to 5.7 s, was declared to be significant.

Furthermore, by including all the 12 trials of the paradigm, there were 185 and 203 trials in the MF and FF conditions respectively. The resulting temporal dynamics of Target Look followed a similar pattern in both of the conditions; in the FF condition, $\text{Mean}_{\text{FF}} = 0.50$, SE = 0.17, t(22) = 0.1, p = 0.92, and in the MF condition, $\text{Mean}_{\text{MF}} = 0.49$, SE = 0.016, t(23) = 0.44, p = 0.66.

Also including all the 12 trials did not make a difference in the temporal dynamics of the FF condition. The Target Look after 1.4 seconds from voice onset, reached zero, and within this period of 1.4 seconds, the Target Look was significantly below chance, $\text{Mean}_{\text{FF - 2s : 3.4s}} = 0.42$, SE = 0.035, t(22) = -2.34, p: 0.029, and after 1.4 seconds from the voice onset to the end of the phase, the Target Look was significantly above chance, $\text{Mean}_{\text{FF - 3.4s : 5.7s}} = 0.57$, SE = 0.02, t(22) = 2.33, p = 0.03.

### 2.4.7. FMF Condition versus the MF Condition

To further clarify the role of the ambiguity phase in the FMF condition, we compared the Target Look of the FMF group versus MF group over the whole duration of the phase (from 2 to 5.7 s), since in both of the groups the distribution of the average Target Look across subjects followed a normal distribution, we applied a *t*test to compare the mean of the two populations, which resulted in t(57) = 2.1, p = 0.039. In this analysis, 34 subjects from the FMF condition and 20 subjects from the MF condition were included. Also a Wilcoxon non-parametric ranksum comparison between the population of trials of the two conditions resulted in z = 2.6, p = 0.009. In this latter analysis 125 and 135 trials from the FMF and MF conditions were included (see Figure 2.11).

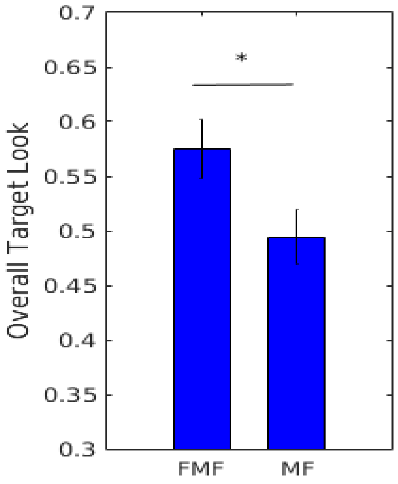

**Fig. 2.11.** The comparison of the overall Target Look across subjects in the FMF group versus the MF group. A two-way *t*test revealed that the Target Look of the FMF group is significantly higher than the MF group, t(57) = 2.1, p = 0.039. The errorbars are one standard error from the mean.

However, the same analysis comparing the Target Look in FFF group versus FF group resulted in z = 0.72 and p = 0.47, In this latter analysis, the two groups consisted of 120 versus 143 trials.

### 2.4.8. Interim Conclusion

In the FF condition, the significant look at the target (the cue face) was followed by a significant look at the novel face. We were not expecting a period of novelty preference in this condition, but we speculate that this initial preference could be related to the fact that the faces were static, and infants first expect to find a talking face rather than a static face while they hear a voice; therefore in the FF condition, they first looked at the novel face when the voice started, and when they found it static similar to the cue face, they fixated back on the correct (cue) face. However, we did not further assess this speculation.

The failure of infants to look at the target in the MF group suggests that a novel to novel mechanism could not be utilized by the infants in this context. Subsequently, this failure suggests an important role for the ambiguity stage of the FMF condition. We can conclude that a novel to novel mapping mechanism similar to N3C is not an alternative explanation for the observed Target Look in the FMF condition.

These results however, are not in line with the results reported in the context of word learning. Mather and Plunkett (2010) suggested that at the age of 10 months, utterance of a novel label, steers infants to look at a novel object rather than a familiar object. The authors considered these results as an evidence for N3C mechanism in the word learning context. Association of novel labels to novel objects has also been reported in younger infants at the age of 4 months (Saksida, 2014), but if this association is supported by a disjunctive inference should further be assessed. However, the results of MF condition suggest that the analogue of the N3C mechanism at this age may not be available in the context of face-voice association, but might specifically be in service of word-learning.

## Exp.4. Inference in Five-Month-Olds

### 2.5.1. Introduction

In order to track the developmental course of the disjunctive inference we sought to explore if at an earlier age, we could observe a Target Look similar to the 10-month-old infants. For this purpose, we chose to test the FMF condition at the age of 5 months.

### 2.5.2. Stimuli

The stimuli we used in this experiment were identical to the stimuli used in experiment one, (see section 2.2.2).

### 2.5.3. Participants

Twenty-five healthy monolingual Italian infants were tested in this condition, from 26 to 30 weeks, with a mean of 27.5, SD = 1.*7* weeks.

Three subjects were excluded due to lack of attention, two more subjects were excluded due to crying at the beginning of the experiment.

All the infants were full term and had an APGAR number over 7 out of 10. The participants were recruited from Trieste, Italy, through sending invitation letters to a random selection of parents, whose babies fit the age range of the study. The parents were all reimbursed for attending the experiment and the infants received a certificate of attendance.

### 2.5.4. Apparatus

The eye-tracker setup was identical to the one explained in Experiment 1. (See section 2.2.4)

### 2.5.5. Procedure

The procedure for calibration and design of the paradigm, was identical as described in the experiment one for the FMF condition (see section 2.2.5).

### 2.5.6. Scoring

The methods and the inclusion criteria used in this analysis were the same as previously explained in experiment one (see section 2.2.6). Two subject were further removed due to side bias.

### 2.5.7. Results

After applying the inclusion criteria, 52 trials from 18 subjects, included in the analyses. We checked the normality of the distribution of Target Looks average over subjects. A Lillie test of normality confirmed a normal distribution, with failing to reject the null hypothesis of normality with a p-values equivalent to 0.50 (0.5 is the largest tabulated number the test can report).

We then calculated the Target Look average across subjects, Mean = 0.60, SE = 0.05. A one-way *t*test based on subject averages resulted in t(17) = 2.18, p = 0.043. The test suggests a marginal deviation of the Target Look from the chance level. We furthermore calculated the Target Look as a function of time in time bins of 200 ms (see Figure 2.12).

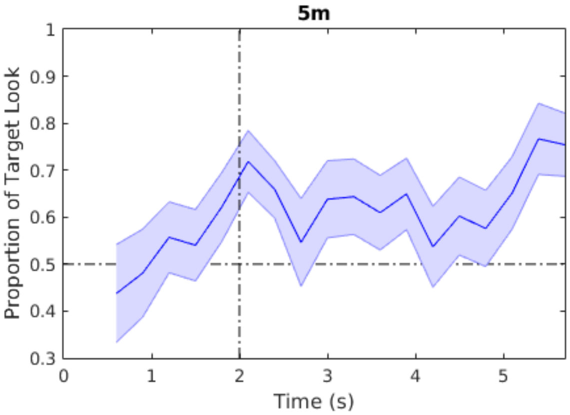

**Fig. 2.12.** Temporal dynamics of the Target Look average over subjects. Fifty-two trials from 18 subjects were included. The errorbar shows one standard error from the mean at each bin. The vertical dashed line indicates the time of voice onset. The horizontal dashed line indicates the chance level at 0.5

As can be seen in Fig. 2.12, the Target Look started at chance level and after the voice onset deviated from the chance toward the correct face until the end of the test phase.

### 2.5.8. Interim Conclusion

The results of this experiment tentatively suggest that 5-month-old infants potentially exhibited a Target Look. However, we do not consider that these results point directly toward a disjunctive inference. First because the results were only marginally significant, so we are continuing testing participants in this group to acquire a higher number of valid subjects. But the main concern about this age group is that approximately 4 seconds of exposure to a voice may not be enough to establish an association. So we are running the FSF and FFF control conditions to assess the possible explanations for the observed performance in this age-group.

## Exp.5. Inference in Adults

### 2.6.1. Introduction

We decided to test our paradigm with an adult control group with two main objectives.

First, to compare the performance of the adults in the FMF condition versus the FFF condition.

And second, to compare the profile of the pupil dilation in the adult group in contrast with the 10-month group during the inference process.

To pursue this aim we tested the FMF, FSF and FFF conditions in each adult subject. The three conditions were tested identical to the 10-month group in order to be able to have qualitative comparisons across ages. The only difference we implemented was to test the three conditions in a within subject design.

In the current section we discuss the behavioral results of the experiment and in the next section we will return to the results of the pupil analysis.

### 2.6.2. Stimuli

In this experiment, we tested 5 trials in each condition of FMF, FFF and FSF, therefore the experimental session consisted of 15 trials. Consequently we designed three more sets of faces and voices in addition to the stimuli used in experiment 3 (which consisted of 12 trials). The criteria to design the faces are explained in section 2.2.2.1, and the details of preparing the voices are explained in section 2.2.2.2.

### 2.6.3. Participants

Eighteen adults were tested with ages from 22 to 27 years. All participants were native Italian speakers. The participants were recruited through a Facebook call, and the majority were from university of Trieste, Italy. One subject was eliminated from the analysis due to a poor eye gaze recording.

### 2.6.4. Apparatus

The eye-tracker set up was identical to the one described in experiment one (see section 2.2.4).

The subjects sat at a distance of approx. 60 cm from the screen, having their head fixed on a chin rest during the entire experiment.

### 2.6.5. Procedure

The design of the experiment was similar to the design in 10-month groups (see section 2.2.5), except with the following variations:

First, the participants passed trials from all three conditions interleavingly. Each condition contained five trials, therefore in total the participants passed 15 trials. The calibration process was identical to the one used in experiment one.

Second, after passing the calibration (with an identical protocol to experiment 1), participants were presented with the following written instruction on the screen “In questo esperimento, ci saranno diverse coppie di facce e voci.

le voci appartengono alle facce che vedi sullo schermo. Ogni volta che senti una voce, scegli la faccia a cui pensi che appartenga. Per fare la tua scelta, guarda piu' a lungo nella direztione della faccia scelta.”

The English translation is: “In this experiment, there will be different pairs of faces and voices. The voices belong to the faces that you see on the screen. Each time you hear a voice, choose the face that you think it belongs to. To make your choice, look longer at the face of your choice.”

The instruction was given in order to have the participants engaged in a simple preferential looking task, although it could alter the natural way they choose to look at the faces.

After reading the instructions, similar to experiment 1, participants passed 6 seconds of initial animation, in order to adjust the pupil baseline (see section 2.2.5).

### 2.6.6. Scoring

The gaze, pupil diameter and the quality of gaze, were collected and analyzed identical to the procedure explained in experiment one (see section 2.2.6). The Target Look in the FSF condition, was analyzed as explained in experiment two (see section 2.3.6).

None of the participants were excluded due to side bias.

## 2.6.7. Results

### 2.6.7.1. Overall Results

After applying the constraints, 79, 77 and 79 trials were included in the analyses of the Target Look. The populations of the Target Look, averaged over subjects, did not follow a normal distribution in the FMF and FFF groups, based on the Lillie test of normality. The reported p-values were 0.001 in both conditions (this value is the lowest tabulated value the test can report). However, in the FSF group, the population of the Target Looks followed a normal distribution, with a p-value equal to 0.5.

The Target Look in the FMF and FFF groups averaged over subjects were as following: $\text{Mean}_{\text{FMF}} = 69\%$, SE = 0.05, and $\text{Mean}_{\text{FFF}} = 79\%$, SE = 0.06 respectively. And similar to the 10-month-olds, the Target Look in FSF condition was at chance. $\text{Mean}_{\text{FSF}} = 0.51$, SE = 0.05, (see Figure 2.13).

We used Wilcoxon non-parametric ranksum test to compare the Target Looks obtained from the different conditions. Comparing the FMF versus FSF conditions resulted in $z = 2.5$, $p = 0.012$, and the same comparison between the FFF versus FSF conditions resulted in $z = 3.75$ and $p = 0.00017$ (corrected for Bonferroni multiple comparison).

An interesting observation was the higher Target Look at the FFF condition compared to the FMF condition ($\text{Mean}_{\text{FMF}} = 69\%$ versus $\text{Mean}_{\text{FFF}} = 79\%$). However, the Wilcoxon rank sum test did not result in a significant difference in the Target Looks of these two conditions. $Z = 1.56$, $p = 0.12$.

This trend although not significant, could be due to the fact that the disjunctive process in the FMF condition was more computationally demanding for the participants, compared to the inference in the FFF condition *(Cheng 1986)*.

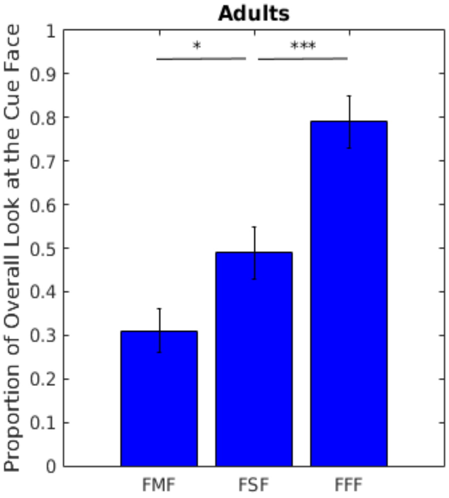

**Fig. 2.13.** Proportion of overall look at the cue face (the face presented in the cue phase) across conditions. The performance in the FMF and FFF conditions were significantly different from FSF condition based on Wilcoxon non-paramteric rank sum test, $z_{\text{FMF-FSF}} = 2.5$, $p = 0.012$ and $z_{\text{FFF-FSF}} = 3.75$, $p = 0.00017$.

We further calculated the dynamics of Target Look as a function of time, similar to the analysis described in the experiment one (see section 2.2.7). We set the onset of the analysis at 300 ms from the face onset and averaged the Target Look over trials, in time bins of 200 ms (see Figure 2.14).

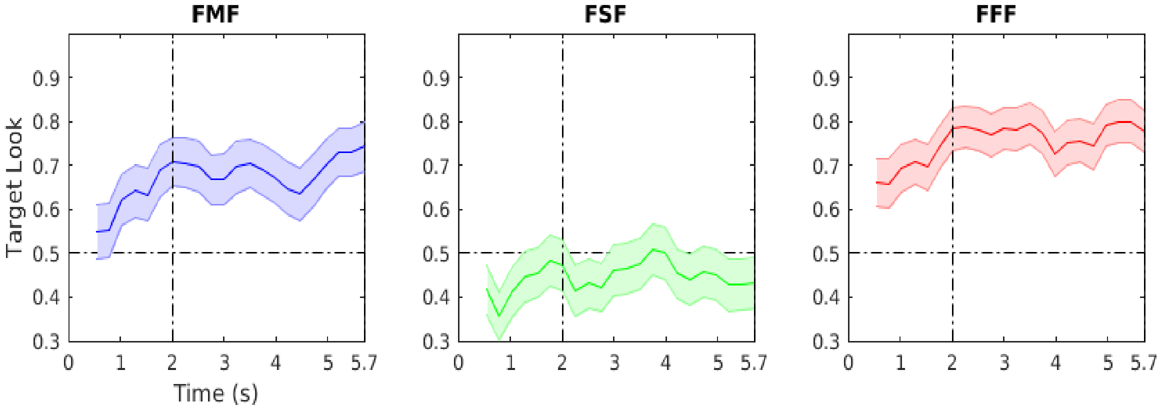

**Fig.2.14.** Target Look dynamics over time, in the adult group across conditions. The average Target Look is calculated at each time bin of 200 ms, the starting time bin is 300 ms from the face onset. The vertical dashed line shows the voice onset, the horizontal dashed line shows the chance level at 0.5. All errorbars are one standard error from the mean.

As the temporal dynamics of the Target Look in Figure 2.14 suggests, in the FFF condition the Target Look started as early as the face onset. However, in the FMF condition the gaze pattern started at chance and then rapidly shifted to the ceiling of the Target Look. This pattern of initial Target Look at the FMF condition may suggest that at least the final step of the disjunctive inference, or the final step and the occlusion step, occurred as early as the face onset in the test phase. We can refer to this pre-voice Target Look as anticipatory look. We would further expand this observation in the next section.

### 2.6.7.2. Anticipatory Target Look

Based on the temporal dynamics of the Target Look, in Figure 2.14, we observed that the FMF and FFF conditions show a tendency to look at the correct face, before the onset of the voice. Mean $_{\text{FMF-Pre-Voice}}$ = 0.64, SE = 0.04, Mean $_{\text{FFF-Pre-Voice}}$ = 0.73, SE = 0.038, and

Mean $_{\text{FSF-Pre-Voice}}$ = 0.44, SE = 0.04 (see Figure 2.15).

Comparing the ratio of looking at the Target Face, before the voice onset, in the FMF and FSF conditions revealed a significant difference across the two conditions, based on Wilcoxon rank sum test, z = 3.43, p = 0.00059. The same analysis comparing the Target Looks within the first two seconds across the FFF and FSF conditions, resulted in z = 2.8, p = 0.0047.

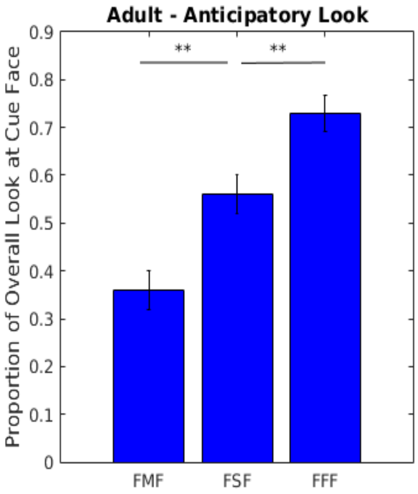

**Fig. 2.15.** Initial Target Look before the voice onset at the test phase of the adult group, across conditions.

An explanation for this observation could be that participants always heard a female voice at the beginning and the same female voice at the end of each trial. Therefore, by the time that subjects received the cue in the cue phase, they had sufficient information to conclude their inference both in FMF and in FFF conditions. As a result, at the onset of faces in the third phase, they showed a preference for looking at the correct item. To visualize the development of the anticipatory look, we obtained the rank number of the valid trials of the participants, and calculated the initial Target Look before the voice onset in the FMF and FFF conditions. As can be seen in Figure 2.16, already by the second trial of a condition (from the 3rd to the 5th trial of the experiment), in both conditions the Target Look was above chance.

