The Development of Melodic Representations at Early Age: Facts Towards a Computational Model.

Inês Salselas Cabral

Master thesis submitted in partial fulfilment of the requirements for the degree:

Master in Information, Communication and Audiovisual Media Technologies

Supervisors: Dr Hendrik Purwins, Dr. Xavier Serra

Department of Information and Communication Technologies
Universitat Pompeu Fabra, Barcelona.

Spain

September 2007

Summary

Music capability is universal across human cultures. Every human being is able to make sense of a sound sequence, and emotional states are evoked. This research work concerns the appearance of musical capacity, particularly, the early melodic representations looking forward its computational modelling. This way, we propose a review of the existent literature intended to contextualize the appearance of musical capacity, particularly, the early melodic representations and also an organization of the information collected looking forward the conceptualization of its computational modelling.

Finally, a general analysis is done over the whole research work, looking for to point a direction towards future work.

This research work was developed in the context of the Master Program in Information, Communication and Audiovisual Media Technologies and the author was integrated in the EmCAP Project.

Acknowledgments

I would like to thank Xavier Serra, for giving me the opportunity to work at the Music Technology Group.

I would also like to thank Hendrik Purwins for his supervising and for giving me the opportunity to integrate the EmCap.

Finally, I would like to show my gratitude to Perfecto Herrera, for his suggestion, ideas and especially for his support.

Contents

1 Introduction

Music capability, by which every human being is able to make sense of a sequence of sounds, and where emotional states are evoked, is universal across human cultures. In this phenomenon, many questions remain open.

The study of how the brain processes music emerged as a rich and stimulating area of research in cognition, perception and memory. Consequently, studies on how these processes could be modeled started to appear.

However, the origins of musical capacity (whether they are a biological adaptation or cultural invention) and biological and psychological principles involved (aspects of music that are innate and the ones that are acquired by learning mechanisms through the exposure to a culture) are subjects that are still in debate. Consequently, little is known regarding the development of the process of musical cognition and its modeling. This means that a lot of work is still to be done, which makes this research area a fertile, fascinating and interesting field where there it is still possible to make new contributions.

On the other hand, the development of music cognition is a very multifaceted phenomenon that involves conjugations of different brain mechanisms like attention, auditory pattern processing, storage, retrieval and memory. Thus, its modeling leads to the use of perspectives of various domains, such as psychology, biology, neuroscience and engineering, which is a very complex task, but also a very challenging research work.

1.1 Context

This research work was developed in the context of the Master Program in Information, Communication and Audiovisual Media Technologies and the author was integrated in the EmCAP Project. EmCAP (Emergent Cognition through Active Perception) is a research project of the Music Technology Group (MTG) in the field of Music Cognition funded by the European Commission. It started in October 2005 and will finish by September 2008. EmCap investigates how complex cognitive behaviour in artificial systems can emerge through interacting with an environment, and how, by becoming sensitive to the properties of the environment,

such systems can develop effective representations and processing structures autonomously. This way, the project investigates the development of music cognition by combining the complementary approaches of perceptual experiments using human subjects, functional and neuro-computational modelling, and the implementation of an interactive embodied cognitive system.
The present research work emerged from the need to gather data from experimental studies using neonates to determine whether certain basic perceptual abstractions, such as pitch, are innate, or whether they develop through early experience.

Thus, this research work concerns the appearance of musical capacity, particularly, the early melodic representations looking forward its computational modelling. This way, we propose a review of the existenting literature intended to contextualize the problem into the approaches of experimental psychology and computational modelling, and also a conceptualization of a developmental model gathering the information previously collected.

1.2 Problem Statement

Music cognition develops simultaneously with the development of other kinds of characteristics, such as brain development, motor development and specially language development. They might share a common process of evolution with common rules that will adapt and develop in a different manner depending on the specific cognition process. This basic common process, for the auditory cognition, would be shaped in a particular mode depending on individual characteristics such as auditory physiologic constrains, in the case of music cognition, depending on the musical culture that the subject is exposed and more other factors. But what is this common structural process of cognition? And how does it adapt to the case of the music cognition? Which are the variables that make this process evolve in an individual and unique (single) way? How could this be modeled using computational methods?
The present research work concerns the appearance and development of the musical perception and cognition in early age from the perspective of its modelling. More specifically, this work looks for new directions for solutions through the answers to questions such as:

- In musical perception and cognition, what is innate and what is learned through the course of development?
- Are there musical predispositions in newborns?
- What are the musical predispositions?
- How do they develop and refine?
- How does musical representation evolve and get more complex in an elaborated way?
- Can answers for these questions be found in the existent experimental studies?
- How can this developmental process be modelled?

1.3 Research Goals

The present research is specifically concerned with the development of the musical capacity in the perspective of its computational modelling.
The aim of this work is to develop research on solutions for modeling the conditions underlying the early development of musical cognition, focusing just on melodic and non rhythmic musical aspects. To achieve this, the work is divided into more specific stages, such as:

- Collect relevant experimental studies in an eclectic fashion.
- Extract from the gathered experimental studies data that can give clues on which behaviours develop at specific times related to specific stimulus in the course of infant development between prenatal and postnatal stages.
- Use the previous collected information to conceptualize the appearance of music cognition in the first year of life in the form of a developmental timetable

This way, the intention is to search for new elements on how can the appearance of melodic cognition in the early development be modelled through its different development stages, looking for solutions to model the way we learn to extract statistical regularities and build the accumulation and reconstruction of memories.

1.4 Organization of the thesis

Knowledge production is a cumulative process with methodological demands. These demands lead us trough a retrospective view on the solid and relevant findings about our study object. Complementary, leads for future studies develop from the permanent need to consider current results provisory and available as tools for new approaches to the problem. These considerations result of the unfinished nature of knowledge and lead to a never-ending will to know more at to find better solutions for the problems we are set to resolve.

These presumptions underlined this thesis, having had a direct influence in the methodological guidelines adopted that, in turn, are reflected on this document structure.

Hence, this research work whose general aim is to look for clues to model computationally the appearance and development of music cognition in early age, started with a review on studies considered relevant, resulting of a large bibliographic research. This survey, which respects the theoretical structure of the problem in study, aims to build a consistent starting point required for the construction of hypothesis. This review is presented in the Section *State of the Art* and can be divided into three stages.

- *Computational Modelling of Music Cognition* looks for to gather information on modelling human cognition and also some work done in the computational modelling of some aspects of the music cognition.

- *Methods and techniques for studying melodic cognition in babies* presents the methodology followed by scientists that perform experimental studies in psychology, as well as other techniques used to study babies' perception and cognition of music. This section aims to frame the following one in order to better understand the methods used in the experimental studies in which this research work will rely and thus know the limitations of each method so that one can extract the reliable information from the experimental studies.

- Finally, a collection of behavioural data from experimental studies is done, in order to filter and extract developmental information on music perception and cognition in an early age so that it can be used flowingly. This gathering is presented in the *Experimental data from psychological experiments* Section.

After exploring the different aspects of an initial direction towards the computational modelling of melodic representations in early age, clues are presented to the construction of a model of the human's early appearance and development of the musical cognition and perception. For that, a rule or an explanation is searched to correlate stimulus with behaviours, looking for to answer the questions proposed in the definition of the problem. This is presented in the Section *Towards a Computational Model.*

Finally, in the *Discussion*, a general analysis is done over the whole research work, looking for to point a direction towards future work.

2 State of the Art

Music is present in all cultures and has a remarkable diversity of forms. Every human being is born with the capacity to apprehend meaning and emotion in music without conscious effort, translating spectral and temporal patterns of acoustic energy into music's basic perceptual elements, such as melody, harmony and rhythm.

Despite its ubiquity, musical capacity has been commonly seen as a cultural invention, studied from a humanistic of historical perspective. However, the perspective that considers musical capacity as a biological function is gaining more and more attention from researchers [37], where they build arguments based on issues such as domain-specificity, innateness, and brain localization, so to offer a unified conceptual basis for the study of music processing.

Cognitive scientists have discovered neural basis, domain-specific and detailed knowledge of musical structure even in individuals without musical training [22], [39]. Trehub (2001) [54] explores the notion of musical innate predispositions. The author discusses the implications that research findings on experimental studies may have in the origins of music, focusing on the early appearance of receptive musical skills. The author concludes that "infants do not begin life with a blank musical state". Instead, they are predisposed to attend to the melodic contour and patterning of sound sequences, whether music or speech.

Peretz and Coltheart (2003) [38] investigate how the human mind processes music. Relying on the notion that the music faculty comprises a set of neuronally isolable processing components having each the potential to be specialized for music, the authors propose a functional architecture for music processing. The modular model proposed is illustrated in Figure 1.

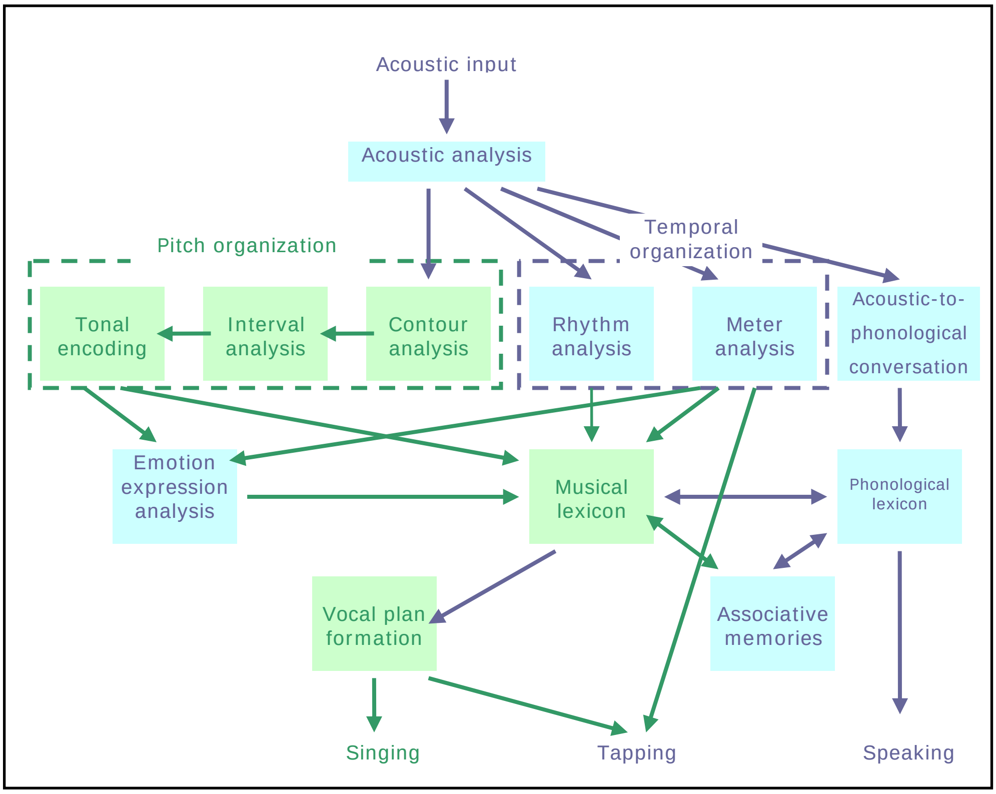


Figure 1 Modular model of music processing proposed by Peretz and Coltheart taken from [38]. Each box represents a processing component and arrows represent pathways of information flow or communication between processing components. Green components represent specific music domain and the others in blue.

This architecture captures typical properties of modular organization, providing valuable guidelines for exploring music capability.

The modules are organized into two parallel independent subsystems, melodic ("pitch organization") and the temporal pathways. The "pitch organization" processing modules have a special interest in the scope of this research The "Contour analysis" module receives the auditory stimuli and processes de pitch trajectories (pitch direction between adjacent tones without regarding the exact pitch intervals) or the speech elements such as intonation and pitch (prosody). This representation of the input will grow in detail in the next modules ("interval analysis" and "tonal encoding"), where pitch intervals and their tonal functions will gain shape. These components send their outputs to the "musical lexicon" that contains all the representations of the specific musical phrases to which one has been exposed during lifetime. But how does this representational system starts to collect information? How do the modules of "pitch organization" appear and evolve? What are their default features, or in other words, what is innate in the human cognition of music? How does music capacity appear?

The following studies have devoted its focus on the process of appearing of musical capacities in different perspectives.

Jackendoff and Lerdahl (2006) [21] explore the capacity for music and its appearance having as a starting point the following five questions:

- What cognitive structures are invoked by music?
- What are the principles that create these structures?
- How do listeners acquire these principles?
- What pre-existing resources make such acquisition possible?
- Which aspects of these resources are specific to music, and which are more general?

Accordingly, the authors examine the previous issues by looking at components of musical organization such as rhythm, tonal organization and affect providing, this way, a synoptic view of the full complexity of the musical capacity.

For Deliège (2001), music perception can be seen as a process of categorization, beginning with motif grouping (rhythm, for example) and grouping of groups, perceiving similarities or dissimilarities between motifs based on the cues they share and culminating through the organization of the traces left by cues into categories organized around prototypes [12].

The author considers categorization a fundamental process by which subjects decrease the complexity and the diversity of the social or physical environment by organizing it. A category can be defined as a class made up of a number of different items that are considered equivalent yet discriminable. The function of category systems is to provide maximum information with the least cognitive effort. The information –processing advantages of conceptual categorization include organized storage of information in memory, efficient retrieval of this information and the capability of responding equivalently to an indefinitely large number of exemplars from multiple categories.

Categorization, or category formation, is involved in the identification and recognition of objects and in the assimilation and organization of new knowledge. Infant categorization refers to infants' ability to respond to stimuli that share common properties.

