Jodivi:
An Application to target sound sensitivity features in People with Autism Spectrum Disorder

Joel Sandé

A thesis submitted in partial fulfillment of the requirements for the

MASc degree in
Biomedical Engineering

uOttawa

University of Ottawa
Ottawa, Ontario, Canada
June 2017

Acknowledgements

This research thesis was done under the supervision of Prof. Dimitrios Makrakis from the Faculty of Engineering, University of Ottawa. Prof. Virginie Cobigo has been an advisor from the Faculty of Social Sciences, University of Ottawa. I would like to express my gratitude to Prof. Dimitrios Makrakis and Prof. Virginie Cobigo for their guidance and encouragements throughout my studies at the University of Ottawa. Without their knowledge, advice and mentoring the completion of this research thesis would not have been possible. I would also like to acknowledge Prof. Hilmi Dajani for his guidance, which provided me inspiration in this thesis. Prof. Virginie is the interface between this work and the medical personnel, which I expect will use it. She helped me to tailor my application to allow its use by families and by medical personnel.

A special mention to M. Upal Mahfuz, for his help, not only to find studies room, but also for some paper editing. Another mention to my colleagues Ronggang Chen and Adam Noel for their knowledge in signal processing, and help with computational issues, and use of Matlab.

I would like to express my gratitude to my parents, Véronique Adegbenlé and Oladélé Sandé, my sister Linda Sandé and her husband Michael Gandonou Migan, and my brother William Joseph Sandé, for their patience and unconditional support throughout my studies. Without their support, the completion of this research thesis would have been impossible.

I finally want to thank all the people who have helped me along the way.

Abstract

The main objective of this research work is to provide a tool to prevent the severe hearing sensitivity to patients with Autism Spectrum Disorder (ASD) experience. The key element in our work is to identify commonalities between sounds that bother an ASD patient, and implement a procedure using PC or smartphone as platform - based on those results, which will lead to the prevention of "bothersome" sounds for the ASD sufferer and later to desensitization. To do so, we implemented a first application that evaluates the auditory sound sensitivity of a person, and a second application that determines those factors that are related to hearing sensitivity of the patient, suggest sounds in the preventive process, and proposes use of appropriate sounds in the desensitization process.

While the current implementation is a prototype, we are determined to pursue the development at professional level and implemented as very user friendly application, which we hope will become a popular tool used by medical personnel and ASD patients for the identification of an individual's specific sound sensitivities and his/her desensitization to those sounds.

Table of Contents

List of Figures

List of Tables

Terminology

Autism/ASD
A profound neuropsychiatric conditions characterized by impairments in social skills and communication, language incomprehension, as well as repetitive interests and activities.

Critical Sound
Wave File which sounds is difficult to bear for certain autists.

Synthetic sounds
Sounds made on computer for which we fix the exact value of frequencies and intensities.

Cross-correlation
Measures similarity between two vectors/signals.

Auto-correlation
It is the cross-correlation of a vector/signal with itself.

Fast Fourier Transform
Algorithm capable of generating the frequency response (spectrum) of a signal.

Decibel
Measurement unit of measurement of the intensity of a signal

Hertz
Measurement unit of frequency.

List of Abbreviations

ASD : Autism Spectrum Disorder

CSW : Correlated Sine Wave

dB : Decibel

DFT : Discrete Fourier Transform

EEG : Electro Encephalo Gram

FFT: Fast Fourier Transform

fMRI : functional Magnetic Resonance Imaging

Hz : Hertz

OSPSD : One Sided Power Spectral Density

PDDs : Pervasive Developmental Disorders

PSD : Power Spectrum Density

SA : Signal Amplitude

SAC : Signal Amplitude Correlated

SW : Signal Wave

1. Introduction

1.1 Research motivation

According to Epidemiological research [1] [2] 1% of youth experience Autism Spectrum Disorder and related pervasive developmental disorders (PDDs). People having those disorders experience "severe neuropsychiatric conditions characterized by impairments in social skills and communication, language incomprehension, as well as repetitive interests and activities" [1] [3] [4]. Also, "signals of Autism Spectrum Disorder (ASD) on 1 and 2 years old children are: reduced social interest and effect, lack of warm, joyful emotional expression, lack of interacting, abnormal language development, lack of coordination, facial expression and gesture during interaction" [5] [6].

Most children suffering with autism are hypersensitive to sounds. Because of this, they tend to cover their ears when they hear certain sounds or certain intensity has been exceeded [7] [8].

90% of individuals with autism experience abnormality in sensory perception, and 15% to 100% suffer from auditory hypersensitivity, which is the most common sensory-perceptual abnormality [9]. "The prevalence of the Autistic Spectrum Disorder varies from 5 to up to 60 for each 10.000" [9] with some variances ranging from 15% to up to 40% in clinical condition, and from 16% to up to 100% using parent questionnaires. As concern the handling threshold, 63% of the autistic individuals did not support stimulations above 80dB. [9].

The most common intervention approach used to address this problem is a by using sound isolation [7]; they use earmuff (insulating earphone) to reduce the intensity of the sound [7]. Due to the gradual human adaptation to different conditions, the devices are likely to worsen sound sensitivity after a period of time, since the hearing threshold can fall and increase its sensibility [7]. Nevertheless, according to Morris, this approach is the fastest to deal with serious conditions that require quick intervention [7]. Another way to address this problem is to use Auditory Integration Training, which is intervention based on sound, by exposing sensitive ears to sweet and usually classical music [10]. The technique employs about 5 to 10 hour sessions during which the patients listen to "electronically modulated music" [11]. Doing that, they use this soft music to re-train the sensitive ear by going through several sessions of training [7].

1.2 Definition of the Research problem

Some people suffering from Autism Spectrum Disorder have auditory sensitivity. This sensitivity is different from one subject to another. This research aims to design a preventive application that detects common points between several sounds bothering ASD subjects in order to find similar sounds in a preventive measure. For a known bothersome sound, parents want to know the evolution of the sensitivity of their child related to that sound; for that, we designed a component of the application that evaluates the sensitivity improvement within the time that we named « direct evaluation ». For parents, or specialists, or therapist that want to go further in more details dealing with sound parameters, dealing with numbers, we designed a component of the application that evaluates the sensitivity to given (varied) frequencies and intensities, that we named « indirect evaluation ». Auditory Integration Training is a component we integrated as complementary to the two evaluation processes, as desensitization to sounds procedure; the AIT will help to fix the sensitivity, while the evaluations will allow seeing the improvement.

In summary, this research aims to design a preventive application that detects common points between several sounds bothering ASD subjects (Section 3.1), to design a component that evaluates the sensitivity improvement to a known (natural) sound (Section 3.2.1), to design a component that evaluates the sensitivity to pure (synthetic) sound (Section 3.2.2), and integrate Auditory Integration Training component as complementary feature (Section 3.3).

1.3 Contribution to the Society

A usually very young person (e.g. toddler) person might be showing sensitivity to certain sounds (e.g. clapping, washing machine's & dryer's sound, vacuum cleaner's sound). This auditory hypersensitivity might be due to the fact that the individual might be suffering from ASD. It should be pointed out that not all ASD sufferers are sensitive to the same sounds [12]. The spectrum of variations to sensitivity is quite wide [8] [12].

While some sound sources producing discomfort might be easy to recognize (e.g. crowd clapping, working vacuum cleaner), others might not be as such (e.g. a "bongo" playing with other organs in a music event). Currently, to identify those noise sources requires lengthy and expensive visits to hospitals and use of expensive highly trained personal and equipment. Our vision is that we wish to reach the point where time consuming and costly process will be replaced by passing a Flash drive to the family containing the application. The medical personnel will contribute in a consultation capacity and instructions how to use it properly. After installing and running the application according to the instructions, the application will not only be able to identify the sounds that are bothersome to the person, but also extract the commonalities between those sounds, which are expected to be the actual sources of discomfort. It will then be able to provide list of various other physical sources of sound, which are expected to be bothersome to the individual, and derive a treatment method to be followed for desensitization. After the derived method is passed (possibly electronically) to and reviewed by the physician in charge, the desensitisation treatment can start for the desensitization of the suffering individual.

1.4 Challenge with the Research problem

As mentioned earlier, ASD patients have different level of sensitivity to sound. The nature of sound source varies, and is depending on the individual ASD patient. Thus generalization is not appropriate. We had to find solutions able to adjust to the needs of each individual. We should also mention that we couldn't access collected data directly from patients, a situation that increases the difficulty of the work; the analysis could not be done with true data and patients, due to the existing restrictive and time consuming procedure, which is concerned with patient confidentiality and the need to have thorough detailed and well executed and carefully reviewed testing protocols, that ensure they will not create any excessive inconvenience to the participating patients. Such tests are to be carried out at a later time, with the participation of the Children's Hospital of Eastern Ontario (CHEO) and the Psychology department of Ottawa U. We are confident the results will verify the usefulness of the developed application.

1.5 Explaining hearing, sound and spectral decomposition of acoustic signals.

Sound is generated by producing pressure changes in the air, which propagate as a wave through the atmospheric medium. As with all signals, a sound wave can be expressed through its spectral content, determined by using Fourier Transform [13]. Depending on the nature of sound, the spectral content of an acoustic signal can differ. Some sounds such as the engine of a vehicle or a dish-washing machine tend to have energy content concentrated at the lower part of the frequency spectrum (low frequencies content). A drill or an alarm tends to generate energy at higher frequencies. The ear of a normal human being can perceive sounds as low as 16 Hz and as high as 20 kHz [14]. Above 20 kHz the human being hears nothing, unlike some animals, such as whales, dogs, and bats, which hear and generate frequencies in the atmospheric medium above 2 kHz. The sound is for the whale a means of communication, while for the bat, a means of displacement, hunting. Sounds below 16 Hz are called ***infrasonic***, whereas those above 20 KHz ***ultrasonic***. Sounds with strong power content within the 10 kHz to 20 kHz spectral band are considered high-frequency. Human voice has content between 300 Hz and 3.4 kHz.

Signals with similar spectral content are expected to have similar acoustic properties. Thus, spectrum properties allow us to identify classes of acoustic signals which will have bothersome impact on the listener. It should be pointed out that in the real world, we tend to be receiving a multitude of acoustic signals at the same time (e.g. during a lecture, we might be receiving the noise of the classroom's air-conditioning device, road traffic related noise, etc.). Decomposition of a multi-sound composed audio signal to the sounds generated by individual sources is quite complex if attempted at the time domain, but easier in the frequency domain. Thus spectral analysis will allow to isolate specific sources and sound's spectral characteristics (frequency content, power density) that are responsible for producing discomfort to a specific ASD patient. The treatment then becomes easier; since the focus of desensitization will be placed on the limited "generic" sounds which have been discover to be responsible for the uncomfortable experience of the patient.

1.6 Thesis outline

The thesis outline is as follows. Chapter 2 presents a multi-field overview on the past 10 years of research on ASD, highlighting the potential origins of the disease, the comparison of attention and brain volume, an overview of the brain network, the genetics and biochemistry of the disease, the low level processing issues, some existing solutions, and current ways used to find similarity between sounds. Chapter 3 presents the methodologies to extract the common point between sounds, followed by the methodologies to evaluate sensitivity to sounds, followed by the methodology used for auditory integration training. Results of simulations are shown and discussed in Chapter 4. Finally, concluding remarks are presented in Chapter 5 followed by directions for future research presented in Chapter 6.

2. Literature Review

Different studies have been done relative to the ASD; from magnetic resonance imaging and brain network analysis, to a deep level in gene analysis. The aim of the present review is to give a multi-field overview of the research that has been done the past 10 years, including research related to the auditory sensitivity. In addition, research related to identifying similarities between sounds is presented.

2.1 ASD origins

ASD is considered a disorder that involves multiple genes [15] [1]. Recent research shows an ongoing or dynamic process characterized by abnormal brain growth and dependency on age [6] [1]. As pointed in [1] [16]. "both, gray and white matter abnormalities have been identified throughout the brain, reflecting the distributed nature of brain involvement in ASD" [1].

2.2 Brain Volume Comparison using MRI or Near-Infrared

The understanding of the neurobiology of ASD and its impact on the brain has increased drastically since the appearance of Magnetic Resonance Imaging (MRI) and Infrared Spectroscopy (IRS) [1]. Currently, we can measure responses to language sounds made from a young individual when s/he is sleeping by using MRI. This eliminates the need to immobilize him during the period the observation is run. IRS is used for observation and data collection when the person is awake. It has been found that an increase in the total brain volume and fast brain overgrowth occurs at the early stages of brain development of ASD subjects [5] [16] [1]. This finding is illustrated in Figure 1.

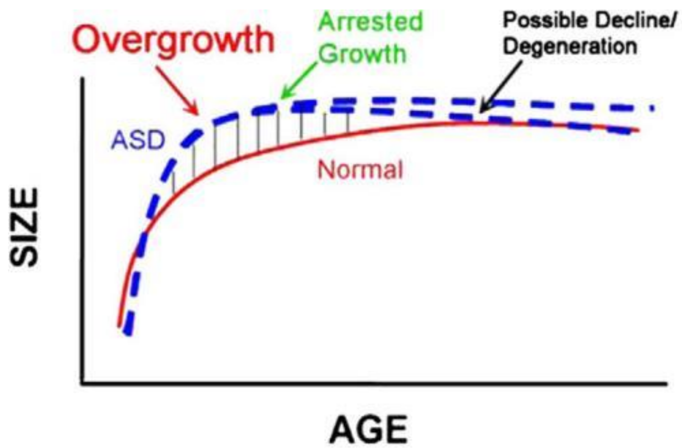


Figure 1: Early Brain Overgrowth in ASD

(*with permission of Eric Courchesne and Karen Pierce* [5]*).* ASD brain (blue curve) is overgrowing compared to a normally developing brain (red curve). This growth is followed by an arrest, while normal brain continues growing. Sometimes, "the arrest of growth is followed by degeneration, indicated by the blue dashed lines that slope slightly downward" [5].

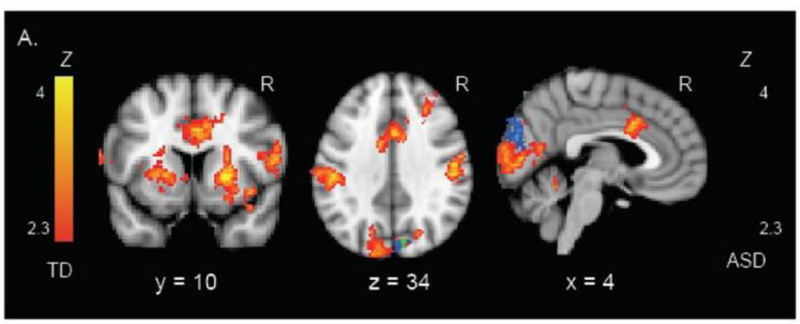
This abnormal brain enlargement may occur right at the birth, or later, after one or two years, and is later followed by degeneration during the next 10 years [5]. This degeneration might result in a "smaller overall brain volume for the autistic patient" [5]. It is also difficult to study the changes from childhood to adulthood because it represents a long period of time [5]. Researchers have been targeting different areas of the brain that expect to be abnormal in autism, e.g. those that support language development such as amygdala, frontal and temporal cortices:

"Amygdala plays an important role in emotional and social behavior" [17]. Kids under 10 years of age with autism have been found to have an amygdala bigger than typically-developing (TD) controls [18] [1] . Relative to that, previous studies found a direct relationship between decreased amygdala volume and "decreased time fixating on the eye region of faces" [19] [1]. This is opposite to another study discussed in this dissertation (Section 2.3), which mentions that typically-developing children pay more attention to face than ASD children.

"Integrity of frontal and temporal cortices is essential for normal language development" [17]. Frontal and temporal cortices "were found to be 13% larger in volume to children with autism spectrum disorder" [3]. In ASD children, left temporal cortex activity was reduced whereas the right side activity was increased. There is also growing difference with age growth [3] [20]. Typically-developing children have stronger activity at the inferior frontal gyrus during imitation of emotional expressions. "A between-groups comparison revealed a significant difference" [1].

2.3 Comparative attention

Attention is a good criterion to discriminate between ASD and TD subjects. Greene et al. revealed "greater brain activities for social cues in normal subject than for non-social cues" [21]; result that corroborates the hypothesis that "social cues are not assigned the same privileged status in the autistic brain as in the typically developing brain" [21].



<u>Figure 2</u>: Z statistic activation map of the 2 (Gase Vs Arrow) X 2 (Directional vs. Neural) interaction for each group.

Color bars indicate *Z* statistic; TD group shown in red-yellow, ASD group shown in blue-light blue, overlap shown in green. "Images were acquired using a Siemens Trio 3.0 Tesla MRI scanner. Two sets of high resolution anatomical images were acquired for registration purposes. Analyses were performed using FSL Version 4.1.4" (Image and caption from [21]).

Figure 2 has been included in the present manuscript to provide a 3D view of the brain's areas of the brain that have been found to be different from those of TD brains Greene et al. study [21]. Bird et al. [22] mention an example related to attention, and for a situation in which "a single stimulus must be selectively attended" [22] i.e. in the presence of single distractors. A typically-developing child would shift his attention and focus rapidly to that distractor. This reaction is executed below normal levels in ASD subject [22]. That indicates "problems with higher-order attentional control network" [22].

In an experiment with control patients to correlate the fusiform gyrus [1]to face-selective regions, Bird et al. [22] and Wang et al. [23] found a significant attentional modulation in the right inferior occipital gyrus during attending a face. This was however not the case for an unattended face. In another related test to show that parahyppocampal regions are place (non-social) selective regions, these researchers found significant modulation in bilateral parahyppocampal regions when house was attended [22] [23]. In ASD subjects, there was no significant difference in the brain activity when faces were attended, suggesting a lack of attention for face-selective regions. In contrast, bilateral parahyppocampal has an increased brain activity in response to house for ASD, [22] [23]. Analysis between normal subjects and ASD subjects revealed no significant difference in attention for non-social (house) selective regions, whereas ASD showed significantly less attention in face-selective areas (left fusiform gyrus) [22].

[1] Fusiform Gyrus is part of the temporal and occipital lobes. Its functionality is not fully understood yet, but it is known it is linked to neural pathways that are related to recognition. It is involved in several neurological deficiencies such as dyslexia, prosopagnosia, and synesthesia.

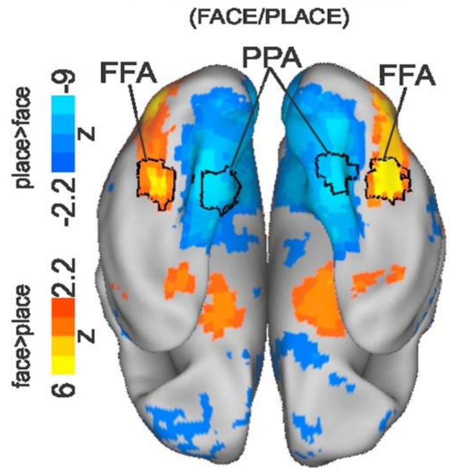


Figure 3: Parahyppocampal Place Area Vs Fusiform Face Areas

Brain map showing Parahyppocampal Place Areas (PPA) and the Fusiform Face Areas. Thoses are areas that control face and place recognition. The image is the summation of a large group of subjects (N=39) collected on the 3 T scanner and displayed on a ventral view of the surface. The black lines highlight the border of the PPA and FFA regions of interest (image from [24]).

Refereeing back to Figure 1 and the growth pattern it describes can be seen as a kind of brain's adaptation, where it seems to try a recovery to the normal stage (as the curve is going downwards).

2.4 Brain Network Overview

2.4.1 Neural Connection

When considering a somato-sensitive ASD subject, *Hughes* [25] showed that "massage of the body muscles for 5 minutes in 13 patients aged 3 – 6 years improved their sensory impairment" [25]. This indicates that the brain's network is flexible and adjustable. Interestingly, this is the same kind of adaptation for hearing circumstances that is known by the Auditory Integration

Training (section 3.3 and 4.4) to stimulate and readjust the auditory network in case of auditory impairment.

In his review paper [26] published in 2011, Wass states a very important fact: There is no relationship between the brain growth trajectories and the quality of the connection, however, the speed of the brain growth will lead to a "different optimal connectivity pattern" [26]. He also states that, according to researcher findings, a network that has grown faster tends to remove inter-hemispheric connections; a sign of removing longer-distance connections, to the benefit of shorter-distance connections [26]. "A larger brain tends to rely more on local than on long-distance connections" [26] because a local connection consumed relatively less bioenergy. In a larger brain, long-distance connections require more energy resources to build [26]. Barttfeld et al. [27] [28] found that there is some difference in the connectivity pattern between control and ASD subjects. One of those is the quantity of long range connections. There are less long range connections in ASD subjects, with a most pronounced deficit in front-occipital connections. Conversely, ASD subjects showed more short-range connections [28]. The problem with that is that different pattern of network connectivity will leads to different efficiency level in term of transmission of information [28]. They also found that, in addition, there are differences in functional connectivity, and differences at the attention level.

2.4.2 Attention

Indeed, there are two types of attention; external world attention, and self-body attention. ASD sufferers tend to have more self-body attention than external world attention, and when an external distraction comes that requires a fast shift of attention, ASD subjects perform much below typical performance [28]. Barttfeld et al. [27] ranked brain regions according to the region's classification power.

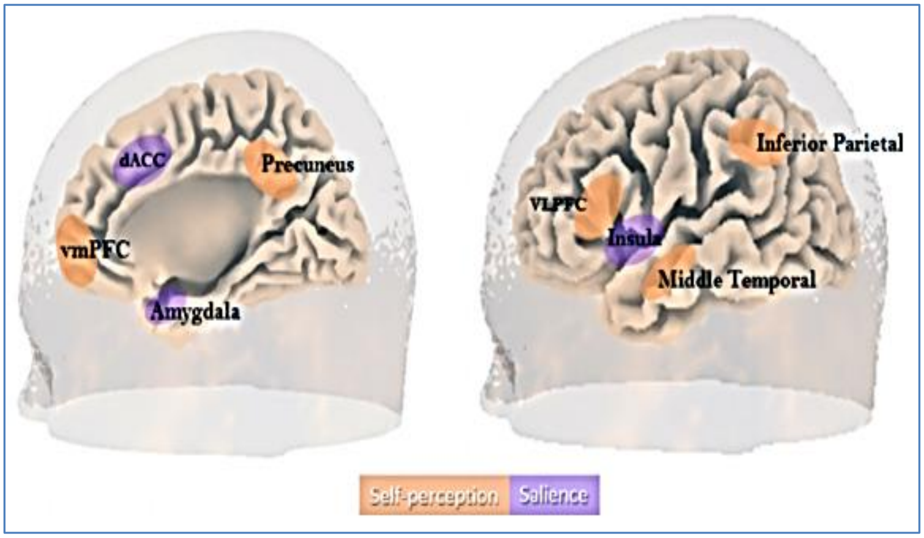


Figure 4: Anterior insula and Dorsal-anterior cingulate (dACC) in salience and self-perceptional network.

When the dorsal anterior cingulate cortex and anterior insula are active, an external auditory stimulus is consciously perceived; when the dACC and insula are deactivated, auditory information is pulled from memory. (Image and caption from **[29]**).

The regions that showed stronger differences of functionality between ASD subjects and typical subjects were the Anterior insula and Dorsal-anterior cingulate, which are identified in purple in Figure 4. Those differences were higher to subjects having higher ASD severity [27].

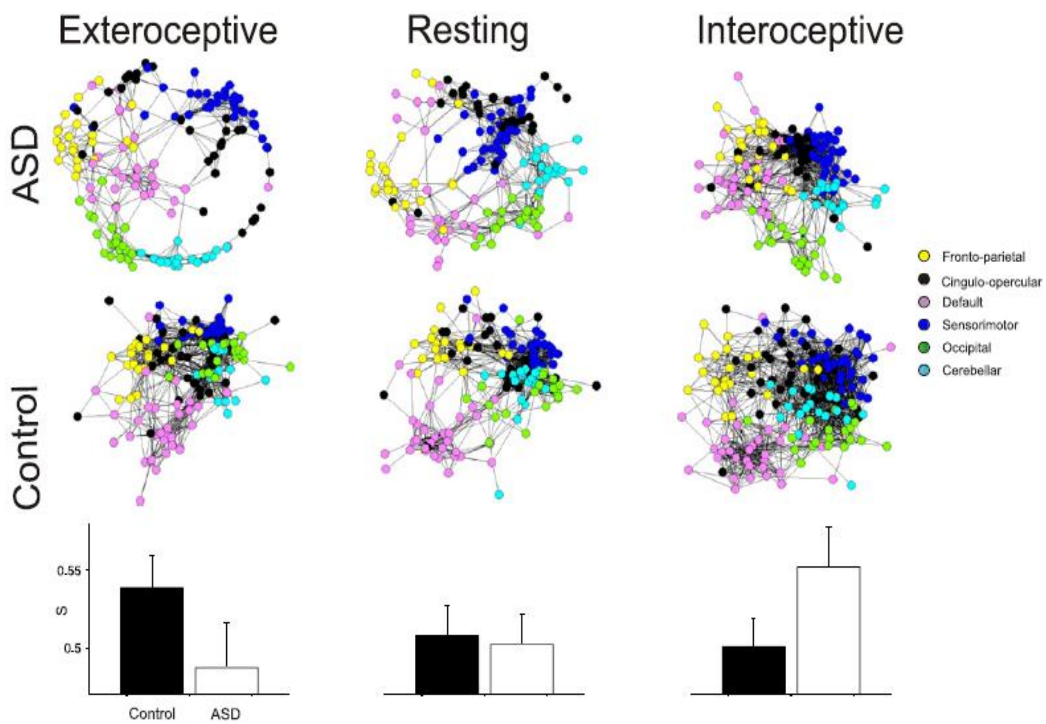


Figure 5: Network comparison at three (3) states (Under permission from **[27]**)

At the left, exteroceptive state, at the middle, resting state, and at the right, interoceptive state. There is a lack in global connectivity in ASD at exteroceptive state. The network is more compact in ASD for interoceptive state. For ASD, the network pattern is completely different from exteroceptive to interoceptive state. In TD brain, the connectivity change from a state to another is not so drastic. **[27]**

Networks obtained in the exteroceptive state condition (external world attention) show more topological difference across both groups. This is because a TD brain is more social than an ASD brain (Figure 5). Barttfeld et al. [28] added that the typical network is more packed, suggesting a shorter diameter of the entire brain network; a statement which corroborates the idea of early larger brain volume of ASD (Figure 1).