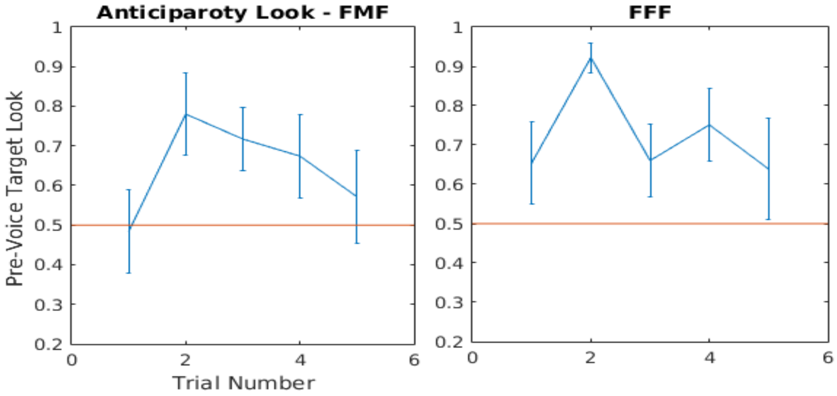

**Fig. 2.16.** Target Look within the first two seconds of the FMF and FFF conditions, averaged across subjects at each consecutive trial number of a condition. Each participant passed five trials of each condition interleavingly. The errorbars show one standard error from the mean. The horizontal solid line shows the chance level at 0.5.

Fig.2.16 provides evidence that except for the first trial in the FMF condition, for the consecutive trials, the inference was established before or within the first 2 seconds of silence at the beginning of the test phase. Suggesting that adults anticipated the presentation of a female voice in the test phase and they fixated on the target face in advance before the voice onset.

### 2.6.8. Interim Conclusion

In the adult group, a persistent look at the target was observed in both the FMF and FFF conditions and similar to the 10-month group, the performance in the FSF condition was at chance.

Furthermore the looking dynamics suggested that since the voice in the third phase was always female, the participants could anticipate the voice, and subsequently showed an early preference for looking at the target, prior to the voice onset.

This anticipatory Target Look in the FMF condition may suggest that the inference process was finalized in the cue phase or as early as the onset of the faces in the test phase. In the next chapter, we use the pupil data of the cue phase, to assess this proposed time-frame more accurately.

## 2.7. Pupil Dilation Correlates of the Inference Process

### 2.7.1. Introduction

Pupil dilation has been shown to trace several different cognitive processes. The list includes memory loads, attention, decision-making, emotional arousals, inconsistencies and surprise, etc. (see Hartmann & Fischer, 2014, for a brief review). In a simple but classical study in 1964 by Hess and Polt, the authors showed in a number multiplication task that pupil dilation increased monotonically with the increase in the difficulty level of a multiplication task. Later Beatty and Kahenman confirmed these results with a comprehensive battery of experiments, showing that the pupil dilates as a response to task difficulty in different types of problem solving tasks, including multiplication of numbers, letter matching or pitch discrimination (Beatty & Wagoner, 1978; Beatty, 1982; Kahneman, 1967).

On the other hand, Beatty showed that in the adult participants the magnitude of pupil dilation does not depend on the pupil baseline for a wide range of initial baselines, therefore the dilation can be reported as percentages (Beatty, 1982). The same patterns of dilation were seen in digit memorization tasks, Kahneman and Beatty famously showed that pupil dilates monotonically up to seven digits, a capacity that is considered to be the limit of short term memory (Beatty, 1966; Kahenman, 1966). And in a beautiful follow up study, Peavler showed that the dilation reaches a plateau when the seven-digit capacity is being reached (Peavler, 1974; see Beatty  Wagoner 2000 for a review).

In a separate line of research, in recent years several studies have reported that when the subjects are exploring the possible options of a task, pupils exhibit a dilation; and on the other hand when the subjects are anticipating an upcoming event, there is a constriction in the pupil diameter. For example Daniels, Nichols, Seifert, and Hock (2012), showed that when the attention is focused on a narrow spatial area, there is a constriction in the pupil compared to when the attention is spread in a wide area. In another study Gilzenrat, Nieuwenhuis, Jepma, and Cohen (2010) showed that the baseline of pupil diameter preceding the task predicts the exploration (finding the options) versus exploitation (anticipating the reward within one option) status of the upcoming behavior. When the subject is attending a goal achieving task or exploiting to maximize the gain, there will be a constriction in the pupil baseline, in line with the results of Daniels (2012); and when the subject is exploring other alternative choices, there will be a higher baseline for the pupil diameter at the onset of the task.

Over the last years, examining the pupil dilation due to cognitive related processes has been extended to infant studies. For example, Jackson and Sirois (2009), provided a linear model in a violation of expectation task to describe the looking time of 8.5-month-old infants, when they were looking at a set of events, that could be plausible/implausible, and novel/familiar. They used their method to dissociate the perceptual factors contributing in the looking time measures from the conceptual factors (possible/impossible scenarios), they extended their model later to study object permanence in 10-month-old infants (Sirois & Jackson, 2012).

In another study, pupil dilation has been shown to index violation of expectation with detection of frequent versus infrequent syllables in 3- and 6-month-old infants (Hochmann & Papeo, 2014). Other studies addressed pupil dilation as an informative parameter in tracking incongruent trials in social or ostensive contexts (Gredeback & Melinder, 2011; Geangu, Hauf, Bhardwaj & Bentz, 2011; Marno et al., 2015).

Since inference process is one of the several factors that triggers pupil dilation, we sought to design the settings of our paradigm to be able to consistently register the dynamics of cognitively induced pupil dilation on the course of the inference process, with an aim to examine if we could observe a component of dilation that could be interpreted as a unique trace of the potential steps in the disjunctive inference process.

As a general pattern, upon the start of a trial phase, the pupil immediately adjusts to the brightness of the stimuli (within a few hundred milliseconds). Then after, the pupil may start to dilate because of various cognitive processes (within several seconds). It may reach a maximum diameter and stabilize or may constrict.

We hypothesized that if there was a disjunctive inference occurring in the FMF group of infants, we should be able to observe a higher dilation corresponding to the trials with a higher Target Look, compared to trials with a lower Target Look. If this modulation of pupil dilation by the Target Look, would be due to the process of disjunctive syllogism, we hypothesized that this modulation should be absent in the control conditions such as the FFF condition. Furthermore we should be able to see a similar pattern of results in the adult group, since in the adult group, we can speculate that the subjects used disjunctive syllogism to reach the correct choice and this process should elicit a pupil dilation.

To explore this hypothesis, we analyzed the pupil dynamics to investigate if we could find traces of such a modulation between the pupil dilation and the Target Look of the trials in the FMF condition.

### 2.7.2. Scoring

Here we provide the results of the cue phase only, and we do not expand the analysis to the phase 1 and phase 3. The reason is that in these two phases there were two faces on the screen, and the continuous saccading of subjects from one side to the other side of the screen, introduces fluctuations in the pupil diameter (Brisson et al, 2013). These oscillations need to be corrected within a richer mathematical framework (a work that we haven't concluded yet). However, in the cue phase, since the participants were fixating on the only face at the center of the screen, we could have a rather stable signal for the pupil diameter.

Since the baseline of the pupil diameter varies across subjects and across ages, and furthermore, the luminosity of the visual stimuli alters the diameter of pupil, we needed to introduced two steps to normalize the pupil diameter.

The pupil diameter was sampled by the Tobii eye tracker, at 60 HZ. For the analyses we included the data only from the valid subjects and from the valid trials, with the conditions specified in the scoring section of experiment 1 (see section 2.2.6, we considered an 80% minimum for Ratio of Valid Look in the cue phase, as part of the inclusion criteria, to guarantee a stable pupil profile in this phase).

We further excluded the time bins with a registered pupil diameter smaller than 1 mm or larger than 8 mm to avoid possible artifacts. Since the diameter of the two eyes dilate or constrict with the same rate, we performed all the analysis based on the data from the right eye only. However, applying the same analyses on the left eye or averaging the data of both eyes did not alter the results.

### 2.7.3. Normalization of the Pupil Diameter

The normalization of the pupil diameter was done in two steps. First step to correct for the inter-trial-variability of the luminosity of the stimuli, and the second step, to correct for cross-subject variability of the pupil diameter.

**Step 1. Trial Specific Baseline**

At the age of 10 months, the reflex of pupil to light in healthy infants has been reported to be within 300 ms from the onset of a brief pulse of light (Nyström, Gredebäck, Bolte & Falck-Ytter, 2015). In our paradigm, since the face onset in the cue phase is 1500 ms before the voice onset, the pupil had enough time to adopt to the luminosity of the screen. Therefore, we corrected the pupil diameter for the variations in luminosity of the stimuli by measuring the baseline of the pupil diameter at the beginning of each trial prior to the voice onset. We performed this by averaging the pupil diameter from 1000 ms to 1500 ms after the face onset. We will refer to this baseline measure as *Phase_Initial_Diameter.*

**Step 2. Subject Specific Baseline**

Since we tested our paradigm both in infants and in adults, and considering that the pupil diameter of individuals differ across age, therefore we corrected the measured pupil dilation by the pupil baseline of each subject.

At the beginning of the experiment, there was 6 seconds of animation (see section 2.2.5). This animation was then followed by the first trial. We set the pupil baseline of each subject as the pupil diameter just after the 6 seconds of animation. We refer to this baseline as the Subject_Base_Diameter.

Measuring a baseline for each trial and a global baseline for each subject, we could then define the normalized value of the pupil diameter at a given time *t* as following:

Normalized Pupil Dilation (t) = (Pupil_Diameter (t) – Phase_Initial_Diameter) /

Subject_Base_Diameter

By performing this normalization procedure, we were able to compare the dilations across conditions and across ages. We further only included the trials that both of the baselines, the Phase_Initial_Diameter and Subject_Base_Diameter, were available.

### 2.7.4. Results

For discussing the dynamics of pupil dilation, we focus on the 10-month group and the adult group. In the following sections by pupil dilation we are referring to the normalized pupil dilation.

#### 2.7.4.1. Qualitative Results

We hypothesized that the potential process of exclusion and disjunctive inference in the FMF condition, could trigger a dilation component that can be traced in the dynamics of the pupil in the second or the third phase of the paradigm. Therefore in the trials with a higher Target Look, we expected to observe a higher dilation of pupil, compared to the trials with a lower Target Look.

To start assessing this hypothesis, we first demonstrate the difference between the pupil profiles corresponding to trials with the higher and the lower 50% of Target Look range. We distributed the trials based on the Target Look to higher than 50% and lower than 50%, and compared the corresponding pupil profile at the second phase of these two clusters of trials.

We observed that in the FMF condition, in both adult and 10-month groups, there was a facilitation of pupil dilation in the higher range of Target Look compared to the lower range (see Figure 2.17).

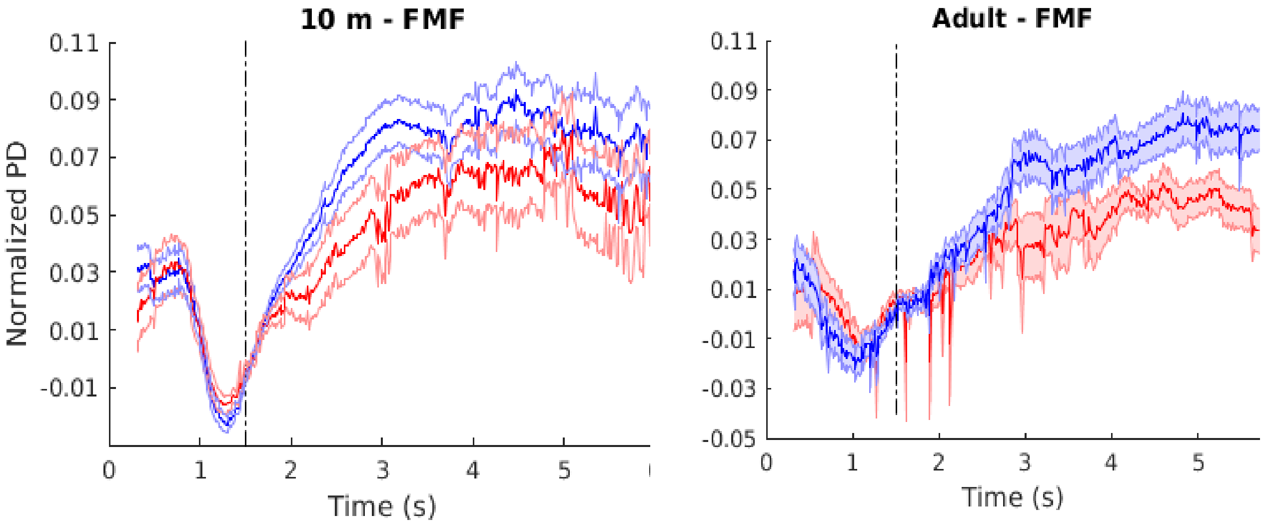

**Fig. 2.17.** Dynamics of pupil dilation in phase two of the FMF condition in adult group and 10-month-old infants. Comparing the trials with upper 50% range of Target Look (blue curve) and with the lower 50% of Target Look (red curve). **Left Panel** demonstrates the results of the 10-month-olds and **Right Panel** demonstrate the results of the adult group. In the 10-month group, 77 trials were included in the higher range versus 52 trials corresponding to the lower range. In the adult group, 45 trials were included in the higher range versus 19 trials in the lower range. Note that value of zero on the vertical axis, corresponds to the time that we set our baseline for normalization (1000 to 1500 ms). The vertical dashed line shows the voice onset. The errorbars show one SE from the mean.

Fig. 2.17, demonstrates that the dissociation of the normalized pupil diameter curves corresponding to the two categories in the adult group, is facilitated toward the end of the second phase; however, in the 10-month group, the dissociation became less salient toward the end of the phase.

We then repeated the same analysis for the FFF condition of the adult group and the 10-month group. The analyses revealed an overall overlapping of the two profiles in the 10-month group (see Figure 2.18).

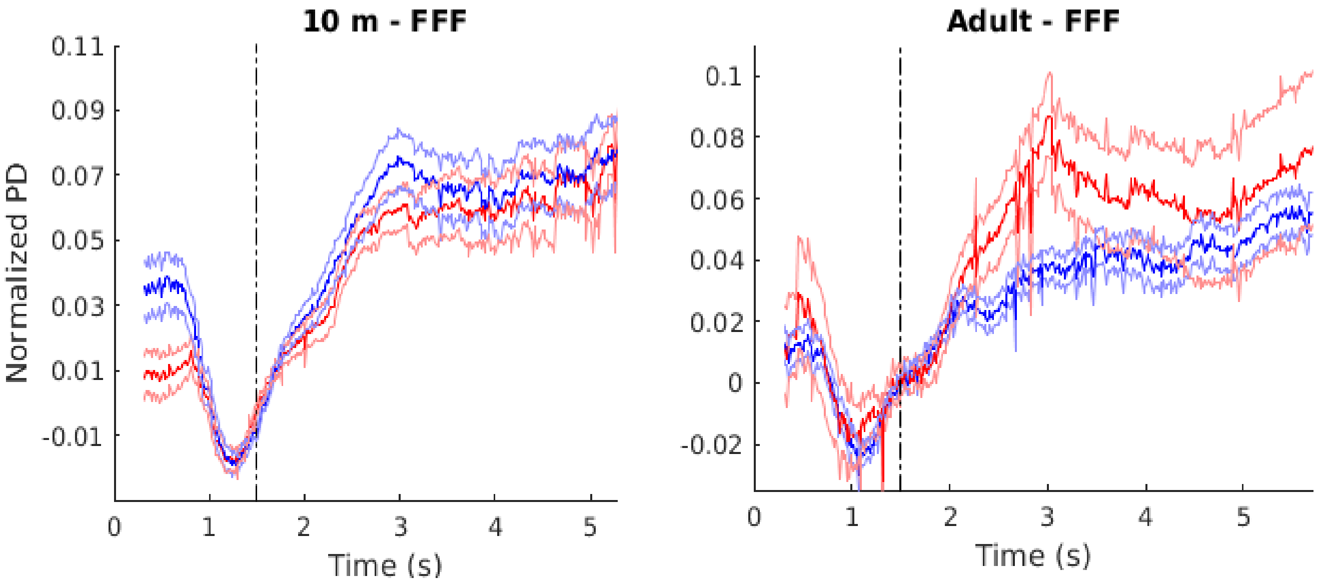

**Fig. 2.18.** Dynamics of pupil dilation in phase two of the FFF condition in adult group and 10-month group. The blue curve corresponds to the trials with a Target Look over 50%, and the red curve corresponds to the trials with a target Look below 50%. **Left Panel** demonstrates the results of the 10-month-olds and **Right Panel** demonstrate the results of the adult group. In the 10-month group, 66 trials were included in the higher range of Target Look versus 59 trials in the lower range, while in the adult group, the number of trials were 54 and 16 for the higher versus the lower range of Target Look. Note that due to a high overall performance in the FFF condition of the adult group, there is a high standard error in this condition for the lower Target Look curve (the red curve). Value of Zero on the vertical axis corresponds to the time that we set our baseline for normalization (1000 to 1500 ms). The vertical dashed line shows the voice onset, the errorbars are one standard error from the mean.

The dissociation in the pupil curves corresponding to the higher and lower 50% of Target Looks in the FMF condition suggests that carrying the inference may have triggered a pupil dilation in the second phase, an effect that might be absent in the FFF condition.

### 2.7.4.2. Interaction between Target Look and the Pupil Dilation

We sought to investigate the interaction between different attributes of the dilation profile and the Target Look in the two conditions of FMF and FFF. The attributes we considered were the mean and the peak of the pupil dilation curve of the cue phase of the trials. Both attributes were calculated for each trial, after the voice onset up to 5.7 seconds.

#### 2.7.4.2.1. Pupil Dilation Correlates of Target Look at 10 Months

To characterize the interaction between the Target Look and the attributes of the pupil dilation, we applied a Pearson linear correlation analysis to measure the magnitude of the linear interaction between the two measures. For these analyses, we excluded the trials with a Target Look at 0% or 100% to avoid biasing the correlation coefficients (26 trials in the FMF condition and 16 trials in the FFF condition were excluded due to this consideration. The calculated coefficient for the mean of pupil dilation in the FMF condition, was more significant by including these trials).

The number of trials included in this analysis were 111 in the FMF condition and 114 in the FFF condition. The Pearson linear correlation between the mean of pupil dilation and Target Look in FMF condition resulted in $r(111) = 0.20$, $p = 0.038$, and in the FFF condition, $r(114)= 0.76$, $p= 0.42$ (see Figure 2.19).

We applied the same analysis for the peak of the pupil curve of each trial at the second phase. The results were similar to the results obtained from the mean of pupil dilation (see Figure 2.20). FMF condition, $r(111)= 0.27$, $p = 0.0040$, FFF condition, $r(114)= -0.0044$, $p = 0.96$.

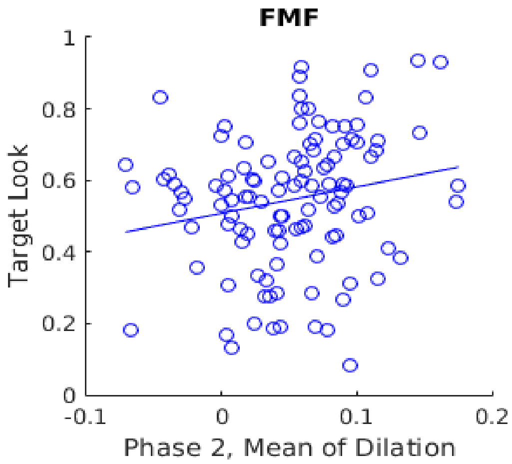

**Fig. 2.19.** Scatter plot showing the mean of pupil dilation in phase two versus Target Look for each trial in the FMF and the FFF conditions. the analysis included 129 and 125 trials in the FMF and FFF conditions respectively. Only in the FMF condition there was a significant interaction r(111) = 0.20, p = 0.038. The linear line shows the best fit based on the least square distance to the data points.