The next concerning refers to how could this process be computationally modelled through the presentation of contributes in the area of the computational modelling of musical cognition.

2.1 Computational modelling of music cognition

Explaining and understanding the music cognition and perception phenomena is interdisciplinary by its very nature. Thus, it involves the contribution of various disciplines such as musicology, psychology, neurosciences, biology and computer science. In the margins where the approaches of these various disciplines overlap, an explicitly interdisciplinary approach can have the virtue of building theories about mental processes which have proven so far to be too difficult to approach in a mono-disciplinary way.

Underlying many of these attempts to cross between disciplines is the idea that the processing and mental representation of musical structures, whether in pitch or rhythm, can provide a common ground for research. Such structures can be stated formally or informally within music theory, their processing can be investigated by experimental psychology. Both of these aspects can be modelled in computer programs, of which architecture are included the contributes of the biological subtract given by neuroscience.

As a result of this process, theoretical predictions are, in principle, much easier to develop and assess [13].

Therefore, computational modelling seen as a scientific approach can be *"useful to specify cognitive mechanisms in detail by using computational systems and to test the sufficiency of those mechanisms for performing some task in ways that are analogous (at some level) to human performance. The goal of providing a detailed mechanistic account (using computational formalisms) distinguishes modelling from some other approaches, such as verbal models, mathematical formalisms, or flowchart models, that may indicate the general sequence and type of processes without specifying their representation or processing characteristics."* ([48] – pp. 246)

This way, computational modelling has been an influence for cognitive science for its broader view towards defining goals. For instance, modelling has made the cognitive architecture of the mind a central concern. *"Computer architecture is an earlier metaphor that referred to the design of a computer system in architectural terms, by specifying the system's basic components or resources (memories, processing units, instruction sets, communication channels) and their organization. Cognitive architecture refers to the design of the mind in analogous terms. Like a computer architecture, a cognitive architecture permits many different procedures and data structures to be constructed and used within it, but the nature of the* *procedures and data structures is constrained by (and cannot be defined independently of) the architecture."* ([48] – pp. 249)

Computational modelling, thus, has become a well-established method in many fields, including music cognition.

According to Honing (2006) [20], there are two approaches might be taken in music research through computational modelling: Modelling musical knowledge and constructing theories of music cognition.

The first attempts to formalize (model) musical knowledge: These models have origin in music theory in which a systematic formalization contributes to an understanding of the theory itself, its predictions and its scope.

The second approach of modelling music cognition has the aim of describing or explaining the mental processes that take place when perceiving or producing music. Here, the intention is to understand music perception and cognition by formalizing the mental processes involved in listening and processing music. For the author, both of these approaches can be complementary.

In [41], a review of facts and models about music cognition is presented, with a special emphasis on their computational implementations. In this review is adopted a multiple-view position, including facts from different disciplines such as neuroscience, psychology, cognition, artificial intelligence and musicology.

Building explanatory and predictive computational models of human music processing can be faced from different perspectives:

- The direct neuro-mimetic approach, based on computational models of neurons and on population dynamics;
- Modelling processes such as categorization and schema activation, unsupervised and recurrent online machine learning algorithms provide useful tools;
- Higher processes, involving the interaction of symbolic units that can be described by cognitive architectures or by grammar-based approaches that make evident the relationship between music and language.

A valuable cognitive model, this way, would be qualified by its predictive and generalization power (ability to explain data), its simplicity and its relation to existing theories. The model should predict behaviour that is displayed latter in an experiment.

For Cross [11], a computational model should form a bridge between research in Artificial Intelligence (AI) and experimental psychology. On the AI side, the program would be developed to perform a certain task and, within the field of psychology the same program would be interpreted as a model of human cognition and its mechanisms would be tested for psychological validity. Both sides have different aims. For technical AI, a computer program should perform well and even the use of an unintelligent tabular or exhaustive search method is not a priori excluded. For psychology, other objectives exist. It is desirable that the program has characteristics (other than input-output relations) that reflect the properties of human cognition as well, such as working memory and attention limitations, slow learning, or distributed knowledge.

In [11], Cross shows an example of artificial intelligence perspective on how humans understand what they hear. This study is intended to illustrate some of the ways in which artificial intelligence has been able to explain aspects of the nature of human musical experience.

Starting from the question "What is that happens when we experience a piece of music?", the author describes the ways in which research that has been carried out over the last thirty years can successfully explain why we experience what we experience in listening to music.

As a conclusion, the author believes that music has to be approached in the light that human experience is a complex, embodied and encultured set of social and cognitive process, rather than being thought of as simply patterns in sound, and as such the study of musical experience can offer much to artificial intelligence, and conversely, artificial intelligence will have considerably more to offer to music.

In [9], a more specific study is presented on a computational model that extracts melodic patterns from a given melodic surface is presented. In other words, it is done a computational attempt for capturing melodic similarity, namely immediate repetition of melodic passages, with a view to achieving melodic segmentation.

Musical parallelism is an important factor for musical segmentation but difficult to formalize. Following the assumption that the beginning and ending points of “significant” repeating musical patterns influence the segmentation of a musical surface, the discovered patterns are used as a means to determine probable segmentation points of the melody. “Significant” patterns are defined primarily in terms of frequency of occurrence and length of pattern. The special status of non-overlapping, immediately repeating patterns is examined. All the discovered patterns merge into a single “pattern” segmentation profile that signifies points in the surface most likely to be perceived as points of segmentation.

From a cognitive point of view, elaborate pattern extraction processes are more likely to be applied to relatively short melodic excerpts due to the heavy computation involved. This activity is usually more intense at the beginning of a musical piece/section where new musical materials are introduced and established. Once a number of such musical ideas have been extracted, links to further new instances (varied or not) can be made more efficiently: once a pattern has been extracted from a local context, it is placed in long-term memory (i.e., it is learned); when the pattern is encountered again, later in the musical surface, it is recognized and used for further parsing of the surface. The proposal here is that pattern extraction takes place primarily within a short temporal window, and it assists chunking the melodic input into meaningful units, thus expanding the storage capacity of short-term memory. Repetition expands the mnemonic capabilities of short- term memory in the sense that more elements/ chunks can be held by short-term memory.

The effectiveness of the proposed melodic representations and algorithms is tested against a series of melodic surfaces illustrating both strengths and weaknesses of the approach.

However, the study is still a long way from providing a robust, flexible, and general model of melodic parallelism. Yet the model shows potential. Further research is necessary to improve the model and to evaluate it on a much larger scale.

"The Music Projector, a Mock-up Expectation and Emergence of categorization of Pitch Sequences." [42] suggests a framework for exploring and evaluating hypotheses, theories and findings in music cognition. Given its modularity architecture, other work and research in the field can easily be integrated. Recombination of building blocks permit flexibility of implementation and change.

Trough concepts such as feature extraction, change/novelty detection, categorization, association, expectation, attention and long term memory, the information flow chart in a listening process can be composed, consisting of the following stages: pre-processing, fusion, grouping and schema. These steps are described as also relevant algorithms that can be used in its implementation.

As a future work, the implementation and testing of the interaction of the mentioned components is pointed, updating the system whenever specific neuroscience and psychology evidences occurs.

Although there is a considerable number of models proposed for many musical domains in the form of working computer programs, the psychological validation of these models is often far from satisfactory.

The most common way of evaluating a computational model is to show a good fit with empirical data (the model should reflect human behaviour). A good fit between a theory and the empirical observations should then be a starting point (but not the end point) of model verification [43]. Moreover, without reliable data about human behaviour and brain functions (against which to compare a model's behaviour), the psychological value of computational models will always be limited, however well constructed. Furthermore, computational models should not exist separated from existing theoretical knowledge about music: musicological knowledge has an important role to play in providing formalized constructs which may be modelled and tested to investigate their cognitive veracity.

Therefore, in order to allow a ground truth on experimental data for model construction and computational modelling, a bibliographic survey has been done, based on experimental studies.

But first, in order to allow a proper analysis and better understanding of the experimental studies, a brief study on the methods used in those experiments was done and is presented next.

2.2 Methods and techniques for studying melodic cognition in babies

It is very important to know and understand how experimental studies are processed. Despite the fact that some of experimental studies seem to get to contradicting conclusions, the apparent contradiction may be due to different experimental designs, namely, differences in stimuli or experimental tasks. Not having knowledge in these issues may lead to wrong conclusions.

Thus, being conscious of how the results were achieved in experimental studies will allow extracting useful information in a critical way, so that cognition based computational models can be constructed.

Moreover, it is essential to rely our research on solid experimental works in order to achieve consistent and coherent results. In the search for rigorous experiments, one should give attention to possible existent pitfalls. But how to look for good significant studies? How to identify a serious experiment with quality?

There are a few aspects in experimental methodology that might help to recognize a good experimental design that will be discussed next.

Experimental psychology looks for to describe, predict, explain and determine the causes of a behaviour. The experimental methodology used is based on the scientific method, exploring new knowledge about common behaviours to a population. In the particular case of musical cognition, experiments are done using infants to test a hypothesis or demonstrate a principle in auditory perception and cognition. For so, the methodology can be divided into the following phases [24]:

1. Prior Hypothesis – Establishing a cause effect relation that explains a certain situation. In other words, the aim is to relate the presence of an evidence (a fact) with the modification of another (behaviour).

2. Experimentation – Verifying the previous hypothesis. In this phase, it is determinant the rigour of the observations and the control of the experimental setting:
    a. The control of variables, the elements that comprise the experimental setting relative to the change of behaviour which nature is unknown. In good experiment design, variables are carefully controlled or accounted so that reasonable conclusions can be drawn from the experiment's outcome. There are three types of variables:
        i. Dependent variables stand for the modification of the behaviour which is aimed to explain. These variables appear as an answer, a consequence of the independent variable.
        ii. Independent variables are the factors responsible for the setting which is about to be manipulated and controlled in order to verify the dependent variable.
        iii. External variables are the elements of an experiment that cannot be controlled, that happen unexpectedly, but may influence the experiment. These variables may well be related to the behaviours and expectancies of the subject observed, to the behaviours and expectancies of the observer and to the circumstances of the experimental setting, such as local or the hour of the day.

In the study of musical cognition, as in most of the studies of general cognition, an independent variable is manipulated in order to look for an effect in the dependent variable (the one that the researcher cannot manipulate).

In a study [56] where Trehub et al investigate in infants the capacity for detecting a semitone change in any position of a five note melody. In this case, the melody or the stimuli, represent the independent variable. The dependent measure, in turn, was the reaction of the infants to the sounds, precisely, the direction to where the infant was looking in response to the melodies presented.

b. Recording of the observations requires the recording of the occurrence and duration of behaviours. This allows latter analysis of data by researchers.

c. Control of the experimental conditions is a very important topic. This corresponds to the validation of the information. One very important issue in quality experiments is the number of subjects involved. That is, the population, or total group of people to which the researcher wishes to generalize her findings. But, what is a representative sample of this population that has sufficient size that one can make correct inferences about the whole population? How many subjects are enough? This number depends on the degree of homogeneity of the studied population, on how many different conditions there are in the study and also depends on what the experimenter expects from the experiment.

   The way that subjects are assigned is also very significant. For example, random assignment of subjects (with same characteristics such as age, sex) to guarantee that they start out the same (ideal experiment would use truly identical subjects in each condition but, because this is impossible, the technique of random assignment is used). The researcher starts with a group of subjects and randomly assigns them to an experimental condition.

   Experimental conditions such as room environment, like lightning, temperature or acoustic characteristics should also be identical.

The stimulus presented to the subjects is another essential issue to have into account. Not only the nature of the stimuli, but also the order and the way it is presented. It is very common to find incoherent results from different studies and this is mainly due to differences in stimuli. For example, in [55], infants were asked to discriminate between a 6 tone melody and one of five transformations. Thus, they were presented with a standard or repeating background melody and then a transformation of the melody would be presented. All melodies, including the transformed, were separated by a 800ms of silence. Infants were treating the transposition or contour preserving transformations as novel, which was not expected at all. The fact that a 3 tone distractor melody was inserted between presentations of standard and transformed melodies changed radically the results, as infant this time didn't discriminate the transposed and contour preserving melodies from the standard.

3. The generalization stage is where conclusions are launched by means of the generalization of the sample data to the universe which one desires to study.

Although research tools have been advancing, only a limited number of methods are appropriate for studying infants' perception and cognition of music. For this specific case, the main behavioural and brain activity methods used will be briefly presented, described and discussed next.

Behavioural measurements

The conditioned head turn procedure is the main procedure in which researchers have relied. This technique is used for conducting auditory assessment with infants (auditory perception and thresholds) and also is used as a standard technique for clinical evaluation of auditory acuity in infants and for psycho-physical research in infant audition. It provides data on detection, discrimination, categorization and perceptual grouping, allowing hypothesising about perception of rhythm and musical structure [61].

This paradigm is based on the principle that humans have a tendency to seek rewards and avoid punishments, establishing predictions and use them to form the basis of decisions that guide behaviour. This way, engaging infants in a task that rewards them for correct responses will lead to a more accurate picture of sensory and perceptual capabilities.

The method is used with infants between 5.5 and 18 months of age, but it is ideally suited for infants between 6 -10 months of age, because beginning at roughly 11 months, infants' mobility increases and they are less content to sit quietly on lap for a very long time.

This method requires an initial training stage, where the infant is taught to turn his/her head to a sound or to a change in a sequence of sounds. In this phase, the infant is familiarized with the visual reinforcer, by activating it immediately after the first presentation of a new stimulus. Following, the conditioning stage takes place, in which experimenters 'shape' the infant to turn towards the reinforcer. As the baby is trained to associate between a behaviour (turning) and a consequence (a change trial) this is called an *Operant* conditioning, also called *response-stimulus* conditioning.