Then, the ASD network is more compact in the case of interoceptive state [27]. They paid particular attention to the black dots (cingulo-opercular). Their conclusion is he following: "black dots (cingulo-opercular) system is more tightly packed in ASD, and connected closely to the yellow (fronto-parietal) and default pink system" [27]. The networks were obtained using fMRI.

For ASD subjects, there is a lack in global connectivity in exteroceptive state. In ASD subjects, the network patterns corresponding to exteroceptive and interoceptive states are very different from each other. On the contrary, in the case of control subjects, such drastic difference between these patterns does not exist. They also mentioned the difference is more functional that direct comparison between the two networks connectivity [27].

2.4.3 Neural Synchronization

Dinstein et al. [30] mention that there is disorder of neural synchronization, and it is unknown when that synchronization abnormality appears. The human brain has two hemispheres linked to each other by the Corpus Collasum, which allows synchronisation between them. Dinstein et al. [30] showed that a weak interhemispheric synchronization occurs (weak functional connectivity across the two hemispheres) in the area that manages language [30]. They found that the strength of hemispheric synchronization was directly correlated with the communication skills level, and was decreasing with increase at the level of autism severity [30].

2.4.4 Overactive Microcircuits and Hyper-perception

Markram and Markram [31] proposed a unifying theory on autism, in which they claim that the pathology is due to overactive local neuronal microcircuits, which in turn is due to hyperplasticity. This is due to the fact that the microcircuits become autonomous and taken in an internal memory. This leads to fundamental cognitive consequences in terms of hyper-perception, hyper-focus, memory and hyper-hyper-emotionality. They centered their theory on neocortex and amygdala, but this potential can be applied to all regions of the brain.

This can lead, as we know very well, to the obsession with detailed treatment of integrated information, applying an unintentional and systemic decoupling of information to the point that it becomes very intense information to manage [31]. Thus "The autistic person locks himself in a limited internal world, but highly secure" [31]. Regional difference from one child to another could be due to their genetic traits, or postnatal experiences [31]. In their study published in 2010, Markram and Markram [31] established an explicit schematic relationship between the reactivity and plasticity, leading to different autistic states.

2.4.5 Inefficiency in Tasks performing Imitation

Noonan et al. [32] suggested the existence of inefficiency in the tasks performing optimization of the network connections. Imitation is a great precursor of socio-commutative development. Neuro-imagery studies have found reduced activation in areas associated with imitation [33]. They reviewed the functional and effective connectivity of these areas. Low oxygen levels suggest an atypical connectivity in the imitation network that can result in behavioral impairments in ASD subjects [33]. Finally the mirror neurons, which are active when a person is moving or observing a moving person, are considered dysfunctional in autism [25].

2.4.6 A principle of Autistic perception

Mottron et al. [34] proposed some principles of autistic perception. First, the perception of an autistic person is more locally oriented when compared to the perception of a non-autist person. Second, in the visual modality, the discrimination threshold of global motion is high in autism [34]. The discrimination threshold of global motion is also known under the name of perception level of movement per second [34], which is high in autism meaning that people with ASD have lower perception of movements per second. These results show a contrast to the evidence that autistic persons perform better in the case of "static object" discrimination. Indeed, superiority has been observed in low-level visual input discrimination occurring during the discrimination processes happening when observing random pattern [34]. Third, Mottron et al. [34] interpreted the high prevalence among ASD suffering individuals to limit excessive amount of information and / or focus on the optimal information for a given task. This was deduced from the long fixation of an autistic child on an object; behavior that is not observed at a normal child. This trait is known to occur in the first year of childhood [34]. Four, in ASD suffering individuals, the primary perceptive and the associative brain regions are abnormally activated for social and non-social tasks. ASD subjects show increased activation of visual-perceptual areas (occipital or occipital-temporal) in combination with reduced-activation in areas that are dedicated to "higher order" (front) or socially relevant (Fusiform Area Face - FFA) [34]. Five, the high-order processing is optional in ASD brain and mandatory in the non-ASD ones. Commenting on the conflicting results, Brosnan et al. [35] noted that people with autism are sensitive to visual illusions (when asked “what line looks more" and not to the question "which line is longer") in case of trap, an autistic person won’t fall in that trap, that can been perceived at a lower level processing [35] ; In the other hand, a typically developing brain tends to hide the trap and get into the illusion. This suggests that ASD persons have access to accurate physical representation or psychologically distorted representation depending on the way the question is asked [34]. Six, perceptual expertise in calculation, calendar, memory, 3D design, and detection of prime numbers, mental arithmetic, memory, music and improvisation are qualities that underlines Savant syndrome. Higher capacity field treatment was demonstrated in autistic musically naive [34] [34] [36]. Savant abilities may represent the autistic equivalent of "expertise" for non-autistic people [36]. The special ability works on a set of defined perceptual units. These units are present in organized patterns (books, calendars, mechanical objects, tonal melodies ...) that share a high degree of perceptual similarity across time and space [36]. At individual level, a logical sequence leading to Savant capacity includes encounting material at a critical period during which the device is selected on the basis of their exposure to the individual [36]. The development of the scholar ability can be understood in the context of brain behavior in which repetitive behavior in a specific area of operation "trains" an expertise processing system, but can hinder the development of other abilities [34] [36]. It is obvious that the capabilities of Savant always involve a pattern of behavior of a single restricted and repetitive interest for a certain class of stimuli such as height, word or letter. This leads to a "rule of judgment" with the majority of scientist having only one or two areas of unusually Savant abilities [36]. Spending a lot of time handling specific material can produce expertise in autistic persons, in multiple ways. Mottron et al. [34] [36] defend the Savant performance by claiming that it far exceeds the memory support and is a manifestation of autistic intelligence [34] [36]. The generalization of structured memory by the same rules as the date recovery by extending the calendar to the past or the future, mathematics inventiveness, is the final stage of the Savant ability [34] [34] [36].

2.5 Biochemistry and Genetics of Autism

One can look at autism at two levels: connectivity between the different cortical parts (network examination); genetic, and neuromodulator aspects (molecular level). In children with autism, abnormal dopamine activity in the prefrontal cortex was found, as well as increased levels of dopamine, urine and cerebrospinal fluid were found in the blood [37]. The biological results confirm the clinical diagnosis of observing mucopolyssacharides (MPS) in an autistic child, detected for the first time at the age of eight years, with disease duration for several years [38]. Note that MPS are caused by a deficiency of lysosomal enzymes involved in the degradation of mucopolyssacharides (glucose-amynoglicans) [38].

The abnormal relationship between the hippocampus and the frontal area disrupts the serotonergic innervation of the hippocampus, which leads to changes in behavior [25]. Abnormal behavior in autism has been viewed as related to the disrupted serotonergic innervation of the hippocampus with the cerebral cortex [25]. In persons with Autism central serotonergic hypoactivity was observed, which reduces the width of the mini-column in the cortex; a situation that can explain the focused attention in Savant abilities [25]. The increase of inhibitory synaptic transmission without change in the excitatory synaptic transmission, or a total lack of integration as a result of enlargement of the brain, or many structural brain abnormalities involving the cerebellum, the limbic system, frontal cortex and temporal, corpus callosum [25]. The most dominant factors in autism are the factors related to genes. They predominate on factors related to the environment. Factors related to the environment would affect mainly the phenotype, secondary factors [37] [39]. Examples of genes and chromosomes are chromosomes 17, 1, 9, 16, 2, 4, 6, 10, 15, 19, 21, [25], and genes RELN, NLGN4, SHANK2 and SHANK3 [39] [40] making Autism a polychromosomal disease [25] [41] [39] [42].

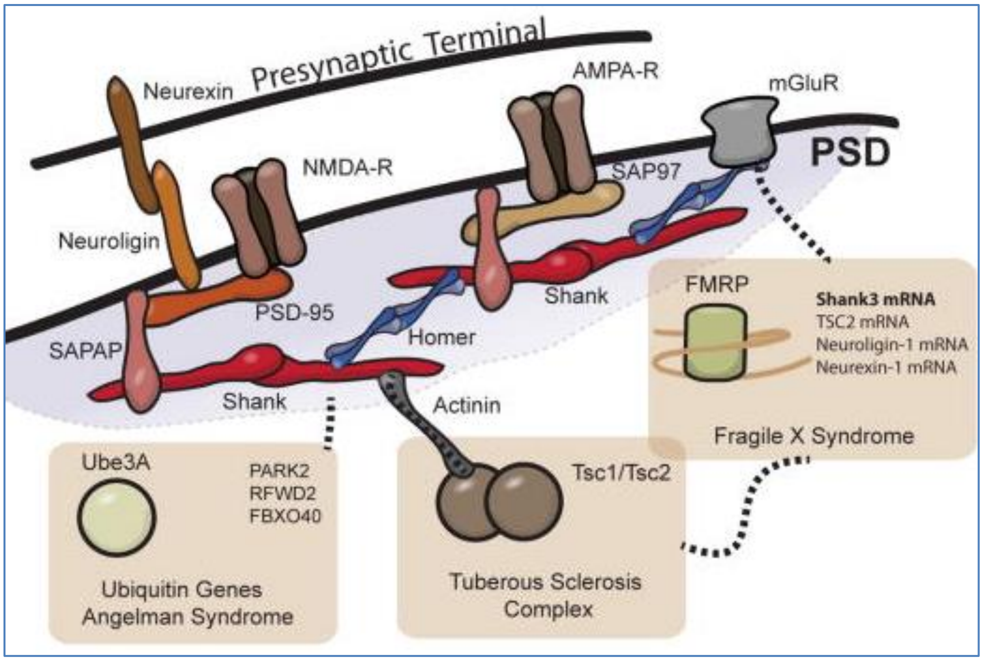


Figure 6: Shank proteins at the center of an ASD disease-module

A model for the overlap between synaptic proteins involved in susceptibility to syndromic and non-syndromic autism. Neurexin and Neuroligins are transsynaptic partners and candidate genes for susceptibility to autism; in the postsynaptic density these bind to the SAPAP family of proteins which have been linked to ASDs, PSD-95 and SAP97 which are involved in intellectual disability and autism. Shank dimers are thought to organize a molecular platform in concert with homer tetramers to stabilize the larger PSD, connecting AMPAR, NMDAR and mGluR into one protein. In the deeper synaptic compartment the control of PSD protein levels may be tightly controlled by independent complexes such as TSC1/2 through mTOR, or via FMRP regulation of synaptic transcripts, and most likely also through synaptic ubiquitin ligases (Image and capitation from [41]).

SHANK3 gene mutations in ASD have been widely studied both in human and mouse genetics, and is one of the most targeted genes of new drugs [40]. The SHANK3 gene maps to the 22q13.3 chromosomal region and encodes for a scaffolding protein in the postsynaptic density (PSD) of excitatory glutamatergic synapses (Figure 6). Mutations in the 22q13.3 lead to diseases like

Phelan–McDermid syndrome, an ASD characterized by hypotonia, cranial dysmorphic features, and language delay. Also, deletions and several smaller mutations such as microduplications, point mutations, and translocations in the SHANK3 gene are strong factors in ASD and intellectual disability, and are thought to be responsible for SHANK3 dysfunction. It has also been reported to significantly alterate metabotropic and ionotropic glutamate receptors at the synaptic levels, which leads to abnormalities in social interaction and affiliation behaviors [40].

As mentioned in the introduction, there is a prevalence of ASD in male than in female, reason why sex chromosome has been explored, and particularly neuroligin (NLGN) genes, the most widely studied. Five of the NLGN genes have been identified in the human genome, which encode a family of cell-adhesion molecules named neuroligins, "essential for the formation of functional neural synapses" [39] which is an important roles in synaptic transmission.

Reelin (RELN) is an extracellular matrix glycoprotein responsible for orienting the migration of several neural cell types and establishing neural connection [39]. It has been attributed an important role in the "positioning of neuronal cells in the inferior olivery complex, cerebral cortex and cerebellum early in embryonic development" [39].

2.6 Speech and Music

A common behavioral characteristic observed in ASD children is that they cover their ears when they hear certain displeasuring sounds. Indeed, if the auditory input is perceived as unpleasant, autistic children learn to avoid it [12] . According to the theories of increasing local cortical activities, people with ASD seem to over-recruit their left primary cortex; fact reveled by MRI studies [12]. This is due either to the inability to properly filter simultaneous processes, visual, auditory and tactile inputs [12].

Recent studies have found differences between individuals with and without autism in terms of how their brains react to stimuli. Early evoked responses recorded with EEG suggest differences in how visual and auditory stimuli are processed at low levels, knowing that "high level" forms of cognition are based on "low level" perceptual processes [43]. The argument is that the perception of the face is a relatively "high level" cognitive process, which includes several components of "low level" processing.

Successful function of the brain system that supports sophisticated language depends on the coordinated activity of the initial generalized networks [3]. It is likely that reported auditory sensory characteristics are related to the difficulties in understanding speech in the presence of background noise (e.g., speech in competition), which are also common symptoms to people with ASD.

To quantify these perceptual difficulties, Alcantana et al. [44] examined whether the poor performance of ASD subjects when listening to speech in the presence of noise is due to poor processing performed to the auditory temporal envelope of speech. They evaluated the speech recognition in the presence of noise for Asperger's ASD [44]. This scenario may be better explained when trying to communicate in a room where other conversations take place at the same time. Sometimes the overall noise may be lower or higher than the conversation of interest. They found that the target speech modulation rates threshold was significantly higher for the ASD group compared to the control group. This can explain why they have difficulty to handle noisy environment. It is easier for the control group to extract module speech in a noisy environment. The autistic brain in a given environment wants to process every coming sound it may hear. It is very good in distinguishing between sounds (the reason why autistics are naturally good in music). The problem is that in a noisy environment, there are too many sounds to process. There is too much to process simultaneously and considerable energy consumption occurs. It is easier for the control person to handle noisy environments because his/her brain automatically isolates and extracts useful sounds and only process what is important for functioning. Other authors such as Price [45] have discussed the human processing of speech and language in general.

2.7 Frequency tone and low level processing

Despite the handicap of language observed in many people with ASD, often, a large number of them have great musical capabilities. Indeed, while activation of the left inferior frontal gyrus is lower in autistic children as compared to controls during stimulation with speech, the opposite occurs when the stimulus is song instead of speech [46].

Functional connectivity between the left inferior frontal gyrus and the superior temporal gyrus becomes elevated in the case of songs as compared to speech for autistic persons. Also, increased activity in frontal-posterior connection is observed [46].

2.8 Existing Solutions

The simplest, and fastest way to deal with the problem of hearing sensitivity is by wearing earmuffs. This way, the sounds reaching the eardrum have lower intensity. These earmuffs are used as insulators reducing the sensation of pain in the perception of a given sound [7]. However, due to gradual human adaptation to different conditions, earmuffs can worsen sound sensitivity in the long-term. Nonetheless, it is the best option when the problem is very serious and requires urgent care.

Another way to address the unpleasant noise during hearing is through a therapy known as the Auditory Integration Therapy (AIT) or sound-based intervention [10] [7] [47] [48] [49] [50]. The method consists of multiple sessions of 20 to 30 minutes in which a patient listens to frequency-modulated sounds. This is to re-train the ears and desensitize them over time [7].

An alternative approach is Exposure Therapy, which consists of gradual re-introduction of the offending sound at progressively closer time intervals, until the patient gets used to it [7]. The problem is that exposure therapy, when it is used, is intended to remedy fear and anxiety, but not pain [7].

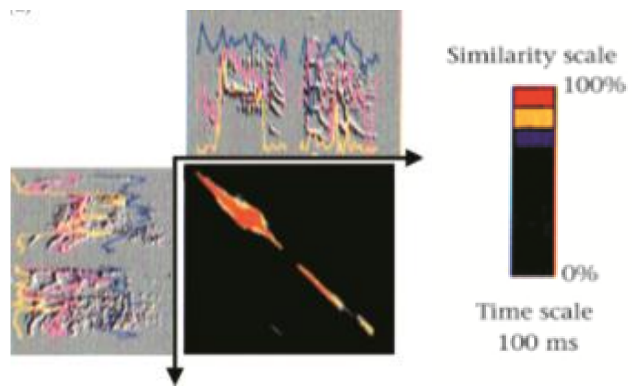
As mentioned earlier in the introduction, the main objective of this research thesis is to provide a tool to prevent the severe hearing sensitivity experienced by patients with ASD. The key element is prevention. The next section provides an otherview on what has been done addressing similarity between sounds.

2.9 Previous works to find similarities between signals

There are many researchers in electro-acoustic and phone applications that have worked on finding similar sounds, music, or similarity analysis of signals. These use the Fourier Transform Algorithm with different complementary object involving Cross-correlation of signals in the time domain, signal segmentation in the frequency domain, and other addition machine learning algorithms we will discuss in this section.

2.9.1 Short Duration Signal (or Short Lasting Signal)

Tchernichovky et al. [51] presented a fully automated procedure that measures parametrically the similarity between songs. They built a procedure that measure similar sections between songs: song of a tutor and that of its pupil. The procedure assigns a numeric score to the pupil's song according to the accuracy of his match with tutor song. High score assigned to songs for which there were close match, and inversely. Obviously, the method uses Fast Fourier Transform, and a time windows of 7 ms [51] (Figure 7).



<u>Figure 7:</u> Similarity matrix between 50-ms intervals across feature

High similarity values are now restricted to the diagonal, indicating that each of the node of the father's song was imitated by his son in a sequential order. Similarity scale: 0-70% (black), 71-80% (blue), 81-90% (yellow), 91-100% (red). (Image and caption from [51])

However, the author recognizes that song is a must complex phenomenon that last in seconds then this method is not suitable for rigorous evaluation of quantitative similarity. The author also mentioned a previous attempt to automate the analysis was based on Sound Spectrographic Cross-Correlation that is not accepted by all [51].

2.9.2 Detection of Similarity in Music File

2.9.2.1 Detection using music parameters

In their work, Thomas et al. [52] performed a comparison method using musical parameters such as tempo, also called speed, key and envelope which are extracted from the music [52]. The tempo is the number of quarter notes measured in beats per minute (BPM). Using tempo as characteristic allows selecting songs which have the same speed. Key is also an important metric that characterizes a song; it is "a group of pitches, or scale upon which a music composition is created" [52]. The key may be of a major or minor mode. Genre is also another music parameter that involves the use of a Support Vector Machine (SVM) classifier that is a machine learning principle [52] which is performed using a large database of sounds whose genre is already known.

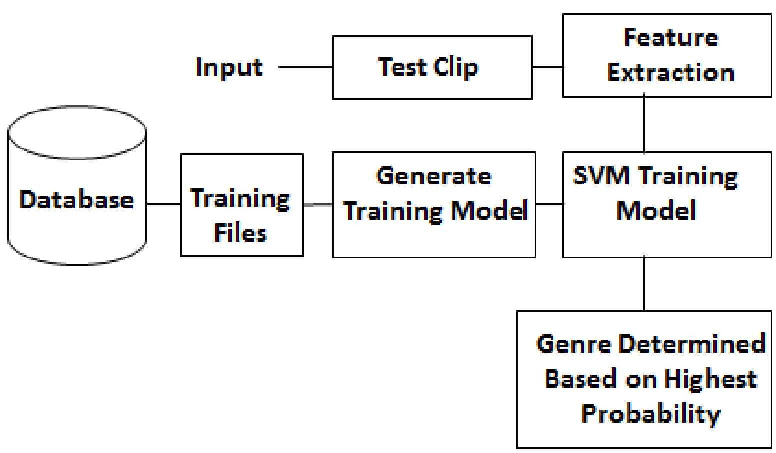
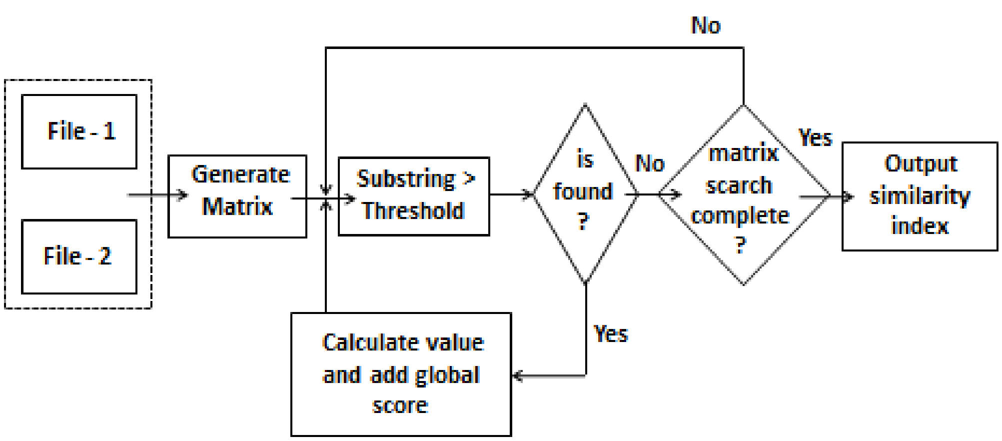


Figure 8: The process of estimating the genre of a song clip using an SVM classifier (Image from [52]).

Genre is more commonly, the category to which a given music belongs, based on certain stylistic properties. It is itself a subjective property to which it is difficult to assign a finite value [52].

Envelope is an important parameter for comparison at the signal level. Envelope may be a function of time, space, and angle. The comparison of the envelopes of two song files allows assessing similarities of their wave files [52]."The similarity factor between two songs has significant commercial application, such as the automatic generation of playlists, as well as copyright protection" [52]. Music Information Retrieval (MIR) from which similarity emerges, is a growing field dealing with pattern recognition and machine learning algorithm. It requires the modelization of music style using machine learning during which the computer is "trained to distinct some properties that are characteristics of different music genres". The process of music same genre extraction is schematized in Figure 8.

As mentioned earlier, MIR uses SVM. The SVM generates a probability value for each potential genre, and the genre with the highest value is selected [52].



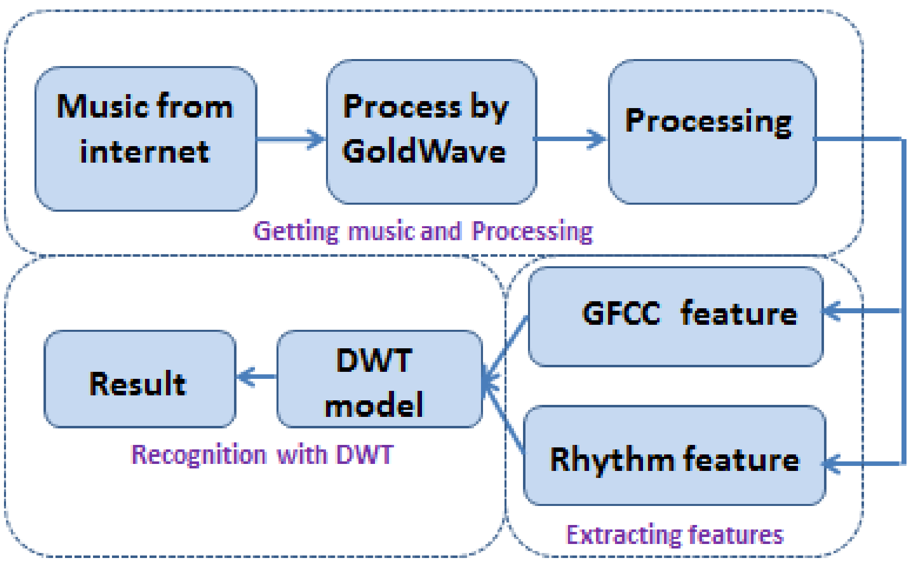
<u>Figure 9:</u> Process extracting all common sub-sequences of two envelope signals (Image from [52])

Multiple algorithms are used to perform SVM. Thomas et al. [52] used envelope and key to perform their extraction. As concern envelope, they performed « All Common Subsequences

(ACS) » which is shown in Figure 9. It is still possible to add more features to the list of parameters to make it more accurate [52].