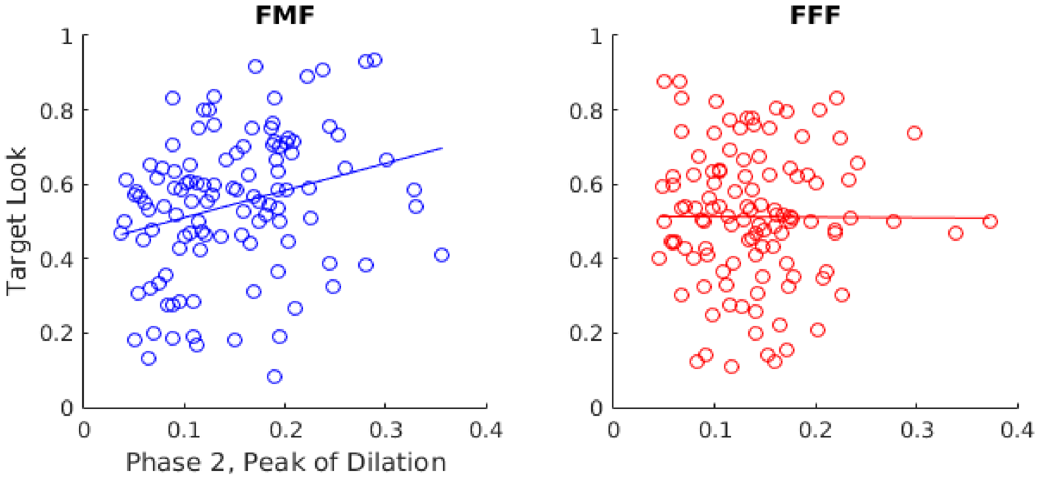

**Fig. 2.20.** Scatter plots showing Target Look versus peak of pupil dilation at the second phase for each trial. Only the FMF condition, showed a significant Pearson correlation, r(111)= 0.27, p = 0.0040, The number of included trials were 129 and 125 in the FMF and FFF conditions respectively. The linear line shows the best fit based on the least square distance to the data points.

The strong linear correlations in the FMF condition between the attributes of the pupil dilation of the cue phase and the Target Look, suggests that the mental processes involved in the cue phase, were directly contributing to the performance, in a way that their absence, led to a lower amplitude in the pupil dilation as well as a lower Target Look. However in the FFF condition, the mental processes in the cue phase, were not entangled with the performance.

#### 2.7.4.2.2. Correlates of Target Look in the Adult Group

We further examined if we could find the same pattern of interactions in the adult group. In the analysis from the FMF and FSF conditions, 16 adults passed the inclusion criteria, providing 48 and 62 trials respectively; and in the FFF condition, 15 adults passed the constraints, providing 50 trials. In this analysis the average of the measures across subjects were analyzed, since several of the trials had a ceiling Target Look (at 0% or 100%), that could subsequently reduce the accuracy of the correlation analysis.

In the adult group, similar to the 10-month-olds, only in the FMF condition, a strong interaction between the pupil dilation and the Target Look was observed. The Pearson correlation analysis corresponding to the mean of pupil dilation resulted in r(16) = 0.54, p = 0.02, in FMF condition, r(15) = -0.06, p= 0.81 in the FFF condition, and r(16) = -0.20, p= 0.41 in the FSF condition.

We repeated the same analysis for the peak of dilation and the Target Look, which resulted in r(16) = 0.50, p = 0.034 in FMF condition, r(15) = -0.34, p= 0.17 in the FFF condition, and r(16) = 0.18, p= 0.64 in the FSF condition (see Figure 2.21).

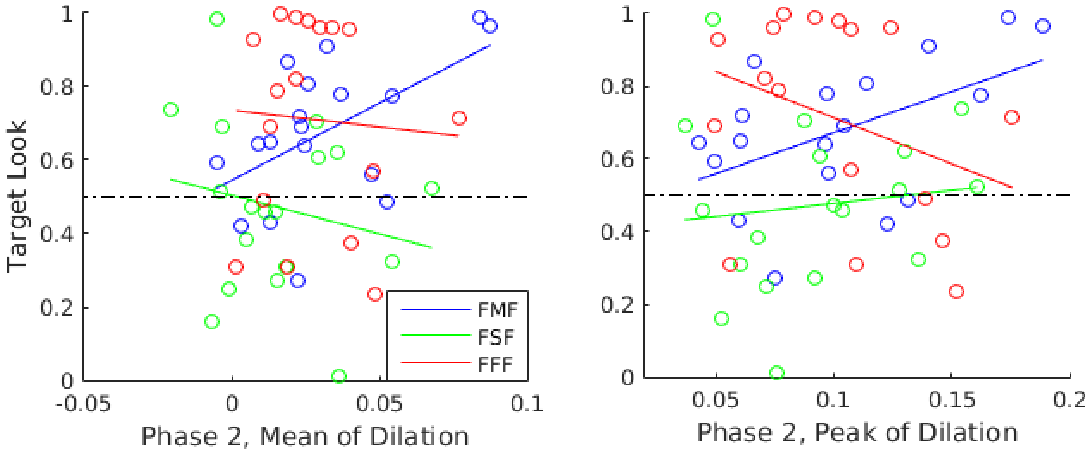

**Fig. 2.21.** Scatter plot of the pupil dilation attributes versus Target Look averaged across subjects in the three conditions. **Left.** Interaction between the Mean of pupil dilation in phase two versus Target Look. **Right.** Interaction between the peak of pupil dilation in phase two versus Target Look. Similar to the 10-month groups, only in the FMF condition a significant positive trend could be seen, $r_{FMF-Mean}(16) = 0.54$, $p = 0.02$, $r_{FMF-Peak}(16) = 0.50$, $p = 0.034$. The linear lines show the best fit based on the least square distance to the data points.

These results suggest that similar to the 10-month group, the Target Look was correlated with the pupil dilation at the second phase, only in the FMF condition.

#### 2.7.4.3. Attention as an Alternative Explanation for Pupil Dilation

In the FMF group of the 10-month-old participants we observed that the trials with a higher Target Look were also exhibiting a higher pupil dilation. An explanation for this observation could be due to a higher state of attentiveness in the FMF condition. Based on this alternative explanation, in the trials with a higher attention, infants performed better and since attention elicits a pupil dilation, subsequently the Target Look was indirectly linked with the pupil dilation.

To address the role of attention in our paradigm, we hypothesized that if attention was the only source of the pupil dilation in trials with a high Target Look, by assuming that the attention is kept constant throughout the cue phase, we then should be able to observe a higher amplitude in the pupil diameter at the starting point of the cue phase (Daniels, et al, 2012).

To assess this hypothesis, we registered the diameter of the pupil within the first 100 ms from the face onset in the cue phase (before stabilization due to luminosity of the face), and calculated the interaction between the subject-normalized measure of this baseline diameter and the Target Look. According to our hypothesis, the trials with a higher baseline diameter at the onset of the cue phase, could indicate a higher level of attention in this phase and could subsequently lead to a higher Target Look.

The Pearson analysis for linear correlation did not show a significant interaction between the two observables. In the FMF condition, $r(111) = -0.034$, $p = 0.72$ and in the FFF condition, $r(114) = 0.076$, $p = 0.42$. The lack of a strong interaction between the Target Look and the pupil diameter at the onset of the cue phase suggests that a contribution of attention in the pupil diameter alone, cannot be predictive of the Target Look in the FMF condition.

This analysis however, is a post-hoc analysis, and suffers from the fact that the transition between the phase 1 and phase 2 (approximately 150 ms), was not long enough to measure the baseline of the pupil diameter independent of the diameter at the offset of the first phase. So the measured baseline at the cue phase, did not merely reflect the level of attention, making it difficult to understand the contribution of attention as an independent factor on the Target Look.

### 2.7.5. Interim Conclusion

Several processes in the cue phase were common in the FMF and the FFF conditions. Among these processes were attention, association of the cue face with the voice and memorizing the association. Considering that the contribution of these processes in the pupil dilation were not predictive of the Target Look in the FFF condition, we could speculate that the correlation between the Target Look and the pupil dilation at the second phase, in the FMF condition, could not be due to these common factors alone; on the other hand the results suggests that the responsible contributor to this interaction should essentially be specific to the computational processes specific to the FMF condition.

Moreover, we can argue that the dilation component elicited by the change in voice, from female in phase one to male in phase two in the FMF condition, could not alone be predictive of the Target Look. The reason is that we only passed trials with a minimum Valid Look of 80%. If we associate the ratio of Valid Look in the cue phase with attending the male voice in this phase, then due to 80% threshold for inclusion in all the trials, subjects were exposed relatively the same amount of time to the male voice (between 80% to 100%), and hence the pupil dilation triggered due to the novelty of the male voice, could not independently correspond to the large variation in the Target Look (between 0% to 100%).

Therefore, we can suggest that the more probable explanation for the observed correlation between the Target Look and the pupil dilation is due to the disjunctive inference process in general, and exclusion process in particular. However, more controlled conditions are needed to further establish this suggestion.

Finally, we are working on concluding the pupil profile of the first and the third phases, to further advance assessing this suggestion.

## 2.8. Discussion

Exploring the functionality of logical operators in early infancy, has been considered to be a formidable task, partly due to relying on language-based tasks that are based on comprehension of the words referring to negation and logical operators and partly because of the nature of logical inferences that are combinatorial in essence and subsequently need a minimum short-term memory capacity to maintain the several alternative mental representations and the provided cues to resolve the ambiguity. These two factors have slowed down the research on logical inferences, particularly in pre-verbal infants.

In this work, we designed a paradigm based on innate tendency of infants in exhibiting a preferential look at a talking face. In our design we presented two static genderless faces at the beginning of each trial while the participants were hearing a female voice, then in the 2nd phase, we presented one of the two faces with a male voice, in this phase infants could infer that the presented face is male and therefore exclude the possibility that it can also be female; so when in the final phase (test phase), we presented both of the faces again with the same female voice, participants could through disjunctive inference, infer that the face that was not presented in the cue phase, can be the female.

In experiment 1, we showed that infants looked longer to the possible face more than the impossible face. In experiment 2, we speculated that this pattern could be explained due to the fact that the cue face was presented for a longer time than the other (target) face, and hence infants showed a tendency for looking at the target face because it was partially novel. Therefore in this experiment, we kept the same settings but removed the male voice in the cue phase and showed that the preferential look at the target face nullified. This experiment ruled out the possibility that the performance in Experiment 1, could be caused by the higher exposure time to the cue face and provided support that the male voice had an essential role in the observed performance. In Experiment 3, we argued that if the infants discarded the first phase, the target face and the female voice both turn into novel stimuli and the infants could potentially utilize a novel to novel mapping mechanism to associate the female voice to the target face. However, by eliminating the first phase, infants failed to perform the association and the performance was at the chance level. This provided an evidence against the capability of infants in using novel to novel mapping in this context, and subsequently confirmed that the ambiguous phase was crucial in the inference process.

The role of the ambiguous stage in the performance, provided an additional support that infants probably made the face-voice association through disjunctive inference.

Furthermore, these results point out that the novel to novel mapping mechanism may not be available to the 10-month-olds in the context of faces and voices. On the other hand, Mather and Plunkett (2010) have reported that the 10-month-olds look more at a novel object, if they hear a novel label, suggesting that the novel-novel mapping might be available in the context of word learning at this age.

More experiments are needed to assess the availability of disjunctive inference in the context of word learning in the first year of life, particularly to rule out the novel to novel mapping or N3C mechanisms by familiarizing the infants to the labels and objects, as we attempted to do for the faces and voices in our paradigm.

However, so far considering our results together with the results reported by Halberda (2006) and

Mather and Plunkett (2010), we can suggest that the N3C mechanism in particular and novel to novel mechanisms in general may be in service of word learning through the first year of life. However, disjunctive inference might have a different developmental time course, being present at least in the domain of face-voice association before the first birthday but may be utilized in word learning contexts, later toward the 14-18 months of age.

Furthermore, the pupil profile and the results of an adult group, suggested a temporal framework for the inference process. We argued that probably the exclusion step of the disjunctive inference is taking place in the cue phase, the two reasons are briefly as follows:

1. In the adult group (Experiment 5), we showed that prior to the onset of the voice in the test phase, the participants exhibited a preferential look at the target.
2. We showed that in the trials with a higher performance, the participants exhibited a higher dilation in the pupil diameter during the cue phase, an interaction that was not observed in the FFF condition. Adult participants tested on the same task, exhibited the same interaction between the amplitude of pupil dilation in the cue phase of a trial and their performance in that trial.

The interaction between the pupil dilation and the performance suggests that the cue phase is computationally more demanding for the participants *only* if they were performing the inference process (performance as measured by showing a tendency to look at the target). We argued that this computational process can be the exclusion step of the disjunctive inference.

However, it should be noted that to confirm the speculation that the exclusion occurs in the cue phase, the experiment should be performed in a more controlled condition and the pupillometry data should be analyzed more comprehensively to completely rule out the alternative explanations for the pupil correlates of the performance.

Finally, these results should be taken with caution especially to generalize to other contexts. One main reason is that in the context of faces and voices, mutual exclusivity bias may probably be utilized by the infants as the base of the exclusion stage of the inference process.

We are running another experiment in which the mutual exclusivity bias (Markman, 1989) is not an inherent part of the stimuli, but its analogy is introduced in a familiarization block at the beginning of the experiment, and we asked if we could observe similar results as we observed for the faces and voices. In this experiment with the exact same framework of the previous paradigm (the FMF condition), we substituted faces with music boxes and male and female voices with easily discriminable sequences of tones (instead of male voice, we introduce a sequence of tones constant in pitch and instead of the female voice, we introduced another sequence of tones, rising in pitch). In this experiment we added 6 trials of familiarization, in which each music box was presented with one tone sequence. In this way, we could attempt to represent to the infants a one to one association between the boxes and the tones, that would subsequently help them to use mutual exclusivity to exclude one alternative. We believe the stimuli used in this experiment could be an extension to the previous stimuli, letting us to investigate the role of mutual exclusivity in the inference process.

# Chapter 3. Integration of Attentional Shift with Icons

## 3.1. Introduction

Visual symbols are widely used cues in the human societies which all are consisted of a simple icon/figure with a reference to a concept (DeLoache, 2004), the concepts can be concrete items or they can be functions or signals. Among the simplest functions that symbols represent, are the functions related to reorientation of attention.

To assess the integration of a function with an abstract icon in early infancy, in a set of studies we we focused on the cues that are related to reorientation of attention. We performed the studies from two complementary perspectives; in one hand we addressed the integration of this function with icons that their visual features could be informative (such as arrow icons) and on the other hand we addressed the capability of infants in associating an attentional shift to arbitrary non-informative icons.

One of the few symbols involved in reorientation of attention is an arrow, which is particularly interesting due to its consistent symbolic meaning across different cultures.

Arrow is considered to be an endogenous cue of attention in contrast to exogenous cues of attention. The exogenous cues are such that a cue, appears at the location where the attention will be redirected to (due to appearance of a target); for instance, a circle appears on one side of the screen that signals the appearance of a target in that location. So the attention of the subject will be redirected to that location before the target onset, due to the fact that the cue appeared in that location. The exogenous cues of attention are considered to be merely visually driven, through a

bottom-up process. On the other hand, arrows and other indirect cues of attention are referred to as endogenous cues, since they do not directly cue the attention, but the participants have an understanding of the meaning of the symbol, and make a volitional reorientation of attention in the direction of the cue (Jonides & Somers, 1977; Posner 1980; Jonides & Irwin, 1981).
Furthermore the shift of attention can be overt, followed by the movement of the eye gaze to the cued location or it can be covert, i.e. having attention in a location without moving the eye-gaze toward that location. The covert orientation of attention facilitates the overt orientation of attention in terms of reaction time and response time in detecting the target that appears in the cued location (Posner, at al, 1985; Posner & Petersen, 1990; Posner 1980).

An important effect to note is that after the shift of attention to a cued location, there can be a prolonged disengagement from that location. In humans from the first days of life, it has been observed that there is an inhibition in returning back to the cued location, after disengaging from that location. Posner and Cohen observed an inhibition of attention to the cued location from 200-300 ms to at least 1500 ms after the disengagement (Posner & Cohen, 1984). They referred to this effect as Inhibition of Return (IOR), and they further discussed that this effect is only elicited by exogenous cues and not by endogenous cues (endogenous cues only elicit this effect, if before the appearance of the target the participant make an overt saccade to the location of the target by following the direction of the cue), (Posner,  Rafal, Choate & Vaughan, 1985). Therefore, as a general scheme, exogenous cues facilitate the shift of attention to the congruent side for a duration of roughly 150 ms and then this effect reverses to inhibition of return (see Klein, 2000, for a review).

In the literature of infancy, the same patterns of covert and overt shifts of attention have been observed. Johnson and Tucker (1996), showed that 4-month-old infants can exhibit a covert shift of attention, after receiving a training to anticipate a target on the opposite side of an exogenous cue. In another study Johnson (1994), extended the cue-target delay interval to 100 ms and 400 ms and he observed that due to the cue, there is a facilitation in reaction time in saccading to the target, that is suggestive of a covert shift of attention. Moreover, Johnson showed while in adults IOR occurs after 200 ms, in 4-month-old infants, IOR can be observed only after 700 ms; and in newborn infants this amount increases to 1000 ms (Valenza, Simion & Umiltà, 1994).

Contrary to the reorientation of attention by exogenous cues that can be observed in newborn infants; the endogenous cues possibly need a further conceptual representation in order to redirect the attention of the subject. For instance pointing has been reported to direct the attention of infants from 4.5 months of age (Rohlfing, Longo & Bertenthal, 2012), however, infants fail to follow the pointing cue, if it is static. They only reorient their attention if the pointing is dynamics and accompanied by movement of the hand.

In the case of arrows, to our knowledge there have not been any study in the first months of life. A study by Ristic, Friesen, and Kingstone in 2002, suggested that toddlers from the age of 3 to 5 years, show a facilitation in reorientation of their attention, if the appearance of a target is accompanied by an arrow icon, and in another study, Jakobsen, Frick, and Simpson (2013) showed that infants failed to comprehend arrows within an age range of 13 months to 41 months (although they only tested 21 infants in the whole age-range), and they argued that not until 4-5 years, infants can follow the direction of an arrow based on its conceptual content. In this experiment the infants received verbal instructions on finding a toy, hidden either on the left side or on the right side of the cue. The authors concluded that the comprehension of arrow, emerges in a later phase of development, through ostensive learning of the symbol.

However, we asked if in the first months of life, the icon of arrow due to its unique combination of perceptual features can trigger a covert shift of attention. We asked this question to assess if the perceptual features of the arrow icon, can potentially consolidate an association between attentional shift and the icon, an effect that potentially can lead the infants to construct a conceptual representation from this abstract icon, later in the toddler years.

