The affect of hearing impairment and hearing aid on music induced emotional responses

Pedro Caetano, Marco Fiorini and Antoni Grabowski

Abstract—**The research explores an effect of hearing loss and hearing aid on music induced emotions, specifically happiness and sadness. Algorithms for hearing loss and hearing aid simulations were implemented and applied to four film score tracks. A listening test that compares listeners' emotional responses to selected music examples in simulating hearing loss and hearing loss with hearing aid was designed and performed on a group of 12 subjects. A dimensional model of music emotions was used to gather data from the participants through a questionnaire. Gathered data showed significant decreases of arousal and valence of perceived emotions, as well as enjoyment of presented music excerpts in hearing loss conditions, which confirmed a hypothesis that was derived from BRECVEMA taxonomy.**

I. Introduction

Hearing loss simulations and hearing aid devices have been a useful tool of research for acoustic and psychoacoustics in past literature [1] [2] [3]. Those can help us further understand the challenges associated with hearing loss and hearing aids designs and implementations. Studies can provide different effects of hearing loss scenarios on normal hearing listeners, as well as evaluating the development of complex audio processing algorithms and other compensatory signal processing schemes. However, their usage in emotional response studies appears to be limited. Despite vast amounts of research examining the influence of hearing loss on speech perception or musical enjoyment, comparatively little is known about its influence on music perception, as no standardized test exists to quantify music perception of hearing impaired persons in a clinically practical manner [4] [5].

The aim of this paper is to implement algorithms for hearing loss and hearing aid simulation, design a hearing test and compare normal hearing listeners' emotional responses to different music examples in various circumstances, including hearing loss and hearing aid - and compare those with non processed musical stimuli.

In doing so, the authors defined the following problem statement: *different levels of hearing loss should affect musical enjoyment, but how does it affect the way we perceive happiness and sadness?*

II. Problem Analysis

A. BRECVEMA taxonomy

The property of music that allows induction of emotion bothered (and perhaps still does) researchers in fields like philosophy, psychology and psychophysiology [6]. The causes of this feature are still a point of dispute. One of the widely used taxonomy of causes of emotional responses to music is a BRECVEMA framework [7], which will be used in the research to theorize a hypothesis. Juslin in this work describes seven mechanisms that produce music induced emotion:

- *Brain stem reflex* describes a response of a brain to the sudden, intense events that urgently require attention like sudden and unexpected loud musical events.
- *Rhythmic entrainment* refers to an influence of strong rhythmic characteristic of music affecting internal rhythm of the listener.
- *Evaluating conditioning* characterises the mechanism that the emotion emerges caused by a element of music that was paired in the past with a specific event in one's life that caused the emotion.
- *Emotional contagion* is a kind of empathetic response to music cue as if it was a voice-like element that expresses curtain emotion.
- *Visual imagery* explains the set of emotion caused by spontaneous mental visual imagery that occurs in listeners mind during listening to music.
- *Episodic memory* Juslin illustrates it as "honey, they're playing our song" phenomenon, in which curtain tune is associated with a specific event in listener's past.
- *Musical expectancy* explains the process in which listener expects curtain event to occur in a piece of music and then the emotion is produced based on whether the expectancy was confirmed, violates or delayed.
- *Aesthetic judgment* refers to evaluating a musical piece in perspective of beauty, complexity novelty etc. and the satisfaction or lack of it related to this.

This framework will help to theorize the changes of emotional response in hearing impaired people, as well as hearing aid users.

B. Hearing loss

It is crucial to characterize changes in hearing loss from sound processing perspective to speculate how they might affect the emotional responses based on framework described in II-A. Mourgela et al. [1] investigate the topic in their development real-time hearing loss simulation plug-in. They list the most significant characteristics: high frequency attenuation, spectral smearing, rapid loudness growth and temporal disruption.

Temporal smearing is characteristic to cochlear hearing loss and effects peoples ability to categorize temporal information [8], [9] in musical pieces. Mourgela et al. [1] represents this phenomenon as manipulation of phase in the lower frequency band. With this in mind, it can be speculated, when referred to BRECVEMA framework, that the emotional response will be altered, based on rhythmic entrainment criterion; if the subject cannot accurately extract temporal information from the piece of music, the emotion associated with this cue will not be as strong as it would be in normal hearing and accurate extraction.

Melodic element is part of the core of musical information. A big part of it is contained in the higher pitch range, including high pitch solo instruments like violins, trumpets etc. as well as human voice, particularly with female singers. Therefore it is crucial for preserving this information by having a good representation of high frequencies. This is not possible in cases of hearing loss; high frequency attenuation is a significant problem in hearing loss. In that case characteristics described by many elements of BRECVEMA like evaluating conditioning, musical expectancy, but also episodic memory and visual imagery will be affected.