Finally, the testing phase begins. Usually, the infant, seated in parent's lap as in the following figure taken from [61], is presented with repetitions of a background or standard stimulus coming from the loudspeaker and will turn when a changed or new stimulus is presented. Turning correctly (when there is a change in the stimulus) is then reinforced by the illumination and activation of toys. Failure to turn to a change or turning at other times has no consequence. The layout of this method can be observed in the Figure 2

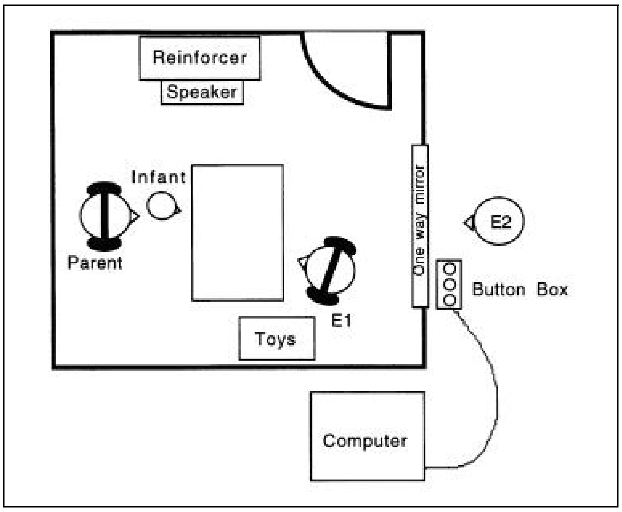


Figure 2 Schematic diagram of the Conditioned Head Turn procedure taken from [61].

Both caregivers and experimenters wear headphones delivering music or noise so they cannot hear the stimuli being presented and this way don't give clues to the infant.

There are some restrictions in the interpretation related to this method. Inferences can be drawn about sophisticated abilities such as categorization. However, results from the experiments using this kind of design can not directly provide information about what the infant perceived in the test procedure. Therefore, it is not made clear if whether the infant recognizes or identifies the target and background stimuli in the same way as adults or if attaches a specific meaning to the stimuli.

Another procedure using head-turn, but not so commonly used in testing infants auditory perception is the Head–Turn Preference Procedure [27]. In the same way as the Conditioned Head Turn procedure, the Head–Turn Preference Procedure takes advantage of facts like children tend to orient visually to an attended sound source and that they also learn to maintain a response (head-turn) when motivating stimulation is contingent on their behaviour.

Particularly, in this method, the goal is to measure the infants' preference (attention) for one kind of auditory stimulus over another. This preference is measured by exposing the infant to one type of sound on half of the trials and the other type of sound on the remainder. Then, the length of time the infant turns his head toward the sound is examined (the infant is allowed to continue to listen to the sound only as long as the head-turn is maintained). The index of preference is then the difference in the average length of the infant's looking time to the two different kinds of stimuli over the test-trial series. The layout of this method can be observed in Figure 3.

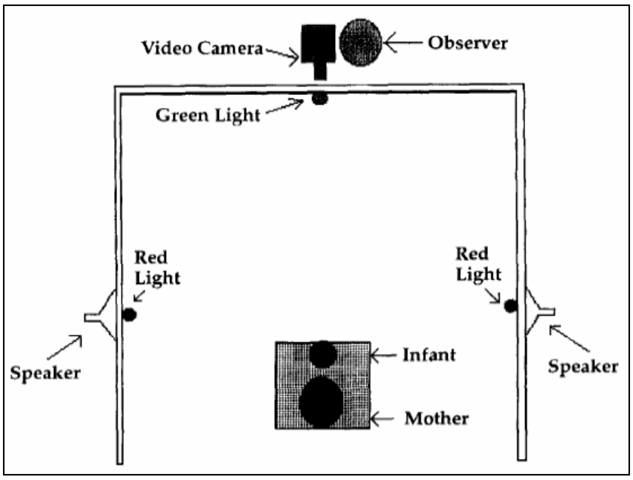


Figure 3 Schematic diagram of the Head Turn Preference procedure taken from [27].

There are also methods that associate infant's reaction to stimuli with differences (changes) in the values such as heart rate (accelerating and decelerating), sucking rate, or changes in infant salivary cortisol. These methods have advantages for being used with foetus and newborns because they still have limitations in body motion capacity.
One of these techniques, the measuring of heart rate accelerating and decelerating was used in [1]. In this experimental study, researchers measure the foetus response to voice and music through the assessment of foetal heart rate change direction using actocardiograms. The foetal heart rate change direction was categorized into a decelerative change, an accelerative change or no change at all.

Brain activity measurements

Another approach to understand and explain musical cognition is to study the brain and its functions [48]. (Chapter 9: pp 295-322)
However, the strategies used in brain studies examining cognition don't differ that much from other type of studies, as the researcher is looking for dissociations on a measure that in this case is brain activation (neuroimaging).

In lesion studies, researchers try to relate behaviours to brain injury. In this case, independent variable is the lesion in the brain of the individual, since it cannot be manipulated. The dependent variables are the same type of behaviours assessed in the traditional cognitive studies. The dissociation in performance between subjects with brain injuries and normal subjects can give clues to the organization of cognitive function.

Event related Potentials (ERP) provide non-invasive measurements of electrical activity on the scalp that are linked to the presentation of stimuli, relating it to the underlying brain function. In this type of study, a number of electrodes are placed on the subject's head in a variety of locations. These electrodes will measure the electrical potentials. In ERP studies, the independent variables are the same as those in behavioural studies but the dependent variable is the change in electrical potentials on the scalp produced by the stimuli. The dissociations in potentials between stimulus types provide clues to the neural processing underlying the cognitive function.

Mismatch Negativity (MMN) is used to investigate sound discrimination in auditory cognitive function. It is assumed that MMN response results from the existence o specific neurons for sound change detection and reflects a comparative process between a novel sound and the auditory memory trace [14]. For so, sounds are presented in a sequence of a standard (frequent) sound intermixed with a deviant (infrequent) sound with different characteristics (frequency, duration or intensity). The MMN is obtained by subtracting the evoked response to the standard from those of the deviant tones. It is called so because it appears as a negative deflection in electroencephalographic (EEG) recordings. The latency of the MMN varies from 100 ms to 200 ms in adults.
MMN provides a tool to access sound change detection capabilities of foetus, infants and children. However, foetal brain is still not completely developed and so this as to be taken into account when interpreting the results achieved using this paradigm.

Functional Neuroimaging is a recent set of techniques developed for studying the relationship between activity in certain brain areas and specific mental functions. There are two main techniques: *Positron-emission tomography* (PET) and *Functional magnetic resonance imaging* (fMRI). Both techniques measure indices of blood flow in the brain, based on the idea that blood flow in the brain is related to neuronal activity. However, the method used in each technique to measure the blood flow differs. In PET, an injection in the subject with a radioactive isotope is needed to measure the blood flow while in fMRI the measure is made taking advantage of the natural magnetic properties of oxygenated and deoxygenated haemoglobin, and so there are no injections or radiation.
Functional imaging provides coloured maps of where in the brain the processing is occurring, providing a great deal of data about cognitive processes. Therefore, this is a safe and non-invasive technique to measure localized brain functioning in subjects.

Methods help the researcher to have a conducting line to imprint coherence to its research. Moreover, the described methods allow testing infants' musical sensitivity before they have the capacity to verbally report no their experiences
Despite this, all methods have limitations. One limitation that affects the methods referred is the difficulty in interpreting the results. Moreover, although the procedures indicate discrimination, they don't tell exactly "what" the infant has perceived. In other words, they don't exactly define whether the infant recognizes or identifies the stimuli in the same way that adults do.

The point is to know the limitations of each method so that one can extract the reliable information from the experimental studies.

2.3 Experimental data from psychological experiments

The aim of this section is to review the literature on experimental studies of music perception and cognition at an early stage of development.

We think that this bibliographic research would benefit from looking into studies done with infants in the first years of life, comparing them with similar studies made with adults. More over, infants' lack of cultural exposure and learning (contrarily of adults) could indicate clues on whether determent features of music cognition are innate or not. For so, in this review, we focused on experimental studies coming from developmental psychology with infants in the first years of life.

The results (data) from psychological experiences are organized in time development, before and after birth, and presented by features: processing of pitch, interval processing, processing of melodic contour and melody, infant directed speech and singing, processing of scale and music and language. An attempt to analyse these features separately will be done. However, this task becomes very difficult to do because the features are not independent from each other as they happen and are perceived simultaneity. For instance, the processing of the melodic contour depends on the mean and the range of the pitch. Even though, an effort of an individual analysis of each feature will be done as follows.

Prenatal Musical experience

Infants appear to be sensitive to music even before birth. The auditory system is functioning by the sixth prenatal month [63], and although filtered substantially at the high-frequency end, environmental sounds do reach the foetus.

Research has shown that prenatal exposure to music and human voice alters the foetal behaviour. In [1], experiments were done with foetuses 37-40 weeks old. Their results showed that when exposed to auditory stimuli, both music and voice at SPL of 105 and 94dB, respectively foetuses responded with an accelerative heart rate change. The experiment also reports that changes in frequency of the sound affected the foetal response.

Another study made with foetus [23] achieves results that indicate fetal voice processing. In this experiment, foetuses are exposed to stimulus consisting on a tape recording of their mother or a female stranger reading a passage. Fetal heart rate increased in response to the mother's voice and decreased in response to the stranger's, which shows a preference of the foetus for the mother's voice.

Though still little is known about prenatal auditory experiences, these studies point towards a promising future in the study of prenatal auditory processing.

The Table 1 presents a summary of the most relevant experimental studies found in prenatal music experience.

Paper	citation	experimental paradigm / kind of experiment	subjects	stimuli	results	Other
Foetal Response to Music and Voice. [1]	2005 Al-Qahtani, N. H.	Computerized assessment of foetal heart rate (acceleration) and activity in response to the stimuli	20 foetus (37-40 weeks)	15s of music and voice at different sound pressure levels via headphone on the maternal abdomen.	Foetus responded with heart rate acceleration and motor response to both music and voice.	Prenatal exposure to music and voice alters the foetal behaviour
Sound frequency change detection in fetuses and newborns, a magnetoencephalographic study. [14]	2005 Draganova, R.; Eswaran, H.; Murphy, P.; Huotilainen, M.; Lowery, C.; Preissl, H.	MMN paradigm in response to auditory stimuli to assess the ability of the fetal brain to detect sound changes	12 foetus between 33 and 36 gestational weeks	Sequence of two complex sounds at 110 dB with varied interstimulus interval between 500ms and 1100ms. Standard sound (prob. of 88%) a 500Hz tone (100ms duration)with additional harmonics at 1000Hz and 1500Hz. "Deviant" sound (prob of 12%) a 750Hz tone with harmonics at 1500Hz and 2250Hz. Amplitude attenuation of harmonics by 3 and 6dB, respectively.	Response to detection of sound changes was found in 60% of the fetal data. Because the fetal brain is not completely developed, the results cannot be interpreted as analogous to those of adults and children studies.	Sound (stimuli) generated by a speaker located outside the room and delivered to the maternal abdomen through a plastic tubing. The distal end of the tube was attached to an inflated bag, placed over the maternal abdomen. 80dB of sound intensity reached the fetus, given the abdominal attenuation.
Effects of Experience on Fetal Voice Recognition. [23]	2003 Kisilevsky, B. S.; Hains, M. J. S.; Lee, K.; Xie, X.; Huang, H.; Ye, H. H.; Zhang, K.; Wang Z.	Computerized assessment of foetal heart rate (acceleration) and activity in response to the stimuli	60 fetus 38,4 week (median)	Two minutes periods with: no stimulus, voice (mother or a female stranger reading an adult poem), and no stimulus.	Foetal heart rate increased in response to the mother's voice and decreased in response to the stranger's.	Precocious language processing abilities observed in newborns may not be due to a hardwired speech-processing module in the brain. Foundation for speech perception and language acquisition may be laid before birth, since in utero experience may play a critical role.

Table 1 Prenatal music experience.

Postnatal musical experiences

Processing of pitch

Pitch is one of the most obvious psychological attributes of sound. This feature is fundamental to melody in music and prosody in speech. Pitch perception by humans can be seen as having two dimensions: the pitch height and chroma [61]. Pitch height is related with the similarity perceived between notes separated by an octave interval and provides a basis for segregation of notes into streams to separate sound sources. The pitch chroma is related with the cycle of notes within the octave and provides information conveyed by a particular sound source.
The Figure 4 illustrates this concept where pitch height is the vertical dimension and the chroma is the horizontal dimension.

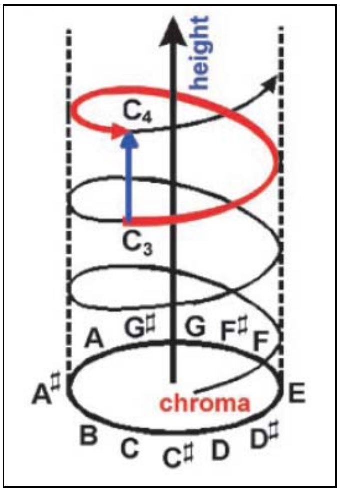


Figure 4 Pitch helix taken from [61].