To assess our hypotheses, we sought to test 4-month-old infants, because of the following reasons:

**1.** At this age with a high likelihood infants have not been introduced to arrow icons, through heuristics.

**2.** It has been shown that the age of 4 months is the earliest age on the course of development, in which infants can show anticipatory looks toward the upcoming targets and can readily disengage from attending a fixation before shifting their attention to a target (Johnson, Posner & Rothbart, 1991; Haith, Hazan & Goodman, 1988).

When closely attending an arrow icon, a naive observer may find the head of the arrow, visually more salient than its tail, due to the sharp edges of the triangle at the head of the icon. This gradient can induce a higher load of attention on the head compared to the tail. This asymmetric distribution of attention on the icon, in an extended (large enough) arrow, can mimic an exogenous reorientation of attention, and could be sufficient in steering the attention of subjects toward the target at the head of the icon, or inversely induce an inhibition of return in the direction of the cue, depending on the delay in appearance of the target (Johnson & Tucker, 1996).

Since we wanted to specifically address the arrow as an endogenous icon, we set to avoid the potential immediate exogenous effects due to low level attentional asymmetries of this icon.

To satisfy this concern, a crucial consideration in our paradigm was to include a delay of 1 second between the disappearance of the arrow and appearance of the targets, and we further reinforced the participants to fixate on an attractor at the center of the screen during this 1 second delay period. So we could speculate that if the infant after the delay period could reorient her attention in the direction of the preceding arrow; this could only be due to a holistic representation of the entire icon, rather than an exogenous cuing of attention.

We hence predicted that if a predisposition exists in the cognition of young infants at the age of 4 months to accommodate them in reorienting their attention based on an arrow icon, we should observe a tendency for following the direction of the arrow, even after 1 second of delay from the disappearance of the cue.

We furthermore, decided to contrast the performance of 4-month-old infants with an older age group, to track the developmental alterations in the feature-based reorientation of attention. For this reason, we tested another group with an age range of 8.5 to 9.5 months, younger than the infants tested in the study by Jakobsen et al (2013).

## Exp.6. Arrow as an Informative Cue

### 3.2.2. Experimental Design

To address if young infants can shift their attention due to perceptual features of arrows, we designed an experiment in which each trial consisted of three phases. In the first phase participants were presented with a horizontal arrow that was pointing to the left or to the right with a random order (see Figure 3.1).

After the presentation of the arrow, a central attractor was presented for a duration of 1 second to reorient the attention of the participants to the center of the screen. Our choice of 1 second, was to choose a value over 700 ms, to assure that there is no facilitation due to a possible exogenous cuing of attention induced by the icon. This duration was long enough to minimized the potential effects that could play a role due to a possible inhibition of return (Johnson et al., 1991; Johnson, Tucker, 1996).

In the last phase of the trials, two identical toys appeared simultaneously on both sides of the screen (see Figure 3.1). In this phase, we predicted that if the attention of the infants were cued congruent with the direction of the arrow, when the two toys appeared, the subjects would look more at the toy congruent with the direction of the arrow, compared to the one at the incongruent side.

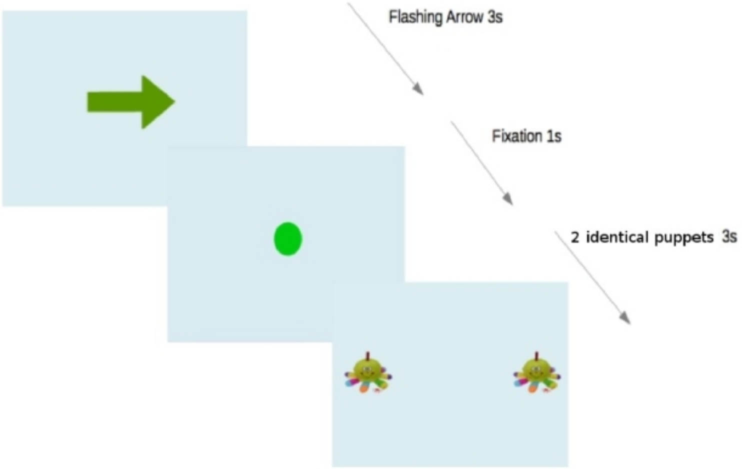

**Fig. 3.1.** Time line of a single trial, showing the stimuli at the phases of arrow presentation, fixation and toy presentation. The duration of the three phases were set to 3 s, 1 s and 3 s correspondingly.

### 3.2.3. Stimuli

In the first phase of each trial, we presented an arrow with a blue background, with an screen resolution of 1280 by 1024 pixels, the arrows were either pointing to the left side or to the right side of the screen (see Figure 3.2). The arrows were confined to an area covering the central 40% of the width of the screen.

In each trial, the arrow was presented for a duration of 3 seconds. To have the arrow attracting more attention, during the first 2 seconds of arrow presentation, we set the arrow to blink 3 times, during each blink the arrow was disappearing for 300 ms, and for the last 1 second of the presentation, the icon was static on the screen.

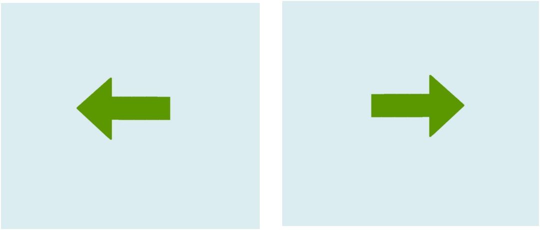

**Fig. 3.2.** The two arrows pointing to the left and to the right. Both the left and the right arrows, where confined within the central 40% of the o horizontal axis, so there was 30% distance from the margins of the screen to the two ends of the arrow. The base of the triangle at the head of the arrow, was covering 30% of the whole length.

The presentation of the arrow was followed by a central attractor. The attractor was meant to redirect the attention of the participant to the center of the screen. For this purpose we presented a green circle, with a diameter of 65 pixels, which during the 1 second period of its presentation, expanded to a diameter of 80 pixels, an increase of 20% (see Figure 3.1).

Immediately after the disappearance of the central fixation, we presented the test phase which was two identical toys shown on the two sides of the screen at the same time.

The toy at each trial was randomly chosen from a set of 8 images. The toys were selected from public domain websites. The toy images were all confined to an area of 300 by 300 pixels, and they were positioned approximately 60 pixels form the edge of the screen; the toys were centered at the horizontal mid-line of the screen.

In all the trials the background color was set to blue, as shown in Figure 3.1. We chose to have the two toys identical to each other, in order to avoid any preferential look to any of the toys, due to the relative saliency of their visual features.

At the onset of the toy phase a short bip sound was being played from two load speakers located on the two sides of the monitor, to attract the attention of the infants.

## 3.2.4. Participants

The participants were tested in two groups of 4 months and 8 months.

### 3.2.4.1. Four-month Group

Thirty-three healthy monolingual Italian infants were tested in this group, from 15 to 19 weeks, with an mean of 17.5 weeks, SD = 1.3 weeks,

Ten subjects were excluded due to lack of attention and fuzziness, and not passing enough number of trials.

### 3.2.4.2. Eight-month Group

Twenty-six healthy monolingual Italian infants were tested in this condition, from 34 to 38 weeks, with a mean of 36.5 weeks. SD = 1.4 weeks,

Four subjects were excluded due to lack of attention or crying at the beginning of the experiment.

All the infants were full term and had an APGAR number over 7 out of 10. The participants were recruited from Trieste, Italy, through sending invitation letters to a random selection of parents, whose babies fit the age range of the study. The parents were all reimbursed for attending the experiment and the infants received a certificate of attendance.

### 3.2.5. Apparatus

The visual stimuli were presented on a 15” monitor with a screen resolution of 1280 pixels by 1024 pixels, the screen was equipped with Tobii eye-tracker T120 that recorded the location of gaze at 60 Hz.

The participants sat on a fixed chair in the lap of their parents, approximately 70 cm from the monitor, with their eye sight being in line with the center of the monitor. Parents were asked to wear opaque sunglasses, so they could not interfere with the performance of the infant and we could further be sure that the eye tracker was only registering the eyes of the infant.

The experimental booth was completely darkened, so the only light was emitted by the screen.

### 3.2.6. Procedure

Before the experiment, the subjects passed a 5 point calibration protocol, included in the Tobii studio. After the calibration infants watched an animation for 3 seconds for the purpose of engaging their attention on the screen. After the animation we presented 8 trials to the participants to control their side bias. Each trial consisted of presentation of two identical toys (the same images as used in the test blocks) on the two sides of the screen. The images were shown for a duration of 3 seconds.

After this phase, we presented the test block with 24 trials. In this phase, each trial consisted of 3 phases:

1. Presentation of the arrow icon with a random left or right direction, for a duration of 3 seconds
2. Fixation lasting 1 second
3. Presentation of two identical toy images for a duration of 3 seconds (see Figure 3.1)

Altogether a single trial lasted for 7 seconds. If the participants were distracted from attending the screen, a few seconds of an animation from Pixar animation production company was shown, until they reoriented their attention back to the screen.

### 3.2.7. Scoring

The coordinates and the quality of the gaze points were collected for further analysis. All the analysis were done using Mathworks Matlab software (v.2015b), Statistics and Machine Learning Toolbox™ of Matlab.

#### 3.2.7.1. Inclusion Criteria

##### 3.2.7.1.1. RoI Settings

Each trial consisted of three phases. In all the phases, the stimuli always appeared at the horizontal mid-line of the screen. In order to eliminate the gazes that were away form the stimuli, we omitted the data points if they were within the upper and lower 20% of the vertical range.

Over the different phases of a trial, we defined different Regions of Interest (RoI) based on the spatial extent of the stimuli.

**Phase 1 RoI:** In the phase 1, the arrow icon was presented at the center of the screen. The arrow was positioned 30% from the left and 30% from the right edge of the screen. Therefore we defined an RoI with an extent of 25% from the edges of the screen, and considered a gaze valid if its coordinates were within the 50% of the horizontal range.

**Phase 2 RoI:** At the 2nd phase, we presented a fixation point; to make sure the subjects were gazing at this point, we defined an RoI confined to the central 20% of the horizontal range.

**Phase 3 RoI:** At the 3rd phase two identical toys were presented symmetrically on the two sides of the screen. All the toys were positioned within the 30% of the left and the right margins of the screen; therefore in this phase, we defined two RoIs, one covering the left 35% of the screen and the other, covering the right 35% of the screen.

#### 3.2.7.1.2. Valid Look

We defined the Valid Look at each phase equivalent to the ratio of time the subject was looking inside the valid RoI(s) during the phase, divided by the total duration of the phase.

Valid Look = Time looked in valid RoI(s) / Total duration of the phase

A gaze point was regarded as valid, if it was inside an RoI and having a quality score registered by the eye-tracker between 0-1 out of 4 (4 is equivalent to poor quality). The quality score was assigned to each point automatically by the eye-tracker if the data of both eyes was available and consistent.

#### 3.2.7.1.3. Side Bias Control

We had two assessments to measure the side bias of the participants;

1. Measuring the side bias during the first 8 trials of the experiment (the RoIs in these trials were defined identical to the test phase of the test trials (see section 3.2.7.1)).

2. Calculating the side bias in the test block, consisting of 24 trials.

In both of the assessments, the side bias was measured based on the overall ratio of looking at the left RoI. If the ratio was more than 85% or less than 15%, we considered a participant as having side bias.

Left Bias = Overall look at the left RoI / (Overall look at the Left RoI + Overall look at the Right RoI)

We generated the side bias list based on the test block of the experiment; since more trials were included in this analysis (24 trials versus 8 trials) and therefore this measure was more reliable.

In total 6 participants were excluded from the 4-month group and 1 participant was excluded from the 8-month group due to side bias.

### 3.2.7.2. Measuring Performance

#### 3.2.7.2.1. Target Look

For measuring the performance, we first defined the correct RoI as the RoI which was congruent with the direction of the arrow. We then defined the Target Look as:

Target Look = Number of time bins within the correct RoI / Total number of valid time bins

**Time bin:** Since the sampling rate of the eye-tracker was at 60 Hz, we measure time, in the units of a time bin, equivalent to 16.67 ms, representing a single gaze point registered by the eye tracker. A valid time bin refers to a valid gaze point (a gaze point within a valid RoI and with a passed quality).

### 3.2.7.2.2. First Fixation

We defined First Fixation as the ratio of first fixations congruent with the direction of the icon. We defined a *fixation* as a continuous look within an RoI, for a minimum duration of 100 ms (equivalent to 6 time bins), with no missing data points.

### 3.2.7.3. Trial Validation and Subject Validation

Our inclusion criteria were as follows:

1. We considered a threshold of 60% as the minimum ratio of Valid Look in each phase.
2. A trial was eliminated from the analyses, if one of the three phases of the trial had less than 60% valid gaze.

A subject was eliminated from the analyses, if providing less than 5 valid trials.

### 3.2.8. Results

We calculated the Target Look and First Fixation as the two measures of performance, as defined in 3.2.7.2. In the 8-month group, 308 trials from 21 subjects were included, and in the 4-month group, 129 trials from 17 subjects were included.

We calculated the Target Look, over the 3-second duration of the 3rd phase, and averaged over the subjects. This analysis in the 4-month group resulted in $\text{Mean}_{4m} = 0.58$, $SE = 0.036$, and in the 8-month group, $\text{Mean}_{8m} = 0.49$, $SE = 0.01$, with 0.5 being the chance level (see Figure 3.3).

Furthermore, we calculated the mean of Target Look across subjects, separately in the trials that the cue was pointing to the left and the trials that cue was pointing to the right. The results were as follows: $\text{Mean}_{4m-Left} = 0.58$, $\text{Mean}_{4m-Right} = 0.59$ in the 4-month group and $\text{Mean}_{8m-Left} = 0.45$, $\text{Mean}_{8m-Right} = 0.57$ in the 8-month group, with the chance level being at 0.5.

The distribution of Target Look across subjects followed a normal distribution, therefore we used a *t*test to compare the distributions against the chance level at 0.5; the test in the 4-month group resulted in $t_{4m}(16) = 2.8$, $p = 0.013$, and in the 8-month group, resulted in $t_{8m}(21) = 0.50$, $p = 0.89$.

Furthermore, we calculated the First Fixation measure, the average across subjects was $\text{Mean}_{4m} = 0.16$, $SE = 0.089$, in the 4-month group, and $\text{Mean}_{8m} = 0.030$, $SE = 0.056$, in the 8-month group, with the chance level being at 0 (see Figure 3.3).

Since the population of the first fixations were binomial, either 1 for correct, or 0 for incorrect; we calculated the 95% confidence interval of the three populations based on the Binomial Proportion Confidence Interval:

$$\hat{p} \pm z\sqrt{\frac{1}{n}\hat{p}\left(1-\hat{p}\right)}$$

With *p* being the estimated success from the sample, *n* the number of data points and *z* equivalent to 1.96 for a confidence interval (CI) of 95%. The calculated 95% CI was 0.09 and 0.05 in the 4-month and 8-month groups respectively (see Figure 3.3).

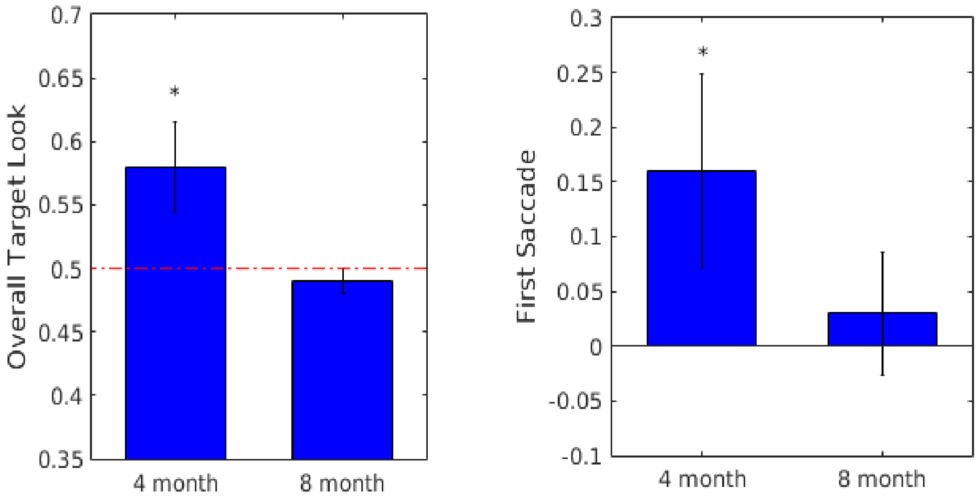

**Fig. 3.3. Left Panel,** the Target Look across the groups of 4-month and 8-month old infants, the horizontal dashed line shows the chance level at 0.5. **Right Panel,** the measure of First Fixation across the two groups. The errorbars in the left panel, indicate one standard error from the mean and in the right panel indicate 95% of confidence interval.

The results of the two measures suggest that contrary to our expectations, at the age of 8 months infants were failing in orienting their attention congruent with the direction of the arrow, however, in the 4-month group, there was a significant Target Look and a significant First Fixation toward the congruent side.

To investigate the duration of looking at the target from the onset of the test phase, we further calculated the dynamics of Target Look as a function of time over the 3-second duration of this phase. For this purpose we calculated the ratio of Target Look, averaged across participants for each time bin of 16.67 ms, over the whole span of the test phase. The results are shown in Figure 3.4.

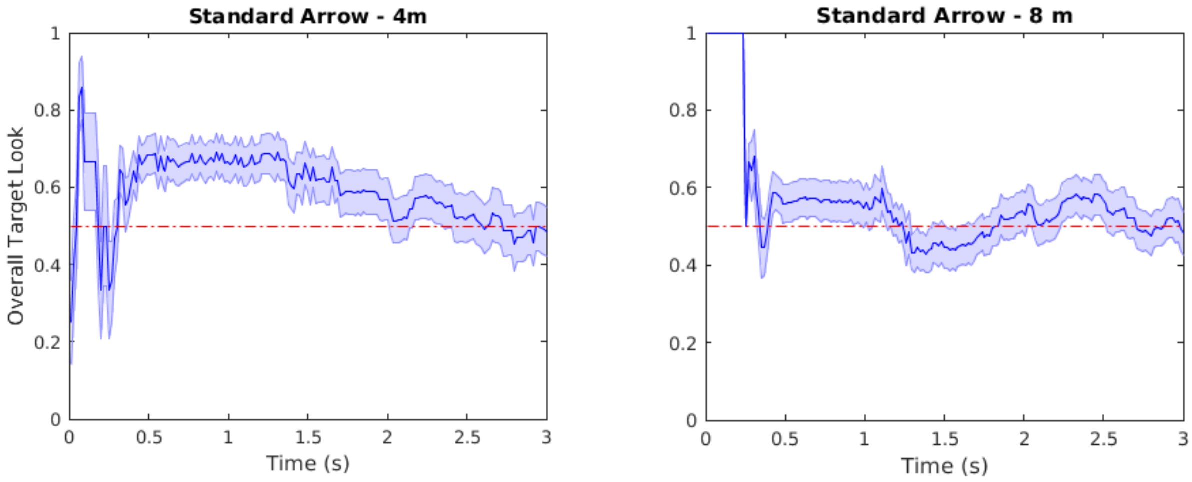

**Fig. 3.4.** The temporal dynamics of Target Look at the 3rd phase, in the 4-month **(Left )** and 8-month **(Right)** groups. Only in the 4-month group, there was an initial tendency for looking at the congruent side. The first few hundred milliseconds were noisy due to lower number of data points. The errorbars indicate one standard error from the mean, the horizontal dashed line indicated the chancel level at 0.5.