2.9.2.2 Detection combining GFCC and DTW

A feature which plays an important role in music feature is rhythm. Ren et al. [53] proposed a method for rhythm retrieval based on addition of Gammatone Frequency Cepstral Coefficients (GFCC) feature (Figure 10). Indeed, they presented in their work technics to measure rhythmic similarity between two or more songs. With their system, similar songs can be retrieved from a large collection. They also used the Dynamic Time Warping (DTW) algorithm to score and rank the distance between the tested music, and any music from the collection; it is a "frame based method which can match two time dynamic series" and minimize the difference between them [53].



<u>Figure 10:</u> Flowchart of the proposed system (Figure from **[53]**).

Method for rhythm retrieval based on addition of Gammatone Frequency Cepstral Coefficients (GFCC). Dynamic Time Warping (DTW) algorithm is used score and rank the distance between the tested music, and any music from the collection.

Rhythm contains information like pitch, tempo, timber, loudness. The flowchart of their system is shown in Figure 10. The Dynamic Time Warping is a frame based method, which can match two time dynamic or speed dynamic time series and finds the time dimension to minimize the difference between these two series [53].

2.9.3 Detection of similarity in Signals

2.9.3.1 Cross-correlation of Time Domain

Cross-correlation is a well-known way to measure the similarity between signals. It helps to identify in what way two signals are related. It also helps to identify a song from a database. Xcorr, a function from the Matlab library (mathworks.com) help to determine if there is a match between two signals (see Appendix). A peak obtained at the cross-correlation implies matching the two signals at the time (see Appendix).

In our work, we cannot use cross-correlation, even if it may give us some useful information, because we cannot do anything with the output obtained from cross-correlation, because it provides information on signal taken 2-by-2. The output may be used for other purposes, but it is not intended to be a signal that can be used to target the database. We can perform cross-correlation for visual purposes, but not to achieve our goal.

2.9.3.2 Frequency Domain Analysis

2.9.3.2.1 Mscore

Frequency domain analysis is a very common and simple way to measure similarity between signals. It is the basis for certain previous methods we discussed earlier in this document. The Power Spectrum Density (PSD) displays the power present in each frequency of the signal. Consider two signals and their respective power spectra. The spectral coherence identifies frequency-domain correlation between signals, in other words, it allows determining visually all components (frequencies with high PSD) that two signals have in common (see Appendix).

2.9.3.2.2 Findpeaks

Another way to find components with high PSD is named « Findpeaks » (mathworks.com). As Spectral coherence, « MatchingFreqs », a function derivated from « Findpeaks » allows getting the common peaks frequencies (see Appendix). It provides the same result as that of Mscore with the advantage of storing the data for use (see Appendix).

2.10 Our Approach for this current work

As you can see from these methodologies extracting similarity, they all attempt to find similarities between two sounds. We need to find a methodology that identifies similarities between 2, 3, 4, 5 input sounds at the same time. In our work, the aim is not to retrieve the same genre of music, because a child can be sensitive to different genre of sounds. The sensitivity factor might not be the tempo, key and signal envelope, ... , but might be something particular to each of them like "hissing", "ringing" a background noise, a feature that we might not be able to characterize using the earlier mentioned criteria. No methodology to find similarities between 2, 3, 4, 5 input sounds at the same time has been proposed before the current work. Then, to find common point between sounds, we propose an approach based on intervals. Such an approach takes every detail into account.

We also mentioned in Section 1.2 the desire of parents to know the evolution of the sensitivity improvement of their child related to natural or synthetic sounds; we designed for that component that evaluate the sensitivities improvement that will be introduced that we named « direct » and « indirect evaluation » or « sensitivity analysis using available information » and « sensitivity analysis without information » . We also mentioned earlier that Auditory Integration Training is a component we integrated as complementary to the two evaluation processes, as desensitization to sounds procedure; that component will also be briefly discussed.

3. Fundamental and Methodology

In this chapter, we are going to introduce the basics of the signal processing tools and the methodology we used to achieve each of the targeted objective. The methodology to find commonalities between the uploaded sounds will be introduced in section 3.1. Direct and indirect evaluation of sensitivity will be introduced in section 3.2. The auditory integration training will be introduced in section 3.3.

3.1 Common Point Extraction (Similar Sound Extraction)

3.1.1 Fast Fourier Transform

Fast Fourier Transform (FFT) is a fast computation algorithm implementing the Discrete Fourier Transform (DFT) [13]. It reduces the growth of the computational complexity of DFT in respect to the processed data size from $O(n^2)$, to $O(n \log n)$, where n is the data size [13].

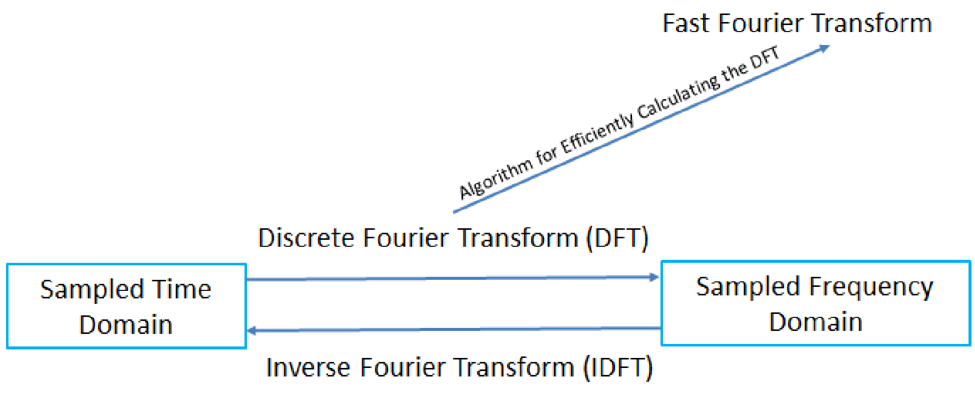
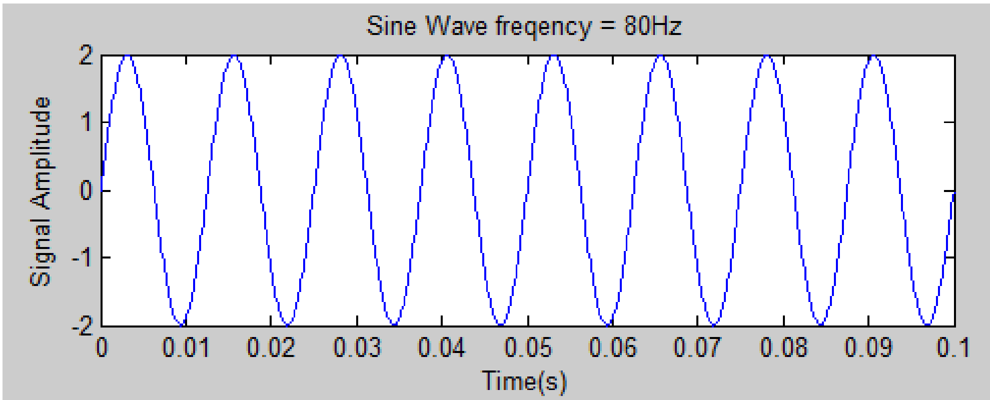


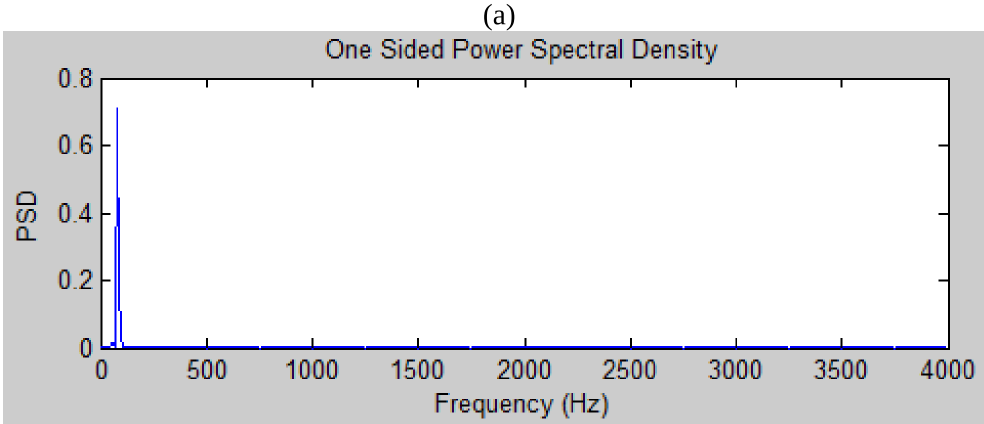
Figure 11: Time domain to Frequency domain transformation through FFT.

(adapted from **[54]**)

The power spectrum of a signal is described by its Power Spectral Density (PSD). PSD gives signal's power distribution over the frequency domain. Please note that from the PSD vs frequency graph, the amplitude vs frequency graph can be derived. We give below the time domain and frequency domain response of a sinusoid signal.



(a)

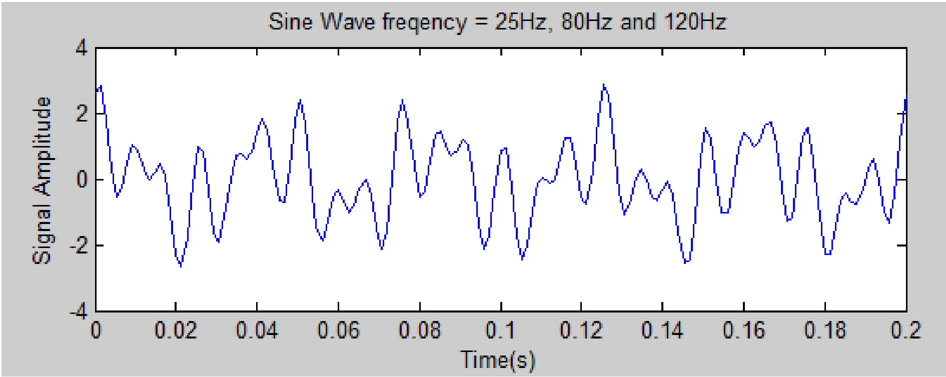


(b)

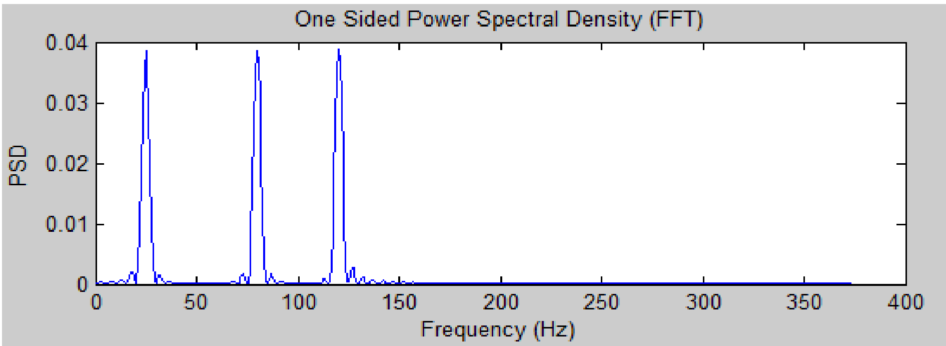
<u>Figure 12</u>: Fast Fourier Transform of a sinusoid waveform

(a)Pure sinusoidal signal of 80 Hz with time duration = 0.1 seconds synthetized on Matlab.
(b) FFT decomposition of the signal shown in (a).

Figure 12 (a) shows the time waveform of summation of 3 sinusoids, at 10 Hz, 80 Hz, 120 Hz.



(a)



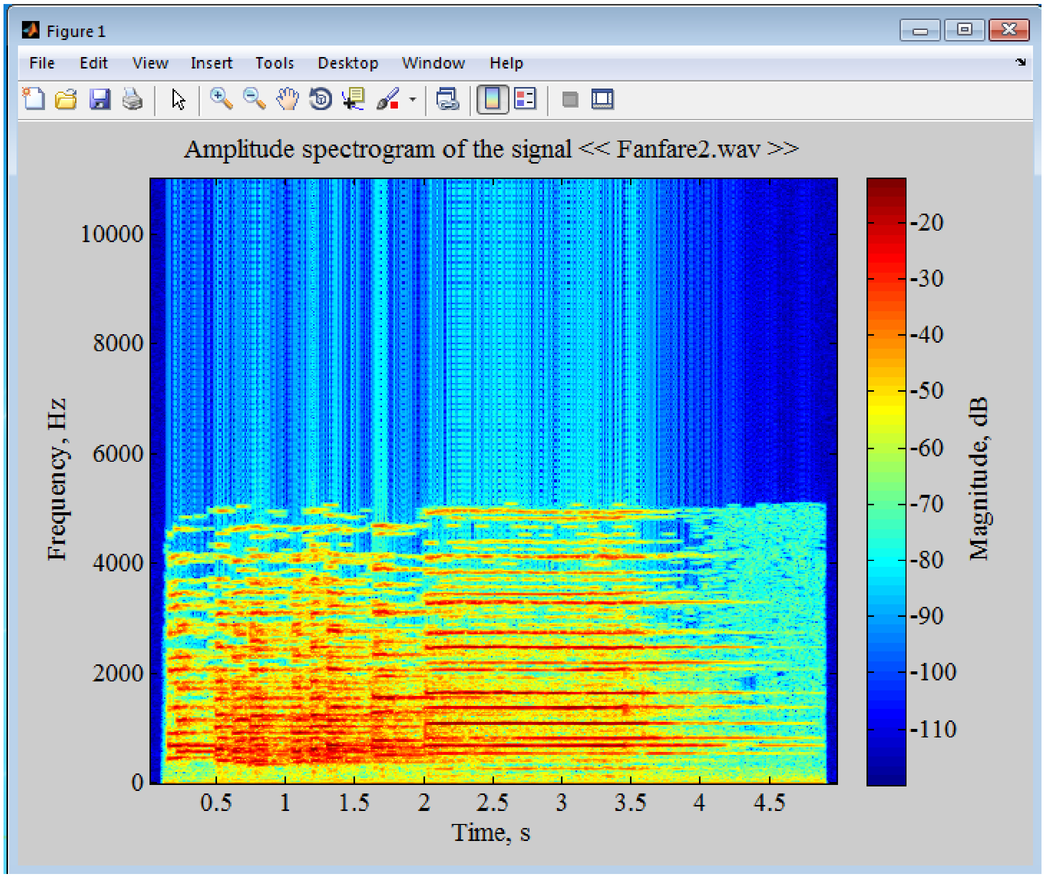
(b)

<u>Figure 13:</u> Fast Fourier Transform of a signal composed by sinusoidal signals of 10Hz, 80Hz, and 120Hz.

From the time domain waveform we cannot identify the three individual sinusoidal signals forming it. This can be done easily however, when we move to the frequency domain, where the three frequency components are shown clearly.

3.1.2 Short Time Fourier Transform

The Short Time Fourier Transform is based on the FFT algorithm. It is more complex than FFT because it takes the time into account, which add one more dimension to the graph. In the Short Time Fourier Transform, the system takes equal time segments of the signal and runs FFT for each of them. The starting time instants of those segments are t = 0, 1, 2, ... , n. Due to the variability of voice the spectral contents of different segments are not identical.



<u>Figure 14:</u> Short time Fourier Transform (STFT) of «Fanfare2.wav»

The colors correspond to the amplitude at certain frequency as has been calculated by the FFT. It ranges from blue (very low) to red (very high). The spectral analysis has been done using a Mathworks code taken from Matlab's website, which we modified (we changed one-or-two lines) to add our own wave file «Fanfare2.wav».

Figure 14 presents the output of the wave file «Fanfare2.wav» in Short-Time-Fourier-Transform. It is a very complex algorithm based on FFT algorithm, which take the time into account; A short period of time (code from **mathworks.com** see annexe).

3.1.3 Power Spectrum Density and Intensity

It is useful to mention here that for the rest of this work, we use the term « intensity » to refer to the strength of the « Power Spectrum Density ». In fact, from the generated graphs, it is easier to extract the PSD, and from PSD, one can evaluate intensities. Each figure used for each of the two Figures (15 and 16) aims to illustrate the target zone that the algorithm will refer to when searching for candidate wave .file from the Database.

3.1.4 Basis of the Two Approaches Used

There is no « a single way » to find the commonality between different sounds. However, the principle of our approaches is based on overlapping areas:

We considered the delimited intervals of all frequencies; lower frequencies to higher frequencies, and lower PSDs to higher PSDs. Then we considered:

- The range between minimum and maximum frequency having power content (i.e PSD value higher than 0) appears in the union of the spectrums of the considered individual signals (Figure 15).
- The range between minimum and maximum values of PSD observed in the signal spectrum is represented by the vertical green arrow (Figure 15)

After that, we have two configurations:

1) The sounds surfaces are very close from each other (overlap each other very well)
2) The sounds surfaces are very far from each other.

We will show later in this work that each of the two approaches has its strengths and its weaknesses.

3.1.5 Inside all Intervals

To define the total interval, we use the « union » function (Figure 15). Then, we look at all the sounds listed in our database, for which the frequencies and intensities are within the yellow rectangle. It is good to note that with the « intersect » function rather than « union » function, we enter the purple rectangle, which decreases our chances to have similarities.

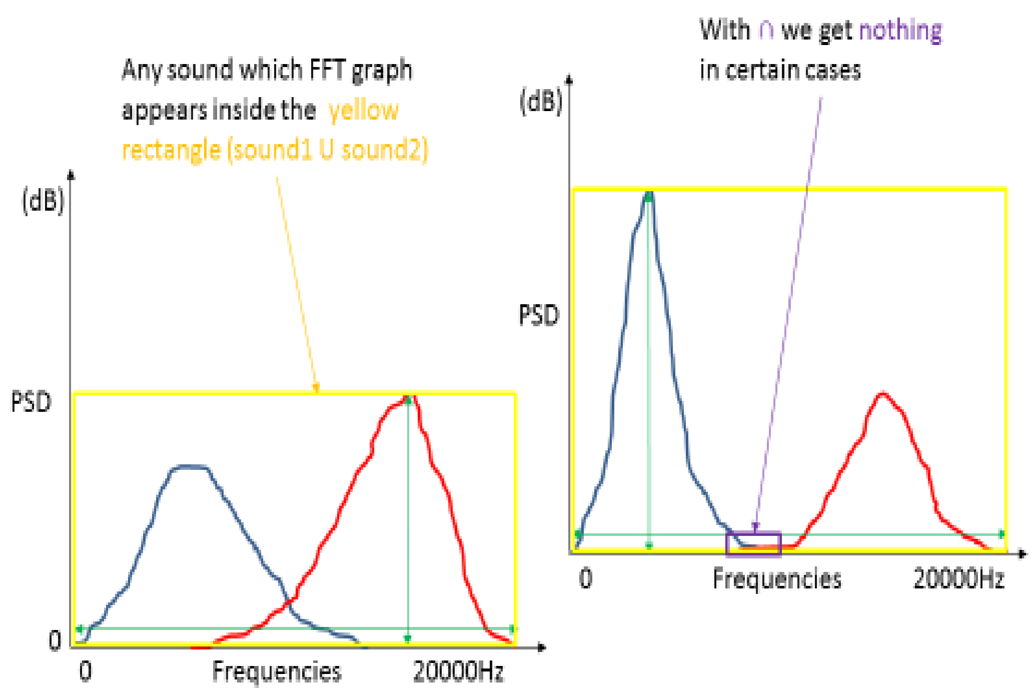


Figure 15: Inside all Intervals

In blue, we provide the PSD content of sound 1, and in red, the PSD content of sound 2. The term « union » " U " means sound1 + sound2, while the term « intersection » " ∩ " operation identify and extract only the area of the spectrum where both signal have energy content. The Union operation combines all frequencies where the signal 1 and signal 2 have power content. The spectral occupancy of the combined signal is shown by the horizontal arrow. If we use the « intersect » function rather than « union » function, we are left with the frequencies in the purple rectangle only, which decreases our chances to have candidate sounds.

In figure 15, we present the case of 2 sounds. With more than two sounds, the process remains the same to extract ranges of intensities and frequencies. In this example, at the left side, both sound1 (in blue) and sound2 (in red) will be selected by the algorithm, while in the example at the right side, none of sound1 or sound2 would be selected.

- ✓ For 2 sounds:

Frequency range = union (F_1, F_2); F_1 and F_2 are 2 vectors.

Intensity range = union (I_1, I_2); I_1 and I_2 are 2 vectors.

- ✓ For n sounds :

Frequency range = union (F_1, ... , F_n); F_1, … , F_n are n vectors.

Intensity range = union (I_1, … , I_n); I_1, … , I_n are n vectors.

3.1.6 Outside intersect intervals.

If we are to be more rigorous with the principle of selecting only similarities between sounds, it is essential to use the « intersect » function (Figure 16). It allows us to consider only the common intervals among the sounds. Under these conditions, we will have two options available to us, but very different from each other: either we restrict ourselves to the sounds of the inner yellow rectangle, or we consider only the sounds outside. Sounds inside the yellow rectangle are very restrictive.

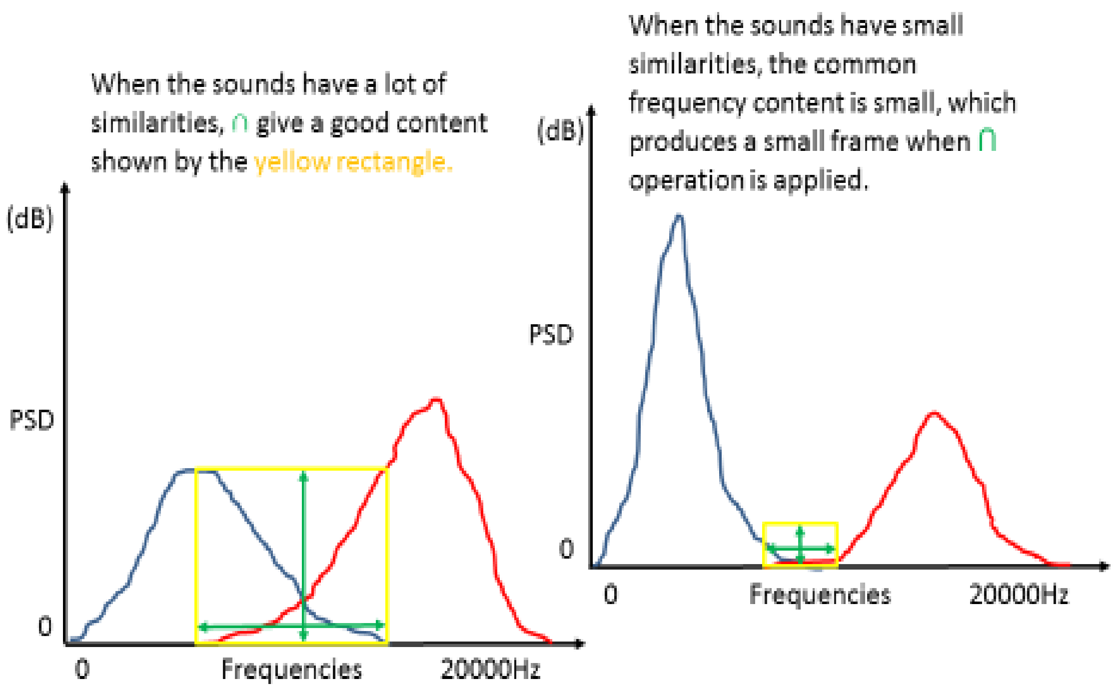


Figure 16: Outside intersect Intervals

In blue, we provide the PSD content of sound 1, and in red, the PSD content of sound 2.

We chose sounds which spectrum have an area inside the yellow rectangle, and which rest of the spectrum goes outside the yellow rectangle as the selected sounds themselves have an area inside that rectangle and spread outside the rectangle. Any other sound (from the database) having a part on this area will be selected by the algorithm. The more the sounds have different frequency bands, the more the yellow rectangle dimensions will be reduced. In the example at the left side, both sound1 (in blue), and sound2 (in red) will be selected by the algorithm, while in the example at the right side, none of sound1 or sound2 would be selected.

3.1.7 Summary of expected strength and weakness of each method

As mentioned previously, each of the approaches used has expected advantages and disadvantages.

- ✓ Inside all intervals : even if all the frequencies and intensities of uploaded sounds are distant from each other (left side, Figure 15) or close to each other (right side Figure 15), this algorithm should always give a good estimate of all candidates. It may give a few « false positive », but we do not expect « false negatives ».

- ✓ Outside Intersect Interval: This algorithm is much more rigorous than the previous because it considers research candidates solely on the basis of the joint surface (Figure 16). We expected not to have false negatives.