Despite extensive research, the representation of pitch by infants is a subject in which still little is known. This is essentially due to most of the experiments on infants' perception of pitch being done with successive pitches, like they occur in melodies of music.
Nevertheless, it is known that infants have the ability to categorize and discriminate single complex tones on the basis of pitch. In fact, in [30], Schoon et al conclude that infants' resolution of frequency is finer than that required for musical purposes. In this study made with infants from 4- to 10-month-old, using an operant head turn technique, thresholds for frequency differentiation were measured. At a standard of 1000 Hz,

presented at 70 dB infants showed thresholds averaged 21.6 Hz while thresholds found for adults were 7.4 Hz.

In [7], Clarkson and Clifton evaluated more sophisticated pitch perception abilities such as perceptual constancy and perception of the missing fundamental. In this study made with 7-8-month-old infants, researchers found that infants learned to categorize spectrally different tonal complexes according to the pitches signalled by their fundamental frequencies and also infants heard tonal complexes that signalled the same pitch categories but for which the fundamental frequency was removed.

Moreover, in [6], experiments with 7-month-old children have shown that their discrimination of sounds, having the fundamental and more than two harmonics was better that that of two-harmonic complex (having fundamental frequencies of 160Hz and 200Hz). Therefore, the results suggest that infants' performance in a pitch perception task deteriorates as the number of harmonics in a tonal complex decreases. The effect of number of harmonics held both for sound containing the fundamental frequency and for sounds lacking energy at that frequency.

An experiment made by Trainor and Zacharias [53] shows that 6-month-old infants prefer higher- over lower-pitched singing. Therefore, infants preferred songs sung one perfect fifth above over the lower pitch version (starting pitches of the lower versions were 220, 204,258, and 204 Hz). Researchers, considering the fact of the preference of higher pitches by infants, refer that the results may be due to the auditory system matures first for high frequencies. Another explanation could be infants are more familiar with women's than men's voices or even that infants prefer the timbre, that is, voice quality changes that inevitably accompany singing or talking at a higher pitch.

Hence, the experimental results shown that six month infants discriminate better high pitches and also have a preference for this type of pitch.

Another study made with 5, 7, 9 and 11 years old children and adults achieves results that indicate that maturation of the child's ability to discriminate frequencies may only reach adult levels until 7 years of age [49].

The Table 2 presents a summary of the most relevant experimental studies found in infant pitch perception.

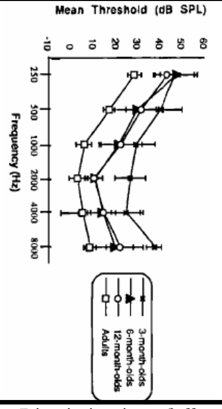
Paper	citation	experimental paradigm / kind of experiment	subjects	stimuli	results	Other
Pure-Tone sensitivity of human infants. [31]	1988 Olsho, L. W.; Koch, E. G.; Carter, E. A.; Halpin, C. F.; Spetner, N. B.	Observer-based procedure[1] to test pure-tone thresholds (dB SPL)	30 3-month old infants that were also tested at 6 and 12 months old and 6 adults 18-26-years old	Pure tone bursts with a duration of 500msand 500ms between bursts. Range frequency from 250 to 8000Hz		Pure-tone thresholds improve between 3 months and adulthood. Between 3 and 6 months, the improvement is greater at high frequencies. Improvements at low frequencies do not occur until after 12 months.
Infants' Perception of Pitch: Number of Harmonics [6]	1996 Clarkson, M. G.; Martin, R. L.; Miciek, S.G.,	Discrimination and categorization of harmonic complexes containing 2, 3 or 5 harmonics using head-turn procedure	36 7-month-old infants	Fundamental frequencies of 160 and 200Hz of 500 ms in duration, containing 2, 3 or 5 harmonics.	Discrimination of all tonal complexes successfully categorized only complexes containing 3 or 5 harmonics and failed to categorize 2 harmonic complexes, even when they contained the fundamental frequency.	Pitch perception task deteriorates as the number of harmonics in a tonal complex decreases. Infants may need more spectral information to perceive pitch than adults.
Discrimination of frequency transitions by human infants [2]	1989 Aslin, R. N.	Obtaining the Difference Limens of frequency discrimination using head-turn procedure.	71 6 to 9-month-old infants and 3 adults	Upwards and downwards sweeps of linear frequency. - Standard repeating un-modulated 1 kHz pulse tone of 300ms and 50ms. - 50ms frequency sweeps and a standard consisting of a 50ms frequency sweep of 350Hz	Infants' performance poorer compared to adults. When sweep duration was constant fast rates were discriminated best.	Rapid frequency transitions are much more difficult to discriminate from frequency transitions of the same category. Extraction of spectral information from a frequency sweep appears to be temporally limited and not determined by a rate invariant mechanism.
Infant Pitch Perception: Evidence for Responding to Pitch Categories and the Missing Fundamental. [7]	1984 Clarkson, M. G.; Clifton, R. K.	Perceptual constancy and perception of the missing fundamental, conditioning head-turn paradigm.	7-8-month-old infants	Seven harmonic tonal complexes for each of three pitch categories corresponding to fundamental frequencies of 160, 200 and 240 Hz. The complexes differed in the number and the frequencies of the harmonics contained.	Infants successfully discriminated the harmonic tonal complexes with different pitch. Infants also categorized missing fundamental stimuli consistently with the pitch of the missing information.	Infants have relatively sophisticated pitch extraction abilities, as they were able to recognize the pitch of tonal complexes having different spectral characteristics (including missing fundamental) and categorize those sounds consistently with their pitch.

Table 2 Infant pitch perception.

Interval processing

Most work on how melody is encoded in the auditory cortex has focused on absolute pitch maps. However, melodic formation is thought to be encoded in the brain in two different relative pitch forms: a domain-general contour code (up/down pattern of pitch changes) and music specific interval code (exact pitch distances between notes) [51].

[1] This method combines features of Conditioned Head turn and Forced-choice preferential looking techniques. For more details see [30] and [31].

Thus, because musical pitch is perceived relationally [32] when listening to music, each tone is perceived in relation to previous heard ones.

Intervals are known as the combination of two tones, either melodic, consisting of two successive tones, or harmonic, consisting of two simultaneous tones. An interval is generally characterized by the pitch distance between its component tones or by its frequency ratio, which relates the frequency of one tone to that of the other. For example, tones of 200Hz and 100Hz have a frequency ratio of 2:1 (or 12 semitones or an octave) [47].

It is possible that some combinations of tones may be more pleasing (consonant) than others (dissonant) to listen to.

Plomp and Levelt [40] developed a theory to explain the origin of the sensation of dissonance. They propose that sensory dissonance emerges from the critical band structure of the cochlea, where the activation patterns on the basilar membrane in the inner ear of two simultaneously frequencies interact when the difference between frequencies don't exceed this bandwidth. Thus, due to the fact that overtone structure of complex tones with energy at components that are integer multiples of the fundamental frequency, that is, pitch, and also due to each component in one tone can potentially interact with each component of the other tone, the net result is that tones whose fundamental frequencies stand in small-integer relations. The Figure 5 taken from [65] may help to understand this previous explanation.

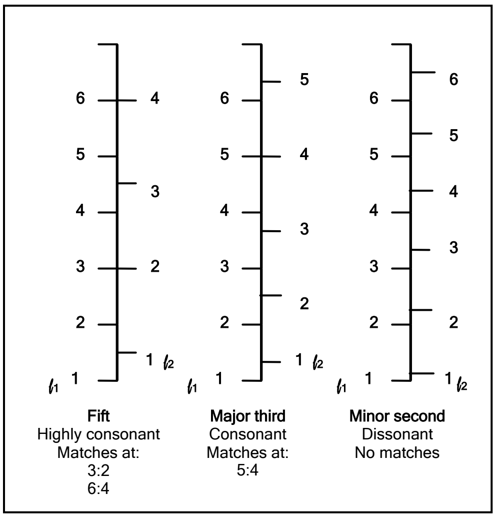


Figure 5 Upper harmonics of representative musical intervals and their superposition taken from [65].

In [52], experiments shown that infants as young as 2 months old are sensitive to the dimension of consonance and dissonance, as they prefer to listen to intervals that are consonant, even before knowledge of scale structure.

Furthermore, in "Natural musical intervals: Evidence from infant listeners" [46], 9-month old infants were tested with harmonic (simultaneous) intervals. Infants could detect changes to simultaneous intervals with small integer frequency ratios (3:2 – perfect fifth; 4:3 – perfect fourth) more readily than to simultaneous intervals with complex large integer ratio (45:32 - tritone). Other experiment in the same study made with 6-month old infants tested subjects with melodic (sequential) intervals. Infants exhibit perceptual biases for simple small integer frequency ratios (3:2; 4:3) compared with complex large integer ratio intervals (45:32). Schellenberg and Trehub find their results consistent with processing predispositions favouring simple frequency ratios over more complex ratios. The authors also suggest that simple frequency ratio intervals may be natural prototypes, being relatively easy to encode, retain and distinguish from other intervals, compared with complex ratio intervals.

Trehub et al [56] investigated the semitone discrimination by 9-month-old infants, in the context of a 5 tone melody. Results show that infants could detect a semitone increment in the context of a musical sequence. The authors find their results supportive with the idea that musical context does not interfere with the frequency resolution of a semitone.

The Table 3 presents a summary of the most relevant experimental studies found in infants' perception of musical intervals.

Paper	citation	experimental paradigm / kind of experiment	subjects	stimuli	results	Other
Preference for Sensory Consonance in 2- and 4-Month-Old Infants. [52]	2002 Trainor, L. J.; Tsang, C. D.; Cheung, V. W.	Preferences for consonant and dissonant two tone melodic intervals using looking-time preference procedure.	20 infants between 8 and 10 weeks 20 infants between 15 and 17 weeks old	2 sets (consonant and dissonant) of 4 simultaneous two-tone intervals	Infants show preference for consonant over dissonant intervals	2-month-old infants are sensitive to the dimension of consonance and dissonance.
Natural musical intervals: Evidence from infant listeners. [46]	1996 Schellenberg, E. G.; Trehub, S. E.	Infants' ability to detect subtle changes to patterns of simultaneous and sequential tones using operant head-turn procedure.	36 9-month old infants and 54 6-month old infants	Repetitions of the standard intervals (simultaneous or sequential with ratio frequencies of 3:2, 4:3 and 45:32) followed by a change trial with subtle changes of semitones, followed by a return to the standard interval.	Infants detected changes to intervals only when the tones were related with simple frequency ratios.	Intervals with simple frequency ratios may be natural prototypes, being relatively easy to encode, retain and distinguish from other intervals, compared with complex ratio intervals.

Table 3 Infants' perception of musical intervals.

Processing of Contour and Melody

Melodic contour seems to be the most salient feature of melodies for infant listeners. Research on melodic perception led to the general conclusion that infants rely more on the melodic contour than on local changes in pitch height or interval size: Infants discriminate melodies if the melodic contour is changed, even if the size of the intervals is kept unchanged; moreover, they do not perceive a transposition as different from an original melody. Finally, infants are able to categorise melodies, irrespective of the changes in their intervals or the absolute pitch height, so long as the melodic contour remains unchanged.

In [55], 8- to 11-month-old infants were presented with 6-pitched melodies in 5 conditions: original, transposed, contour preserved but with a few pitch changes, octave change with contour preserved, and octave change with contour violations. They found that infants could not discriminate the transposed or contour preserved melodies but could discriminate all the other altered versions of the melody. They concluded then that infants encode melodic information by focusing on contour.

Another study made by Ferland and Mendelson [15] achieve results suggesting that 10-month old infants can form contour categories even in the presence of considerably distracting information (wave form of the tones, that is, timbre, and frequency) and respond differentially to members and non-members of the category. Thus, authors conclude that infants, to process auditory signals, must attend to contour and ignore significant differences across signals.

The general conclusion is that infants rely preferably on a global processing strategy, that is, the global melodic movements made of unison-ascending-descending intervals, rather than on local processing strategy (exact interval size or absolute pitch height). Adults recognize a melody effortlessly even if manipulated in terms of altering the absolute pitch but preserves the relative pitch distances. This centrality of relative pitch suggests a role for an innately specified auditory mechanism for encoding stimuli in terms of distances between pitches [28].

The Table 4 presents a summary of the most relevant experimental studies found in infants' perception of melodic contour.

Paper	citation	experimental paradigm / kind of experiment	subjects	stimuli	results	Other
Infants' perception of melodies: the role of melodic contour [55]	1984. Trehub, S.; Bull, D.; Thorpe, L.	Melodic contour perception and discrimination of transpositions and other transformations using head-turn procedure	97 infants 8-11 month-old	6-tone melody and 5 types of transformations of the original melody.	Children could not discriminate the transposition or contour-preserving transformations from standard melody but could discriminate all other transformations	Infants treated transposed and contour retaining melodies as similar or equivalent to the original - Encode and remember information about contour as opposed to absolute intervals or specific frequencies - Extraction of information about the overall frequency range of heard melodies. Infants' ability to categorize auditory sequences on the basis of melodic contour formation
Infants' Categorization of Melodic Contour [15]	1989 Mark B. Ferland and Morton J. Mendelson	Melodic Contour discrimination and categorization using preference paradigm	32 10-month old babies (16 boys and 16 girls)	Six rising and six falling contour sequence with 5 tones that differed in frequency, wave form or both	Infants discriminated sequences that differed only in the waveform of their tones Infants associated same contour patterns, despite the presence of distracting dimensions such as frequency and wave form (timbre)	Infants may equate sequences on the basis of melodic contour. Infants may attend to contour and ignore significant differences across signals.

Table 4 Infants' perception of melodic contour.