As Figure 3.4 demonstrates, only in the 4-month group we observed a significant deviation from the chance level toward the direction congruent with the direction of the arrow, and this tendency continued until around 2 seconds after the toy onset.

### 3.2.9. Interim Conclusion

In this experiment the results suggest for the first time, that infants at 4 months of age, cue their attention toward the direction of an arrow after 1 second of delay. This evidence provides a support for the hypothesis that arrows merely due to their perceptual features, can be considered as an endogenous cue of attention by young infants. It is very unlikely that at the age of 4 months, this effect could be learned through ostention; however, the more plausible scenario is that the combination of the perceptual features of the arrow triggers a covert cuing of attention that lasted for more than 1 second after the disappearance of the arrow.

In the next experiment, we will explore if other icons that share some features with arrows, also trigger such a reorientation of attention and what could be the minimum visual feature that can cue the attentional.

Furthermore the fact that we did not observe an effect in the 8-month group, opens two possibilities:

First. The arrow icon still redirects the attention of infants, but because the two toys that appeared on the screen were identical to each other and they appeared at the same time, may have disengaged them in following the direction of the cue, since they immediately understand that the cue is not *informative*. However, in the 4-month group, the informativeness of the cue may not be assessed by the subjects.

Second. The 8-month-olds are not sensitive to the perceptual features of the arrow, and since the icon has not been conceptualized yet, the failure of attentional shift based on the perceptual features subsequently lead to a failure in following the icon. However, later in the development the function of an arrow is heuristically delivered to the infant, introducing the icon as a symbol, regardless of its perceptual features.

We did not further explore these two scenarios at the age of 8 months beyond this point. However, based on the failure observed in other studies for comprehension of arrows until the age of 3-4 years, we speculate that the second hypothesis for the failure of 8-month-olds is the more plausible one.

## Exp.7. What is Especial About an Arrow?

### 3.3.1. Introduction

An arrow icon is consisted of 2 parts, a triangle and a rectangle. The triangle introduces a sharp edge at one end of the icon in contrast to opposite end. Furthermore, it introduces a gradient of visual saliency toward the head of the icon (since there are acute angles toward the head). Each of these two features alone may have the potential in cuing the attention of the infants.

In a follow up of experiment 6, we sought to explore how visual features of the arrow may lead infants to reorient their attention, and what are the minimum features that can accommodate this reorientation. To persuade this goal, we aimed to present arrow-like icons instead of a standard arrow, with specific modifications in each icon.

For this purpose we tested infants at the age of 4 months with a similar paradigm; however, as a substitute to arrow, we introduced three new icons that represented three types of modifications (see Figure 3.5).

**1. *T* Icon:** The *T* icon was introduced to investigate if the observed tendency of the infants in following the direction of the arrow, could be explained due to the gradient of visual saliency from one end to the other end of the icon. For this purpose we replaced the triangle at the head of the arrow with a vertical rectangle. This icon in terms of the distribution of its salient features, mimics that of an arrow (see Figure 3.5).

**2. *Balanced* Icon:** We decided to explore the role of gradient in saliency in a different way. We introduced the *Balanced* icon that mimics an arrow, but the number of acute angles are two at the tail of the icon, compared to one in its front (see Figure 3.5). In this icon, it is not the case anymore that its head is more salient than its tail.

**3. *Triangle* Icon:** We further included an acute triangle to investigate if this icon alone can reorient the attention toward its head.

By studying the performance elicited by these three icons, we could try to characterize the minimum combination of features in an arrow that makes it a potential abstract symbol to cue the attention of the observer.

### 3.3.2. Experimental Design

We designed the paradigm similar to experiment 6. Each trial consisted of three phases:

1. Presentations of the cue with a random left or right direction.

2. Presentation of a dynamic attractor as a central fixation for a duration of 1 second.

3. Presentation of two identical toys on the two sides of the screen at the same time (see Figure 3.6).

In the last phase, if the participants were cued by the icon, we expected to measure a higher ratio of looking time at the toy in the side of the screen congruent with the direction of the icon.

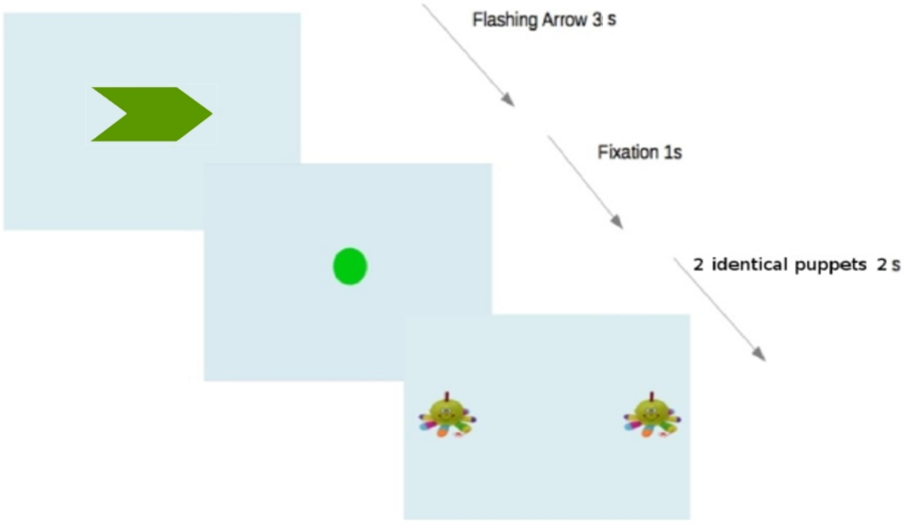

**Fig. 3.6.** The design of a single trial showing the three phases. In phase one, the stimuli was presented for a duration of 3 seconds. In phase 2, the central attractor appeared for 1 second, and in the third phase, two toy images appeared on both sides of the screen at the same time. This phase, contrary to experiment 6 that lasted for 3 seconds, in this experiment lasted for 2 seconds.

In order to mimic the design of experiment 6, we presented the different icons in a block design.

Within one block the infants were presented with only one type of cue.

### 3.3. 3. Stimuli

The stimuli at the phase 2 and phase 3 of the trials were identical to the stimuli design in experiment 6 (see section 3.2.3). The only difference was that 8 toys were used in the test phase of experiment 6 however, in this experiment we used 30 toys; all designed with the same criteria as explained in experiment 6. In the whole experiment, the background and foreground colors were the same as in experiment 6 (see section 3.2.3).

The *T* icon was designed with the same length as the arrow in experiment 6 (see Figure 3.7).

The *Balanced* icon, was a rectangle with the same length as the *T* icon, we then joined it to the base of a triangle, and subtracted the area of that triangle from the opposite end of the icon.

The *Triangle* icon was the same size as the triangle used to design the *Balanced* icon.

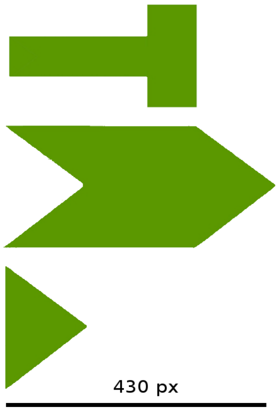

**Fig. 3.7.** The three icons used in the experiment. *T* (upper icon), *Balanced*, (middle icon) and *Triangle* (Lower icon).

In the cue phase, the icons were all positioned at the center of the screen during the presentation. Furthermore, similar to experiment 6 we made the icons flashing in order to make them more attractive for the participants. During the 3-second period of the 1st phase the icon disappeared twice, each time for a duration of 330 ms. The first disappearance was after 600 ms from the stimuli onset, and the second disappearance was at 2000 ms from the stimuli onset.

### 3.3.4. Participants

Forty healthy monolingual Italian infants at the age of 4 months were tested in this condition, with an age range between 14 to 18 weeks, with a mean of 16.5 weeks, SD = 1.2 weeks.

Six subjects were excluded from the analysis due to lack of attention and fuzziness and therefore not passing enough number of trials.

The infants had an APGAR number over 7 out of 10 and they were declared by their parents to be full term. The participants were recruited from Trieste, Italy, through sending invitation letters to a random selection of parents, whose babies fit the age range of the study. The parents were all reimbursed for attending the experiment and the infants received a certificate of attendance.

### 3.3.5. Apparatus

The apparatus was identical to the apparatus described in experiment 6 (see section 3.2.5).

### 3.3.6. Procedure

The calibration procedure was identical to the procedure described for experiment 6 (see section 3.2.6).

After a successful calibration, infants watched a short animation for 3 seconds and then proceeded to the experiment.

The experiment consisted of 4 blocks, the first three blocks were test trials, each block assessing one icon, and the last block was for side bias control. Each of the test blocks consisted of 10 trials, with presentations of only one of the icons. The direction of the cues were counterbalanced in a random order. The order of the three blocks were randomized across subjects.

The last block consisted of 6 trials, with an identical trial design to the test trials. However, in these 6 trials the stimulus at the first phase was always a horizontal rectangle with the same dimension as the *T* icon. We included these extra 6 trials, in order to have an extra measure for the side bias control. Therefore, in total we had 30 test trials and 6 trials for the side bias control.

### 3.3.7. Scoring

For each of the three cues we performed a separate analysis. Every participant could provide a maximum of 10 trials for each analysis (equivalent to number of trials in a block).

#### 3.3.7.1. Inclusion Criteria

The constraints to include a trial was identical to the constraint explained in experiment 6 (see section 3.2.7, considering a threshold of 60% Valid Look for each of the three phases as the inclusion criteria).

We included a subject in the analyses of a cue, if provided at least 1 valid trials in the corresponding block.

##### 3.3.7.1.1. RoI Settings

The RoI settings were identical to the settings described in experiment 6 (section section 3.2.7).

##### 3.3.7.1.2. Valid Look

The calculation of the Valid Look was identical to the analysis described in experiment 6 (see section 3.2.7).

#### 3.3.7.1.3. Side Bias Control

We had two assessments to measure the side bias of each participant.

The first assessment was based on the 6 trials provided at the end of the experiment.

The second assessment was based on the test blocks. Similar to experiment 6, we considered the second assessment, since 6 trials were few compared to the trials provided in the three test blocks and moreover, some participants did not reach the end of the experiment to pass this last block of trials.

The criteria to exclude a participant based on side bias, was identical to the conditions explained in experiment 6 (see section 3.2.7). In total 12 participants were excluded due the to side bias.

### 3.3.7.2. Measuring Performance

The performance was measured identical to experiment 6 (see section 3.2.7).

### 3.3.8. Results

After applying the inclusion criteria, 21 participants were included in the dataset of *T* icon, 21 subjects were included for *Balanced* icon, and 22 subjects for the *Triangle* icon.

These subjects provided, 83 valid trials for *T* icon, 77 trials for *Balanced* icon, and 103 trials for the *Triangle* icon. We calculated the Target Look averaged over subjects for each of the cues. The results were as follows: Mean $_{T}$ = 0.41, SE = 0.027, Mean $_{\mathit{Balanced}}$ = 0.65, SE = 0.031, Mean $_{\mathit{Triangle}}$ = 0.63, SE = 0.032.

To assess if infants reoriented their attention in the same extent in both of the left and right directions, we calculated the average Target Look for each direction across subjects. The results were:

Mean $_{T\text{ - Left}}$ = 0.35, Mean $_{T\text{ - Right}}$ = 0.41 for the *T* cue, Mean $_{\mathit{Balanced}\text{ - Left}}$ = 0.69,

Mean $_{\mathit{Balanced}\text{ - Right}}$ = 0.61, for the *Balanced* cue and Mean $_{\mathit{Triangle}\text{ - Left}}$ = 0.57, Mean $_{\mathit{Triangle}\text{ - Right}}$ = 0.65 for the *Triangle* cue.

We further tested the normality of the population of Target Looks for each cue. All the three populations were confirmed to be normal based on Lillie test of normality, with p-values to reject the hypothesis of normality, equivalent to 0.5, 0.11 and 0.14 corresponding to *T*, *Balanced* and the *Triangle* cues (0.5 is the largest tabulated value the test can provide).

Therefore, we could use a one-way *t*test to compare the population of Target Looks corresponding to each icon, with the chance level at 0.5. The test resulted in t(20)= -0.20, p= 0.24 for the *T* icon, t(20) = 3.01, p = 0.006 for the *Balanced* icon and t(21) = 9.67, p = 0.014 for the *Triangle* icon (see Figure 3.8).

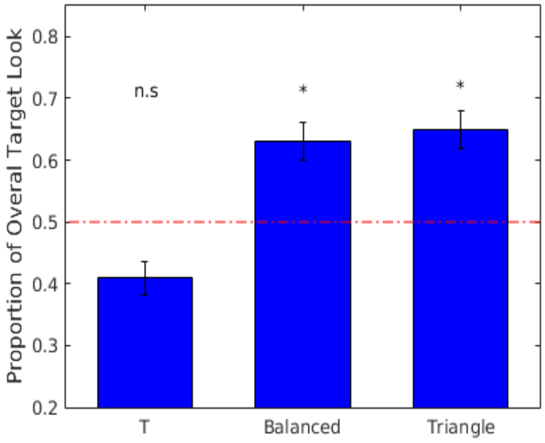

**Fig. 3.8.** The mean of Target Look averaged across subjects for the three icons. The errorbars indicate one standard error from the mean. The horizontal dashed line shows the chance level at 0.5.

We applied the Bonferroni-Holm Correction for the multiple comparison problem of having three *t*tests analyses. The correction resulted in declaring the p-value of *Balanced* and *Triangle* cues as significant and the *T* cue not being significant.

We then calculated the ratio of First Saccades to the correct RoI. The results are shown in table 1. We calculated the 95% confidence interval of the three populations based on the Binomial Proportion Confidence Interval:

$$\hat{p} \pm z\sqrt{\frac{1}{n}\hat{p}\left(1-\hat{p}\right)}$$

With *p* being the estimated success from the sample, *n* the number of data points and *z* equivalent to 1.96 for a confidence interval (CI) of 95% (see table 1 and Figure 3.9).

| | First Fixation (Chance level at 0) | 95% CI |
|---|---|---|
| *T* | -0.11 | 0.10 |
| *Balanced* | 0.07 | 0.076 |
| *Triangle* | 0.28 | 0.12 |

Table. 1. The average First Fixation across subjects and 95% confidence intervals for each icon.

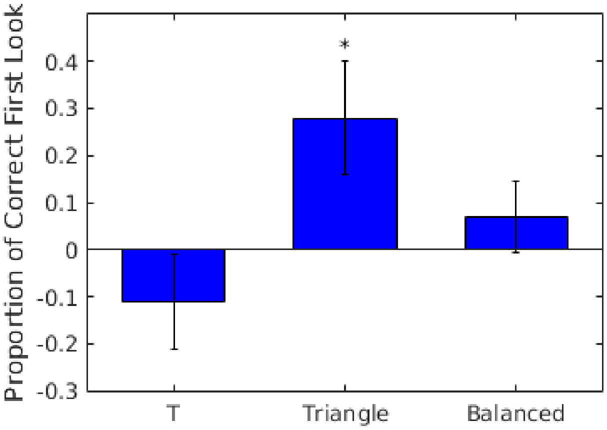

**Fig. 3.9.** The calculated measure of First Fixation across the three datasets. The errorbars indicate 95% confidence intervals. As can be seen, the First Fixation in the *Triangle* icon was significant toward the correct RoI, with the other two conditions, being only marginally significant.

Furthermore, to shed light on the duration of the cuing effect of the icons, we calculated the performance as a function of time, by averaging the Target Look across participants in each time bin (see Figure 3.10).

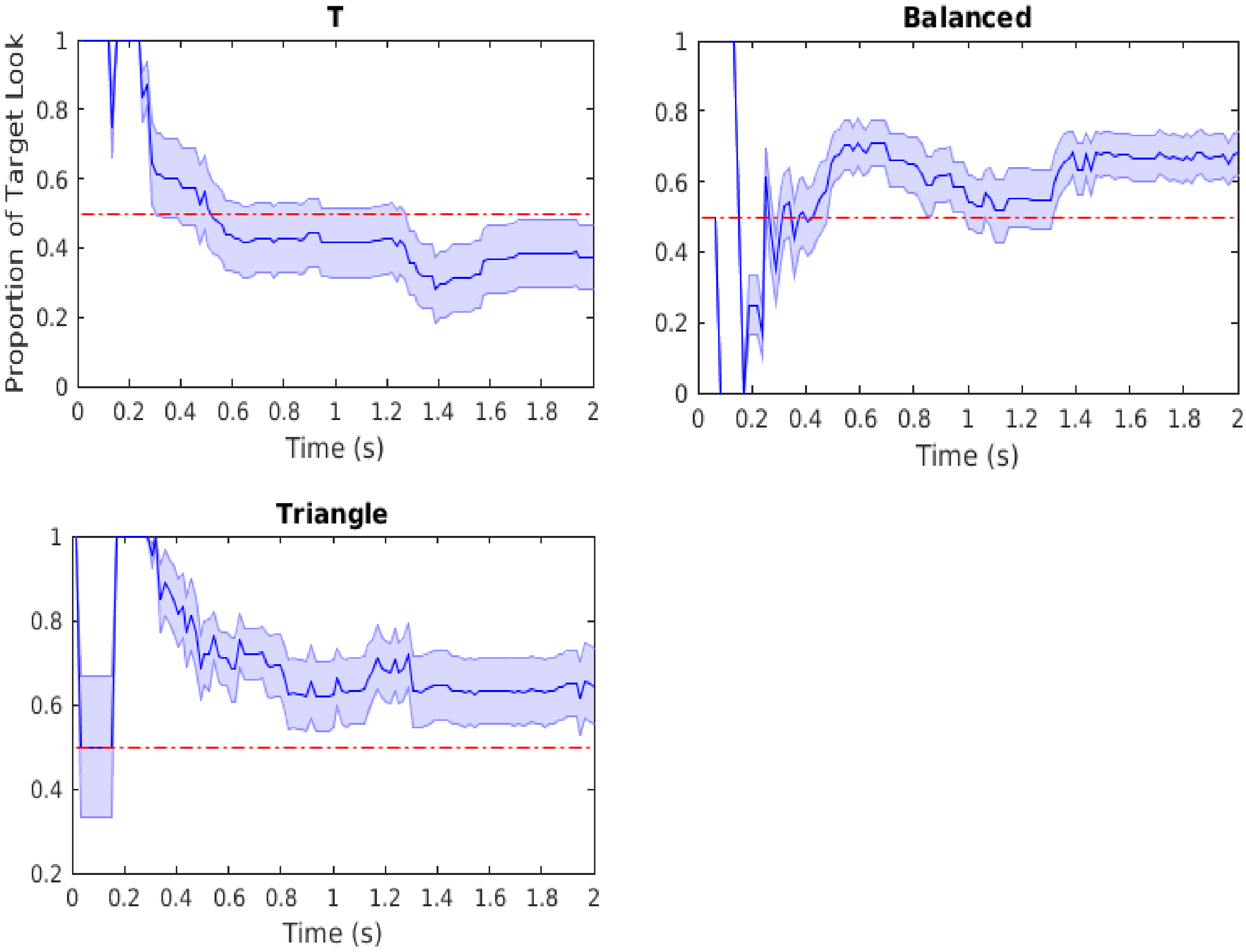

**Fig. 3.10.** The temporal dynamics of the Target Look over the 2-second duration of the 3rd phase. Each data point is an average across participants at each time bin. Since in the 2nd phase, the participants were fixating on the attractor ball at the center of the screen, at the onset of the third phase there were fewer data points and therefore a higher initial fluctuation. The horizontal dashed lines show the chance level at 0.5.