Table 1: Expected Advantages and Disadvantages of each of the two Methods

Method	Advantages	Disadvantages
Inside all intervals	Should always generate a result to have at least an idea on what the user is looking for. Should be the best choice when the sounds are	Can give false positives.

	very different.	
Outside intersect intervals	Much more specific than the first in term of research of common area between sounds	No result if there is no intersect area, which generally never happens

Algorithm 1 and 2 are inverses and complementary to each other. Table 1 gives a summary of expected advantages and disadvantages for each of them.

3.1.8 Input generation for tests

3.1.8.1 Internal Database Generation

Any wave file from external source is supposed to work as long as the length it is not too long (1 to 15 seconds sound). Normally, 1 to 10 seconds is enough for a user to identify a sound; also the wave files used to write the codes and perform the simulations all have lengths varying between 1 to 20 seconds. But for more sounds options, we got permission from this website **FindSounds.com** that the user can use too (see appendix for link and description).

The author has carefully collected the most diverse and imaginable sounds from different websites on his website. From that, our Internal Database consists of a Folder named « Sound_Data_Base » that contains a variety of different sounds found in daily life. Ninety percent of the wave files of our database came from the website. Autistic person experience different levels of sensitivity to certain sound, thus we have to include any type of sounds in our database. One cannot generalize, so there is not specific sound that we should collect; here is no sound that is generally known to bother autistic patients in general, a reason why the algorithms should take all possibilities into account.

3.1.8.2 Input generation by users

Let us say we have in mind two sounds we want to test: «fan» and «alarm clock». We have 2 choices: either we get the file from the Direct Evaluation record, or we download it online. From online, a good example of website is FindSounds.com that we used to develop this application.

3.2 Sensitivity Evaluation

We evaluated the sensitivity by two different methods: « direct evaluation » also called « sensitivity analysis using available information », and « indirect evaluation » also called « sensitivity analysis without information ». The difference between direct and indirect evaluation of sensitivity is that in direct evaluation, a few sounds that are problematic to the ASD individual are known; e.g. it is known that the individual is sensitive to the sound of the sound of a fan, or an electrical razor. Then the operator will record the sound using the software, or download it from the sound database and integrate it into the software; « sensitivity analysis using available information » is a time based evaluation on a known sound (Figures 17, 31 and 32). In « sensitivity analysis without information » the operator does not know what sounds are bothersome to the individual, nor he/she is aware of what the problem is. The operator will work with the software, and the information regarding the individual's sensitivity will be recorded (Figures 18, 34 and 35). In the case of « sensitivity analysis without information » the only data are those the application will generate. The data are raw (e.g: 100 Hz, 2dB). They may have to convert it into something they know, e.g: Drying. The question now is: how does a user converts a frequency/intensity to an interpretable sound such as 'drying machine'? For now the interpretation is purely visual; the user will be seeing his sensitivity profile through « sensitivity analysis without information » (Section 3.2.2); this provides him informations on frequencies and intensities. For all sounds in the database, s/he can also access information on frequencies (Figures 23, 25 and 30). From that, he can visually associate it with a recorder sound. However in section 6, we suggest another solution which will future work.

3.2.1 Sensitivity Analysis Using Available Information: Critical sounds (Natural sounds).

3.2.1.1 Principle and Interface

As mentioned earlier, « Sensitivity Analysis Using Available Information » also called « Direct evaluation » is based on sounds that the user has been exposed to and knows he/her reaction to them. It allows the user to know the evolution of his ability to tolerate a sound that has been registered in the application by a relative of the user, or therapist expert (Figures 17, 31 and 32).

The application allows storing critical sounds (which are difficult for the ASD patient to bear), listen to them as long as the user wants to, and evaluate the listening time duration. To acquire such sounds, the user has 2 choices. Either s/he records the sound generated by such source, (e.g. record at home the sound generated by the vacuum cleaner) and input it in the software through «Record a new sound», or download the sound from Internet, if available (e.g. from findsound.com); S/he will have to save it on the desktop and integrate it in the application through «use the Browser to Upload a new sound».

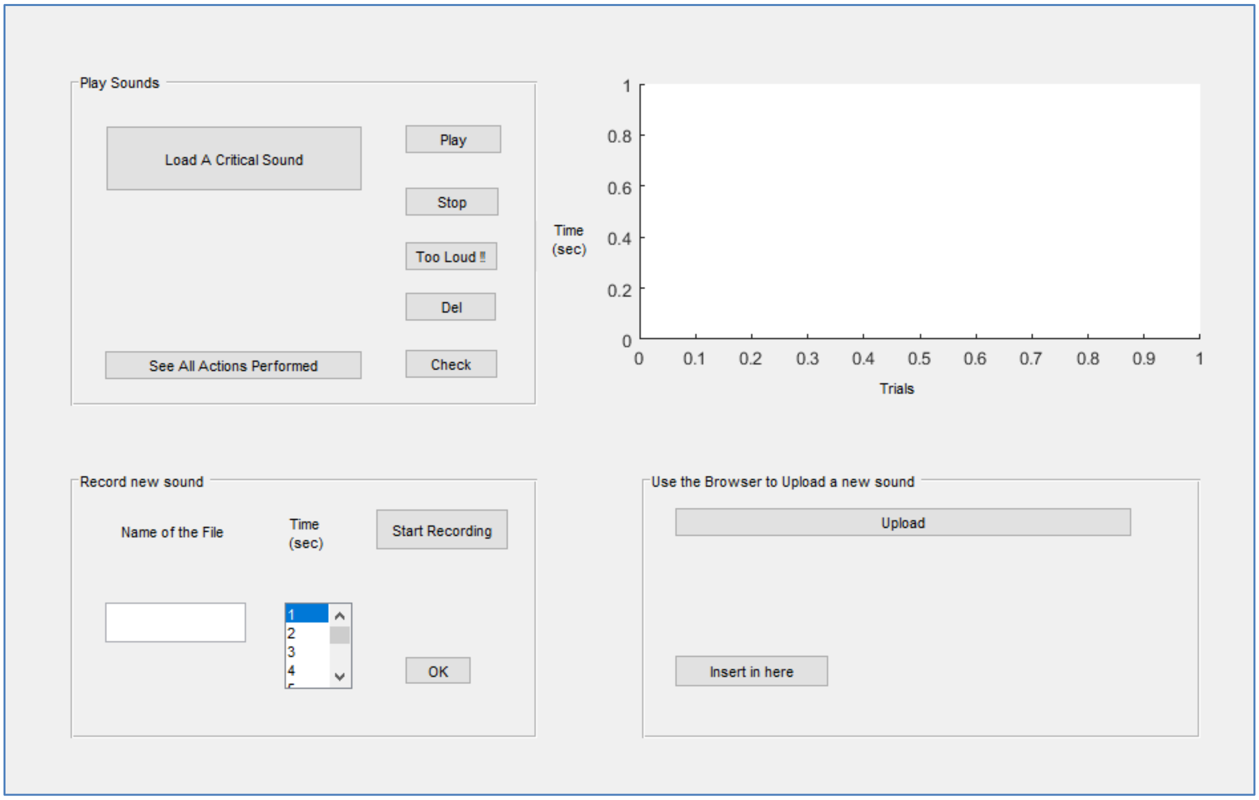


Figure 17: Direct Evaluation of the Sensitivity

This part of the software is designed to assess the sensitivity of autistic person to sounds that are bothersome to them. "Play Sounds" allows listening to sounds. "Record new Sound" allows recording a new sound. "Use the browser to Upload to upload a new sound" allows the user to input in the application, sounds from an external resource, and save them in the desktop. The most important is the graph that allows the overview of the improvement when pressing the "Check" button for a selected sound.

3.2.1.2 Input generation for tests

3.2.1.2.1 Recording through the application

The input data are generated directly through the application « Evalutation of sensibility »: they consist of Wave files that are created inside the application when the user presses the « Record

New Sound » button (Figure 17). When the user presses the button, the software records surrounding sound, which is saved into a wave file.

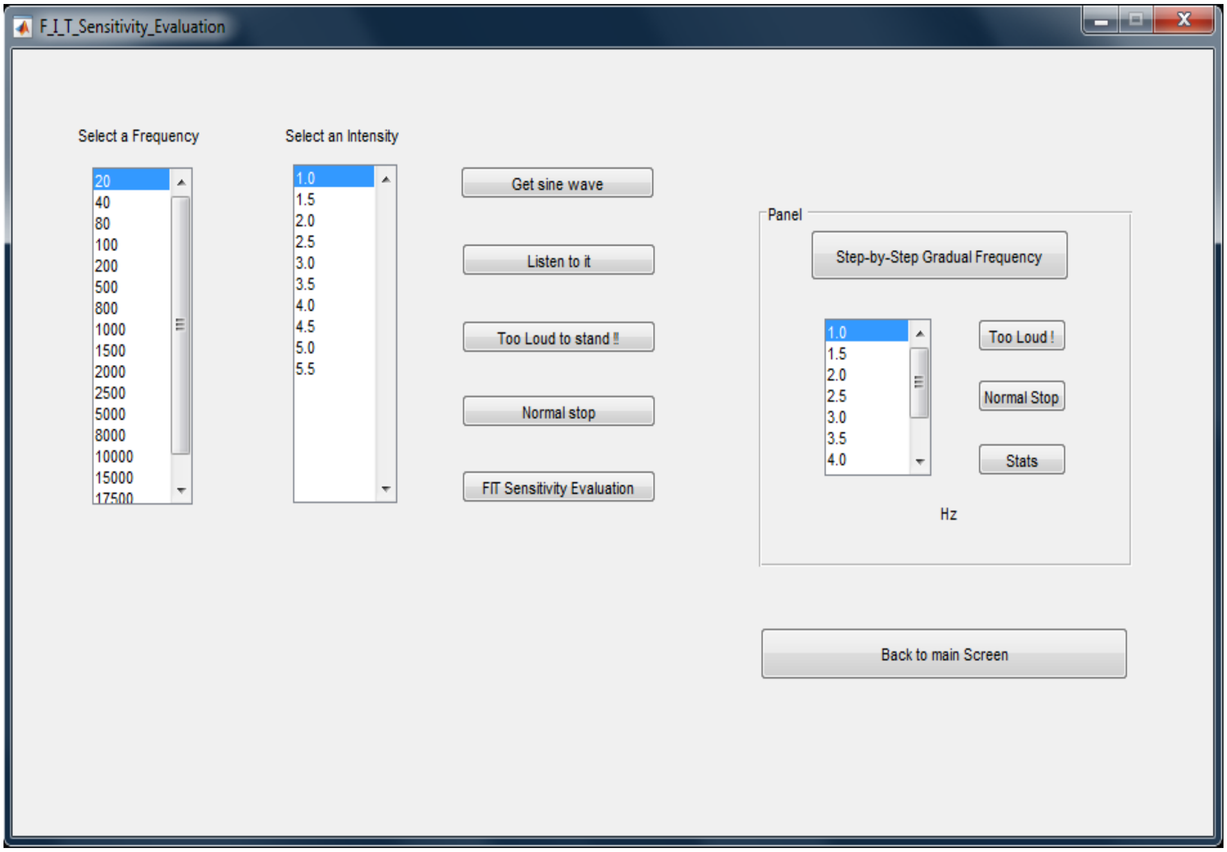
3.2.1.2.2 Download from the online Database

The file can be downloaded from the sound database website (findsound.com see appendix) and then inserted in the application using « Use the Browser to Upload a new sound » (Figure 17). The application can currently read only wave files; later, we are going to add mp3 and mp4 file readers.

3.2.2 Sensitivity Analysis Without Information: (Synthetic Sounds ie: frequency, intensity and time).

3.2.2.1 Principle and Interface

In « Sensitivity Analysis Without Information » also called « Indirect Evaluation », the user deals with numbers, values of sounds components, then it is synthetic sounds; all sounds can be decomposed to sinusoids of certain frequencies, intensities and phase/time lags. Implemented using Matlab, the purpose of this component is to assess the evolution of an Autistic person's ability to withstand pure synthetic sounds (Figure 18). With this component, the application will identify the frequencies and their intensity levels to which the user is susceptible. This information helps to identify the type of sound (s) to which the user is sensitive. We get the precise values for each parameter.



<u>Figure 18:</u> Indirect Evaluation of the Sensitivity

We evaluate the frequency and intensity sensitivity. For the « constant » evaluation (on the left), we set different range of frequencies and also different range of intensities; the user is going to select the frequency and intensity. « Get sine wave » allows the user to see the wave form of the sound he is hearing. « Listen to it » allows to user to listen to the sounds. « Too loud to stand!! » allows the user to stop the sound instantly. The parameters, Frequency and Intensity, will then be recorded in a file to draw a graph of sensitivity. « Normal stop » allows the user to stop the sound not because it is loud but because it is boring. « FIT Sensitivity Evaluation » allows the user to see visually the statistics. For the « gradual » evaluation (on the right), we set different range of intensities; and for each selected frequency, when the « step-by-step Gradual Frequency » button is pressed, the sound keeps growing from 20 Hz to 20 000 Hz as long as the user doesn't press the « Too loud » or « Normal Stop » button.

For frequencies, we had to put frequencies ranging from 10 to 20000 Hz which is the upper human limit. Using the term low, medium, high could be appropriate here, but we believe that 3 options are not enough because there is a huge difference between 20 Hz and 20 000 Hz. The same situation applies to intensities.

A similar non-invasive psychoacoustic test has been done by Khalfa et al. [8] in order to measure the auditory dynamic range of ASD versus typically developing brain. In their first test, their frequencies ranged from 0.25 to 8 kHz and the tone level was decreased and increased in 5-dB steps. "The procedure involved pure tones of 500 ms duration where stimuli characteristics were either controlled by audiometer, or an experimenter" [8]. They showed that in both cases (frequency increase versus intensity increase), there is hyperacusis with people having ASD from 2.5kHz to 8kHz and from 15dB to 80dB [8]. Their second test was qualitative; it consisted of "random presentations of the 1 kHz tone at various intensities in multiples of 5 dB, ranging from the Quiet Threshold to the Loudness Discomfort Level determined for each subject in the first phase" [8]. When the subject heard the sound, s/he qualitatively rates the tone’s loudness using a numeric scale where 1 was "low"; 2 was "medium"; 3 was "loud"; and 4 was "too loud". [8]. They showed that children with ASD rated the 1-kHz pure tone as "Loud" at much lower intensities than TD [8].

3.2.2.2 Input generation for tests

The input data are the selected data on the listboxs (Figure 18). Every time the user uses it, if s/he is sensitive, the sensitivity information is stored in the application. For constant evaluation (left side) s/he just has to select a value for frequency and a value for intensity. For gradual evaluation (right side) s/he just has to select intensity in the listbox.

3.3 Auditory Integration Training

Auditory Integration Training (AIT) is an additional feature we added to the application. AIT is the current way to treat the impairment based on music therapy [47], [48] [49], [50] [48].

Auditory integration therapy has been developed by Berard and Tomatis in 1993 [48]. It typically involves "10 hours of listening to electronically modified music delivered via headphones during 2 half-hour daily sessions over 10 days" [48]. Its device uses filters to remove peak frequencies to which the individual is "hypersensitive". The filtered music is modulated by a random low and high pass filter. [48]. The intervention, according to Gold et al. [49]¸ should work as a non-verbal and pre-verbal language activity, which means interaction activities without words, that is more oriented toward emotion, relationship with others. The intervention also include the selection of "music that is meaningful for the person" [49].

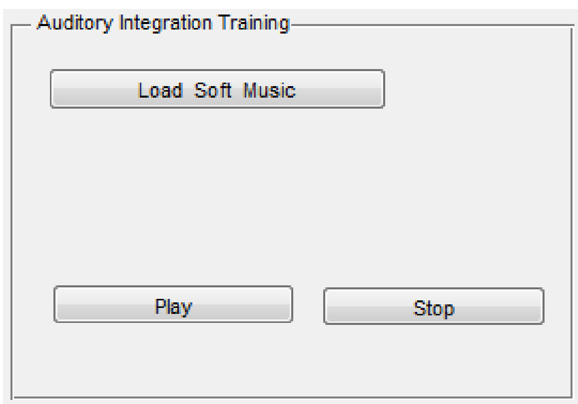


Figure 19: Auditory Integration Training (AIT)

Interface with 3 main buttons: "Load Soft Music" to search for the music of interest in the Software internal Database; "Play" to play that music and "Stop" to stop it any time. Those music file have a medium length of about 10 min.

We built a very simple interface that allows listening to soft music. In general, most of the soft music is coming from classical genre. Any type of soft music can be added to the application.

In their investigation, Bhatara et al. [55] reported that among all genre of music, the favorite of people with ASD are respectively Rock music and Classical music. But the between group comparison in music genre showed a net preference of ASD individual to classical music genre. This prevalence to classical music according to Bhatara et al. [55] might be due to social affiliation or "increased complexity of classical music relative to other genres" [55].

4. Generated Results and Discussion

In this chapter, we will present all the results obtained starting with the extraction of common sounds in a preventive measure (Section 4.2), results for sensitivity evaluation (Section 4.3), and results for Auditory Integration Training (Section 4.4). The interface of the application will be presented in Section 4.1.

4.1 Jodivi Interface

The application has been named "Jodivi" as for Joel Sandé, Dimitrios Makrakis, and Virginie Cobigo. It is possible to load one, two, three, four or five sounds, five being the maximum capacity of sounds that can be analyzed at the same time.

Figure 20 is the main figure of the present project. It shows the main interface of the application. The interface contains:

1. Main application that finds common point between uploaded sounds and filters sounds from the database. It is divided into 2 parts :
   1.1 See Analysis Graphs: It gives mathematical information on the uploaded sounds.
   1.2 Common Feature Analysis: The key-button of the application that allows the filtrating process. It gives two outputs explained in details in the next paragraph.
2. Sensitivity Evaluation:
   2.1 Direct Evaluation: Sensitivity analysis using available information
   2.2 Indirect Evaluation: Sensitivity analysis without information.
3. Auditory Integration Training: It allows listening to some soft music (generally classical music).

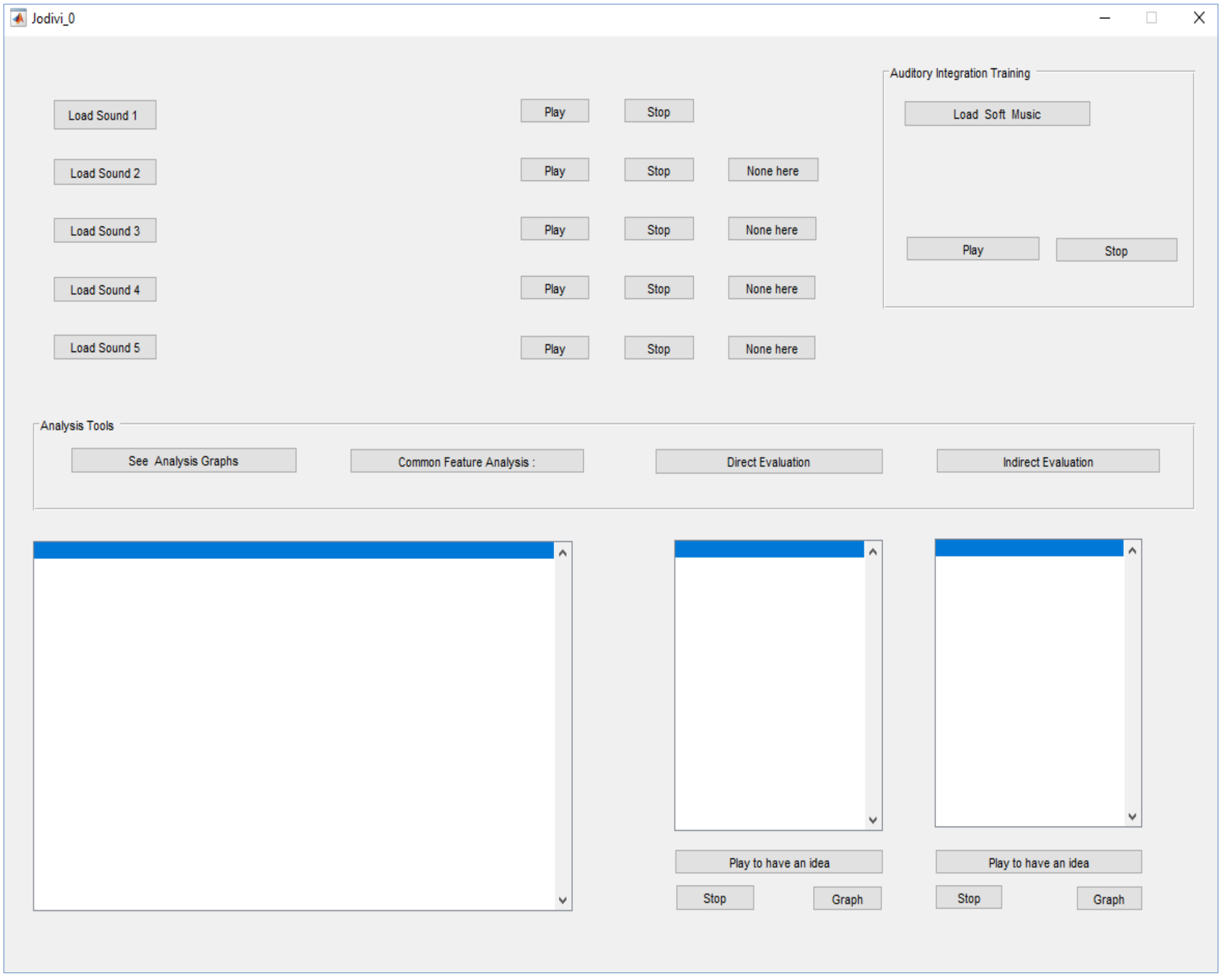


Figure 20: Jodivi

4.2 Results for the research of Common sounds and extracting Common points between sounds

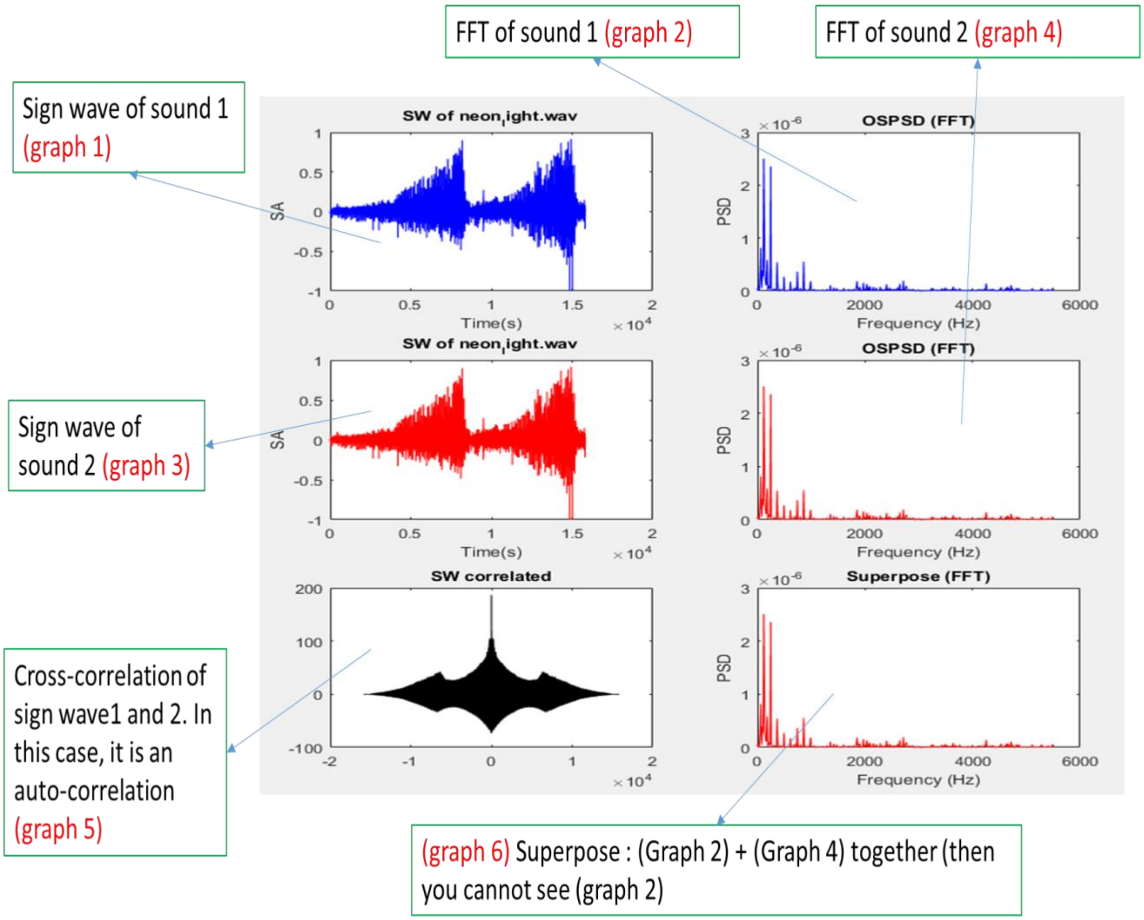
The main application searches for common features and extracts common sounds between 2, 3, 4 or 5 wave files, depending on what the user wants. If only one wave sound is uploaded, the application skips some steps of pre-treatment between files.