Infant Directed Speech and Singing

Melodic contour may also be the most salient feature of mothers' speech to pre-linguistic infants and not the semantic content. In this special register, the infant directed speech, prosody (phonetics and melody) is characterized by acoustic adjustments in speech elements and the most salient concern the melodic contours: hyper-articulation, with more extreme vowel formant structure, higher mean pitch, wide pitch range, longer pauses and shorter phrases. In [34], it is developed a quantitative acoustic analysis of contours of the maternal utterances in tone and stress languages (mandarin Chinese and American English), employing a computerized colour-sound spectrographic system. Features were extracted like duration, peak frequency F_0, mean frequency F_0 and frequency range. Results allowed researchers to build different contour categories relating them to specific caregiving contexts, resumed in the Table 5.

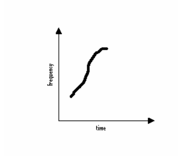
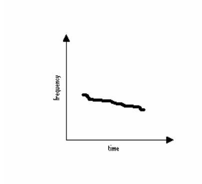
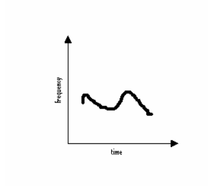
Caregiving Context	Contour shapes	
Encouraging infant to turn	Rising	
Soothing a distressed infant	Falling	
Contingent reward	Bell	
Discouraging unfavourable behaviour	Complex	

Table 5 Contour shapes corresponding to caregiving contexts.

These acoustic properties of melodic contours and the modes in which they are displayed can help the infant to detect, categorize, and abstract elementary holistic units in the flow of speech and also facilitate processing of linguistic and contextual information. For example, in expressions of enjoyment and happiness mean F_0 and F_0 range increase while in expressions like irritation and controlled anger they decrease.

There is also a hypothesis that proposes that this "baby talk" points to biological pre-programmed parenting behaviour [35] that has been selected as a preadapted parental capacity during human evolution and also because of infants' innately sensitiveness to the meaning of emotional expression of vocal affect. Motherese is used extensively in parental speech across a variety of languages and cultures [34]. According to [36], 7-week old infants have preference on this type of speech and show more responsiveness to their mother's own voice. According to [16], the melody carries the message in speech addressed to infants. This conclusion was achieved in an experiment where adult directed and child directed speech were electronically filtered to eliminate linguistic content, in 5 standardized interactional contexts: Attention-bid, approval, prohibition, comfort and game/telephone. In this content-filtered stimuli (prosodic information), the communicative intent of the speaker was correctly identified by subjects with significantly higher accuracy in infant-directed speech than in adult directed speech.

The Table 6 presents a summary of the most relevant experimental studies found in infant directed speech.

Paper	citation	experimental paradigm / kind of experiment	subjects	stimuli	results	Other
Intonation and communicative intent in mothers' speech to infants: is the melody the message? [16]	1989 Anne Fernald	Subjects had to identify the communicative intent using only prosodic information, given 5 forced choice alternatives.	80 adults, 40 experienced parents and 40 inexperienced with infants.	Records of natural samples of infant and adult directed speech from 5 mothers of 1-year old infants in 5 standardized interactional contexts and then electronically filtered to eliminate linguistic content.	Listeners were able to identify the speaker's intent with significantly higher accuracy in ID speech than in AD speech.	There is a relation of prosodic form to communicative function is made uniquely salient in the melodies of caregiver speech and these characteristic prosodic patterns are potentially meaningful to the preverbal infant.
Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages [17]	1993 Anne Fernald	Head turn preference (looking time and facial affect measured) ID Speech – cross cultural languages	20 5-month old (English)	Approval and prohibition vocalizations in ID and AD English. ID speech in non sense English, German, Italian and Japanese.	Respond with more positive affect to approvals and more negative affect to prohibitions. ID vocalizations elicited differential emotional responses in infants, while AD vocalizations did not. Infants showed evidence of differential responsiveness to affective vocal expressions in German, Italian, non sense English and English but not in Japanese.	Not responding to Japanese may be related with cultural differences of not expressing emotion.
The meaning of melodies in motherese in tone and stress languages. [34]	1991 Papousek M.; Papousek, H.,	Analysis of the meaning of melodic contours in maternal speech in relation to eight interactional caregiving contexts and compare contour – context relations across tone and stress languages	10 American and 10 Chinese mothers and their 2-month old infants	Recordings of spontaneous interactions between 10 American and 10 Chinese mothers and their 2-month old infants	Melodic contours in parental communication with pre-syllabic infants represent cross-linguistic universals which may function as guiding messages in communication.	Maternal melodies were categorized by contour type and were subjected to digital acoustic analysis.
Infant responses to prototypical melodic contours in parental speech. [33]	1990 Papousek, M.; Bornstein M. H.; Nuzzo, C.; Papousek, H.; Symmes, D.	Melodic prototypes' influence in infants measuring their visual behaviour (looking time and photography of the infant's face)	32 4-month-old infants	Female and male vocal sounds without any linguistic information in an approving and disapproving infant directed speech context. Also the reverse copies of the contours.	Approving contours recruited infant looking and disapproving contours inhibited it.	Melodic prototypes may function as didactic caregiving messages for infants.

Table 6 Infant Directed Speech – Prosody.

There are also differences between Infant Directed Singing and non-infant-directed singing. Not only do songs for infants differ structurally from other song categories, but parents also perform songs differently when singing to their infants than they do otherwise.

In [50], acoustic differences between infant-directed and non-infant-directed singing were examined. They found that these types of singing differed in a number of respects, such as slower tempo, relatively more energy at lower frequencies, lengthened inter-phrase pauses and higher pitch and jitter factor (measure associated with increased intensity of emotion). In addition, Pitch variability was higher and the rhythm exaggerated in infant-directed play-songs, but not lullabies.

These acoustic modifications in infant directed singing likely attract infants' attention and may be used by adults to regulate infants' states to communicate emotional information [44].

The Table 7 presents a summary of the most relevant experimental studies found in infant directed singing.

Paper	citation	experimental paradigm / kind of experiment	subjects	stimuli	results	Other
The Acoustic Basis of Preferences for Infant-Directed Singing [50]	1997 Trainor, L. J.; Clark, E. D.; Huntley, A.; Adams, B. A.	Analysis of acoustic differences between infant and infant-absent directed singing.	-	Pairs of recordings (infant and infant-absent directed singing) of 6 play-songs and 4 lullabies	Slower tempo, relatively more energy at lower frequencies, lengthened inter-phrase pauses and higher pitch and jitter factor. In addition, Pitch variability was higher and the rhythm exaggerated in infant-directed playsongs, but not lullabies.	Acoustic modifications in infant directed singing likely attract infants' attention and may be used by adults to regulate infants' states to communicate emotional information.
Distinctive Messages in Infant-Directed Lullabies and Play Songs. [44]	1999 Rock, A. M. L.; Trainor, L. J.; Addison, T. L.	Adults were presented to the stimuli and had to identify the different versions of singing.	72 undergraduate students 18-30-years old	Recordings of 19 mothers singing to their 6-7-month-old (lullabies and play-songs)	Adults identified the play-song-style and lullaby-style versions with 100% accuracy.	Play-songs were more brilliant, clipped and rhythmic. Lullabies were more airy, smooth and soothing.
Mothers' singing to infants and preschool children. [3]	1999 Bergeson, T. R.; Trehub, S. E.	Mothers were recorded singing two versions of the same song, one to their infants and other to their preschool children. Then adults listen to the records and try to identify the infant directed version. Acoustic measures like pitch, tempo or intensity were measured	Mothers: 18 with infant 6 to 11 months and preschoolers 2 to 11 years. Adult listeners: 31 women and 35 men 19 to 66 years	Lullabies and play-songs.	Adults accurately Identified the infant directed version from each pair of mother's songs.	Mothers sing at high pitch level for their infants than for their preschoolers. The lyrics are more clearly pronounced when sang for preschoolers.
Maternal Singing in Cross-Cultural Perspective [60]	1993 Trehub S. E.; Unyk A. M.; Trainor L. J.	Adult listeners had to identify the infant directed song version from songs sung to infants and in infants' absence.	American and Indian adult listeners.	Pairs of songs sung by mothers (American and Indian) once to their infants and once in the infant's absence.	Listeners identified infant directed songs.	Findings suggest a distinctive style of singing to infants across cultures and musical styles.

Table 7 Infant Directed Singing.

Processing of scale structure

A musical scale could be defined as a set of pitches, usually with an interval of repetition. This interval appears divided into unequal steps. A study made by Trehub et al [57] investigates if this unequal division arises due to confer processing advantages. Thus, adults and 9-month-old infants were required to detect mistuned tones in one of the following ascending-descending scales (15 tones):

a. A major scale, potentially familiar, with unequal steps.
b. An artificial analogue of the major scale, unfamiliar, with unequal steps.
c. An artificial scale with equal steps, unfamiliar.

Adults' performance was better on the major familiar unequal-step scale and equally poorly on both unfamiliar scales. Infants, in turn, detected mistuned tones only in the scales with unequal steps, performing equally on the major and artificial unequal step scales. Authors, thus, conclude that this feature of unequal step division of scales, present in an almost universal way, facilitates pitch processing in infancy.

Despite these common features among scales, they appear in diverse forms across cultures. And the question that arises is if our receptiveness to these scales is innate or rather a product of acculturation. Trehub et al [56] investigated the development of sensitivity to the semitone and diatonic structure, two musical relations considered significant in western tonal music. For that, they examined 9 to 11-month old infants and 4 to 6-years old children, requiring them to detect a semitone change in any position of two five-note melodies. One with diatonic tones only and other that included a non-diatonic tone. Preschool children were superior in detecting the semitone change in the diatonic context, compared with the non-diatonic context while infants' performance was not influenced by the diatonic context, having detected the semitone change equivalently in all positions. These results obtained offered no support for an innate bias favouring diatonic over non-diatonic structures but indicated priority o diatonic structure at early ages.

The Table 8 presents a summary of the most relevant experimental studies found in infants' perception of musical scales.

Paper	citation	experimental paradigm / kind of experiment	subjects	stimuli	results	Other
Infants' and adults' perception of scale structure. [57]	1999 Trehub, S. E.; Schellenberg, E. G.; Kamenetsky, S. B.	Detection of mistuned tones in multi-tone sequences	Adults and 9-month-old infants	A major scale, potentially familiar, with unequal steps. An artificial analogue of the major scale, unfamiliar, with unequal steps. An artificial scale with equal steps, unfamiliar.	Adults' performance was highly accurate on the familiar major scales and poorly on unfamiliar scales. Infants showed better performance on unequal step scales and significantly more poorly on the equal-step scales.	Unequal step scales may facilitate pitch processing in infancy.
Development of the perception of musical relations: semitone and diatonic structure. [56]	1986 Trehub S. E.; Cohen, A. J.; Thorpe, L. A.; Morrongiello, B.A.	Subjects were tested for their detection of a semitone change in any position of a five-note melody (composed of diatonic tones or containing a non-diatonic tone), using conditioned head turn procedure (infants).	45 9 to 11 month-old infants and 34 4 to 5 years old children	Five-note melody. Ascending and descending major triad and augmented triad.	Infants and preschool children detected a semitone increment in the context of a musical sequence, but only children showed superior performance for the diatonic sequence.	Diatonic structure may only emerge by 4 to 6 years of age.

Table 8 Infants' perception of musical scales

Music and language

Although the existing imaging and neurological evidences [38] pointing for a cortical separation of the faculties of music and language, some interesting parallels can be drawn between both, during the development. Music and language may seem very different systems on the surface, used by humans to communicate and express distinct information with different meaning. However, from the perspective of an infants that must learn about each system before understanding its communicative intent, music and language may be very similar and, therefore, at an early development stage, infants may use a single mechanism underlying learning both domains [29]. Aspects in the developmental process such as early perception of sound, pre-musical and pre-linguistical vocalization and the emergence of singing and speech may give some interesting hints that music and language should be closely related.

Subsequently, some similarities were investigated between music and language in order to establish a link between both and shed light on the understanding and explanation of the emergence and development of musical capabilities during early development.

The first common aspect to highlight in music and language development is the perception of sound. Both are perceived as sounds with elements such as frequency, intensity, duration and tempo. Furthermore, events are organized temporally, with relevant structures unfolding in time and spoken languages, similar to music, reach our perceptual system as frequency spectra, arrayed as pitches.

With the Figure 6 taken from [10], Chen-Hafteck shows links between the developmental process in music and language.

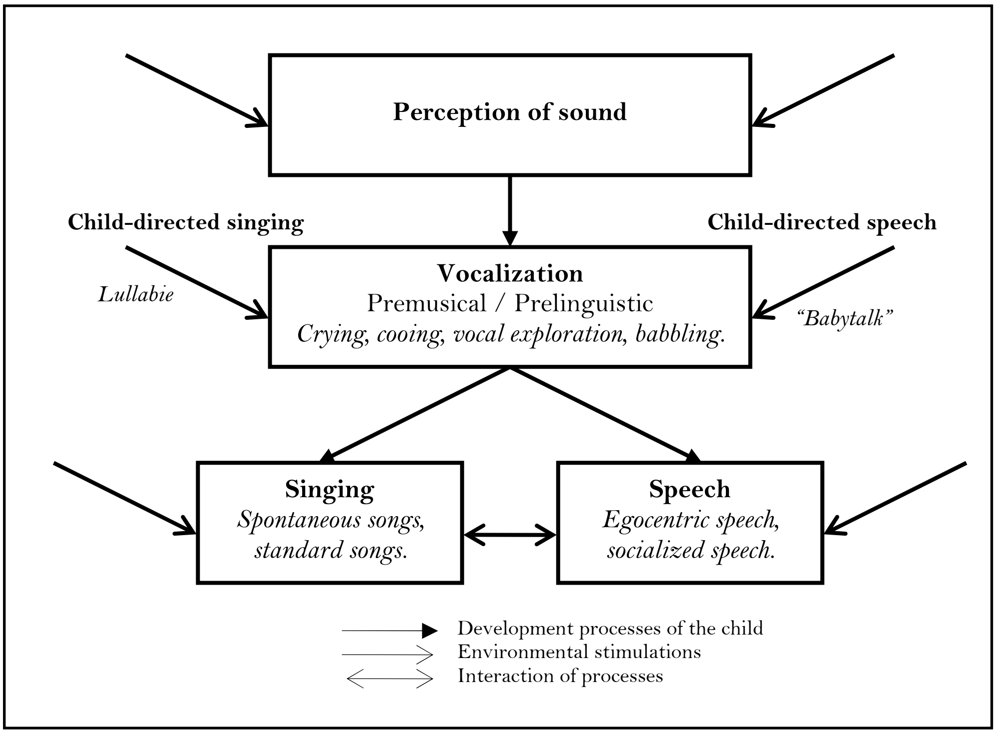


Figure 6 Developmental processes in music and language taken from [10].