As the temporal dynamics of the *Triangle* and the *Balanced* cues suggest, the shift in attention toward the congruent side was preserved over the whole time span of the test phase. However, the Target Look of the *T* icon, showed a non-significant tendency to the opposite direction.

### 3.3.9. Interim Conclusion

The results of this experiment demonstrated that a *Triangle* as well as a *Balanced* icons, could trigger a shift in the attention of infants in 4-month-olds. Furthermore the infants did not significantly reorient their attention by the *T* icon toward the congruent side; on the other hand, the measures of the First Fixation and Target Look were observed to be below the chance level, although not significant.

These observations together with the results of experiment 6 for standard arrow, suggest that probably a triangle shape alone can induce a reorientation in the attention of infants. We based this speculation on the fact that a triangular area was the common feature shared with the standard arrow icon in experiment 6 and the *Triangle* and *Balanced* icons in this experiment. This speculation can further potentially explain the negative trend of performance observed for the *T* icon. Since the *T* icon can be considered as a partial inversed-triangle, with the base of this triangle being at the vertical bar of the *T* icon.

Moreover, as we noted earlier, the observed attentional shift could not be explained as an exogenous reorientation of attention, particularly due to the fact that the icons did not reach to the sides, and the consideration that the infants were reinforced to make a 1 second fixation on the center of the screen, at the interval between the cue and the targets. These two factors together, suggest that the icons were perceived by the subjects as a single unit that are associated with an attentional shift.

Further experiments are needed to confirm these results and address if any shape mimicking the surface gradient of a triangle can induce a reorientation of attention or the sharp acute edges of the triangle are playing the crucial role. A control experiment can be to replace the *triangle* icon with 3 dots, corresponding to the three acute edges of the triangle, and assess if 4-month-olds still redirect their attention.

Furthermore, we are analyzing the gaze pattern during presentation of standard arrow, *Balanced* icon and the *T* icon, to investigate if considering the accuracy of the registered data, we can comment on the differences in the looking pattern in the correct versus incorrect trials; with an aim to characterize how the salient areas of these icon, play a role in cuing the attention.

Finally, we note that at the moment we do not have enough evidence to provide a theory to explain how the visual features of a triangle can turn the icon into an endogenous cue and induce a prolonged covert reorientation of attention.

# Exp.8. Associating Attentional Shift to Arbitrary Icons

## 3.4.1. Introduction

In experiment 7, we showed that a set of arrow-like icons can tentatively trigger a reorientation of attention in 4-month-old infants. The observed reorientation of attention was based on the perceptual features of the icons.

As a next step, we sought to compare the capability of 4-month-olds with 8-month-olds in associating an attentional cuing function to arbitrary icons; icons that in contrary to arrow-like icons, do not have any perceptual feature that could be a hint for the infants to a reorientation of attention. To our best knowledge there are few studies in this context. It has recently been shown that 8-month-olds can learn to associate a cuing function to geometrical shapes and also they can choose to rely on a specific attribute of the cue (shape or color for instance) based on a second cue (Werchan, Collins, Frank & Amso, 2015), but in another recent study, Tummeltshammer, Wu, Sobel, and Kirkham (2014), showed that 8-month-olds can assign an attentional shift to a shape that morphs into an arrow-like icon, in a contingency paradigm, however the subjects failed to anticipate the location of the target, when they were presented with a novel direction of the cue. Moreover, in 4-month-olds, the study by Johnson and colleagues (1991), suggested an inability of the subjects to associate attentional cues to the congruent side of the screen in a contingency paradigm.

In this experiment, we sought to extend these studies by comparing the learning trend of 4-and 8-month-olds in an anticipatory looking paradigm. We designed a contingency paradigm, in which a target was appearing first at the center of the screen and then disappearing and reappearing on a random side of the screen.

In the occlusion time, if the target was moving toward left, one icon was appearing to signal the location of the target to the infant, and if the target was appearing on the right, another icon was being presented to signal the location. Between the offset of the cue and the onset of the target on the side of the screen, there was a random delay period, in which the participants could make an anticipatory look to either of the sides of the screen.

In this design, we were reinforcing the infants to associate two cuing functions to the two arbitrary icons, based on the synchrony of the events over the course of the trials.

We tested the same age groups as in experiment 6, 4- and 8-month-old infants, so we could make a comparison between the results of the attentional cuing triggered by arrow icons with the results of this experiment.

### 3.4.2. Experimental Design

In this experiment, we developed a paradigm in which we were introducing the participants to two actions. The actions were a toy initially positioned at the center of the screen, was disappearing and reappearing at either of the sides of the screen.

Between the disappearance and reappearance of the toy, there were two phases;

**Cue phase:** In this phase, one of the two icons were shown at the center of the screen, these icons were set to cue the participants to the side of the screen where the toy would reappear.

The icons were presented in a random order.

**Occlusion phase:** At this phase, the cue disappeared and the participants were anticipating the appearance of the toy. The participants could anticipate the side of the screen where the toys were set to appear, based on the presented cue. The duration of the occlusion phase was randomized to make the participants more inclined to look at the side of the screen, where they predicted the toy to appear.

After this phase, there was the reinforcement phase in which the toy was appearing on the side of the screen based on the presented icon at the cue phase.

Therefore, we predicted if infants were able to learn the regularity between the side of the screen in which the toy appeared and the icon, after passing several trials, during the occlusion period, they could show an anticipatory look toward the side of the screen where they anticipated the toy to appear (see Figure 3.11).

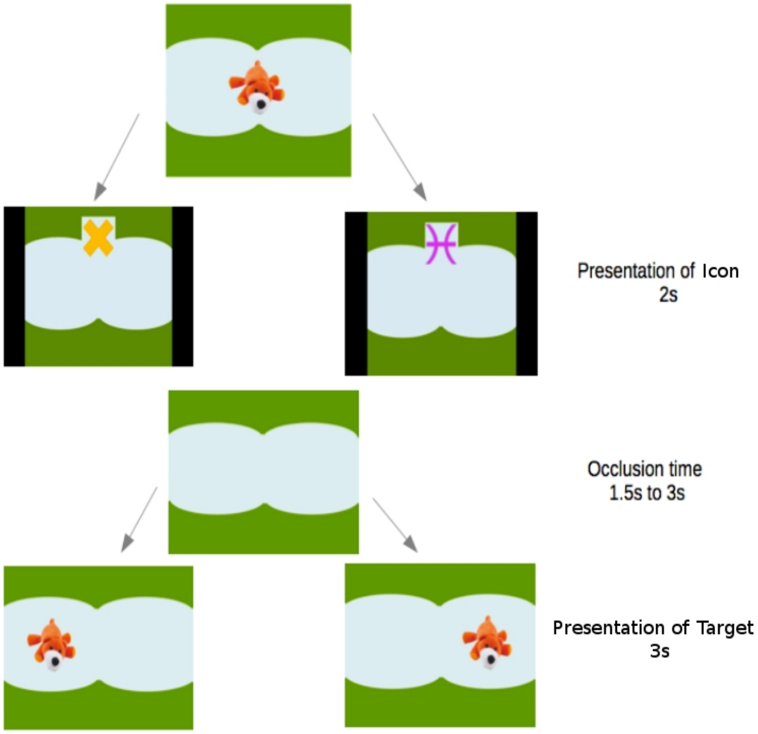

**Fig. 3.11.** The design of a single trial. In the first phase a toy appeared at the center of the screen, in the next phase, the toy disappeared and one of the two icons appeared randomly at the top center of the screen. In the occlusion phase, the icon disappeared and there was a random delay until the final phase, when the toy reappeared on one side of the screen based on the presented icon.

### 3.4. 3. Stimuli

Each trial of the experiment consisted of 4 main phases (see Figure 3.11). All the 4 phases had a common background image that consisted of two elliptical cells. The two cells were the locations in which the toys were appearing in the last phase (the reinforcement phase). The two cells were helping the infants in confining their search area during the occlusion time to the regions inside the cells.

**Target Stimuli:** The stimuli consisted of 8 toys, all confined to an area of 300 by 300 pixels.

**Icon Stimuli:** For the icon stimuli, we chose two icons from a public domain on-line symbol database. Our considerations were to choose two icons that were clearly different from each other, and moreover, lack any cuing feature (see Figure 3.11). We further set the color of the icons as yellow and purple to make them easier to differentiate.

### 3.4.4. Participants

Participants were tested at two different ages, at 4 months and at 8 months.

#### 3.4.4.1. Four-month Group

Twenty-seven healthy monolingual Italian infants were tested in this condition, from 14 to 18 weeks, with a mean of 15.8 weeks, SD = 1.9 weeks.

Two subjects were excluded due to lack of attention.

#### 3.4.4.2. Eight-month Group

Twenty-six healthy monolingual Italian infants were tested in this condition, from 30 to 34 weeks, with a mean of 32.3 weeks, SD = 2.2 weeks, two subjects were excluded due to lack of attention.

All the infants were full term and had an APGAR number over 7 out of 10. The participants were recruited from Trieste, Italy, through sending invitation letters to a random selection of parents, whose babies fit the age range of the study. The parents were all reimbursed for attending the experiment and the infants received a certificate of attendance.

### 3.4.5. Apparatus

The apparatus was identical to the apparatus explained in experiment 6 (see section 3.2.5).

### 3.4.6. Procedure

At the beginning of the experiment, the participants passed a calibration session. The calibration procedure was identical to the procedure explained in experiment 6 (see section 3.2.6).

After the calibration, the participants watched a 3-second animation, to reorient their attention to the screen. After this session, the test block started. The test block consisted of 24 trials. Each trial consisted of 5 phases.

**Phase 1.** Presentation of the background image without any stimuli for a duration of 750 ms.

**Phase 2.** Presentation of a toy at the center of the background image, for a duration of 1700 ms. The toy was chosen randomly (see Figure 3.11).

**Phase 3.** Presentation of one of the two icons at the top center of the screen, for a duration of 2700 ms. To make the icons more attractive, we animated the size of the icons. During the 2700 ms, they shrunk from 100% to 50% of their initial scale, and then back to 100% of the initial scale. The shrinking and expansion times were each lasting for 1350 ms. The initial dimension was 210 by 210 pixels.

The icons were presented in a random order.

**Phase.4.** At this phase, the background image with no stimuli was presented to the participant, with a duration varying between 1500 to 3000 ms. This duration was chosen randomly from a uniform distribution of delay periods generated by Matlab software. In this phase the participants could demonstrate an anticipatory look to the side of the screen, where they predicted the toy to reappear. We will refer to this phase as the occlusion phase.

**Phase.5.** At this phase, the same toy as in phase.2. reappeared at one of the two sides of the screen. In order to increase the anticipatory look of the infants during the occlusion time, similar to the icon presentation, the dimensions of the toys, were animated to decrease to 50% of their initial size, and then increased back to 100%. The duration of toy presentation was 3000 seconds. At the onset of the toy phase a short bip sound was being played from two load speakers located on the two sides of the monitor, to attract the attention of the infants.

The side of the screen corresponding to each icon was counterbalanced across subjects.

### 3.4.7. Scoring

The coordinates of the gaze and the gaze quality of each eye, were collected for further analysis.

All the analysis were done using Mathworks Matlab software (v.2015b), Statistics and Machine Learning Toolbox™ of Matlab.

#### 3.4.7.1. Inclusion Criteria

##### 3.4.7.1.1. RoI Settings

We analyzed the gaze profile of the 3rd phase of the trials (icon presentation), the 4th phase (occlusion period) and the 5th phase (toy presentation).

For the analysis of the gaze, we defined three Regions of Interest (RoIs), each covering 33% of the width of the screen.

#### 3.4.7.1.2. Valid Look

Valid look at each phase is equivalent to the ratio of time the subject was looking inside the valid RoI(s) divided by the total duration of the phase.

Valid Look = Time looked in RoI(s) / Total duration of the phase

We set a minimum of 60% for the minimum Valid Look in all the phases.

A gaze point was regarded as valid, if it was inside an RoI and having a quality score registered by the eye-tracker between 0-1 out of 4 (4 is equivalent to poor quality). The quality score was assigned to each point automatically by the eye-tracker if the data of both eyes was available and consistent.

#### 3.4.7.1.3. Side Bias Control

We measured the side bias based on the total ratio of looking at the left RoI over the course of the trials.

Left Bias = Overall look at the left RoI / (Overall look at the Left RoI + Overall look at the Right RoI)

If the ratio was more than 85% or less than 15%, we considered a participant as having side bias. In total 2 participant were excluded from the 4-month group and 4 participants were excluded from the 8-month group.

### 3.4.7.2. Measuring Performance

#### 3.4.7.2.1. Target Look

To measure the performance, we defined the correct RoI, as the RoI in which the toy was appearing in the last phase.

We then defined the Target Look as:

Target Look = Number of time bins within the correct RoI / Total number of valid time bins

A **time bin**, is the unit of time, equivalent to 16.67 ms, that represents a single gaze point registered by the eye tracker.

#### 3.4.7.3. Trial Validation and Subject Validation

We set the following criteria to validate a trial:

1. We considered a threshold of 60% as the minimum ratio of Valid Look for each of the 3 phases (cue presentation, occlusion phase and the toy presentation).
2. A trial was eliminated from the analysis, if one of the three phases of the trial was having less than 60% valid gaze.
3. A subject was eliminated from the analysis, if providing less than 5 valid trials.

## 3.4.8. Results

### 3.4.8.1. Overall Results

We analyzed the total looking time during the occlusion phase to assess the anticipatory look of the participants. We predicted that if the participant is learning the association between the icons and the location of the screen where the target would appear, we should be able to see a positive trend in anticipatory look at the correct side, over the course of the trials.

For this analysis, 18 subjects passed the criteria in the 8-month group and 23 subjects in the 4-month group.

To proceed with the analysis we acquired the sequence of valid trials of each subject and measured the proportion of Target Look at the occlusion phase for these trials. We assigned a rank to the ordinal number of each valid trial. We then averaged the Target Look for all the trials with the same rank (see Figure 3.12). Higher ranks which corresponds to the last trials, could include lower number of data points. We set a constraint that the last rank should contain at least 3 trials, i.e. the subjects included in this analysis, had at least 3 trials until the last included rank.

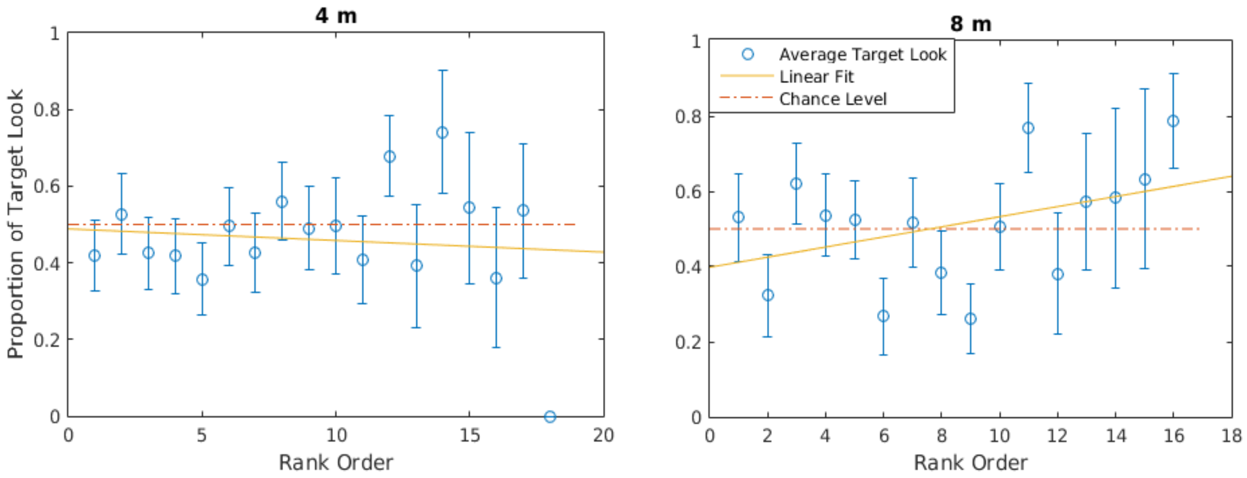

**Fig. 3.12.** The average proportion of Target Look at each rank of the valid trials. Each point represents the mean Target Look of the valid trials, with the corresponding rank number (their ordinal number in the set of valid trials of the subject). The fitted line shows a linear fit based on least square distance from the data points. The Spearman correlation in 8-month group revealed a positive, but non-significant trend, ρ = 0.41, p = 0.12, however, in the 4-month group, this trend was at chance. ρ = 0.05, p = 0.810. The errorbars show one standard error from the mean. The horizontal dashed line shows the chance level at 0.5.

We applied Spearman rank correlation to measure the dependency of the performance as indicated by Target Look on the rank order of the trials. Pearson rank correlation is the favorable statistical test, since it does not consider any assumption on distribution of the variable under study. The test for the 8-month group resulted in ρ = 0.41, and p = 0.12 and for the 4-month group, ρ = 0.05 and p = 0.810.

We further applied the same test on the population of the valid trials sorted by their rank order, without averaging out across subjects (see Figure 3.13). For the analysis we only included trials in which there was at least one fixation on both sides of the screen during the occlusion time (excluding the ceiling trials, to avoid biasing the correlation analysis). The test in the 8-month group resulted in $\rho = 0.25$ and $p = 0.046$ and in the 4-month group, $\rho = 0.06$ and $p = 0.60$.

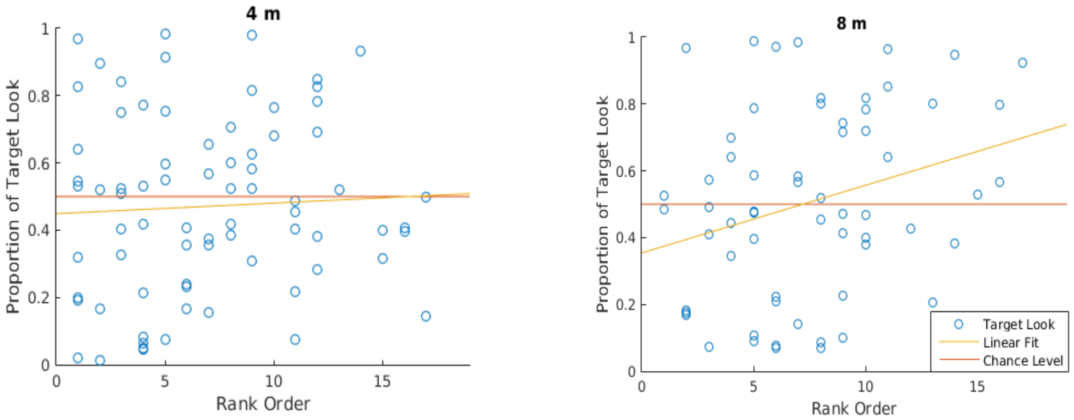

**Fig. 3.13.** Scatter plot showing the distribution of the Target Look of valid trials, indexed by their rank order. The fitted line indicates the linear fit based on least square distance form the data points. At the 8-month group, $\rho = 0.25$ and $p = 0.046$ and at the 4-month group, $\rho = -0.06$ and $p = 0.60$.