It gives three outputs:

1. List of all songs that correspond to the selective criteria (On the Listbox at the right Fig. 16).
2. Assessment: On the Listbox at the left (Figure 20). It is the assessment after running each algorithm. It informs us which uploaded sounds are similar, and which the common intervals are based on the « Yellow Rectangle » of the algorithm.
3. Graphs visualization: It allows the mathematical interpretation of the sounds.

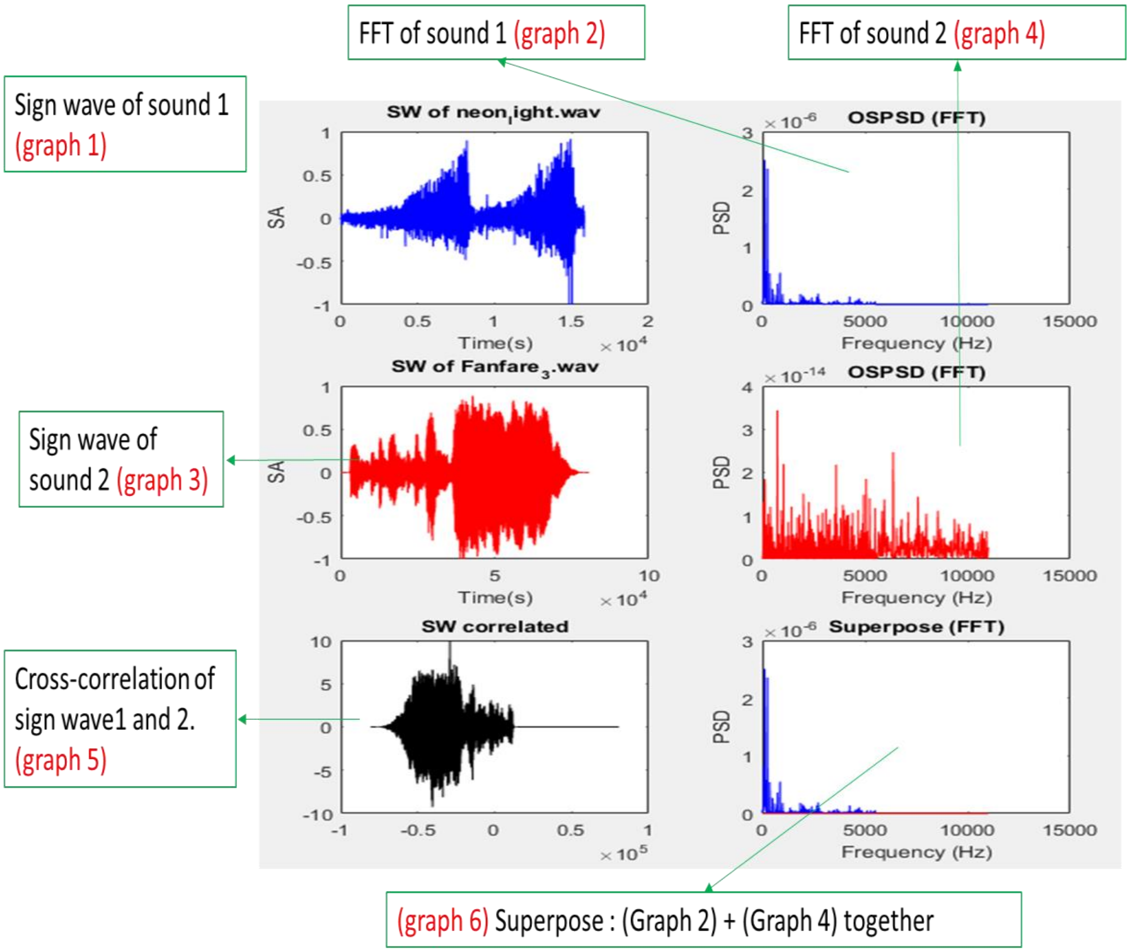
We have already downloaded a large number of sounds from an online database (findsound.com see appendix), which we incorporated into the Matlab software. This is our internal database. Let us say the user is known to be sensitive to both, fan and clock sounds. S/he wants to know the other sounds s/he may be also sensitive to. Then when each of the two algorithms run, they will search using the specific criteria linking those 2 sounds. The application will then go to the database, and select among all sounds those that meet those criteria.

Figures 21 and 22 aim to highlight the role played by each graph in the « See Analysis Graph ». Note that the « See Analysis Graph » interface is intended for specialist, pediatricians, health professional that can understand and interpret the meaning of values in the output. It first of all allows the developers to have a view on data while running the application. The plots in Figure 21 were generated using the same wave file «Neon-ligth.wav», and those in Figure 22 were generated using «Neon-ligth.wav» and «Fanfare_3.wav».



<u>Figure 21:</u> Simulation with «neon_ligth.wav».

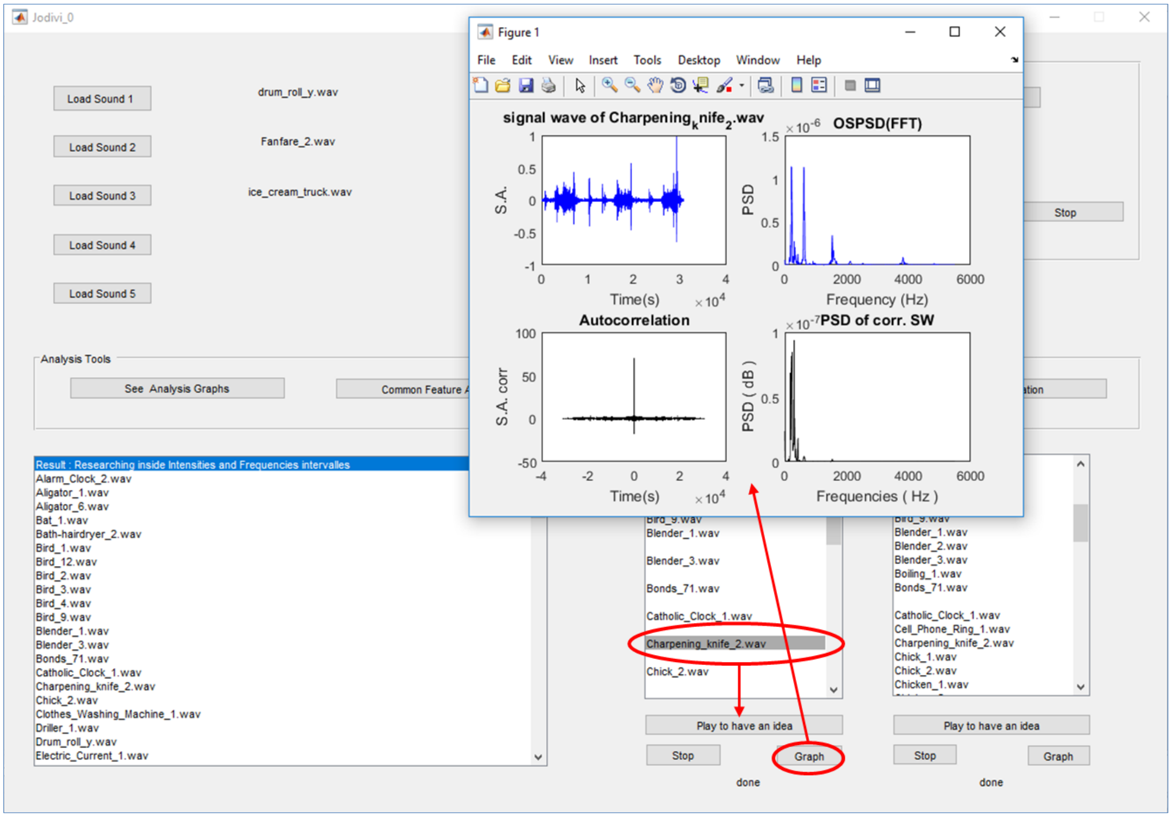
The left column shows the signal wave for each sound. The bottom left is the signal obtained from cross-correlating the 2 signals on the top left. The column at the right shows the FFT graphs of signals at the left, except for the graph at the bottom right, which is all FFT graphs superposed.



<u>Figure 22</u>: Simulation with «neon_Light.wav» and «Fanfare_3.wav» sound File.

As can be seen in Figure 22, PSD of «Fanfare_3.wav» has lower values that PSD of «neon_light.wav», which is the reason why in the "Superpose (FFT)" you can see only a red line.

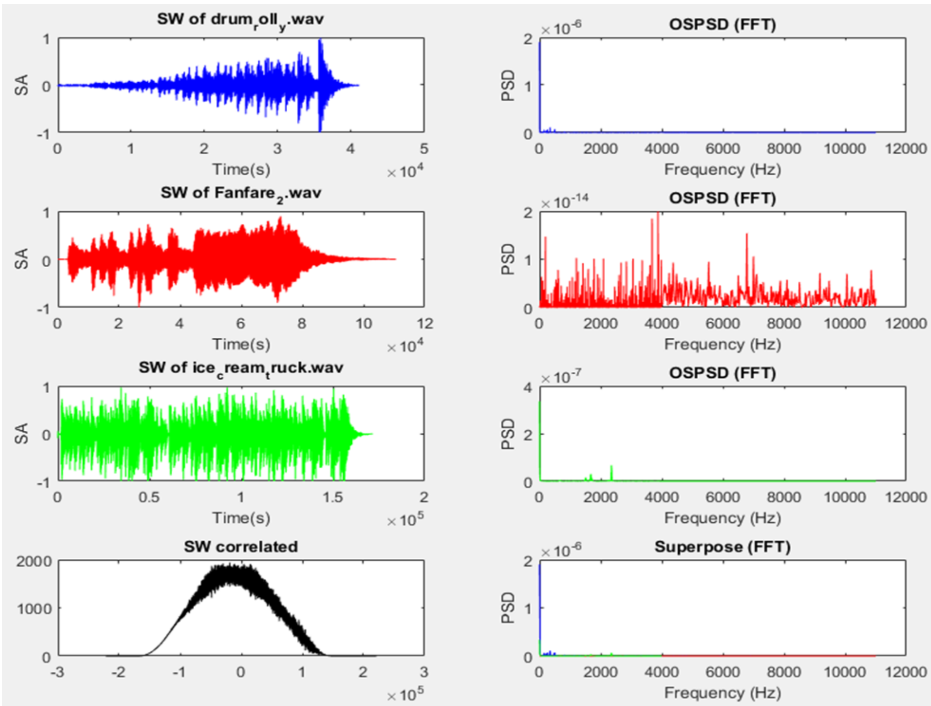
After running the application, the user can select one sound from the methods output and listen to it (Figure 23). The user can select any output sound s/he wants; it is up to the user. S/he can also access the graph of that sound by pressing the « Graph » button. Everything will be up to the user.



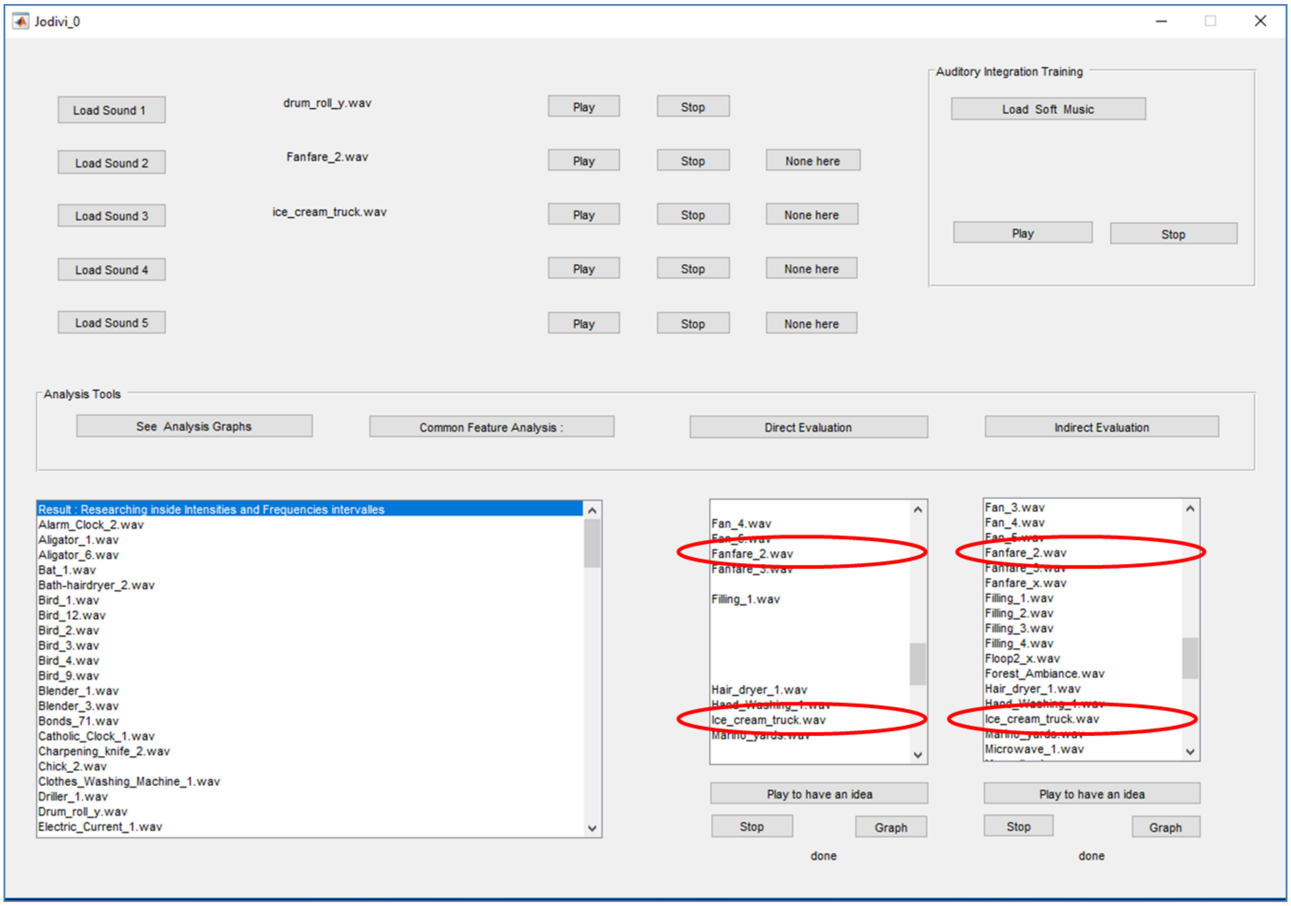
<u>Figure 23:</u> Example of a selected sound

The user can select one sound name and play it to see what that sound looks like. Eventually, s/he can stop the sound at any time. S/he can also press the « Graph » Button to get a graphical view of the selected sound. Here, graph 1 is the sound signal, graph 2 is its power spectrum which is extracted by running FFT, graph 3 is the autocorrelation of the sound signal.

The first simulation for "extraction of sounds for preventive measure" was done with three wave files named «drom_roll_y.wav», «Ice_cream_truck.wav» and «Fanfare2.wav». Prior to performing a simulation, we put all uploaded sounds in the database. We expected all two algorithms « Inside all intervals » and « Outside Intersect Intervals » to give us at least one of the uploaded sounds as response, and this is what happened. There is something notable with our algorithms in that they are complementary: Algorithm 1 « Inside all intervals » targets toward the inside area, while algorithm 2 « Outside Intersect Intervals » target beyond the overlapping area. Then some of true positives will be generally those selected by both algorithms. Unless we perform the test with real users, One cannot state with 100% certitude which other sound selected by algorithm 1 and 2 is a true positive or false positive.



(a)



(b)

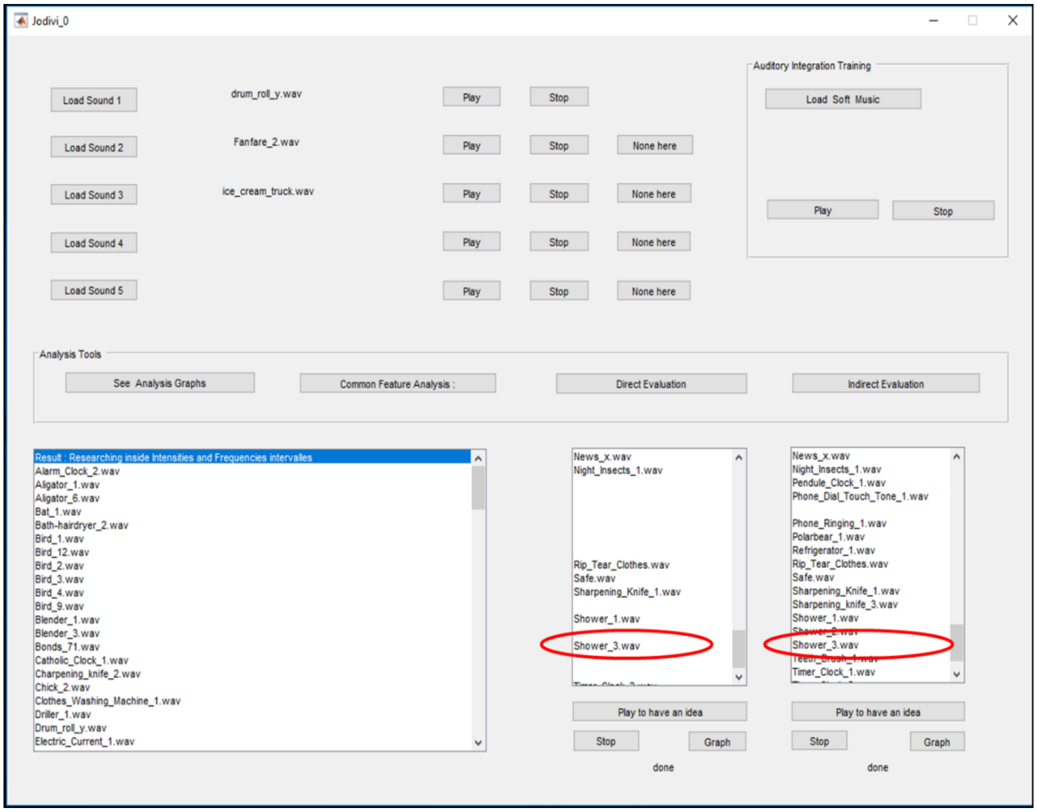
Figure 24: Simulation 1 with 3 sound files uploaded

Simulation for “extraction of sounds for preventive measure” was done with three wave files named «drom_roll_y.wav», «Ice_cream_truck.wav» and «Fanfare_2.wav». «Outside Intersect Interval» has been selecting more sounds than algorithm 1 «Inside all Intervals». A red circle indicates that the algorithm found an input file as output candidate.

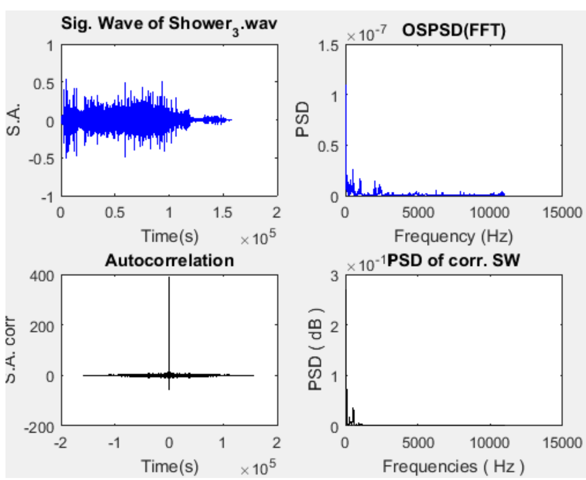
At this stage, one can see that algorithm 2, « Outside Intersect Interval » has been selecting more sounds than algorithm 1 « Inside all Intervals » (Figure 24 b); this is because algorithm 2 searches its candidates on a much bigger area.

There is also in the database, a sound named «Fanfare_3.vaw» which is also a Fanfare sound. This sound has been selected by both algorithms (show below as Fanfare_3 in Figure 24 b), which was expected as «Fanfare_2» and «Fanfare_3» sound very similar.

Let us take a look at a random sound that was selected by both algorithms: «Shower_3.wav» (Figure 25). Let us also remember that «Fanfare_2.wav», «Drum_roll_y.wav» and «Ice_Cream_Truck.wav» have more chance than «Shower_3.wav» to be selected by each of the algorithms. «Shower_3.wav» has mainly low frequencies, and its highest PSDs are located below 2500 Hz. This is true for 2 out of 3 sounds, namely «Drum_roll_y.wav» and «Ice_Cream_Truck.wav». When looking at the graph «Superpose» (Figure 24 a), which PSD scale is (10^{-6}) and looking at the frequency domain of the sound «Shower.wav» which PSD scale is (10^{-7}) (i.e. smaller), one can see that there is a net framing of «Shower.vaw» by all three sounds. Also, if all high frequencies due to the Fanfare sound could be removed, Fanfare would align perfectly with the other two sounds; «Drum_roll_y.wav» and «Ice_Cream_Truck.wav». This same pattern can be seen on some selected sounds by all two algorithms.



(a)



(b)

<u>Figure 25:</u> View of a selected sound

In (a) we randomly selected the sound « Shower3.wav ». (b) At the top left, the signal; at the top right, the FFT of that signal; at the bottom left, the autocorrelation of the signal.

We performed another simulation in which we removed the file «Ice_cream_truck.wav» (Figure 26). For this simulation, both algorithms selected «Drum_roll_y.wav» and «Fanfare_2.wav».

One can notice fewer outputs for method 1 and 2 « Inside All Intervals » and «Outside Intersect Intervals». For algorithm 1, this depends on the position of the removed sounds as related to the others; the more sounds exist, outputs will be generated, but if an added sound has its frame located in-between, there are no changes. In the current case, removing one sound has reduced the interval, which has led to fewer outputs. For algorithm 2, the larger the common area is, the fewer outputs there will be, and if the removed sounds were surrounding the common area, there will be no changes. In the current case, removing one sound has led to the enlargement of the common area, which leads to fewer outputs. In fact, in an intersect condition, the fewer uploaded sounds, the more there will be matches between them.

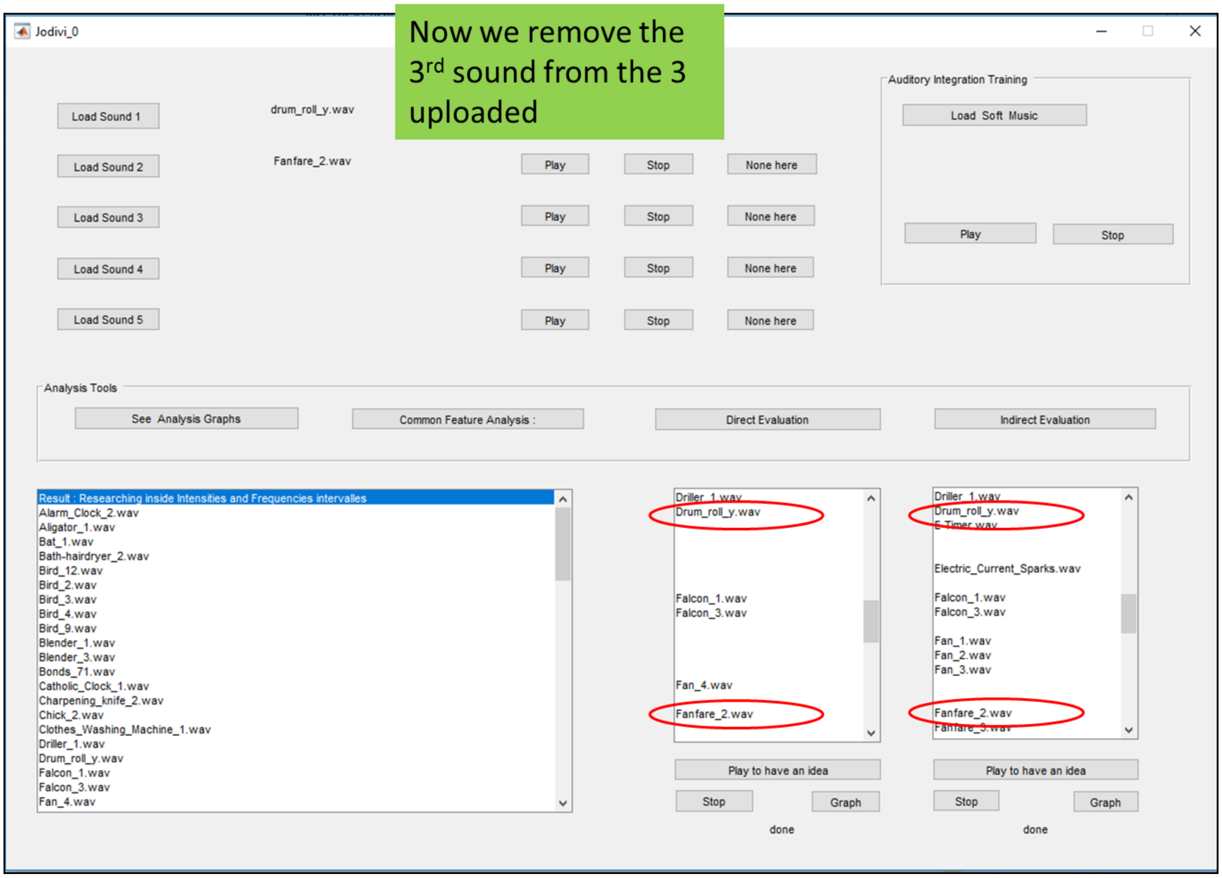
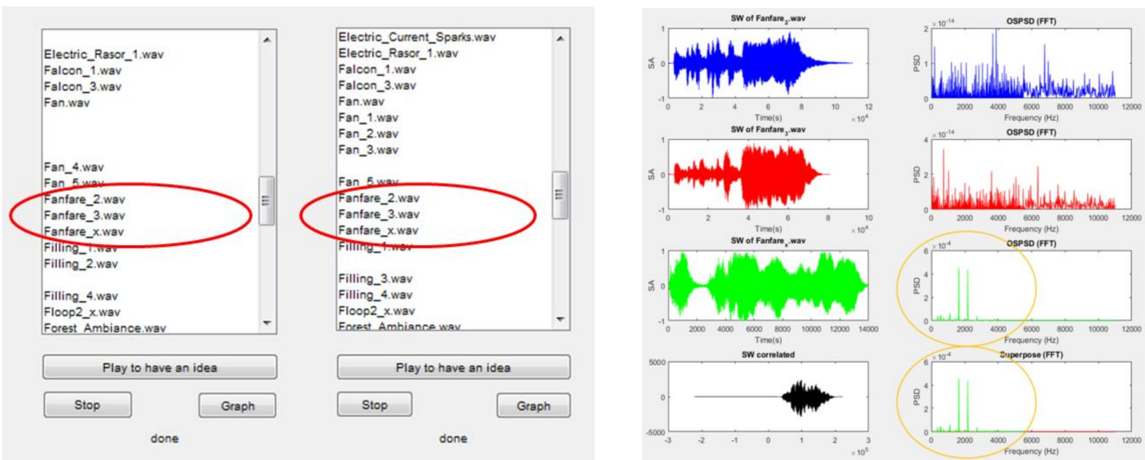


Figure 26: Impact of adding or removing one sound from the ones previously uploaded

From the previous simulation, we removed the file «Ice_cream_truck.wav». A red ellipse indicates that the algorithm found an input file as output candidate. Both algorithms selected «Drum_roll_y.wav» and «Fanfare_2.wav».