Chen-Hafteck believes that music and language development proceed in close relationship with each other and that the two domains are indistinguishable during the early stages, as the figure indicates.

Another interesting study made from McMullen and Safran explores the developmental comparison between music and language [29]. Some interesting clues to the link between music and language are given, and can be briefly summarized as following.

- Prosodic structure of music and language, characterized by patterns of rhythm, stress, intonation, phrasing and contour are highly salient to infants and most likely are responsible for the early processing in both systems. The exposure to this kind of stimuli starts even before birth. The filtering properties of the utero (to which the fetus is) leave the rhythmic cues intact relative to high frequency information and thus, foetuses start to receive and learn the incoming rhythmic patterns. The fetal learning also includes the rhythmic patterns of the mother's native language, making possible to newborns infants to use this experience to differentiate between languages.
- After birth, infants continue to be exposed to prosodic information in both domains. Both linguistical and musical stimuli (infant directed speech and singing) are modified by caregivers to be more attractive to infants. These modifications are mainly characterized cross-linguistically by slower rate of

speech, higher fundamental frequency, greater range of pitch variations, longer pauses and characteristic repetitive intonation contours. Learning appears to be facilitated by these exaggerated prosodic contours used in infant directed style and additionally, the prosodic cues may also play a role in delineating the structural information that infants must learn to process in language and music.

- Similar strategies are used by infants to represent auditory experiences in memory in both linguistic and musical systems in order to learn successfully, permitting the subsequent accumulation and manipulation of acquired knowledge. That is, some aspects of music and language may be acquired via the same learning mechanism, for example, the statistical learning, that means that the acquisition of knowledge is made through the detection of patterns of stimuli like sounds (musical tones), words or other units in the environment that cue underlying structure.

There is still an open question whether infants use the same mechanisms to detect the parallel cues discussed across domains, or whether instead they have learned about these similar prosodic properties independently. Despite this, many interesting suggestions are given that point to a better understanding of the emergence and early development of music cognition.

There are other kinds of experiments that can give clues for computational modelling of musical emergence. An example is Psyche Loui et al [25] which examine the ability of humans to acquire knowledge via passive exposure to new musical grammars based on a non-western tuning system, creating melodies as legal exemplars of each grammar. These musical grammars were constructed such that they obeyed some psychoacoustic and Gestalt Principles[2], but were completely novel to all participants. This way, the learning of artificial (new) musical grammars were investigated, with the specific intention of probing for recognition, expectation and preference formation in the new musical system.

After participants (adults) were exposed repeatedly to five melodies for 25 minutes, they recognized and preferred melodies they had heard, but could not generalize their knowledge to new melodies composed in the same grammar. In other experiment done, exposition was made to a larger number of melodies (15 instead of five) for the same overall length of time. This increase in exemplars and reduction in repetition affected

[2] Gestalt Principles such as Similarity, good Continuation, Closure or Proximity are based on the theory that human brain tends to organize elements into groups when these principles are applied.

participants' performance in the forced-choice tests. Participants showed in this case not only reliable recognition, but also some evidence of generalization.
The results achieved suggest that the subjective experience of musical preference is not directly related to the increase of statistical sensitivity or to the understanding of grammatical structure. Instead, preference may be a result of recognition and familiarity.

"Statistical learning of tone sequences by human infants and adults" by Saffran et al [45], is another example for providing clues for computational modelling of musical emergence. In this study, researchers investigated whether statistical learning ability is uniquely tied to linguistic materials. For that, subjects were exposed to continuous non-linguistic auditory sequences whose elements were organized into 'tone words'. This way, statistical information was the only word boundary cue available to learners. Both adults and 8-month-old infants succeeded at segmenting the tone stream, with performance indistinguishable from that obtained with syllable streams. These results suggest that a learning mechanism previously shown to be involved in word segmentation can also be used to segment sequences of non-linguistic stimuli.

As a synthesis, this survey provided an exploration of information that could shed light on a way to understand how to approach the modelling of the appearance of music perception and cognition. For that, an inventory of several dimensions of the problem was done through: gathering research produced on computational modelling of music cognition; studying methods and techniques used in experimental research; and finally, collecting data from experimental studies. These procedures look for to allow the construction of a developmental timetable to give clues in the modelling of the development of melodic representations at early age.

3 Towards a computational model

Modelling a theory and implementing it as a computer program may be of an extreme importance, because its constructing involves level of detailed explanation that is not required by verbal description of the behaviour. This way, going deep into the understanding of a theory in order to model it, leads to explicitness or, in other words, a clear and exact description of the phenomena.

Also, the modelling of a theory may lead to consequences, for example, interactions between different parts of the model, that otherwise could be difficult to predict.

Finally, unexpected behaviours that may arise from simulations can suggest new experiments to be run with humans and, thus, contribute to new steps in the production of knowledge.

In the present work, the aim is to present clues to the construction of a model of the human's early appearance and development of the musical cognition and perception. For that purpose, experimental studies were collected in order to extract useful data so to investigate if there are capabilities present at birth or even before, what is learned through exposition or what develops with the unfold of time.

The search for data from experimental studies as an iterative process, in which studies found would lead to new questions, new researches and consequently to new studies. The process of collecting behavioural data from experimental studies is a hard task whose limits are difficult to establish and for so, it seems always to be unfinished.

In order to facilitate the analysis of the gathered studies, a summary of each of the considered relevant papers was done and organized in several tables. These tables contain the title of the paper, the citation, the kind of experiment, subjects and their age, the stimuli, the results achieved and other information considered relevant. Each table gathers papers that refer to one of the following topics: Prenatal music experience; perception of pitch; perception of musical intervals; perception of melodic contour and melody; Infant directed speech; Infant directed singing; and perception of scales. These tables (Table 1 to Table 8) were presented throughout the Section 2.3 (*Experimental data from psychological experiments*)

The survey done provided information about musical capacities in infants and also about the development of the musical cognition. Despite, after the exploration stage, it still is necessary to describe the collected data, as well as finding an explanation to it. For that, the next stage is to build hypothesis about what associates a certain stimuli with a behaviour. The aim is to find a rule, an explanation that correlates stimulus with behaviours, looking for to answer the questions proposed in the definition of the problem.

- Infants have as a basis the physical ability to hear sounds (music or language) and discriminate them in different levels and ways through its elements such as frequency, intensity, duration and tempo (innate mechanisms).
    - The foetus's alteration in behaviour facing exposure to music reveals that these physical abilities are present before birth specifically auditory discriminative ability.
    - The preference for intervals with simple frequency ratios over complex ratio intervals observed in infants relates with the fact that these (simple frequency ratio intervals) are natural prototypes and thus, relatively easy to encode, retain and distinguish from other intervals, compared with complex ratio intervals.
    - The sensitiveness verified in 2-month-old infants for the dimensions of consonance and dissonance reveals that consonance perception may provide a bootstrap into the task of learning the pitch structure of the musical system to which the infant is exposed.
- The auditory system of infants and adult's manifest common mechanisms, despite the infants' auditory system may not be totally developed and so showing an overall deficit in frequency resolution and constrains in basic psychoacoustic abilities, specifically:
    - Pure tone thresholds improve between 3 months and adulthood. Between 3 and 6 months, the improvement is greater at high frequencies. Improvements at low frequencies do not occur until 12 months.
    - Pitch perception at 7 months deteriorates as the number of harmonics in a tonal complex decreases. Infants need more spectral information to perceive pitch than adults.
    - Rapid frequency transitions are much more difficult to discriminate from frequency transitions of the same category. At 6-month-old, the extraction of spectral information from a frequency sweep appears to be temporally limited and not determined by a rate invariant mechanism.

- The capacity observed in infants of:
    - Equate sentences on the basis of melodic contour;
    - Attend to melodic contour ad ignore significant differences across signals (such as timbre or interval formation);
    - Treat (at 8-month old) new melodies or tone sequences as familiar if these sequences have the same melodic contour and frequency range as a previously heard sequence and as a novel if either the contour or range differs;

  Is related with the fact that, in general, infants take on a global-processing strategy, in which perception is based on larger units such as melodic contour.
- Infant's sensitiveness and responsiveness to features of infant directed speech, such as:
    - prototypical approving and disapproving contours;
    - affective vocal expressions in German, Italian, non sense English and English;
    - high pitch and expanded melodic contours;

  Means that melodic prototypes have didactic caregiving messages (maintain infant's attention) and these characteristics are related with the sound features that are fundamental to music processing across cultures.
- The system (infant) starts by integrating simple structures sequences (learn) that correspond to simplified representations (example of melodic contour) that, with development will grow in complexity and in different ways, depending on individual characteristics, surrounding culture and musical exposure (diatonic structure, in western culture, only emerges by 4 to 6 years of age).
- The innate capacity of acquiring knowledge (contributes from biology), has been developed throughout the human species evolution and develops with the individual experience.

In order to conceptualize the stimuli and behaviours in terms of inputs and outputs, the available data was reorganized with the aim of gathering the following:

- A set of input stimuli
- A set of rules (hypotheses) that link the input with the output
- Output correctly paired with each input

The built table can be seen flowingly, presented in Table 9 to Table 12.

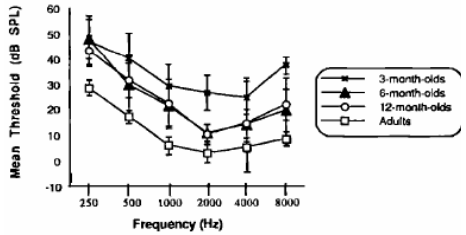
<table>
<tr><th>Paper (Behavioural experiment)</th><th>Topic</th><th>Age</th><th>Stimuli - INPUT</th><th>Behaviour OUTPUT</th><th>Hypothesis</th></tr>
<tr><td rowspan="3">Foetal Response to Music and Voice.
[1]</td><td rowspan="3">Prenatal music experience</td><td rowspan="3">foetus (37-40 weeks)</td><td>Music – Spanish guitar (15s) at 105 dB.</td><td>Heart rate acceleration</td><td rowspan="3">Prenatal exposure to music and voice alters the foetal behaviour</td></tr>
<tr><td>Voice – Nursery rhymes recited by a female (15s) at 94 dB.</td><td>Heart rate acceleration</td></tr>
<tr><td>Sham</td><td>No significant heart rate acceleration</td></tr>
<tr><td>Sound frequency change detection in fetuses and newborns, a magnetoencephalographic study.
[14]</td><td>Prenatal music experience</td><td>Foetus between 33 and 36 gestational weeks</td><td>Sequence of two complex sounds at 110 dB with varied inter-stimulus interval between 500ms and 1100ms.
Standard sound (prob. of 88%) a 500Hz tone (100ms duration) with additional harmonics at 1000Hz and 1500Hz.
“Deviant” sound (prob. of 12%) a 750Hz tone with harmonics at 1500Hz and 2250Hz. Amplitude attenuation of harmonics by 3 and 6dB, respectively.</td><td>Response to detection of sound changes was found in 60% of the foetal data.</td><td>Auditory discriminative ability may start before birth.</td></tr>
<tr><td rowspan="4">Pure-Tone sensitivity of human infants.
[31]</td><td rowspan="4">Pitch Perception</td><td>3-months old</td><td colspan="2" rowspan="4">
</td><td rowspan="4">Pure-tone thresholds improve between 3 months and adulthood. Between 3 and 6 months, the improvement is greater at high frequencies. Improvements at low frequencies do not occur until after 12 months.</td></tr>
<tr><td>6-months old</td></tr>
<tr><td>12-months old</td></tr>
<tr><td>Adults 18-26 years old</td></tr>
<tr><td rowspan="3">Infants’ Perception of Pitch: Number of Harmonics
[6]</td><td rowspan="3">Pitch Perception</td><td rowspan="3">7-month-old infants</td><td>Complex tone of 160 Hz fundamental containing either 2, 3 or 5 consecutive harmonics versus complex tone of 200 Hz fundamental containing either 2, 3 or 5 consecutive harmonics, respectively.</td><td>Discrimination</td><td rowspan="3">Pitch perception task deteriorates as the number of harmonics in a tonal complex decreases. Infants may need more spectral information to perceive pitch than adults.</td></tr>
<tr><td>Random sequence of a pitch represented by sets of three harmonic complexes as a background. The frequencies of the harmonics in the sounds shifted up and down in a random fashion. Infants had to respond on the basis of the fundamental frequency, while the harmonic content of the sounds constantly varied.</td><td>- Categorization only complexes containing 3 or 5 harmonics.
- Fail the categorization of 2 harmonic complexes.</td></tr>
<tr><td>Missing fundamental containing 2, 3 or 5 consecutive harmonics of the fundamental frequency.
Infants were required to turn when the pitch of the missing-fundamental changed as the harmonic content of the sounds constantly shifted.</td><td>- Categorization only complexes containing 3 or 5 harmonics.
- Fail the categorization of 2 harmonic complexes.</td></tr>
<tr><td rowspan="3">Discrimination of frequency transitions by human infants
[2]</td><td rowspan="3">Pitch Perception</td><td rowspan="3">6 to 9-month-old infants</td><td>300 ms frequency sweeps and a 1kHz un-modulated standard</td><td>Difference Limens of approximately4.3% (0.97% for adults)</td><td rowspan="3">Rapid frequency transitions are much more difficult to discriminate from frequency transitions of the same category. Extraction of spectral information from a frequency sweep appears to be temporally limited and not determined by a rate invariant mechanism.</td></tr>
<tr><td>50 ms linear frequency sweeps appended to the beginning or the end of a 250 ms segment of 1kHz un-modulated standard</td><td>Difference Limens of approximately 7.3% (1.98% for adults)</td></tr>
<tr><td>50 ms frequency sweeps and a standard consisting of a 50 ms frequency sweep of 350Hz.</td><td>Difference Limens of approximately 80% (28 % for adults)</td></tr>
</table>

Table 9 Representation of the gathered information (Part 1).