C. Hearing aids

Despite great technological complexity, hearing aids do not restore hearing loss to the point of normal hearing. Moreover, they are not designed for music listening purposes but rather to allow information extraction from speech, as described in Vaisberg et al. 2019 [10].

Due to limited power of the device the sampling frequency is only $16kHz$ which, according to the Nyquist frequency law (1), allows to represent frequencies until $F_N = 8kHz$. This suggests that hearing aids cannot greatly compensates for high frequency attenuation caused by the previously mentioned hearing loss Nyquist band.

$$f_N = \frac{f_s}{2} \tag{1}$$

Another important function of hearing aid is noise reduction. This algorithm utilizes compression to denoise the incoming signal after processing. The algorithm will be explained in more details in III-B1. The algorithm utilizes compression, which affects the dynamics of the signal. Those changes can be significant, from BRECVEMA point of view, when we consider elements like brain stem reflex for example, since the sudden increases of volume will not be accurately translated to the listener.

BRECVEMA offers interpretations as to why the hearing loss and hearing aid processing might change peoples' perceived emotions of musical pieces. This creates the need for a model for actually measuring such changes.

D. The VAHS-model

Measuring emotions felt when listening to music (perceived emotions) has been considered a challenging endeavour. As discussed previously, hearing loss and hearing aid devices and simulations affect frequency response, rhythmic information and other vital parameters for music perception. Therefore, BRECVEMA leads us to conclude that they might contribute to a change in perceived emotions - but how do those changes affect the way we perceive emotions?

A range of different discrete and dimensional models have been previously used to measure perceived emotions - or 'emotions that are represented by music and perceived as such by the listener' [11]. This research focuses specifically on the discrete emotions *happiness* and *sadness*, based on previous studies with reliable musical examples to elicit those perceived emotions [12]. The authors used a hybrid approach between a dimensional and a discrete model, similar to Taruffi.

We opted to use a variation of the two-dimensional circumplex model to shine a light on the perceived emotions when listening to different music excerpts [13] [14]. The original model will be referred to as the VA-model. For this research, the standard dimensions *valence* and *arousal* are utilised, with the addition of extra happiness and sadness dimensions, as used in discrete models and previous relevant research [12]. Therefore, it will be referred to as the VAHS-model.

Valence is defined as a 'pleasure-displeasure continuum' and arousal an 'activation-deactivation' one [11]. Valence can also be understood by the following adjective-extremes: pleasant-unpleasant, good-bad and positive-negative [15], whereas arousal by alertness-sleepiness.

This model gives us access to the reliability to define discrete emotions from the dimensional models, specifically the VA-model. The use of this two-dimensional model means that 'all emotions can be understood as varying degrees of both valence and arousal' [11]. Paired with that, the directness of the discrete model can be used as a way to further compare results, if needed. This is because 'the ratings of discrete emotions may be recovered to a large extent by a two dimensional model ($\approx$ 80%) and vice versa ($\approx$ 90%)' [11].

Finally, it must be emphasized that the emotions that are being measured are limited to happiness and sadness, and that the tracks selected have been previously assessed to induce those emotions [12].

E. Conclusions

By using a BRECVEMA taxonomy, it can be speculated that emotional responses might be altered by the sound processing present in the simulation of hearing loss as well as hearing aid. Both present significant problems in representation of musical information, namely temporal information, high-pitch melodies, as well as dynamics. Therefore, we can put forward the hypothesis:

Both hearing loss and hearing aids simulations decrease listeners' emotional response to musical stimuli.

The hypothesis will be validated or invalidated through the VAHS-model, specifically if the processed tracks alter arousal and/or valence ratings significantly. Obvious changes in happiness and sadness ratings can also confirm the hypothesis, but they will receive less emphasis, as the main method used for this paper is based on the dimensional VA-model.

III. Implementation

A. Hearing Loss Simulation

Trying to replicate specific perceptual aspects of hearing loss described in Section II-B, the MATLAB implementation focuses on three different processes: frequency attenuation, spectral smearing on high frequencies and temporal disruption on low frequencies.

1) Frequency Attenuation: In order to cover different options of audiogram attenuation, three different settings are available: mild, moderate and severe. The signal is attenuated in five different frequency bands (1000, 2000, 4000, 6000 and 8000 Hz) [1]. A low pass filter is also applied with cutoff frequency of 8000 Hz and a slope of 48 dB/octave. Second order Biquadratic filters have been used to implement the multi-band, with a Q factor of 2.3 and different gain settings, according to the desired setting.