The following simulation was done with three sounds of the same category whose names are respectively «Fanfare_2», «Fanfare_3» and «Fanfare_x» (Figure 27). They are all sounds of Fanfare; it is the same musical instrument that played 3 different tones. On Figure 27 (b), one can see a high similarity between «Fanfare_2» and «Fanfare_3». The graph « Superpose » at the bottom middle is almost identical to the FFT graph of the «Fanfare_x» sound, because of the PSD scale of the «Fanfare_x». When multiplying PSDs (illustration not shown here), the winner frequency zone is the one between 0 Hz and 5000 Hz, seen on the graph at the right side (Figure 27 b).

Both algorithms 1 and 2 selected «Fanfare_2», «Fanfare_3» and «Fanfare_x» (Figure 27 a). which is not quite the case with «Fanfare_x» because of the distribution of PSD (Figure 27 b). Once again, our algorithms provided all known true positive, «Fanfare_2», «Fanfare_3» and «Fanfare_x».



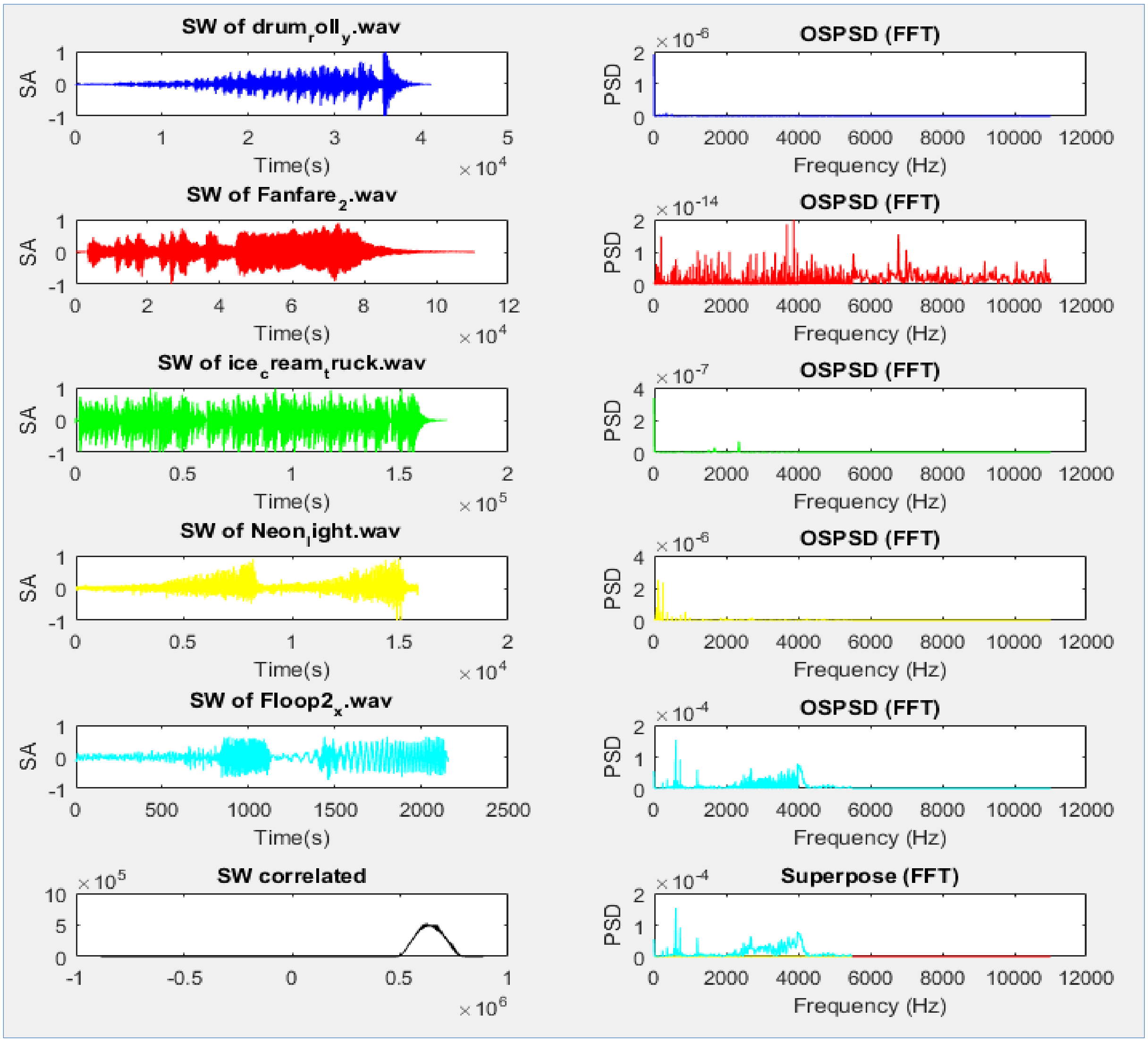
(a) (b)

<u>Figure 27:</u> Same Category of Sounds

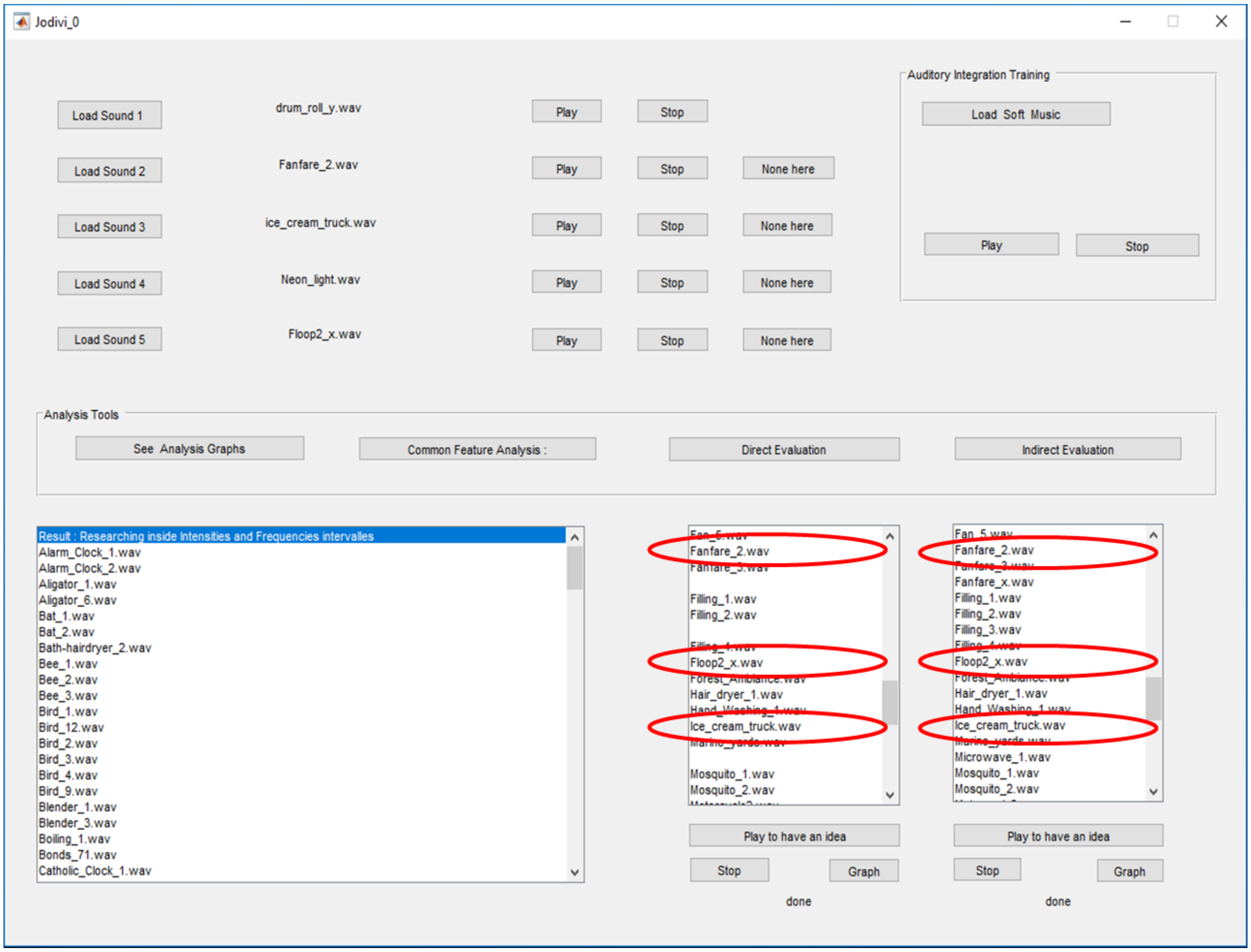
(a)The output from algorithm 1 and 2. «Fanfare_2», «Fanfare_3», «Fanfare_x» have been selected by both algorithms. More output are generated by algorithm 2 as compared to algorithm 1. (b) The graph shows that «Fanfare_x» has a significantly higher PSD, reason why it is predominant in the graph at the superposition.

We added two more sounds to the three previous sounds from the first run (Figure 28). The more we upload sounds, the more we approach a background noise (Figures 28 a, 28 b). A constant zero dB at the bottom right means that there is almost no PSD variation.

Algorithm 1 and 2 «Inside All Intervals» and «Outside Intersect Intervals» selected all five sounds.



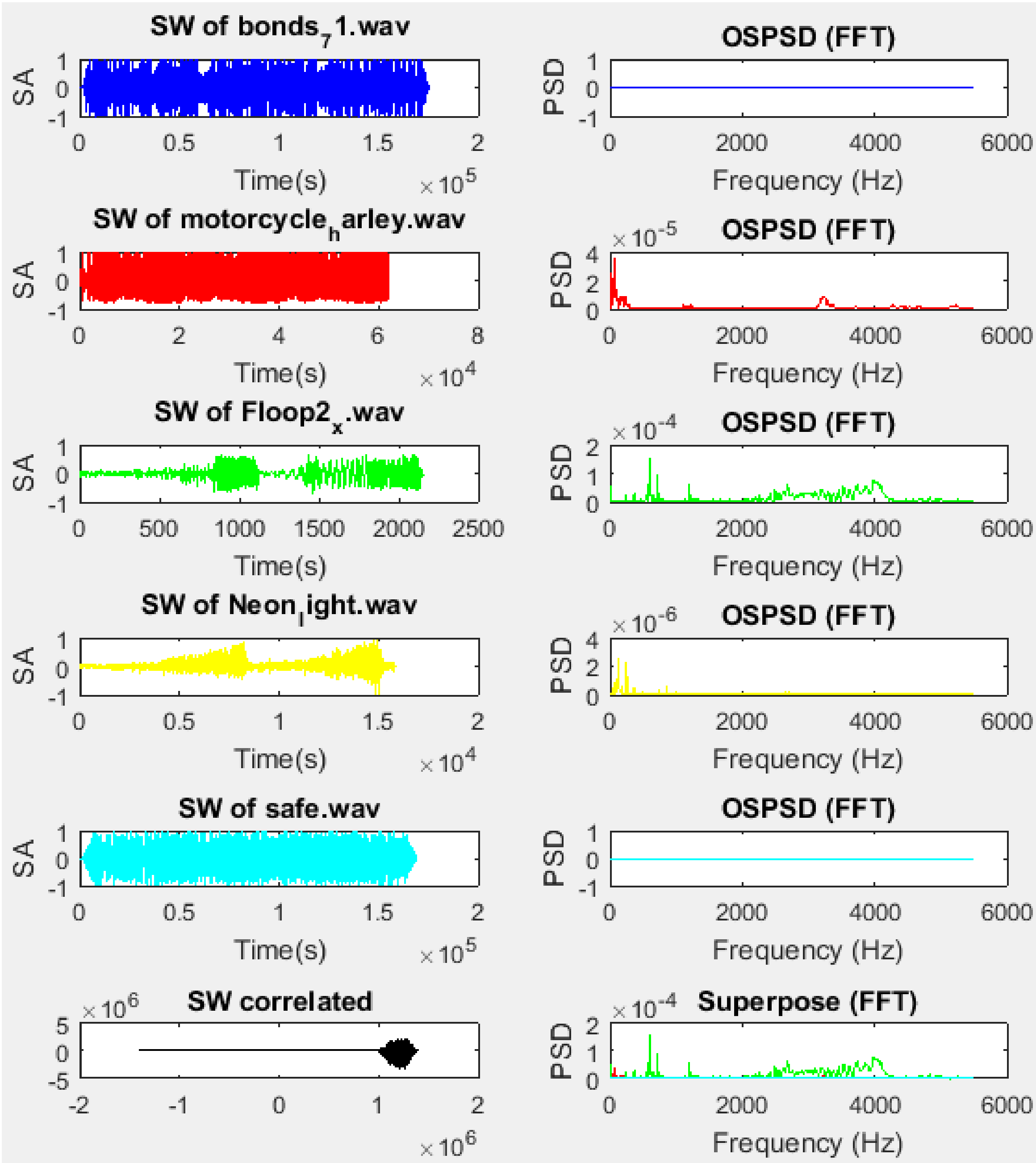
(a)



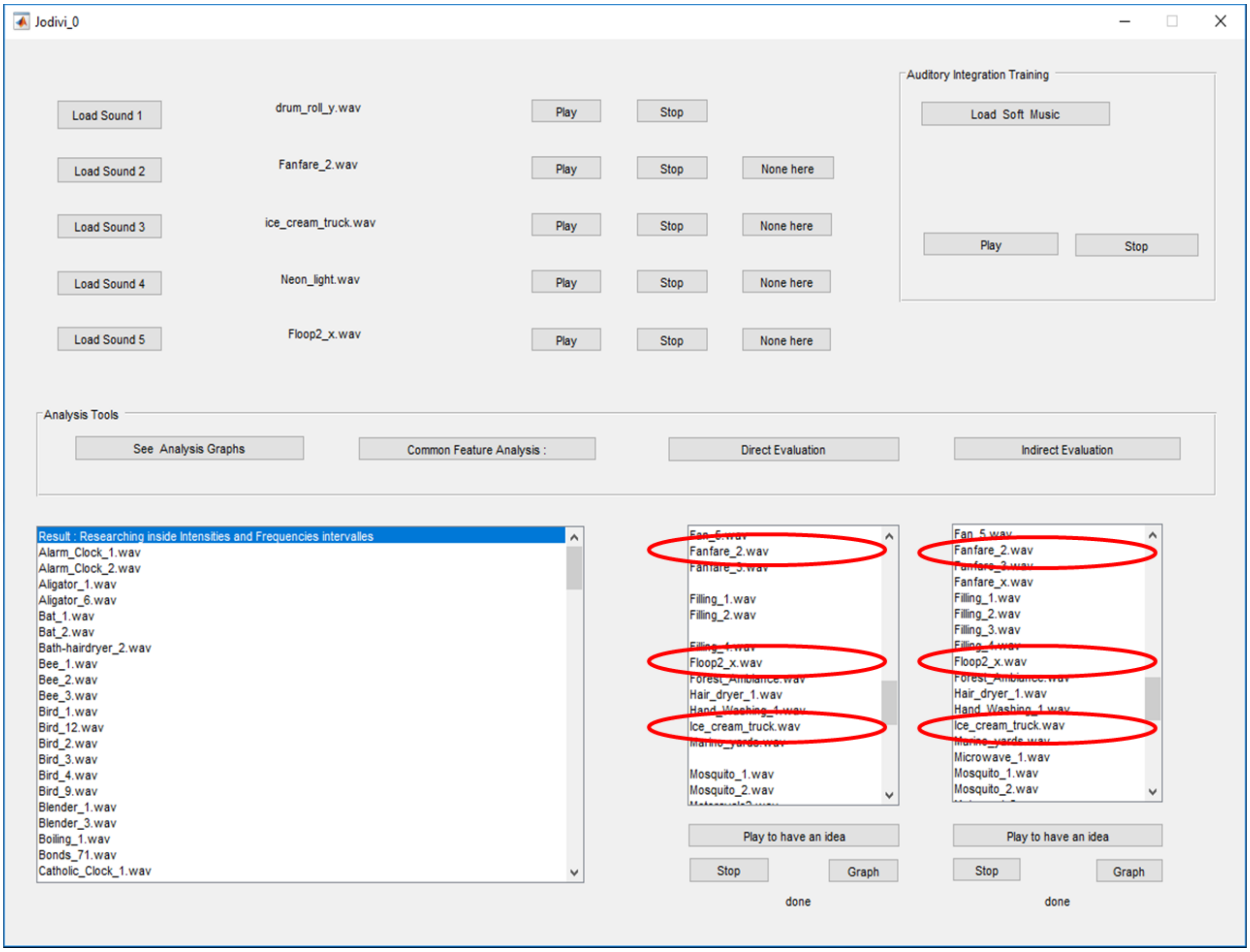
(b)

Figure 28: Simulation 1 for five sounds uploaded

Five sounds have been uploaded for this simulation. (a) Shows the graphical view of the sounds. (b) A red circle indicates that the algorithm found an input file as output candidate. The two others sounds are hidden by the current view of the Listboxs.


SW of bonds7 1.wav
OSPSD (FFT)
SW of motorcycle h arley.wav
SW of Floop2 x .wav
SW of Neon l ight.wav
SW of safe.wav
SW correlated
Superpose (FFT)
SA
PSD
Time(s)
Frequency (Hz)

(a)



(b)

Figure 29: Simulation 2 for five sounds uploaded

Five sounds have been uploaded for this simulation. (a) Shows the graphical view of the sounds. (b) A red ellipse indicates that the algorithm found an input file as output candidate. The two others sounds are hidden by the current view of the Listboxs.

Unless we perform the test with real users, one cannot state with 100% certitude which other sound selected by both algorithm (1 and 2) is a true positive or false positive. But based only on the generated actual quality of the results, we can state with confidence that the real true positives will be those selected by both algorithms (See appendix for another example).

Table 2: Advantages and Disadvantages of each of the two Methods after test results

Method	Advantages	Disadvantages
Inside all intervals	Always generates a result so that the user has an idea of what he is looking for. It is the best choice when the sounds are very different.	Covers only sounds inside the intervals May give a false positive, as it is just a framing.
Outside intersect intervals.	Much more specific than the first in terms of searching for common area between sounds. Covers all possible sounds from the database.	Gives many results, so it may introduce false positives. Hard to select complex sounds, or sounds having short length.
Conclusion	The best outputs are those which are selected by both algorithms.	

Note that two people with Autism that upload the same files in input will have a significant chance to get the same outputs; that would mean that it is the same case of sensitivity. There will be difference only if at least one of the uploaded sounds is different.

For most of the simulations we performed in this work, results were close to our attendance: when only low frequency sounds were uploaded in the input, we got in the output mainly low frequency sounds in synthesis (sounds that are recognized in both algorithms). The same was observed when we uploaded only high frequency sounds. The output was consisted mostly in high frequency sounds.

- Statistic results of Common intervals

As mentioned earlier, our purpose was to extract common frequencies and intensities from the uploaded sounds. For each method, when selecting sounds from the database, the application revels in output, the range of frequency and intensity of selected sounds. This information may not make sense for some user unfamiliar with frequency and intensity, but it could be useful for doctors and medical professionals.

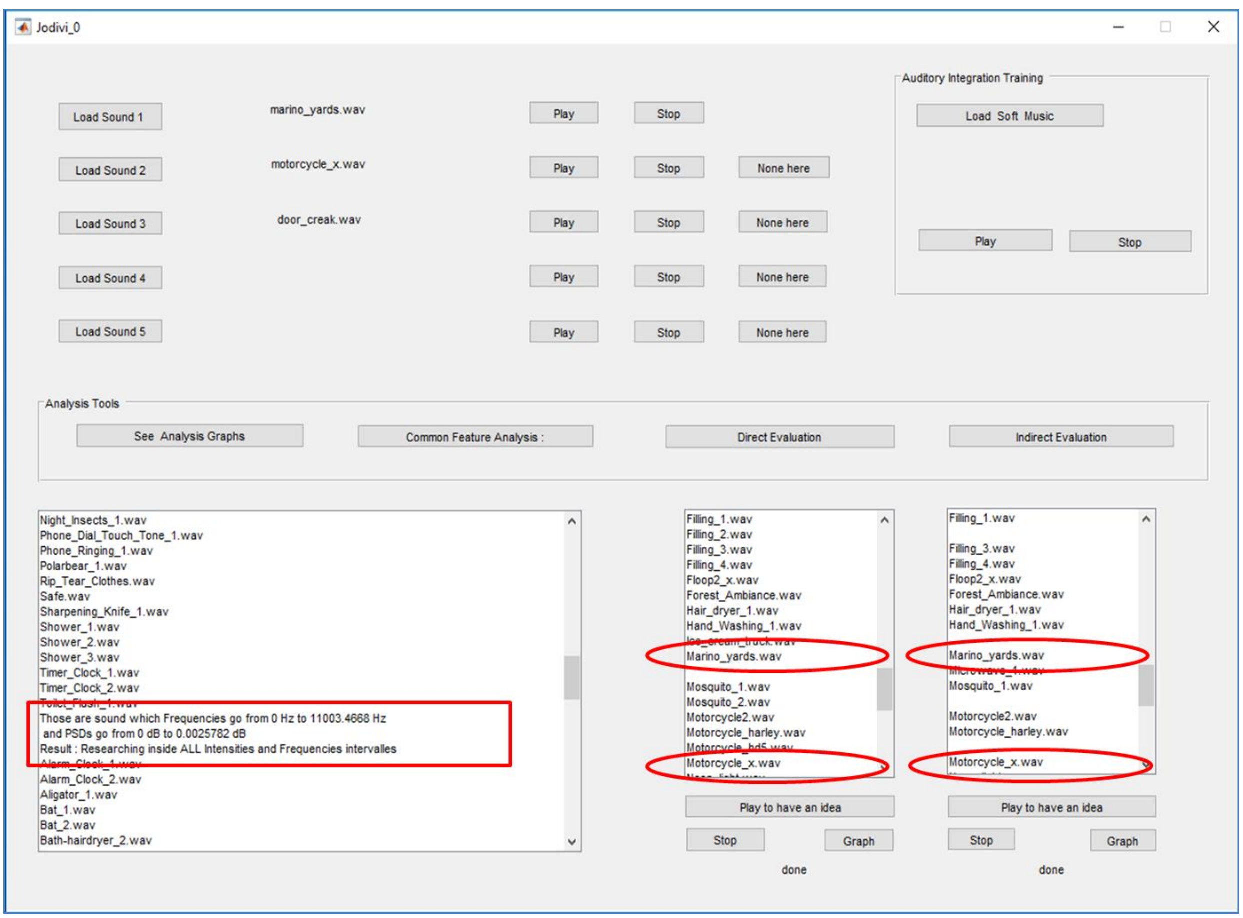


Figure 30: Intervals values of Frequencies and PSDs.