Paper (Behavioural experiment)	Topic	Age	Stimuli - INPUT	Behaviour OUTPUT	Hypothesis
Preference for Sensory Consonance in 2- and 4-Month-Old Infants. [52]	Intervals perception	infants between 8 and 10 weeks and between 15 and 17 weeks old	Set of consonant intervals perfect fifths [$A_3 – E_4$ and $C_4 – G_4$] octaves [$C_4 – C_5$ and $E_4 – E_5$]	Preference approximately 72s looking time	2-month-old infants are sensitive to the dimension of consonance and dissonance. Consonance perception may provide a bootstrap into the task of learning the pitch structure of the musical system to which the infant is exposed.
			Set of dissonant intervals Tritones [$Bb_3 – E_4$ and $E_4 - F_5$] Minor ninths [$Bb_3 – B_4$ and $E_4 – F_5$]	Approximately 58s looking time	
Natural musical intervals: Evidence from infant listeners. [46]	Intervals perception	6-month-old infants	Standard pattern presented repeatedly with consecutive presentations transposed upward or downward by 2 semitones. Melodic (sequential) intervals [300 Hz – 450 Hz] – 3:2 [300 Hz - 422 Hz] – 45:32 [300 Hz - 400 Hz] – 4:3	Mean discrimination: 3:2 – **0.3** 4:3 – **0.35** 45:32 – **0.03**	Intervals with simple frequency ratios may be natural prototypes, being relatively easy to encode, retain and distinguish from other intervals, compared with complex ratio intervals.
		9-month-old infants	Repetitions of the standard Harmonic (simultaneous) interval followed by a change trial in which the top tone of the interval is displaced by ¼ semitone for 3 repetitions, followed by a return to the standard interval. [$C_4 - G_4$] – 3:2 [$C_4 - F_4\#$] – 45:32 [$C_4 - F_4$] – 4:3	Mean discrimination: 3:2 – **0.5** 4:3 – **0.35** 45:32 – **0.1**	
Infants' perception of melodies: the role of melodic contour (experiment 1) [55]	Melodic contour - perception of transpositions and other transformations	8-11 month old	six tone sequence (C4, A4, G4, B4, G4, C4) tone:200ms; pause:200ms; time between melodies: 800ms	---------------------	Infants treat new melodies or tone sequences as familiar if these sequences have the same melodic contour and frequency range as a previously heard sequence and as novel if either the contour or range differs
			sequence transposed 3 semitones (D#4, C5, A#4, D5, A#4, D#4)	Discriminate - novel melody	
			contour preserving +-+-- (C4, B4, F4, A4, E4, C4)	Discriminate - novel melody	
			Octave change/contour preserving +-+-- (C4, A5, G4, B6, G5, C4)	Discriminate - novel melody	
			octave change/contour violation -+--+ (C4, A3, G5, B4, G3, C4)	Discriminate - novel melody	
Infants' perception of melodies: the role of melodic contour (experiment 2) [55]			six tone sequence (C4, A4, G4, B4, G4, C4) tone:200ms; pause:200ms; tbm: 800ms +3X 262Hz tone (200msdur and inter-tone) distractor melody	--------------------	
			sequence transposed 3 semitones (D#4, C5, A#4, D5, A#4, D#4)	Don't Discriminate - Same melody	
			contour preserving +-+-- (C4, B4, F4, A4, E4, C4)	Don't Discriminate - Same melody	
			Octave change/contour preserving +-+-- (C4, A5, G4, B6, G5, C4)	Discriminate - novel melody	
			octave change/contour violation -+--+ (C4, A3, G5, B4, G3, C4)	Discriminate - novel melody	

Table 10 Representation of the gathered information (Part 2).

Paper (Behavioural experiment)	Topic	Age	Stimuli - INPUT	Behaviour OUTPUT	Hypothesis
Infant responses to prototypical melodic contours in parental speech. [33]	Infant Directed Speech - Prosody	4-month-old	Female Approving Contours (Peak frequency: 831 Hz; frequency range: 21 semitones; Duration: 1287 ms)	Looking Time **77seconds**	Prototypical approving and disapproving contours from Infant directed speech affect preferential visual behaviour. The high pitch and expanded melodic contours elicit and maintain infant's attention. Melodic prototypes may function as didactic caregiving messages for infants.
			Female Disapproving Contours (Peak frequency: 848 Hz; frequency range: 21 semitones; Duration: 331 ms) and reverse (approving and disapproving) Contours	Looking Time **50 seconds**	
			Male Approving Contours (Peak frequency: 344 Hz; frequency range: 17 semitones; Duration: 1267 ms)	Looking Time **63 seconds**	
			Male Disapproving Contours (Peak frequency: 345 Hz; frequency range: 15 semitones; Duration: 235 ms) and reverse (approving and disapproving) Contours	Looking Time **50 seconds**	
Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages [17]	ID Speech cross cultural languages	5-month old (English)	Approval vocalizations in English, German, Italian, Non sense English	Show positive facial affect / mean looking time – **6.45s**	Infants respond with more positive affect to approvals and more negative affect to prohibitions across the presented languages. Infants showed evidence of differential responsiveness to affective vocal expressions in German, Italian, non sense English and English but not in Japanese.
			Prohibition vocalizations in English, German, Italian, Non sense English	Show negative facial affect / mean looking time – **5.38s**	

Table 11 Representation of the gathered information (Part 3).

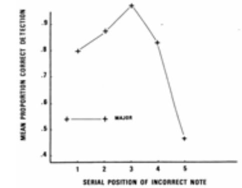
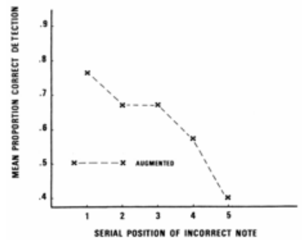
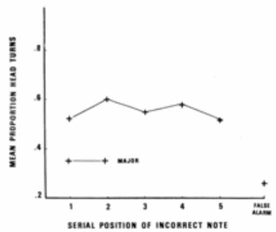
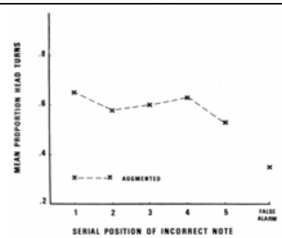
Paper (Behavioural experiment)	Topic	Age	Stimuli - INPUT	Behaviour OUTPUT	Hypothesis
Development of the perception of musical relations: semitone and diatonic structure. [56]	Perception of musical scales.	9 to 11 month-old infants	Sequence major triad [C4 E4 G4 E4 C4] with semitone changes in any position.		Diatonic structure may only emerge by 4 to 6 years of age.
			Sequence augmented triad [C4 E4 G#4 E4 C4] with semitone changes in any position.		
		4 to 5 years old children	Sequence major triad [C4 E4 G4 E4 C4] with semitone changes in any position.		
			Sequence augmented triad [C4 E4 G#4 E4 C4] with semitone changes in any position.		

Table 12 Representation of the gathered information (Part 4).

The organization of data presented in the previous tables allows a systematized analysis and connected perspective over the experimental studies.

Music cognition involves genetic and experiential factors and, as a developmental phenomenon, there is an age window during which experience can have more effect than at other times. Thus, the development of musical perception and cognition should be expected to be different, depending on the individual experience. However, it is possible, by looking at data resulting from experimental studies indicating early musical experience, to establish a developmental timetable of auditory processes and identify periods where the perception and cognition of determent features can be identified.

The Figure 7 illustrates a developmental scheme in which the inputs and outputs were arranged in terms of age (growth).

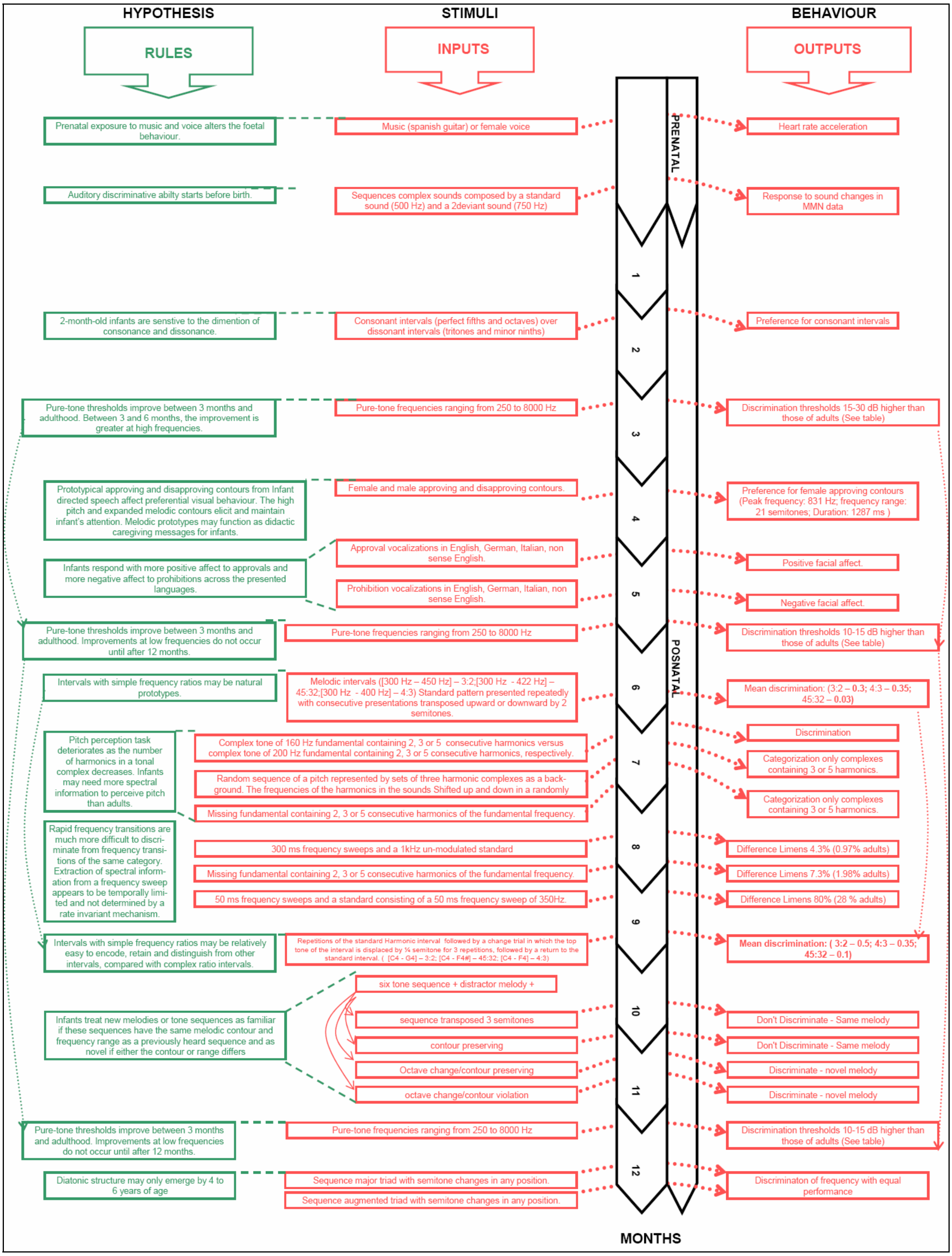


Figure 7 Evolution of the musical features' perception during childhood development.

The disposition of the data in function of age allows observing the evolution of behaviours given the stimuli. This organization can also give clues on the kind of mechanisms involved in the appearance of music cognition.

4 Discussion and conclusions

As it was mentioned before, the aim of this research work is to review of the existent literature intended to contextualize the appearance of musical capacity, particularly, the early melodic representations and also an organization of the information collected looking forward the conceptualization of its computational modelling.
As in [41], "a model is qualified by its predictive and generalization power, its simplicity and its relation to existing theories." A model should generalize, being able to explain data (e.g. results of experiments) and also build relations between existing theories. This should help to understand a phenomenon "instead of just reproducing input-output relation inherent in the data" [41]. Having this into account, a survey of the literature has been done about the music cognition capacity, the modelling of music cognition, and experimental studies done on the music cognition in the first years of life. In the survey we framed the study of music cognition mainly into two perspectives:

- The computational modelling where facts and models about music cognition are presented. The view of multiple disciplines over the music cognition modelling is exposed. The study of music cognition modelling imprinted in the elapse of our research (in which concerns the search for experimental studies and data on them) the carefulness of a perspective in terms of inputs and outputs.