The two analysis suggest, there is a positive trend, although marginally significant, in the Target Look over the course of trials at the 8-month group, indicating an overall learning trend, but in the 4-month group this trend was absent.

### 3.4.8.2. The Learning Trends

We further asked if the subjects at the age of 8-month, were learning both of the icons or the observed positive trend in the Target Look over trials' rank, was due to learning only one of the icons; of course if the subject was learning only one icon, that icon could be different from one subject to the other.

A positive learning trend signifies an overall increase in the Target Look on the course of valid trials. For instance, it can be the case that a subject exhibited a positive Target Look uniformly from the beginning of the trials to the end, whenever one specific icon appeared. In this case, although there is a high Target Look for that icon, but this pattern of Target Look does not signify a learning trend, since there is no improvement from the first trial to the last trial.

To further address our question we performed a subject by subject analyses, and obtained a learning trend for the sequence of trials of each icon separately, by measuring the Spearman's rank correlation coefficient of the two icons, independent of the other icon.

For the analyses, we only accepted the coefficient of an icon, if at least 5 valid trials were passed for that icon, and we accepted a subject as valid, if provided at least 5 valid trials for each icon. Only 12 subjects passed this criteria in the 8-month group and 16 subjects in the 4-month group. Based on this criteria, we obtained one coefficient for each icon for each subject.

In the next step, to assess if both of the icons were having a significant learning trend, we created two clusters of coefficients; one cluster consisted of the higher learning coefficient of a subject and another cluster consisted of the lower coefficient of the subject. Figure 3.14 shows the mean of the two clusters for the 8-month group.

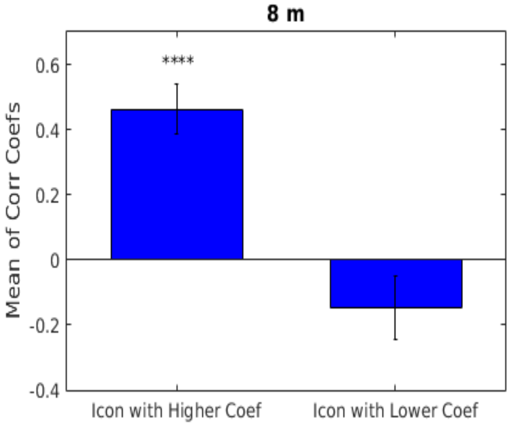

**Fig. 3.14.** The means of the set of higher correlation coefficients and the set of lower correlation coefficients, in 8-month group, obtained from the Spearman's rank correlation coefficients for the two icons of each subject.

The analysis in the 8-month group, resulted in a mean coefficient for the higher set equivalent to r = 0.46, SE = 0.076 and the mean of the set with lower coefficients was, r = - 0.14, SE = 0.096. We applied a *t*test to compare the two sets of coefficients versus the chance level at 0.0, the higher set resulted in t(11) = 6.04, p = 0.000083, and the lower set resulted in t(11) = - 1.5, p = 0.15.

We repeated the same analyses in the 4-month group, the higher and the lower correlation coefficients were r = 0.34, SE = 0.063 and r = - 0.21, SE = 0.076. The *t*test comparing the population of coefficients versus chance resulted in t(15) = 5.4, p = 0.000070, for the higher coefficient and t(15) = -2.8, p = 0.012 for the lower coefficient.

If in our paradigm the infants were learning both of the icons, we could expect to see a significant positive coefficient for both of the icons, however, in the 8-month group, the results were significant only for the icons indexed with a higher coefficient. These results provided a strong evidence of a positive learning trend for only one of the icons and not both.

Furthermore, in the 4-month group, the significant positive trend for one icon, and a significant negative trend for the other icon, suggests that by the progression of the trials, the subjects as a general pattern, were increasing their anticipation to the targets on one side, and reducing their anticipation to the targets of the opposite side. This pattern is compatible with a scenario, in which the participants were essentially developing a side bias, that lead to a higher Target Look for one icon and the lower Target Look for the other icon (note that a full side-bias results in 100% performance for one icon and 0% performance for the other icon). This pattern of learning trend suggests that participants were not learning any of the icons until the last valid trial, or this learning was inhibiting them to disengage when the opposite cue was appearing.

We further calculated the Target Look over the course of valid trials separately, for the icons with the higher learning coefficients and the icons with the lower learning coefficients (see Figure 3.15).

For both icons in the 8-month group the average trend measured by Spearman's rank correlation was positive, however, only for one icon the trend was significant.

$\rho_{\text{Higher Coef}} = 0.95$, $p = 0.00035$, $\rho_{\text{Lower Coef}} = 0.31$, $p = 0.41$.

The overall results suggest that at the start of the experiment, 8-month-olds had a tendency to look only at one side, resulted in a high Target Look for one icon and a low Target Look for the other icon. Then, over the course of trials, they started to learn one icon and showed a positive but non-significant learning trend for the other icon.

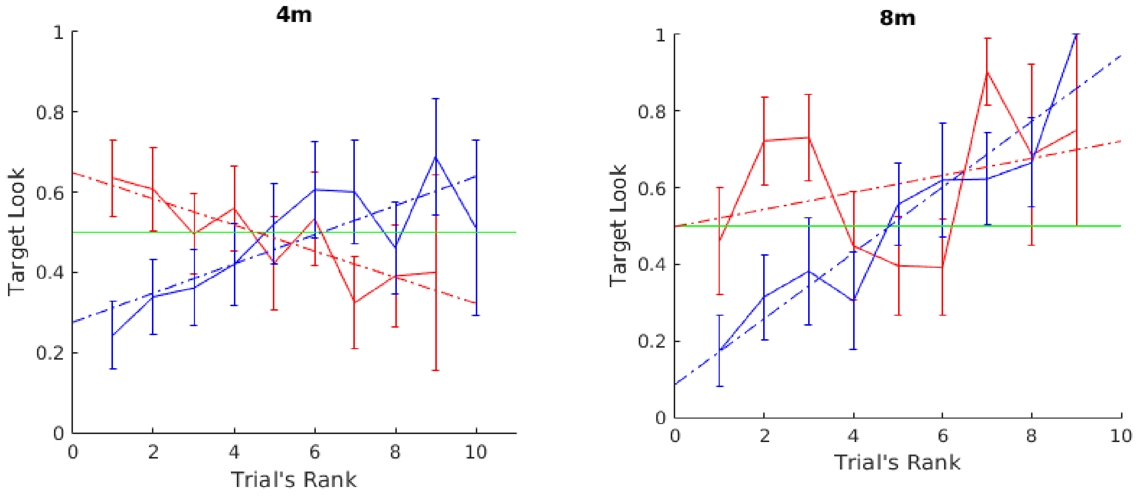

**Fig. 3.15.** Mean Target Look at each trial's rank for the icons with higher correlation coefficient (blue), and the icons with lower correlation coefficient (red), in both the 4-month **(Left)** and 8-month group **(Right)**. The dashed lines show a linear fit based on least square distance from the data points. The errorbars are one standard error from the mean. The green horizontal line shows the chance level at 0.5. Note the two positive trends in the 8-month group and the positive and negative trend in the 4-month group.

We repeated the same analyses in the 4-month group, which resulted in:

$\rho_{\text{Higher Coef}} = 0.80$, $p = 0.005$, $\rho_{\text{Lower Coef}} = -0.85$, $p = 0.006$. The results of the 4-month group, confirms that the increase in Target Look in one icon is accompanied by the decrease in Target Look of the other icons, a trend that is suggestive of development of a side bias over the course of trials.

### 3.4.9. Interim Conclusion

In two age groups of 4-month and 8-month, we tested the capability of infants in associating two cuing functions to two abstract icons in an interleaving sequence. The functions were reorientation of attention to the left or to the right side of the screen based on the presented icon, so the infants could anticipate the appearance of a toy in that side of the screen. We showed that only at the age of 8-month, participants showed a reliable positive trend in the facilitation of their anticipatory look toward the correct side. Furthermore, we argued that this trend is due to learning one of the two icons and not both of them.

These results are in line with the failure of 4-month-olds reported by Johnson et al. (1991) and the success of 8-month old infants reported by Werchan, et al., (2015).

Future experiments could provide infants with a higher number of trials, to assess the learning trends in an extended duration, particularly it would be interesting to explore at what point 8-month-olds start to show a significant learning trend for the other icon, and if the fact that in the 4-month group we did not observe a reliable learning trend is due to the few number of trials or they essentially lack the capability of establishing a link between the icon and the location of the target.

## 3.5. Discussion

In three experiments, we attempted to study the association of reorientation of attention to abstract icons in pre-verbal infants, with a goal to shed light on how internal operators in the nervous system can be conceptualized in a non-lingual domain.

We approached this goal from two parallel perspectives; in two studies we studied arrow-like icons, as icons that their perceptual features may be suggestive of the function they are referencing and in another study, infants attended a contingency paradigm to associate a left or a right cue of attention to two arbitrary icons that did not have any relevant perceptual feature.

In experiment 6, 4- and 8-month-old infants attended a paradigm in which an arrow cue was being presented. The results suggested that not at the age of 8-month but at the age of 4-month, infants could follow the direction of arrow. We speculated that the failure in the 8-month group could be due to the presentation of two identical targets in the paradigm, which correspondingly renders the arrow icon as uninformative. Or on the other hand, it could be due to a lack of sensitivity to the perceptual features of the arrow icon at the age of 8 months.

In experiment 7, we showed that the triangular area of an arrow icon may be sufficient in signaling the reorientation of attention of 4-month-olds. These results together suggest that either the gradient of the sharp edges of a triangle or the gradient in the surface area of a triangle, may trigger an attentional shift in the 4-month-olds.

However, the failure of older infants in following the direction of an arrow, as reported here as well as elsewhere (Ristic, 2002; Jakobsen, Frick & Simpson, 2013), discourages us to consider that the association between the attentional shift and the triangular-based cues turn into a sustained concept at the age of 4-months. But currently we do not have a framework to interpret how and why representation of a triangle might induce a sustained reorientation of attention in the 4-month-olds. These results together suggest that although an attentional shift might be triggered based on the detailed asymmetries in an icon, but conceptualization of this process may need an exhaustive ostensive learning process that would span through the first years of life.

In Experiment 8, in a contingency paradigm, we asked if infants can assign a cuing function to two arbitrary icons that lack any relevant perceptual feature. We designed a paradigm, in which two icons were cuing the appearance of the targets either on the left or on the right side of the screen. We discussed that on a course of 10 trials, as a likely scenario, infants at the age of 4-month, did not show a consistent learning trend. However, in the 8-month group infants showed a learning trend at least for one of the icons.

Moreover, it would be informative to compare the learning trends of assigning a cuing function to an icon, to a rule or to a lexical item (see Waxman & Gelman, 2009, for a discussion on word learning; Kovacs and Mehler, 2009), through the same paradigm, to address if different learning trends might suggest the presence of different mechanisms involved in early associative learning.

Finally, it is important to note that the cuing functions in the context of arrows and in the context of arbitrary icons, do not necessarily need to be the same processes. One reason can be that in the case of arbitrary icons, the reorientation of attention is due to anticipation (of appearance of a target), and the direction of anticipation is acquired from and dictated by the synchrony of the icon-target co-occurrences on the course of the trials. However in the case of arrows, at the age of 4 months, it is a mere assumption to speculate that infants anticipate the appearance of *something* in front of the icon. For now, we can ask if the reorientation of attention due to arrows, is a volitional shift of attention due to an anticipation, which is abstracted from or facilitated by social cues (such as pointing); or an automatic activation of a pathway in the visual cortex that directly triggers an attentional shift due to the sharp edges of the icon.

Further experiments are needed to clarify the underlying processes of the observed performance.

## 4. Overall Discussion

In a series of studies, we attempted to explore the application of two different functions on mental representations in the pre-verbal infants. Disjunctive inference applied on propositions in one hand and attentional shift applied on visual icons, on the other hand. In the context of disjunctive inference, we set to explore the utilization of this process with the propositions obtained from an ambiguous scenario to make a deductive inference in 10-month-olds; and in the context of attentional shift, which is reported to function from the first hours of life, we set to explore how it associates with visual icons at 4 months and 8 months of age. The icons were in two classes, arbitrary icons that did not have any perceptual feature relevant to attentional shift and arrow-like icons that their perceptual features could induce an attentional shift.

We showed that at the age of 10 months, infants may be able to apply disjunctive inference to resolve an ambiguity in the context of face-voice association.

On the other hand, the results from referencing an attentional shift to an icon was more intriguing. Two experiments on 4-month-olds, showed that infants at this age are able to follow the direction of triangular-based icons. In the two experiments we minimized the possibility that this effect could be due to any exogenous pathway, and we concluded that the observed attentional shift could only be due an endogenous cuing of the icons. These results together with a failure of infants in following the direction of an arrow in the 8-month group are suggestive of an early endogenous cuing attribution based on the perceptual features of a triangle icon, an attribution that does not develop into a sustained referencing concept, as probably from 4 months to 8 months, infants loose their sensitivity to the perceptual features of a triangular-based icon, which abandons the referentiality to attentional shift observed in triangular-based icons. However, at the age of 8 months, infants showed that they can readily associate a cuing function to one out of two arbitrary icons in a contingency paradigm in a very few number of trials.

The results from chapter 2, demonstrates that the mental processes involved in disjunctive inference, can contrast two mental images and change their validity. However, we can not rule out that this level of operation is dedicated to face-voice association, until we observe similar results in other contexts (see the discussion of chapter 2, for an alternative paradigm); but if further studies confirm that this process can be utilized in other domains, it can provide support for the presence of domain general logic operations that can act on mental representations independent of their specific content (as described in *language of thought hypothesis* by Fodor, 1975).

These results together with other studies on abstract domains, support a dynamics framework of interactions between the mental representations and mental processes in the first year of life. In this framework, the logic processes can combine and change the validity of mental representations, and on a different domain, mental representations of abstract icons can readily integrate with attentional cues.

An interesting question to ask is if an arbitrary operator including the logical operators, can be represented by an arbitrary icon, and to what extent this process alters the mental representation of the operator itself.

To partly answer this question we are running a contingency paradigm, in which each of our two arbitrary icons, cue a switch in either the direction (vertical/horizontal) or the color (purple/yellow) of a reference rectangle. Eight-month-olds are presented side by side, with a possible and an impossible modification of the trial-initial state (color + direction) of the reference rectangle. Subsequently based on the preceding cue, participants observe the appearance of a target on the possible state. This let's us to measure the anticipatory look toward the possible state before the onset of the target. As a continuation of experiment 8, this experiment can help us understanding if two novel arbitrary functions can be associated with two arbitrary icons.

At the end, we believe that the studies on interaction of different operations with arbitrary mental representations in pre-verbal infants, can help us to pave the way toward a more comprehensive theory on the early representation and conceptualization of abstract operators in the first year of life.

## 5. Bibliography

Austin, K., Theakston, A., Lieven, E., & Tomasello, M. (2014). Young children's understanding of denial. *Developmental Psychology, 50*(8), 2061.

Bahrick, L. E., Hernandez-Reif, M., & Flom, R. (2005). The development of infant learning about specific face-voice relations. *Developmental Psychology*, *41*(3), 541.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*(2), 276.

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. *Handbook of Psychophysiology*, *2*, 142–162.

Beatty, J., & Wagoner, B. L. (1977). *Pupillometric signs of brain activation vary with level of cognitive processing.* DTIC Document.

Bennett, M. R., & Hacker, P. M. S. (2008). *History of cognitive neuroscience*. John Wiley & Sons.

Bion, R. A., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30months. *Cognition*, *126*(1), 39–53.

Bloom, L. (1993). *Language development from two to three*. Cambridge University Press.

Bloom, L. M. (1968). Language Development: Form and Function in Emerging Grammars.

Bowerman, M. (1986). First steps in acquiring conditionals. In *On conditionals* (pp. 285–308). Cambridge University Press.

Brisson, J., Mainville, M., Mailloux, D., Beaulieu, C., Serres, J., & Sirois, S. (2013). Pupil diameter measurement errors as a function of gaze direction in corneal reflection eyetrackers. *Behavior Research Methods*, *45*(4), 1322–1331.

Brookes, H., Slater, A., Quinn, P. C., Lewkowicz, D. J., Hayes, R., & Brown, E. (2001). Three-month-old infants learn arbitrary auditory-visual pairings between voices and faces. *Infant and Child Development*, *10*(1-2), 75–82.

Bushneil, I. W. R., Sai, F., & Mullin, J. T. (1989). Neonatal recognition of the mother's face. *British Journal of Developmental Psychology*, *7*(1), 3–15.

Call, J., & Carpenter, M. (2001). Do apes and children know what they have seen? *Animal Cognition*, *3*(4), 207–220.

Cameron-Faulkner, T., Lieven, E., & Theakston, A. (2007). What part of no do children not understand? A usage-based account of multiword negation. *Journal of Child Language*, *34*(02), 251–282.

Carey, S. (2009). *The origin of concepts*. Oxford University Press.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word.

Cheng, P. W. (1986). *Pragmatic versus syntatic approaches to training deductive reasoning*.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, *17*(4), 391–416.

Choi, S. (1988). The semantic development of negation: a cross-linguistic longitudinal study. *Journal of Child Language*, *15*(03), 517–531.

Daniels, L. B., Nichols, D. F., Seifert, M. S., & Hock, H. S. (2012). Changes in pupil diameter entrained by cortically initiated changes in attention. *Visual Neuroscience*, *29*(02), 131–142.

Decasper, A. J., & Prescott, P. A. (1984). Human newborns' perception of male voices: Preference, discrimination, and reinforcing value. *Developmental Psychobiology*, *17*(5), 481–491.

DeLoache, J. S. (2004). Becoming symbol-minded. *Trends in Cognitive Sciences*, *8*(2), 66–70.

Ditz, H. M., & Nieder, A. (2015). Neurons selective to the number of visual items in the corvid songbird endbrain. *Proceedings of the National Academy of Sciences*, *112*(25), 7827–7832.

Fodor, J. A. (1975). *The language of thought* (Vol. 5). Harvard University Press.

Fodor, J. A. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind.* The MIT Press.

Frank, M. C., Vul, E., & Johnson, S. P. (2009a). Development of infants' attention to faces during the first year. *Cognition*, *110*(2), 160–170.

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science*, *291*(5502), 312–316.