In this example, the wave files «Mario_yards», «Motorcycles_x» and «door_creak» have been uploaded. The red rectangle on the left side is the frame delimiting the lowest frequency 0 Hz, the higher frequency 11, 003.48 Hz, the lowest PSD 0 dB and highest PSD of 0.0025782 dB for all 3 uploaded sounds. These values may not be useful for users, as user want to know what does this mean in practice directly by reading the 2 listboxs.

4.3 Results for Sensitivity Evaluation

4.3.1 Direct Evaluation: sensitivity analysis using available information

All action performed on « direct evaluation » is registered in the application (Figure 31). The dates and time of creation, insertion and deletion are stored. It is also possible to extract stored improvement information for a desired sound.

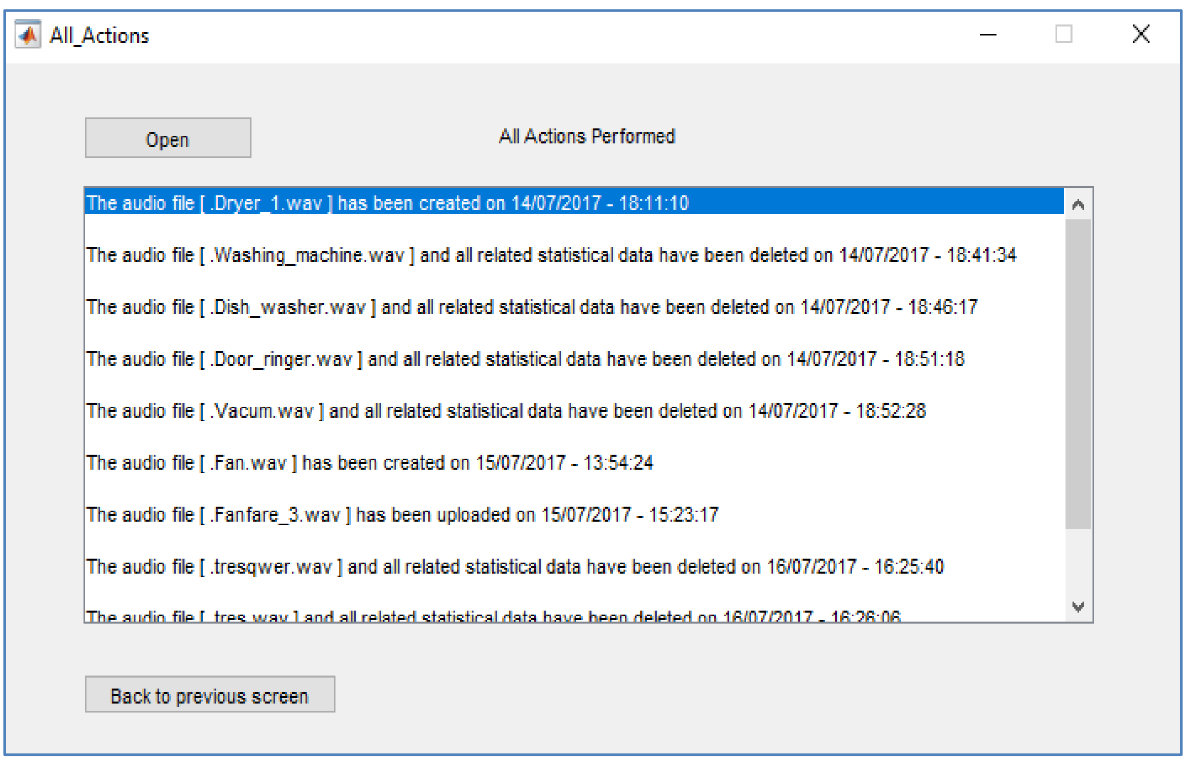
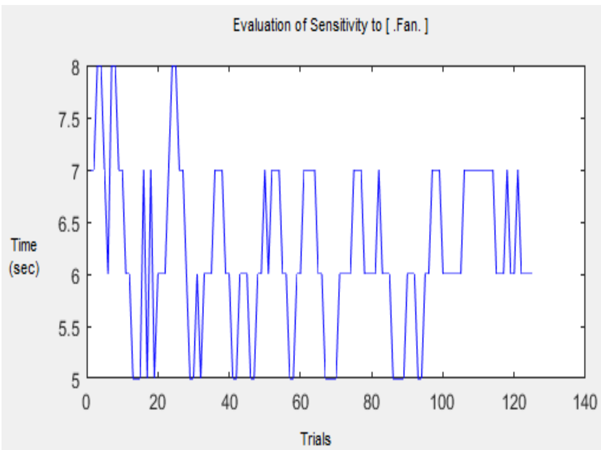


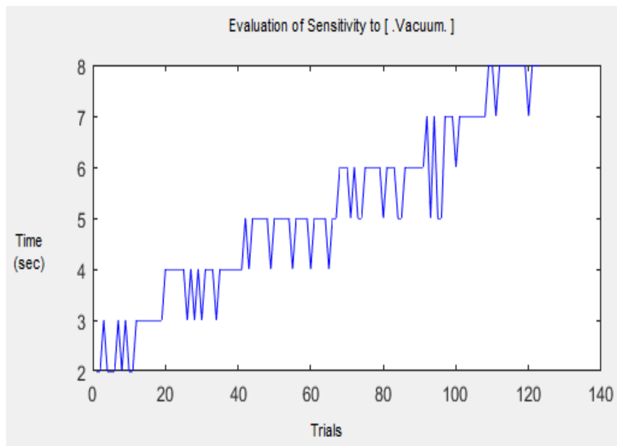
Figure 31: All Actions Performed on Direct Evaluation

All actions performed are stored in the application. Creations, insertions, and deletions of sound files including the related date and time.

Figure 32 shows informations on the improvement for the sounds of «Fan» and «Vaccum». In those simulation examples, the sensitivity level keeps increasing and decreasing for the sound of «Fan», which means that there is no sensitivity improvement to the sound of «Fan» (Figure 13 a). As contrast, there is a notable decrease of sensitivity to the sound of «Vaccum» (Figure 13 b).



(a)



(b)

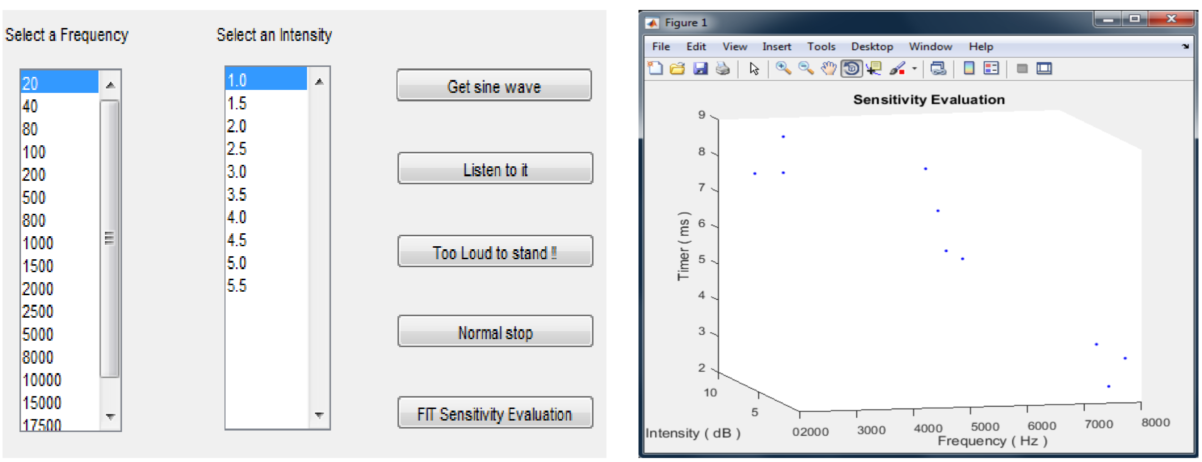
Figure 32: Direct Evaluation of Sensitivity to Fan and Vacuum.

(a) Evaluation of sensitivity to the sound of «Fan». The sensitivity goes up and down and stabilizes between 6 and 7, which indicates that there is no improvement for the sound of «Fan». (b) Evaluation of sensitivity to the sound of «Vaccum». The sensitivity curve goes up, which means that the user is becoming less and less sensitive to the sound of «Vaccum».

If the curve keeps increasing until reaching 10 sec, the button will remain unpressed then, and the curve will stay the same.

4.3.2 Indirect Evaluation: sensitivity analysis without information

Results of the indirect evaluation are shown in Figures 33 and 34. We have the indirect sensitivity with constant parameters (Figure 33), and the indirect sensitivity with increasing frequency over time (Figure 34). For the first case, we see a 3D graph in which x-axis is the frequency, y-axis is the intensity, and z-axis is elapsed time. The frequencies range from 20 to 20 000 Hz, the intensity ranges from 1 dB to 10 dB, and the time from 1 second to 10 seconds. This simulation, according to the presented axis shows a user sensitive to frequencies between 2000 and 5000 Hz and more sensitive to frequencies above 7000Hz. Rotations are possible to have best visualization on each axis.



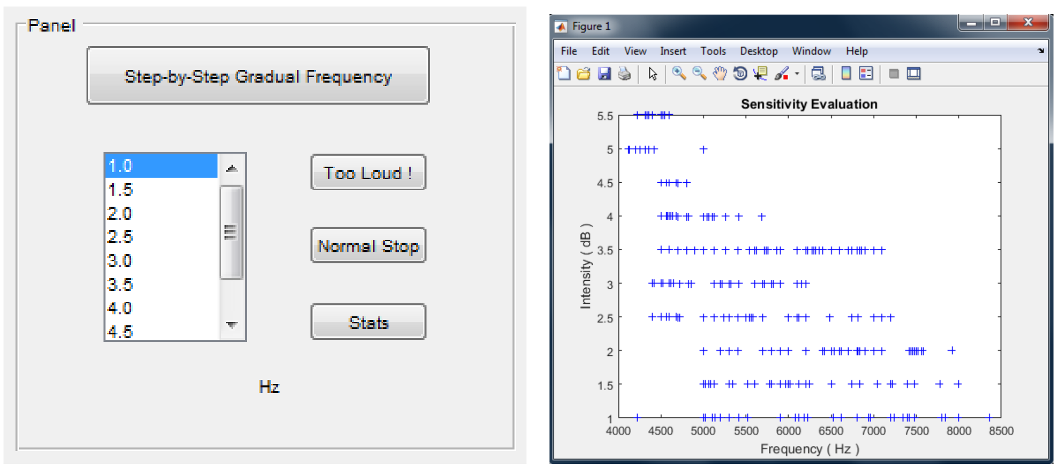
(a) (b)

<u>Figure 33</u>: Indirect evaluation with constant parameters

At the top (a) the frequency and intensity are selected by the user. The frequencies range from 20 to 20 000 Hz, intensities from 1 dB to 5.5 dB, and the time from 1sec to 10 seconds. In Figure 15 (b), the author performed simulations on himself. It shows that he is sensitive to frequencies between 2000 Hz to 8000 Hz, with a higher sensitivity between 7000Hz and 8000Hz.

One starts to lose the hearing sensation above 8000 – 10 000 Hz depending on age and hearing impairment. Children typically, can hear frequencies above 10 000 Hz.

The system is not giving any sound beyond 17500 Hz - 20000 Hz. The feeling at 20Hz is a pure vibration, while the feeling at 15000 Hz and more is very sharp (as if a needle was entering the ears). For the second case (Figure 34) a 2-dimensional graph is shown, whose x-axis represents the frequency and the y axis the intensity.



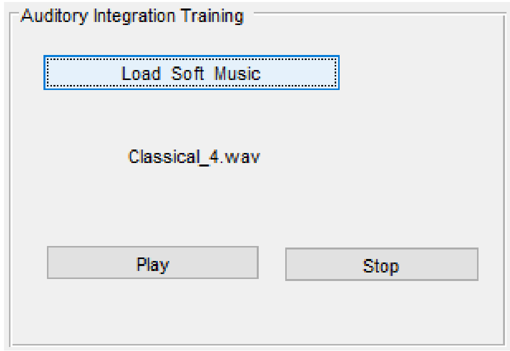
(a) (b)

Figure 34: Indirect Evaluation. Step-by-step Gradual Frequency Increase

Interface (a) allows targeting more specifically the right sensitivity frequency for a given intensity. (b) The simulation example shows peak sensitivity is between 5000Hz and 8000Hz for 1 to 2 dB, and between 4500 Hz and 7250 Hz for 2.5 dB to 3.5 dB.

4.4 Results for Auditory Integration Training (AIT)

As mentioned in Section 1 and 3, Auditory Integration Training (AIT) is an additional feature we added to the application. AIT is the current way to treat the impairment based on music therapy [10] [47] [48] [49] [50].



<u>Figure 35</u>: Classical music selected from AIT interface

As mentioned in Section 2, the technique employs multiple thirty minute sessions where the patients are exposed to personalized frequency-filtered music. The goal here is to use the designed music to re-train the ear and cause the patient to overcome sound sensitivity after multiple sessions of treatment [10] [11]. Regarding the subject, some text from [10] is announced below.

" THE LISTENING PROGRAM (TLP; ADVANCED BRAIN TECHNOLOGIES, OGDEN, UT, USA) is one of several SBIs used by paediatric occupational therapists to address behavioural disturbances related to sensory processing in children with ASD*. According to Advanced Brain Technologies (ABT, 2013), TLP uses psycho-acoustically modified classical music targeting certain frequency ranges that claim to impact functional capabilities including sensory processing, balance, learning, language, play and executive function. An individualized listening schedule is determined by the treating practitioner, usually with listening sessions occurring one to two times per day for 15 minutes per session for 5 days followed by a 2-day break. The listener listens to a series of 10 numbered CDs or tracks in numerical order over a 10-week period. After the first 10-week cycle, the listener reverses the order of the CD – starting with CD#10 and ending with CD#1 for a second cycle. A minimum of two cycles of listening is recommended by the manufacturer for the client to demonstrate moderate improvement (ABT, 2009). TLP uses a specialized set of headphones and a CD or MP3 player. Equipment costs for TLP run between $300.00–2,000.00 (ABT, 2009), yet TLP is typically not reimbursed through* *public or private insurance (Gee, et al., 2013). Therapists who wish to become an authorized provider must complete a provider training through ABT (2013) ".* [10].

Data sources could also include varied sources as non-standardised or standardised instruments, parent or teacher report, school records, as long as the source can be helpful for the training [49]. Gold et al. [49] listed the type of output measures for AIT that are presented in Table 3.

Table 3: Type of output measures (data from **[49]**)

Type of output measures	
Communicative and social skills, social interaction,	Hyperacusis (hypersensitivity to sound)
Quality of social interaction	Activity level
Behavioural problems (e.g. stereotypic behaviour)	Quality of life in both school and home environments
Attention and concentration	Stress in the family
Cognitive ability	Adverse events

Music therapy for individuals with ASD is usually provided as individual therapy [49] as they experience different level of sensitivity.

5. Conclusion

5.1 Research contribution

We designed a preventive application that detects commonalities between several sounds bothering ASD subjects in order to find similar sounds diagnostics and to use them for therapy purposes. For a known bothersome sound, parents would like to know the evolution of the sensitivity of their child related to that sound. For that, we designed a component in the application that evaluates the sensitivity improvement within the time that we named « direct evaluation ». For parents, or specialists, or therapist that want to go further in more details dealing with sound parameters and/or dealing with numbers, we designed a component in the application that evaluates the sensitivity to given (varied) frequencies and intensities, that we named « indirect evaluation ». Auditory Integration Training is a component we integrated as complementary to the two evaluation processes. The AIT will help to fix the sensitivity, while the evaluations will allow to see the improvement.

The sensitivity evaluation is time, frequency and intensity based algorithm computed in Java and Matlab whereas the common point filtering is based on a Fast Fourier Transform algorithm computed with Matlab. For the time sensitivity evaluation, we evaluate the listening time capacity to a given audio file. For the frequency and intensity sensitivity evaluation, we perform two types of evaluations: constant and gradual. For the constant evaluation, we choose at random a frequency and intensity, and apply the resulting sound from those two parameters. From that, the application records the time delay hearing capacity. For the gradual evaluation, the application increases the frequency for a chosen intensity, from low to high, and records the frequency and intensity at the stop point. Concerning the common point filtering, we used the Fast Fourier Transform algorithm that represents frequencies and intensities of a given signal, and implemented two complementary methods based on intervals to find common features between different uploaded sounds.

We developed a computer tool accessible to the patients's parents that will allow doctors and medical staff to better monitor their patients. Before this application, there was no algorithm to find a solution to the problem unless we use a database after the pre-treatment of targeted sounds. This prototype is then, to the best of our knowledge, the only existing tool of its kind today. However, there are a lot of smartphone applications that search for a given song, other songs that are similar in order to play to same kind of song in loop. Yet, they have nothing to do with scientific evaluation, and treatment.

5.2 Validation Tests

Given that we have not done this work on real subjects, the main component of our work « Analysis of common points » would have limited acceptance. We therefore cannot say with absolute certainty what would be the rate of accuracy of this component. However, based on the different results obtained, during our many simulations, apart from a few false positives which we know they occurred, we can affirm that the results are promising.

The other components of the application are the sensitivity and adaptability tests for which it is not necessary to establish the statistic at results.

6. Future research direction

6.1 Application Improvement

To improve this application, an immediate next step is to introduce more advanced "Digital Signal Processing" techniques to perform the processing of different signal. That would allow us for example to remove background noise, some frequencies or intensities, that are too high or too low, from uploaded sounds files. This should be done very carefully because it is not necessarily a good idea to clean the signal before extracting common points, as some features that bother children might be related to "rigging" or "hissing" that is part of background noise and thus, by removing them impact negatively the ability of the tool. It will be great to integrate some pattern search capability, in the time domain of the sound, using artificial intelligence algorithms. We have already implemented some functions to do some cleaning, but it was uncertain which limit to set to avoid removing significant PSDs.

The methods used need to be adaptive to the user; methods should be able to detect pattern for each user. This can be done with the use of artificial intelligence. As we mentioned earlier in the literature review, artificial intelligence algorithms have been used for music retrieval. Then, it is something doable.

We can use the « Findpeaks » function to get a vector for each sound and perform the inside and outside interval algorithms proposed in this work. This will correspond better to what we are looking for, instead of using any algorithm to retrieve a similar song file.

We mentioned that the user can interpret the indirect evaluation results to name known sounds. For now the interpretation is purely visual. It may be something difficult to do it visually. A future work will be to implement the conversion directly through the application based on « Findpeaks »: The Indirect evaluation provides specific frequencies and intensities. Each sound from our internal database has also been decomposed to get frequencies and PSDs. We can then find the mean peaks. Now we just have to find sounds in the database that have those peaks.

The Auditory Integration Training we proposed at the end of the Chapter 4 is generic; we just included a list of soft music, such us classical music, with no real participation of any therapist or specialist. A good improvement will be to get such expert involved to provide us exactly the type of soft music they use in their practice, and protocols for implementation.

6.2 Complementary work

Regarding Auditory Integration Training, every child has his/her own sensitivity pattern. A significant improvement might be to build the AIT specifically for each sensitivity pattern. If this is possible, that will be much more efficient than any AIT music module currently used. In addition, one can design a device to filter frequencies and intensities in real time, to clean listened sound.

References

[1] K. Stigler, B. McDonald, A. Anand, A. Saykin and C. McDougle, "Structural and functional magnetic resonance imaging of autism spectrum disorders," *Brain Research,* vol. 1380, no. 1, pp. 146-161, 2011.

[2] H. Brentani, C. Paula, D. Bordini, D. Rolim, F. Sato, J. Portolese, M. Pacifico and J. McCracken, "Autism spectrum disorders: an overview on diagnosis and treatment," *Revista Brasileira Psiquiatria,* vol. 35, no. 1, pp. 62-72, 2013.

[3] L. T. Eyler, K. Pierce and E. and Courchesne, "A failure of left temporal cortex to specialize for language is an early emerging and fundamental property of autism," *Brain: A journal of Neurology,* vol. 135, no. 1, p. 949–960, 2012.

[4] B. Haesen, B. Bart and W. Johan, "A review of behavioural and electrophysiological studies on auditory processing and speech perception in autism spectrum disorders," *Research in Autism Spectrum Disorders,* vol. 5, no. 1, pp. 701-711, 2011.

[5] E. Courchesne, K. Campbell and S. Solso, "Brain growth across the life span in autism: age-specific changes in anatomical pathology," *Brain Research,* vol. 1, no. 1, pp. 138-145, 22 March 2011.

[6] K. Pierce, S. Glatt, G. Liptak and L. McIntyre, "The power and promise of identifying autism early: insights from the search for clinical and biological markers," *Ann Clin Psychiatry,* vol. 21, no. 3, pp. 132-47, 2009.

[7] R. Morris, *Managing Sound Sensitivity in Autism Spectrum Disorder: New Technologies for Customized Intervention,* MASSACHUSETTS: Massachusetts Institute of Technology, 2009.

[8] S. Khalfa, N. Bruneau, B. Rogé and N. Georgieff, "Increased perception of loudness in autism," *Hearing Research,* vol. 198, p. 87–92, 2004.

[9] E. Gomes, F. S. Pedroso and W. M., "Auditory hypersensitivity in the autistic spectrum disorder," *Pró-Fono Revista de Atualização Científica,* vol. 20, no. 4, pp. 279-284, 2008.

[10] B. Gee, K. Thompson and S. John H, "Efficacy of a sound-based intervention with a child with an autism spectrum disorder and auditory sensory over-responsivity," *Occupational Therapy International,* vol. 21, no. 1, pp. 12-20, March 2014.

[11] S. M. Edelson, D. Arin, M. Bauman, S. E. Lukas, J. H. Rudy, M. Sholar and B. Rimland, "Auditory Integration Training: A Double-Blind Study of Behavioral and Electrophysiological Effects in People with Autism," *Focus on Autism and Other Developmental Disabilities,* vol. 4, no. 2, pp. 73-81, 1999.

[12] E. Marco, L. Hinkley, S. Hill and S. Nagarajan, "Sensory processing in autism: a review of neurophysiologic findings," *Pediatr Res,* vol. 69, no. 5, pp. 48-54, 2011.

[13] M. Iwen, A. Gilbert and M. and Strauss, "Empirical evaluation of a sub-linear time sparse DFT algorithm," *Commun. Math. Sci.,* vol. 5, no. 4, pp. 981-998, 2007.

[14] J. Blauert, Spatial Hearing: The Psychophysics of Human Sound Localization, MIT Press, 1997.

[15] J. Carayol, G. Schellenberg, B. Dombroski, C. Amiet, B. Génin, K. Fontaine, F. Rousseau, C. Vazart, D. Cohen, T. Frazier, A. Hardan, G. Dawson and F. T. Rio, "A scoring strategy combining statistics and functional genomics supports a possible role for common polygenic variation in autism," *Frontiers in Genetics,* vol. 5, no. 33, pp. 1-9, 18 Feb 2014.

[16] M. Herbert, "Large brains in autism: the challenge of pervasive abnormality," *Neuroscientist,* vol. 11, no. 5, pp. 417-40, 2005.

[17] B. Baars and N. Gage, Cognition, Brain and Consciousness: An Introduction to Cognitive Neuroscience, Academic Press Chapter, 2007.

[18] M. Mosconi, H. Cody-Hazlett, M. Poe, G. Gerig, R. Gimpel-Smith and J. Piven, "Longitudinal study of amygdala volume and joint attention in 2 to 4-year-old children with autis," *Arch Gen Psychiatry,* vol. 66, no. 5, p. 509–516, 2009.

[19] K. Dalton, B. Nacewicz, A. Alexander and R. Davidson, "Gaze-fixation, brain activation, and amygdala volume in unaffected siblings of individuals with autism," *Society of Biological Psychiatry,* vol. 61, no. 4, pp. 512-520, 15 Feb 2007.

[20] H. C. Hazlett, M. Poe, G. Gerig, M. Styner, C. Chappell, R. G. Smith, C. Vachet and J. and Pive, "Early Brain Overgrowth in Autism Associated with an Increase in Cortical Surface Area Before Age 2," *Arch Gen Psychiatry,* vol. 68, no. 5, p. 467–476, May 2011.

[21] D. J. Greene, N. Colich, M. Iacoboni, E. Zaidel, S. Bookheimerd and M. and Dapretto, "Atypical Neural Networks for Social Orienting in Autism Spectrum Disorder," *Neuroimage,* vol. 56, no. 1, p. 354–362, 1 May 2011.

[22] G. Bird, C. Catmur, G. Silani, C. Frith and U. Frith, "Attention does not modulate neural responses to social stimuli in autism spectrum disorders.," *Neuroimage,* vol. 31, no. 4, pp. 1614-1624, 15 Jully 2006.