- The experimental studies allowed collecting data on the development of music cognition in the first year of life. This collection gave contributes on indicating signs of the appearance of music cognition having in view its modelling.

The collected data was analysed in a developmental perspective, reflected in the timetable constructed. In this timetable is possible to observe the evolution of the representations in music cognition through the signs collected on the results of the experimental studies.
This information collected may contribute to reduce lack in gathering data from experimental studies on the early music cognition with the aim of its modelling.
In a global view of the research work done, we think that a reconstruction of the review of the literature through the articulation of their several perspectives, looking towards

the main goal of the computational modelling. Having underlined the fact that a model should generalize, and build relations between existing theories, its conceptualization has to be done under a solid theoretical framing. From this articulation, a more rich perspective over the computational modelling of the appearance musical cognition should emerge. This would lead to a more explicative capacity of the phenomenon.

The gathering of experimental studies aiming to analyse musical features (pitch, contour or musical scales) separately was important in order to identify critical periods where specific features are perceived and how their perception develops. But, in fact, these musical features appear simultaneously in music and also this overlap distorts the processing of a particular feature. Therefore, a next phase of work, relying on solid previous findings, would have to have in account the complexity of the music cognition phenomenon. Because of this complexity, its study would have to be also a complex analysis, correlating emotion, learning, development, aesthetic, and genetics.

This approach would focus on the analysis of the mechanisms underlying the processing of the information gathered in this research.

5 References

[1] Al-Qahtani, N. H., 2005. Foetal response to music and voice. *Australian and New Zeeland Journal of Obstetrics and Gynaecology*, Vol. 45: 414-417.

[2] Aslin, R. N., 1989. Discrimination of Frequency Transitions by Human Infants. *Journal of the Acoustical Society of America*, Vol. 86 (2): 582-590.

[3] Bergeson, T. R.; Trehub, S. E., 1999. Mothers' Singing to Infants and Preschool Children. *Infant Behaviour & Development*, Vol. 22(1): 51-64.

[4] Blood, A. J.; Zatorre, R. J., 2001. Intensely Pleasurable Responses to Music Correlate With Activity in Brain Regions Implicated in Reward and Emotion. PNAS, Vol. 98(20): 11818-11823.

[5] Boening, J.A.L, 2001. Neurobiology of an Addiction Memory. *Journal of Neural Transmission*. Vol 108: 755-765.

[6] Clarkson, M. G.; Martin, R. L.; Miciek, S.G., 1996. Infants' Perception of Pitch: Number of Harmonics. *Infant Behaviour and Development*, Vol. 19: 191-197.

[7] Clarkson, M. G.; Clifton, R. K., 1984. Infant pitch perception: Evidence for responding to pitch categories and the missing fundamental. *Journal of the Acoustical Society of America*, Vol. 77 (4): 1521-1528.

[8] Cambouropoulos, E., 2001. Melodic Cue Abstraction, Similarity, and Category Formation: A Formal Model. *Music Perception*, Vol. 18 (3): 347-370.

[9] Cambouropoulos, E., 2006 Musical Parallelism and Melodic Segmentation: A Computational Approach. *Music Perception*, Vol. 23 (3): 249-268.

[10] Chen-Hafteck, L., 1997. Music and Language Development in Early Childhood: Integrating Past Research in the Two Domains. *Early Child Development and Care*, Vol. 130: 85-97.

[11] Cross, I., 1999. AI and Music Perception. *AISB Quarterly*, Vol. 102: 12-25.

[12] Deliège, I., 2001. Prototype effect in music listening: An empirical approach to the notion of imprint. *Music Perception*, 18: 371-405.

[13] Desain, P.; Honing, H.; Thienen, H. van; Windsor, W. L., 1998. Computational Modeling of Music Cognition: Problem or Solution? *Music Perception*, Vol. 16 (1): 151-166.

[14] Draganova, R.; Eswaran, H.; Murphy, P.; Huotilainen, M.; Lowery, C.; Preissl, H., 2005. Sound frequency change detection in fetuses and newborns, a magnetoencephalographic study. *Neuroimage*, Vol. 28(2):354-361.

[15] Ferland, M.B.; Mendelson, M. J., 1989. Infants' Categorization of Melodic Contour. *Infant Behaviour and Development*, Vol. 12: 341-355.

[16] Fernald, A. 1989. Intonation and Communicative Intent in Mothers' Speech to Infants: Is the Melody the Message? *Child Development*, Vol. 60, No6, 1497-1510.

[17] Fernald, A. 1993. Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development*, Vol. 64, No3, 657-674.

[18] Gopnik, A. et al, 2004. A theory of causal learning in children: Causal maps and bayes nets. *Psychological Review*, Vol. 111 (1): 3-32.

[19] Gutkin, B. S.; Dehaene, S.; Changeux, 2006. A neurocomputational hypothesis for nicotine addiction. PNAS, Vol. 103(4): 1106-1111.

[20] Honing, H., 2006. Computational Modeling of Musical Cognition: A Case Study on Model Selection. Music Perception, Vol. 23 (5): 365-376.

[21] Jackendoff R.; Lerdahl, F, 2006. The capacity for music: what is it, and what's special about it? *Cognition*, 100: 33-72.

[22] Janata, P., 2005. Brain Networks That Track Musical Structure. *Annals of the New York Academy of Sciences*, Vol. 1060: 111-124.

[23] Kisilevsky, B. S.; Hains, S. M.; Lee, K.; Xie, X.; Huang H.; Ye H. H.; Zhang K.; Wang Z., 2003. Effects of experience on fetal voice recognition. *Psychological Science*, Vol. 14(3): 220-224.

[24] Levitin, D. J., 1999. Experimental Design in Psychoacoustic Research, From the book "Music, Cognition and Computerized Sound: An Introduction to Psychoacoustics." Editor: Cook, P. R. M.I.T. Press.

[25] Loui, P.; Wessel, D.; Kam, C., H., 2006. Acquiring New Musical Grammars: a Statistical Learning Approach. Proceedings of the 9th International Conference for Music Perception and Cognition.

[26] Marr, D., 1982. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. Henry Holt and Co., Inc. New York, USA.

[27] Mandel, D.; Myers, J.; Turk, A.; Gerken, L., 1995. The Head-Turn Procedure for Testing Auditory Perception. *Infant Behaviour and Development*, 18: 111-116.

[28] McDermott, J.; Hauser, M., 2005. The Origins of Music: Innateness, Uniqueness, and Evolution. Music Perception, Vol. 23(1):29-59.

[29] McMullen, E.; Saffran, J.R., 2004. Music and Language: A Developmental Comparison. *Music Perception*, Vol. 21(3): 289-311.

[30] Olsho, L. W.; Schoon, C.; Sakai, R.; Turpin, R.; Sperduto, V., 1982. Preliminary data on frequency discrimination in infancy. *Acoustical Society of America*, Vol. 71: 509-511.

[31] Olsho, L. W.; Koch, E. G.; Carter, E. A.; Halpin, C. F.; Spetner, N. B., 1988. *The Journal of the Acoustical Society of America*, Vol. 84(4): 1316-1324.

[32] Plantinga, J.; Trainor, L. J.; 2005. Memory for melody: Infants use a relative pitch code. *Cognition*, Vol. 98: 1-11.

[33] Papousek, M.; Bornstein M. H.; Nuzzo, C.; Papousek, H.; Symmes, D., 1990. Infant responses to prototypical melodic contours in parental speech. *Infant Behaviour and Development*, Vol. 13:539-545.

[34] Papousek M.; Papousek, H., 1991. The meaning of melodies in motherese in tone and stress languages. *Infant Behaviour and Development*, Vol. 14: 415-440.

[35] Papousek M.; Papousek, H.; Haekel, 1987. Didactic Adjustments in Fathers' and Mothers' Speech to Their 3-Month Old Infants. *Journal of Psycholinguistic Research*, Vol. 16(5): 492-516.

[36] Pegg, J. E.; Werker, J. F., 1992. Preference for Infant-Directed Over Adult-Directed Speech: Evidence from 7-Week-Old Infants. Infant *Behaviour and Development*, Vol. 15: 325-345.

[37] Peretz, I., 2006. The nature of music from a biological perspective. *Cognition*, Vol. 100, 1-32.

[38] Peretz, I.; Coltheart, M., 2003. Modularity of Music Processing. *Nature Neuroscience*, Vol. 6(7): 688-691.

[39] Peretz, I.; Zatorre, R. J., 2005. Brain Organization for Music Processing. *Annual Review Psychology*, Vol. 56: 89-114.

[40] Plomp, R.; Levelt, W.J. M., 1965. Tonal Consonance and Critical Bandwidth. *Journal of the Acoustical Society of America*, Vol. 38(4): 518-560.

[41] Purwins, H.; Hazan, H.; Marxer R.; Herrera P.; Grachten M.; Solis, H., 2006. Review of Music Cognition Models.

[42] Purwins, H., Marxer R., Herrera P., 2007. The Music Projector, a Mock-up Expectation and Emergence of categorization of Pitch Sequences. Music Technology Group, Universitat Pompeu Fabra.

[43] Rodgers, J. L.; Rowe, D. C., 2002. Theory Development Should Begin (But Not End) With Good Empirical Fits: A Comment On Roberts and Pashler (2002). *Psychological Review*, Vol. 109 (3): 599-604.

[44] Rock, A. M. L.; Trainor, L. J.; Addison, T. L., 1999. Distinctive Messages in Infant-Directed Lullabies and Play Songs. *Developmental Psychology*, Vol. 35(2): 527-534.

[45] Saffran, J., R. et al, 1999. Statistical learning of tone sequences by human infants and adults. *Cognition*, Vol. 70: 27-52.

[46] Schellenberg, E. G.; Trehub, S. E., 1996. Natural musical intervals: Evidence from infant listeners. *Psychological Science*, Vol. 7 (5): 272-277.

[47] Schellenberg, E. G.; Trehub, S. E., 1996. Children's discrimination of melodic intervals. *Developmental Psychology*, Vol. 32 (6):1039-1050.

[48] Sternberg, R. J., 1998. The nature of cognition. *Cambridge*, *The MIT Press*.

[49] Thompson, N. C., Cranford, J. L, Hoyer, E., 1999. Brief-Tone Frequency Discrimination by Children. Journal of Speech, *Language and Hearing Research*, 42: 1061-1068.

[50] Trainor, L. J.; Clark, E. D.; Huntley, A.; Adams, B. A., 1997. The Acoustic Basis of Preferences for Infant-Directed Singing, *Infant Behaviour and Development*, Vol. 20(3): 383-396.

[51] Trainor, L. J.; McDonald, K. L.; Alain, C., 2002. Automatic and Controlled Processing of Melodic Contour and Interval Information Measured by Electrical Brain Activity. *Journal of Cognitive Neuroscience*, Vol. 14 (3): 430-442.

[52] Trainor, L. J.; Tsang, C. D.; Cheung, V. W., 2002. Preference for sensory consonance in 2- and 4-month-old infants. *Music Perception*, Vol. 20 (2): 187-194.

[53] Trainor, L. J.; Zacarias, C. A., 1998. Infants' prefer higher-pitched singing. *Infant Behaviour and Development*, 21: 799-806.

[54] Trehub, S. E., 2001. Musical Predisposition in Infany. *Annals of the New York Academy of Sciences*, Vol. 930: 1-16.

[55] Trehub S. E.; Bull D.; Thorpe L. A., 1984. Infants' Perception of Melodies: The Role of Melodic Contour. *Child Development*, Vol. 55 (3): 821-830.

[56] Trehub S. E.; Cohen, A. J.; Thorpe, L. A.; Morrongiello, B.A., 1986. Development of the perception of musical relations: semitone and diatonic structure. *Journal of Experimental Psychology: Human Perception Performance*. Vol. 12(3): 295-301.

[57] Trehub, S. E.; Schellenberg, E. G.; Kamenetsky, S. B., 1999. Infants' and adults' perception of scale structure. *Journal of experimental psychology: Human perception and performance*. Vol. 25(4): 965-75.

[58] Trehub, S. E.; Schneider, B. A.; Henderson, J. L., 1995. Gap detection in infants, children, and adults. *The Journal of the Acoustical Society of America*, Vol. 98(5): 2532-2541.

[59] Trehub, S. E. ; Thorpe, L. A. ; Morrongiello, B. A., 1985. Infants' perception of melodies: changes in a single tone. *Infant behavior & development*, Vol. 8 (2): 213-223.

[60] Trehub S. E.; Unyk A. M.; Trainor L. J., 1993. Maternal singing in cross-cultural perspective. *Infant behavior & development*, Vol. 16 (3): 285-295.

[61] Warren, J. D.; Uppenkamp, S.; Patterson, R. D.; Griffiths, T. D., 2003. Separating pitch chroma and pitch height in the human brain. PNAS, Vol. 100(17): 10038-10042.

[62] Werker, J. F.; Polka, L.; Pegg, J., 1997. The Conditioned Head Turn Procedure as a Method for Testing Infant Speech Perception. *Early Development and parenting*, Vol. 6(3): 171-178.

[63] Werner, L. A.; VandenBos, G. R., 1993. Developmental psychoacoustics: What infants and children hear. *Hospital and Community Psychiatry*, 44: 624-626

[64] Zatorre, R., Peretz, I., "The Biological Foundations of Music", *Annals of the New York Academy Sciences*, vol. 930, 2001.

[65] Zetner, M.R.; Kagan, J., 1998. Infants' Perception of Consonance and Dissonance in Music. *Infant Behaviour and Development*, Vol. 21(3): 483-492.