Friedlander, B. Z. (1968). The effect of speaker identity, voice inflection, vocabulary, and message redundancy on infants' selection of vocal reinforcement. *Journal of Experimental Child Psychology*, *6*(3), 443–459.

Geangu, E., Hauf, P., Bhardwaj, R., & Bentz, W. (2011). Infant pupil diameter changes in response to others' positive and negative emotions. *PloS One*, *6*(11), e27132–e27132.

Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 252–269.

Glenn, S. M., & Cunningham, C. C. (1983). What do babies listen to most? A developmental study of auditory preferences in nonhandicapped infants and infants with Down's syndrome. *Developmental Psychology*, *19*(3), 332.

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology*, *28*(1), 99.

Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, *56*(4), 544–549.

Gredebäck, G., & Melinder, A. (2011). Teleological reasoning in 4-month-old infants: pupil dilations and contextual constraints. *PLoS One*, *6*(10), e26487.

Haith, M. M., Hazan, C., & Goodman, G. S. (1988). Expectation and anticipation of dynamic visual events by 3.5-month-old babies. *Child Development*, 467–479.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, *87*(1), B23–B34.

Halberda, J. (2006). Is this a dax which I see before me? Use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cognitive Psychology*, *53*(4), 310–344.

Hartmann, M., & Fischer, M. H. (2014). Pupillometry: the eyes shed fresh light on the mind. *Current Biology*, *24*(7), R281–R282.

Hespos, S. J., Ferry, A. L., & Rips, L. J. (2009). Five-month-old infants have different expectations for solids and liquids. *Psychological Science*, *20*(5), 603–611.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, *143*(3611), 1190–1192.

Hill, A., Collier-Baker, E., & Suddendorf, T. (2012). Inferential reasoning by exclusion in children (Homo sapiens). *Journal of Comparative Psychology*, *126*(3), 243.

Hochmann, J.-R., & Papeo, L. (2014). The Invariance Problem in Infancy A Pupillometry Study. *Psychological Science*, *25*(11), 2038–2046.

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, *13*(2), 128–157.

Hume, D. (1978). *A treatise of human natureOxford University press*. Oxford.

Hummer, P., Wimmer, H., & Antes, G. (1993). On the origins of denial negation. *Journal of Child Language*, *20*(03), 607–618.

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences*, *106*(25), 10382–10385.

Jackson, I., & Sirois, S. (2009). Infant cognition: going full factorial with pupil dilation. *Developmental Science*, *12*(4), 670–679.

Jakobsen, K. V., Frick, J. E., & Simpson, E. A. (2013). Look here! the development of attentional orienting to symbolic cues. *Journal of Cognition and Development*, *14*(2), 229–249.

Johnson, M. H. (1994). Visual attention and the control of eye movements in early infancy. *Attention and Performance Vol. XV: Conscious and Unconscious Processing*, 291–310.

Johnson, M. H., Posner, M., & Rothbart, M. K. (1991). Components of visual orienting in early infancy: Contingency learning, anticipatory looking, and disengaging. *Cognitive Neuroscience, Journal of*, *3*(4), 335–344.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1), 1–19.

Johnson, M. H., & Tucker, L. A. (1996). The development and temporal dynamics of spatial orienting in infants. *Journal of Experimental Child Psychology*, *63*(1), 171–188.

Jonides, J., & Irwin, D. E. (1981). Capturing attention. *Cognition*, *10*(1), 145–150.

Jonides, J., & Somers, P. (1977). Voluntary control of the allocation of attention in the visual field. In *meeting of the Midwest Psychological Association* (p. b53).

Jordan, K. E., & Brannon, E. M. (2006). The multisensory representation of number in infancy. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(9), 3486–3489.

Kahnemann, D., & Beatty, J. (1967). Pupillary responses in a pitch-discrimination task. *Perception & Psychophysics*, *2*(3), 101–105.

Kaup, B., Yaxley, R. H., Madden, C. J., Zwaan, R. A., & Lüdtke, J. (2007). Experiential simulations of negated text information. *The Quarterly Journal of Experimental Psychology*, *60*(7), 976–990.

Kenny, A. (2010). Concepts, brains, and behaviour. *Grazer Philosophische Studien*, *81*(1), 105–113.

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, *4*(4), 138–147.

Kovács, Á. M., & Mehler, J. (2009a). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, *106*(16), 6556–6560.

Kovács, Á. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, *106*(16), 6556–6560.

Kuhn, D. (1977). Conditional reasoning in children. *Developmental Psychology*, *13*(4), 342.

Lin, L., Chen, G., Kuang, H., Wang, D., & Tsien, J. Z. (2007). Neural encoding of the concept of nest in the mouse brain. *Proceedings of the National Academy of Sciences*, *104*(14), 6066–6071.

Locke, J. (n.d.). *1690/1975. An Essay Concerning Human Understanding, ed. PH Nidditch*. Oxford: Oxford University Press.

Macchi, C. V., Turati, C., & Simion, F. (2004). Can a nonspecific bias toward top-heavy patterns explain newborns’ face preference? *Psychological Science*, *15*(6), 379–383.

Margolis, E. (1998). How to acquire a concept. *Mind and Language*, *13*(3), 347–369.

Margolis, Eric and Laurence, Stephen, "Concepts", *The Stanford Encyclopedia of Philosophy*

(Spring 2014 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2014/entries/concepts/>.

Markman, E. M. (1989). Categorization and naming in childrenMIT Press. *Cambridge, MA*.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*(2), 121–157.

Marno, H., Farroni, T., Dos Santos, Y. V., Ekramnia, M., Nespor, M., & Mehler, J. (2015). Can you see what I am talking about? Human speech triggers referential expectation in four-month-old infants. *Scientific Reports*, *5*.

Mather, E., & Plunkett, K. (2010). Novel labels support 10-month-olds' attention to novel objects. *Journal of Experimental Child Psychology*, *105*(3), 232–242.

Mather, E., & Plunkett, K. (2012). The role of novelty in early word learning. *Cognitive Science*, *36*(7), 1157–1177.

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, *283*(5398), 77–80.

Mervis, C. B., & Bertrand, J. (1994). Acquisition of the novel name–nameless category (N3C) principle. *Child Development*, *65*(6), 1646–1662.

Moher, M., Feigenson, L., & Halberda, J. (2010). A One-to-One Bias and Fast Mapping Support Preschoolers' Learning About Faces and Voices. *Cognitive Science*, *34*(5), 719–751.

Nordmeyer, A. E., & Frank, M. C. (2013). Measuring the comprehension of negation in 2-to 4-year-old children. In *Proceedings of the 35th Annual Meeting of the Cognitive Science Society. Austin, TX: Cognitive Science Society*.

Nyström, P., Gredebäck, G., Bolte, S., & Falck-Ytter, T. (2015). Hypersensitive pupillary light reflex in infants at risk for autism. *Molecular Autism*, *6*(1), 10.

Pascalis, O., & de Schonen, S. (1994). Recognition memory in 3-to 4-day-old human neonates. *Neuroreport*, *5*(14), 1721–1724.

Peacocke, C. (1992). *A study of concepts*. The MIT Press.

Pea, R. D. (1978). *The Development of Negation in Early Child Language*. University of Oxford.

Peavler, W. S. (1974). Pupil size, information overload, and performance differences. *Psychophysiology*, *11*(5), 559–566.

Piaget, J., Cook, M., & Norton, W. W. (1952). *The origins of intelligence in children* (Vol. 8). International Universities Press New York.

Piaget, J. (1953). Logic and Psychology. Manchester: Manchester University Press.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*(1), 3–25.

Posner, M. I., & Petersen, S. E. (1990). Annual review of neuroscience. *Annu Rev Neurosci*, *13*(1), 25–42.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention and Performance X: Control of Language Processes*, *32*, 531–556.

Posner, M. I., Rafal, R. D., Choate, L. S., & Vaughan, J. (1985). Inhibition of return: Neural basis and function. *Cognitive Neuropsychology*, *2*(3), 211–228.

Premack, D., & Premack, A. J. (1994). Levels of causal understanding in chimpanzees and children. *Cognition*, *50*(1), 347–362.

Quinn, P. C., Uttley, L., Lee, K., Gibson, A., Smith, M., Slater, A. M., & Pascalis, O. (2008). Infant preference for female faces occurs for same-but not other-race faces. *Journal of Neuropsychology*, *2*(1), 15–26.

Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *PERCEPTION-LONDON-*, *31*(9), 1109–1122.

Quiroga, R. Q. (2012). Concept cells: the building blocks of declarative memory functions. *Nature Reviews Neuroscience*, *13*(8), 587–597.

Ramsey-Rennels, J. L., & Langlois, J. H. (2006). Infants' differential processing of female and male faces. *Current Directions in Psychological Science*, *15*(2), 59–62.

Ristic, J., Friesen, C. K., & Kingstone, A. (2002). Are eyes special? It depends on how you look at it. *Psychonomic Bulletin & Review*, *9*(3), 507–513.

Rohlfing, K. J., Longo, M. R., & Bertenthal, B. I. (2012). Dynamic pointing triggers shifts of visual attention in young infants. *Developmental Science*, *15*(3), 426–435.

Saksida, A. (2014). Word learning in the first year of life.

Sirois, S., & Jackson, I. R. (2012). Pupil dilation and object permanence in infants. *Infancy*, *17*(1), 61–78.

Smith, K., Smith, A. D. M., & Blythe, R. A. (2011). Cross-Situational Learning: An Experimental Study of Word-Learning Mechanisms. *Cognitive Science*, *35*(3), 480–498.

Spelke, E. (1976). Infants' intermodal perception of events. *Cognitive Psychology*, *8*(4), 553–560.

Spelke, E. S., & Owsley, C. J. (1979). Intermodal exploration and knowledge in infancy. *Infant Behavior and Development*, *2*, 13–27.

Standley, J. M., & Madsen, C. K. (1990). Comparison of infant preferences and responses to auditory stimuli: Music, mother, and other female voice. *Journal of Music Therapy*, *27*(2), 54–97.

Tam, C. W., & Stokes, S. F. (2001). Form and function of negation in early developmental Cantonese. *Journal of Child Language*, *28*(02), 373–391.

Thompson, V. A. (2000). The task-specific nature of domain-general reasoning. *Cognition*, *76*(3), 209–268.

Tighe, T. J., & Leaton, R. N. (1976). *Habituation: Perspectives from child development, animal behavior, and neurophysiology.* Lawrence Erlbaum.

Tummeltshammer, K. S., Wu, R., Sobel, D. M., & Kirkham, N. Z. (2014). Infants track the reliability of potential informants. *Psychological Science*, *25*(9), 1730–1738.

Valenza, E., Simion, F., & Umiltà, C. (1994). Inhibition of return in newborn infants. *Infant Behavior and Development*, *17*(3), 293–302.

Vouloumanos, A., Druhen, M. J., Hauser, M. D., & Huizink, A. T. (2009). Five-month-old infants' identification of the sources of vocalizations. *Proceedings of the National Academy of Sciences*, *106*(44), 18867–18872.

Wallis, J. D., Anderson, K. C., & Miller, E. K. (2001). Single neurons in prefrontal cortex encode abstract rules. *Nature*, *411*(6840), 953–956.

Walton, G. E., Bower, N. J. A., & Bower, T. G. R. (1992). Recognition of familiar faces by newborns. *Infant Behavior and Development*, *15*(2), 265–269.

Wason, P. C. (1968). Reasoning about a rule. *The Quarterly Journal of Experimental Psychology*, *20*(3), 273–281.

Wason, P. C., & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *The Quarterly Journal of Experimental Psychology*, *23*(1), 63–71.

Watson, J. S., Gergely, G., Csanyi, V., Topal, J., Gacsi, M., & Sarkozi, Z. (2001). Distinguishing logic from association in the solution of an invisible displacement task by children (Homo sapiens) and dogs (Canis familiaris): Using negation of disjunction. *Journal of Comparative Psychology*, *115*(3), 219.

Waxman, S. R., & Gelman, S. A. (2009). Early word-learning entails reference, not merely associations. *Trends in Cognitive Sciences*, *13*(6), 258–263.

Werchan, D. M., Collins, A. G., Frank, M. J., & Amso, D. (2015). 8-Month-Old Infants Spontaneously Learn and Generalize Hierarchical Rules. *Psychological Science*, 0956797615571442.

Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *43*(2), 230.

## S.1. Interaction of Valid Look and Target Look in Chapter 2

### S1.1. Introduction

In experiment 1, the potential inference process in the FMF condition of our paradigm consisted of 3 phases, which we referred to as ambiguity phase, cue phase and inference phase, corresponding to the phases 1 to 3.

For our analysis, we considered an inclusion criteria for each trial based on the minimum ratio of looking at the valid RoI(s), in each of the three phases of a trial.

The thresholds we considered, were 50%, 80% and 50% for the phases 1 to 3 correspondingly.

The main reason we considered a higher threshold in the phase 2 of the trial, was to have a consistent gaze profile for further assessment of the pupil dilation. However, here we try to characterize the sensitivity of Target Look to alterations in the ratio of Valid Look in different phases of the trials. We asked this question to assess if our criteria for the thresholds, could influence the overall measure of the Target Look. We addressed this question, by calculating the overall interaction between the ratio of Valid Look of the different phases of the trial and the corresponding Target Look of the trial.

### S1.2. Methods

Previously in each trial, we registered a Ratio of Valid Look (RVL) at each phase. This measure as defined earlier, is equivalent to the ratio of time during a phase, that the participant was looking at any of the valid RoI(s) divided by the total duration of the phase.

For the purpose of our analysis we studied the interaction between the Target Look and the RVL of each phase separately. To perform this analysis we included all the trials that had an arbitrary

RVL for the phase under study, and a minimum 50% RVL for the other two phases.

For instance, to explore if the variations in the RVL of phase 1 had an impact on the Target Look, we included all the trials with a minimum RVL of 50% at the phases 2 and phase 3 regardless of the RVL in phase 1.

Furthermore, for these analyses we only included subjects that provided at least 3 valid trials.

We repeated the same analysis for all the three phases, and for both the FMF and FFF conditions to be able to make a general picture on the dependency of the Target Look on RVL.

### S1.3. Results

To quantify the dependency of the Target Look on the measured RVL, we used the Pearson linear correlation analysis to obtain a correlation coefficient for each dependency test. Since in order to assess the dependency of Target Look on RVL at each phase, we reset the minimum RVL of the other two phases, after applying this criteria, the number of included trials varied for different phases.

In Table S1, the results of the Pearson correlation analyses have been included for both FMF and FFF conditions.

Furthermore to avoid biasing the correlation analysis, we did not include the trials with a ceiling performance (0% or 100% Target Looks).

| | Phase 1 | Phase 2 | Phase 3 |
|---|---|---|---|
| FMF | r(131)= -0.01, p= 0.82 | r(132)= 0.15, p= 0.09 | r(150)= 0.16, p= 0.05 |
| FFF | r(156)= 0.075, p= 0.35 | r(134)= -0.02, p= 0.81 | r(149)= 0.077, p= 0.35 |

Table S1. Dependency of the Target Look on the ratio of Valid Look at various phases of the task in both FMF anf FFF conditions. The coefficients and p-values are resulted from the Pearson correlation analysis. The numbers in the parenthesis refer to the number of trials included.

As can be seen in the table S1, only the interaction between the RVL at the phase 3 of the FMF condition showed a marginally strong interaction with the Target Look. However, the rest of the interactions were non-significant. Figure. S1. shows the scatter plots corresponding to the 3 phases of the FMF condition.

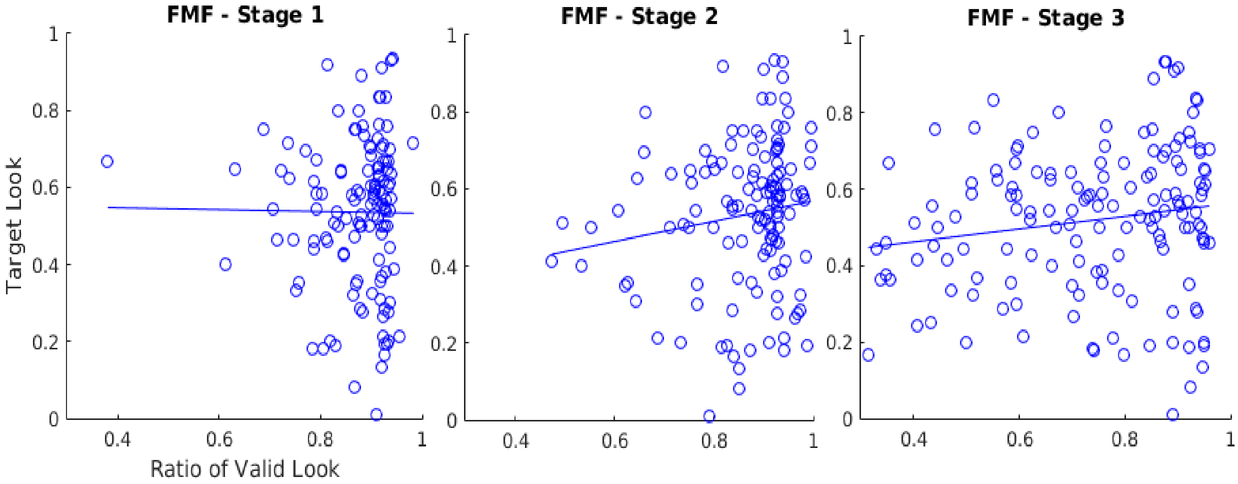

Fig. S.1. The scatter plots of different phases of the trials of the FMF condition, with the ratio of Valid Look of each phase, as the x coordinate and the Target Look of the trial as the y coordinate. The solid lines show the fit based on minimum square root of the distances from the data points, the trials with a Target Look of 0% or 100% are removed from the analysis. In none of the plots we could observe a significant correlation, only the coefficient of the 3rd phase was marginally significant.

These results suggest that the ratio of Valid Look at different phases of the trials, does not have an influence on the measured Target Look. Furthermore, we can conclude that the thresholds we considered in experiment 1, do not alter negatively the measured performance (the marginally significant interaction observed in the 3rd phase of the FMF condition in Table 2, suggests that by increasing the threshold from 50% to a higher percentage in the phase 3 of the FMF condition, we might observe a facilitation in the measured Target Look. However, we kept the threshold values as before).

### S1.4. Assessing the Control Groups

For additional confirmations on the role of RVL on Target Look, we repeated the same post-hoc analysis of dependency of Target Look on the RVL in the FSF, MF and FF conditions. The results are shown in table S2.

| | Phase 1 | Phase 2 | Phase 3 |
|---|---|---|---|
| FSF | r(121)= -0.06, p=0.51 | r(120)= 0.025, p=0.78 | r(181)= 0.01, p=0.90 |
| MF | | r(163)= 0.013, p=0.83 | r(163)= 0.088, p=0.33 |
| FF | | r(149)= 0.065, p=0.44 | r(149)= -0.15, p=0.097 |

Table S2. The dependency of Target Look on RVL at the FSF, MF and FF conditions in the different phases of the trial.

Similar to the FMF and FFF conditions, we did not observe a strong interaction in these three conditions between the RVL and the Target Look of different phases, again suggesting that our choices for the thresholds could not potentially altering the measured Target Look.