[23] B. Wang, T. Yan, J. Wu, K. Chen, S. Imajyo, S. Ohno and S. Kanazawa, "Regional neural response differences in the determination of faces or houses positioned in a wide visual field.," *PLoS One,* vol. 8, no. 8, p. e72728, 2013.

[24] A. Tosoni, R. Guidotti, D. Gratta, G. Committeri and C. Sestieri, "Preferential coding of eye/hand motor actions in the human ventral," *Neuropsychologia,* vol. 93, no. 116-127, 2016.

[25] J. Hughes, "A review of recent reports on autism: 1000 studies published in 2007," *Epilepsy Behav,*

vol. 13, no. 3, pp. 425-37, Oct 2008.

[26] S. Wass, "Distortions and disconnections: disrupted brain connectivity in autism.," *Brain and Cognition,* vol. 75, no. 1, pp. 18-28, 2011.

[27] P. Barttfeld, B. Wicker, S. Cukier, S. Navarta, S. Lew, R. Leiguarda and M. Sigman, "State-dependent changes of connectivity patterns and functional brain network topology in autism spectrum disorder.," *Neuropsychologia,* vol. 50, no. 14, pp. 3653-3662, Dec 2012.

[28] P. Barttfeld, B. Wicker, S. Cukier, S. Navarta, S. Lew and M. Sigman, "A big-world network in ASD: dynamical connectivity analysis reflects a deficit in long-range connections and an excess of short-range connections.," *Neuropsychologia,* vol. 49, no. 2, pp. 254-263, 2011.

[29] D. Ridder, S. Vanneste, N. Weisz, A. Londero, W. Schlee, A. Elgoyhen and B. Langguth, "Review: An integrative model of auditory phantom perception: Tinnitus as a unified percent of interacting separable subnetworks," *Neuroscience and Biobehavioral Reviews,* vol. 44, pp. 16-22, 2014.

[30] I. Dinstein, K. Pierce, L. Eyler, S. Solso, R. Malach, M. Behrmann and E. Courchesne, "Disrupted neural synchronization in toddlers with autism," *Neuron,* vol. 70, no. 6, pp. 1218-1225, 23 June 2011.

[31] K. Markram and H. Markram, "The intense world theory - a unifying theory of the neurobiology of autism," *Front Hum Neurosci,* p. 4:224, 2010.

[32] S. Noonan, F. Haist and R. Müller, "Aberrant functional connectivity in autism: evidence from low-frequency BOLD signal fluctuations," *Brain Res.,* vol. 1262, no. 1, pp. 48-63, 2009.

[33] P. Shih, M. Shen, B. Ottl, B. Keehn, M. Gaffrey and R. Müller, "Atypical network connectivity for imitation in autism spectrum disorder," *Neuropsychologia,* vol. 48, no. 10, pp. 2931-2939, 2010.

[34] L. Mottron, M. Dawson, I. Soulières, B. Hubert and J. Burack, "Enhanced perceptual functioning in autism: an update, and eight principles of autistic perception.," *Journal of Autism and Developmental Disorders,* vol. 36, no. 1, pp. 27-43, 2006.

[35] M. Brosnan, F. Scott, S. Fox and J. Pye, "Gestalt processing in autism: failure to process perceptual relationships and the implications for contextual understanding," *Journal of Child Psychology and Psychiatry,* vol. 45, no. 3, pp. 459-469, 2004.

[36] L. Mottron, M. Dawson and I. Soulières, "Enhanced perception in savant syndrome: patterns, structure and creativity.," *Philos Trans R Soc Lond B Biol Sci,* vol. 364, no. 1522, pp. 1385-91, 2009.

[37] N. Demonceau, A. Roubertie, R. Cheminal, J. Leydet, F. Rivier and B. Echenne, "Apport de la génétique à la compréhension de l'autisme," *Archive de Pediatrie,* vol. 12, no. 10, pp. 1509-1512, Oct 2005.

[38] A. Jabourian, J. Turpin, I. Maire and N. Baumann, "Autisme de l’enfant lié à une mucopolysaccharidose," *Annales Médico-psychologiques, revue psychiatrique,* vol. 160, no. 5-6, p.

421–426, 2002.

[39] X. Li, H. Zou and W. Brown, "Genes associated with autism spectrum disorder," *Brain Res Bull.,* vol. 88, no. 6, pp. 543-52, 1 Sept 2012.

[40] D. Smith and M. Ehlers, "Mining and modeling human genetics for autism therapeutics," *Current Opinion in Neurobiology,* vol. 22, no. 5, pp. 902-910, 2012.

[41] F. G. Peça J., "Cellular and synaptic network defects in autism," *Current Opinion Neurobiology,* vol. 22, no. 5, pp. 866-872, 2012.

[42] S. Uchino and C. Waga, "SHANK3 as an autism spectrum disorder-associated gene," *Brain and Development,* vol. 35, no. 2, pp. 106-110, 2013.

[43] S. E. Donohue, E. F. Darling and S. R. Mitroff, "Links between multisensory processing and autism," *Experimental Brain Research,* vol. 222, no. 4, pp. 377-387, Oct 2012.

[44] J. Alcántara, T. Cope, W. Cope and E. Weisblatt, "Auditory temporal-envelope processing in high-functioning children with Autism Spectrum Disorder.," *Neuropsychologia,* vol. 50, no. 7, pp. 1235-1251, 2012.

[45] C. Price, "A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading," *Neuroimage,* vol. 62, no. 2, pp. 816-47, 2012.

[46] G. Lai, S. P. Pantazatos, H. Schneide and J. and Hirsch, "Neural systems for speech and song in autism," *Brain,* vol. 135, no. 3, p. 961–975, 2012.

[47] N. Kraus and C. Bharath, "Music training for the development of auditory skills," *Nature Reviews Neuroscience,* vol. 11, no. 1, pp. 127-138, 2010.

[48] G. W. R. Dawson, "Interventions to facilitate auditory, visual, and motor integration in autism: a review of the evidence.," *Journal of Autism and Developmental Disorders,* vol. 30, no. 5, pp. 415-421, 2000.

[49] C. Gold, T. Wigram and C. Elefant, "Music therapy for autistic spectrum disorder.," *Cochrane Database of Systematic Reviews,* 2006.

[50] Y. Sinha, N. Silove, D. Wheeler and K. Williams, "Auditory integration training and other sound therapies for autism spectrum disorders: a systematic review.," *Archives of Disease in Childhood,* vol. 91, no. 12, pp. 1018-1022, 2006.

[51] O. Tchernichovski, F. Nottebohm, C. Ho, B. Pesaran and P. Mitra., "A procedure for an automated measurement of song similarity," *The Association for the Study of Animal Behaviour,* vol. 59, no. 0, 2000.

[52] M. Thomas, M. Jothish, N. Thomas, S. G, S. Koolagudi and Y. Murthy, "Detection of Similarity in Music Files using Signal," in *IEEE Region 10 Conference (TENCON) — International Conference*,

2016.

[53] Z. Ren, C. Fan and Y. Ming, "Music Retrieval Based on Rhythm Content and Dynamic Time Warping Method," in *IEEE 13th International Conference on Signal Processing (ICSP)*, 2016.

[54] "www.analog.com," [Online]. Available: ftp://ftp.analog.com/pub/cftl/ADI%20Classics/Mixed%20Signal%20and%20DSP%20Design%20Techniques,%202000/Section_5_Fast_Fourier_Transforms.pdf.

[55] A. Bhatara, Q. Eve-Marie, E. Fombonne and D. Levitin, "Early Sensitivity to Sound and Musical Preferences and Enjoyment in Adolescents With Autism Spectrum Disorders," *Psychomusicology: Music, Mind, and Brain,* vol. 23, no. 2, p. 100–108, 2013.

[56] http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD003681.pub3/full, "Auditory integration training and other sound therapies for autism spectrum disorders (ASD)," *Cochrane Developmental, Psychosocial and Learning Problems Group,* 2011.

[57] O. Perche, F. Laumonnier, L. Baala, M. Ardourel, A. Menuet, V. Robin, S. Mortaud, C. Montécot-dubourg, O. Richard, J. Pichon and S. Briault, "Autisme, génétique et anomalies de la fonction synaptique," *Pathologie Biologie,* vol. 58, no. 5, pp. 381-386, 2010.

[58] D. Geschwind and P. Levitt, "Autism spectrum disorders: developmental disconnection syndromes," *Curr Opin Neurobiol.,* vol. 17, no. 1, pp. 103-11, 1 Feb 2007.

[59] J. Isler, K. Martien, P. Grieve, R. Stark and M. Herbert, "Reduced functional connectivity in visual evoked potentials in children with autism spectrum disorder," *Clin Neurophysiol,* vol. 121, no. 12, pp. 2035-43, Dec 2010.

[60] A. Manzardo, R. Henkhaus, S. Dhillon and M. Butler, "Plasma cytokine levels in children with autistic disorder and unrelated siblings," *Int J Dev Neurosci,* vol. 30, no. 2, pp. 121-7, 2012.

[61] N. Gage, S. B. and T. Roberts, "Cortical auditory system maturational abnormalities in children with autism disorder: an MEG investigation," *Brain Research. Developmental Brain Research,* vol. 144, no. 2, pp. 201-209, 10 Sept 2003.

[62] Y. Funabiki, T. Murai and M. Toichi, "Cortical activation during attention to sound in autism spectrum disorders," *Research in Developmental Disabilities,* vol. 32, no. 2, pp. 518-524, Mar-Apr 2012.

[63] J. Feng, R. Schroer, J. Yan, W. Song, C. Yang, A. Bockholt, E. Cook, C. Skinner, C. Schwartz and S. Sommer, "High frequency of neurexin 1beta signal peptide structural variants in patients with autism," *Neuroscience Letters,* vol. 409, no. 1, pp. 10-13, 26 Nov 2006.

[64] F. Tecchio, F. Benassi, F. Zappasodi, L. Gialloreti, M. Palermo, S. Seri and P. Rossini, "Auditory sensory processing in autism: a magnetoencephalographic study," *Biological Psychiatry,* vol. 54, no. 6, pp. 647-654, 2003.

[65] H. Wei, K. Chadman, D. McCloskey, A. Sheikh, M. Malik, W. Brown and X. Li, "Brain IL-6 elevation causes neuronal circuitry imbalances and mediates autism-like behaviors," *Biochimica et Biophysica Acta,* vol. 1822, no. 6, pp. 831-842, 2012.

[66] D. Wynne, F. Zeng, S. Bhatt, H. Michalewski, A. Dimitrijevic and A. Starr, "Loudness adaptation accompanying ribbon synapse and auditory nerve disorders.," *Brain,* vol. 136, no. 5, pp. 1626-1638, 2013.

[67] Y. Yoshimura, M. Kikuchi, K. Shitamichi, S. Ueno, T. Munesue, Y. Ono, T. Tsubokawa, Y. Haruta, M. Oi, Y. Niida, G. Remijn, T. Takahashi, M. Suzuki, H. Higashida and Y. Minabe, "Atypical brain lateralisation in the auditory cortex and language performance in 3- to 7-year-old children with high-functioning autism spectrum disorder: a child-customised magnetoencephalography (MEG) study.," *Molecular Autism,* vol. 4, no. 1, 2013.

Web references

http://www.analog.com/media/en/training-seminars/design-handbooks/MixedSignal_Sect5.pdf

Last visit on 30/06/2017

http://www.mathworks.com/help/signal/ref/xcorr.html

Last visit on 30/06/2017

http://www.investopedia.com/terms/c/crosscorrelation.asp#ixzz4J8GJ53O2

Last visit on 30/04/2016

http://www.mathworks.com/matlabcentral/fileexchange/45197-short-time-fourier-transformation--stft--with-matlab-implementation/content//example.m

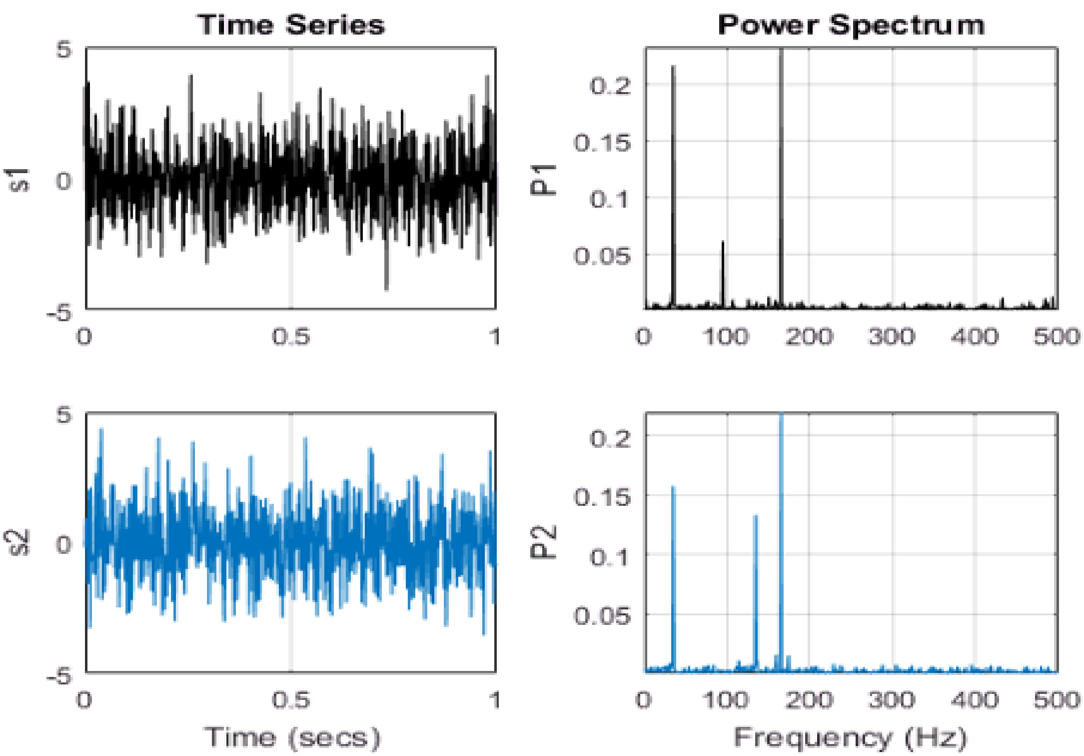
Last visit on 30/04/2016

Appendix

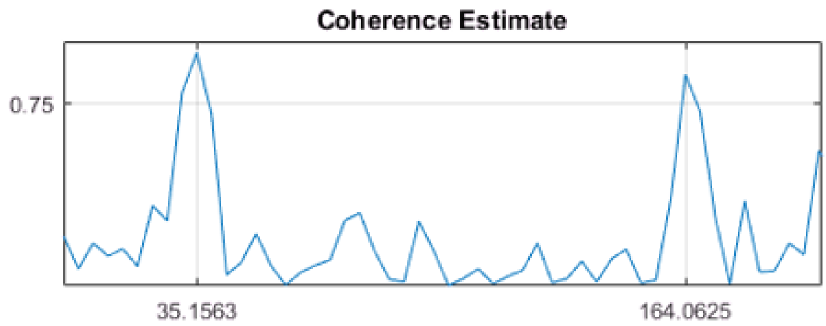
✓ Mathworks comparison of the frequency content of two signals

1) mscohere

A power spectrum displays the power present in each frequency. Consider two signals and their respective power spectra.



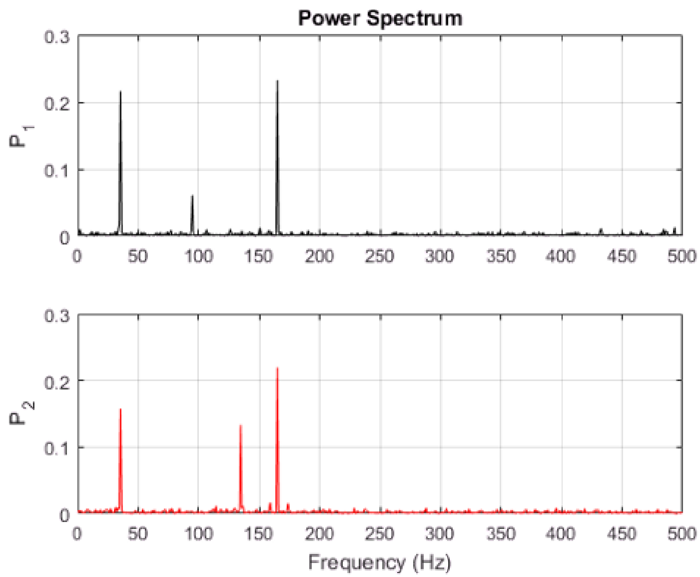
Spectral coherence identifies frequency-domain correlation between signals. Coherence values tending towards 0 indicate that the corresponding frequency components are uncorrelated while values tending towards 1 indicate that the corresponding frequency components are correlated. The mscohere function calculates the spectral coherence between the two signals. It confirms that sig1 and sig2 have two correlated components around 35 Hz and 165 Hz.



(mathwork.com)

2) findpeaks

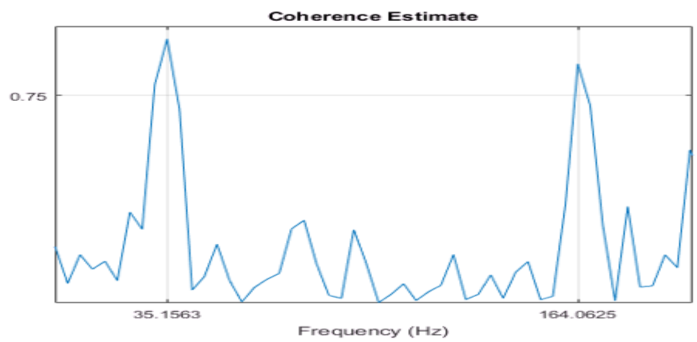
findpeaks allows finding the corresponding frequencies using.



In this example, each signal has three frequency components with significant energy. Two of those components appear to be shared. Find the corresponding frequencies using findpeaks.

```
[pk1,lc1] = findpeaks(P1,'SortStr','descend','NPeaks',3);

P1peakFreqs = f1(lc1)

P1peakFreqs =


  165.0391

   35.1562

   94.7266


[pk2,lc2] = findpeaks(P2,'SortStr','descend','NPeaks',3);
```

```
P2peakFreqs = f2(lc2)
P2peakFreqs =

  165.0391
   35.1562
  134.7656
[Cxy,f] = mscohere(sig1,sig2,[],[],[],Fs);

thresh = 0.75;
[pks,locs] = findpeaks(Cxy,'MinPeakHeight',thresh);
MatchingFreqs = f(locs)
MatchingFreqs =

   35.1562
  164.0625
```



We end up with the same output us it can be seen in the figure above.

✓ Sound Database

FindSounds.com which link is http://www.findsounds.com/types.html is a free site for finding sound effects on the Web. It is a Web search engine, like Google and Yahoo, but with a focus on sounds. It provides powerful features, yet is simple and easy to use, and suitable for all ages. Note to parents: audio files containing obscenities are filtered out so this site is safe for children. To learn how to search the Web using **FindSounds.com**, visit the Help page.

✓ Gaussian Wave

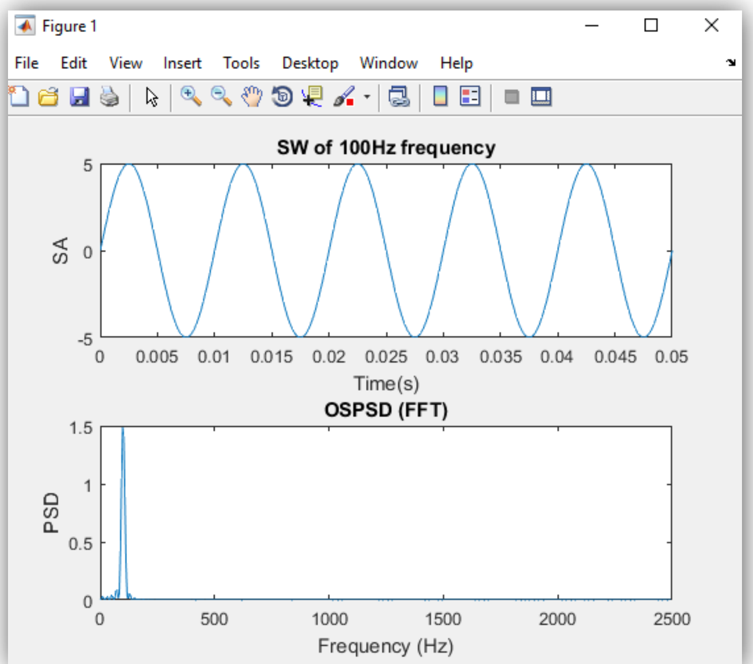
```
function GaussianWave( frequency, intensity, T )
%----------------------------------------------------%
close all
clear variable
clc
%----------------------------------------------------%
% http://www.gaussianwaves.com/2014/07/how-to-plot-fft-using-matlab-fft-of-basic-signals-sine-and-cosine-waves/
%----------------------------------------------------%

overSampRate = F_frequency(frequency); %frequency of sine wave

% To generate a sine wave of the desired frequency f at    % those times the sampling rate
% must be far higher that the prescribed minimum required % % sampling rate which is at least twice the frequency as per % Nyquist Shannon Theorem..

% at the same time, the Over sampling rate should not be   % very high when the frequency is close to 20000 Hz

fs = overSampRate * frequency; %sampling frequency

figure
% Generate time values from 0 to T seconds at the desired rate.
t = 0:1/fs:T; %time base
x = intensity * sin (2 * pi * frequency * t); %replace with cos if a cosine wave is desired
subplot(2, 1, 1);
plot(t,x);
sound(x, fs);
title(['SW of ', num2str(frequency), 'Hz frequency']);
xlabel('Time(s)');
ylabel('SA');

% Power Spectrum – One-Sided frequencies Representing the given signal % in frequency domain is done via Fast Fourier Transform (FFT) which % % implements.
% Discrete Fourier Transform (DFT) in an efficient manner. Usually, power spectrum is desired for analysis in frequency domain. In a power spectrum, power of each frequency component of the given signal is plotted against their respective frequency.
% The command FFT(x,N) computes the \(N\)-point DFT. The number of points – \(N\) – in the DFT computation is taken as power of \(2\) for facilitating efficient computation with FFT. A value of \(N = 1024\) is chosen here. It can also be chosen as next power of 2 of the length of the signal.
% In this type of plot, the negative frequency part of x-axis is omitted.
% Only the FFT values corresponding to 0 to  N/2 sample points of N-point DFT are plotted. Correspondingly, the normalized frequency axis
% runs between 0 to 0.5. The absolute frequency (x-axis) runs from
% 0 to f_s/2.

L = length(x);
NFFT = 1024;
X = fft(x,NFFT);
Px = X.*conj(X)/(NFFT*L); % Power of each freq components
fVals = fs*(0:NFFT/2-1)/NFFT;
subplot(2, 1, 2);
plot(fVals, Px(1:NFFT/2) );
title('OSPSD (FFT)');
xlabel('Frequency (Hz)')
```

```
ylabel('PSD');
end
```

For more informations, please follow this link

http://www.gaussianwaves.com/2014/07/how-to-plot-fft-using-matlab-fft-of-basic-signals-sine-and-cosine-waves/

Here is an example GaussinWave (100, 5, 0.5)



✓ Short Time Fourier Transform

```
clear, clc, close all

% load a .wav file
[x, fs] = wavread('Fanfare.wav');   % get the samples of the .wav file
x = x(:, 1);                 % get the first channel
xmax = max(abs(x));                % find the maximum abs value
x = x/xmax;                  % scalling the signal

% define analysis parameters
xlen = length(x);                % length of the signal
wlen = 1024;                     % window length (recomended to be power of 2)
h = wlen/4;                      % hop size (recomended to be power of 2)
nfft = 4096;                     % number of fft points (recomended to be power of 2)

% define the coherent amplification of the window
K = sum(hamming(wlen, 'periodic'))/wlen;
```

```
% perform STFT
[s, f, t] = stft(x, wlen, h, nfft, fs);

% take the amplitude of fft(x) and scale it, so not to be a
% function of the length of the window and its coherent amplification
s = abs(s)/wlen/K;

% correction of the DC & Nyquist component
if rem(nfft, 2)                    % odd nfft excludes Nyquist point
   st(2:end, :) = s(2:end, :).*2;
else                               % even nfft includes Nyquist point
   s(2:end-1, :) = s(2:end-1, :).*2;
end

% convert amplitude spectrum to dB (min = -120 dB)
s = 20*log10(s + 1e-6);

% plot the spectrogram
figure(1)
imagesc(t, f, s)
set(gca,'YDir','normal')
set(gca, 'FontName', 'Times New Roman', 'FontSize', 14)
xlabel('Time, s')
ylabel('Frequency, Hz')
title('Amplitude spectrogram of the signal')

handl = colorbar;
set(handl, 'FontName', 'Times New Roman', 'FontSize', 14)
ylabel(handl, 'Magnitude, dB')
```

Code from:

http://www.mathworks.com/matlabcentral/fileexchange/45197-short-time-fourier-transformation--stft--with-matlab-implementation/content//example.m