To ensure frequency band separation corresponding to the separation typically observed in the human cochlea, Gammatone Filterbank filters have been used. This process divides the audio signal into 32 equivalent rectangular bandwidth, ranging from 20 to 16000 Hz.

After the band separation, the high frequencies are passed through the spectral smearing processing, while the low frequencies are temporally disrupted.

2) Spectral Smearing: The spectral smearing processor multiplies the high part of the audio signal with a low-passed white noise, according to the three different settings: mild (100 Hz), moderate (200 Hz) and severe (250 Hz) [1].

The result is a frequency smeared signal replicating the observed decreased frequency selectivity in hearing loss [3].

3) Temporal Disruption: According to [9], people with hearing loss are unable to utilize temporal fine structures (TFS) as well as normal hearing listeners. To achieve this result, this part of the process applies random phase shifts to the low register of the audio signal, after a transformation from the temporal domain to the frequency domain.

Specifically, the signal is converted in frequency domain using a Short Time Fourier Transform (STFT) with a 512 sample Hann periodic window. While the magnitude of the signal remains untouched, the phase is randomly shifted within the range $[-\pi/2, \pi/2]$. The signal is then recombined in the time domain using the Inverse Short Time Fourier Transform (ISTFT).

B. Hearing Aid Simulation

This section of the MATLAB implementation takes as an input file the .wav generated after the Hearing Loss Simulation. This input is passed through three subsequent processes i.e. noise reduction, frequency shaping and amplitude shaping and is then written as another .wav file.

1) Noise Reduction: After experimenting with implementing an FIR filter, the authors decided to perform noise reduction using wavelets.

Wavelets are nonlinear functions and do not remove noise by low-pass filtering like many traditional methods. Low-pass filtering approaches, which are linear time invariant, can blur the sharp features in a signal and sometimes it is difficult to separate noise from the signal where their Fourier spectra overlap. For wavelets the amplitude, instead of the location of the Fourier spectra, differ from that of the noise. This allows for thresholding of the wavelet coefficients to remove the noise. If a signal has energy concentrated in a small number of wavelet coefficients, their values will be large in comparison to the noise that has its energy spread over a large number of coefficients. These localizing properties of the wavelet transform allow the filtering of noise from a signal to be very effective. While linear methods trade-off suppression of noise for broadening of the signal features, noise reduction using wavelets allows features in the original signal to remain sharp. A problem with wavelet denoising is the lack of shiftinvariance, which means the wavelet coefficients do not move by the same amount that the signal is shifted. This can be overcome by averaging the denoising result over all possible shifts of the signal [16].

2) Frequency Shaping: This section applies a frequency adjustment filter designed to enhance certain frequency regions in order to correct the hearing loss. To do so, different gains are applied to small regions, defined by a frequency vector, and generated a transfer function in the frequency domain [17]. For this curve, the input frequency vector is $[1000, 2000, 4000, 5000]$. The frequency shaping curve is shown in Figure 1.

The regions will then be 0 to 1000 Hz; 1000 to 2000 Hz; 2000 to 4000 Hz; 4000 to 5000 Hz and 5000 to 22000 Hz. For the first band, a linearly increasing function is generated with

$$firstC = \frac{0.3(g-1)}{first} \tag{2}$$

$$gain = firstC\frac{k}{NT} \tag{3}$$

where $first = 1000$; g is the highest gain value set as an input parameter to 90; k is used as pointer for numbers of samples for modification; N is the total number of samples and T is sampling period. For the second region, an exponential rising curve is generated

$$secondC = firstC \cdot first + 1 \tag{4}$$

$$secondC_2 = \frac{second - first}{5} \tag{5}$$

$$\begin{aligned} gain(k+1) &= 1 + (secondC - 1) \\ &\cdot exp\left(-\frac{\frac{k}{NT} - first}{secondC_2}\right) \end{aligned} \tag{6}$$

where $second = 2000$.

From 2000 to 4000 Hz a constant gain is necessary with

$$thirdC = 1 + (secondC) \cdot exp\left(-\frac{second}{secondC_2}\right) \tag{7}$$

$$thirdC_2 = \frac{third - second}{5} \tag{8}$$

$$gain = g \tag{9}$$

where $third = 4000$.

For the next region the curve is linearly decreasing with

$$fourthC = \frac{0.6(g-1)}{fourth}; \tag{10}$$

$$gain = g - fourthC\left(\frac{k}{NT} - third\right); \tag{11}$$

with $fourth = 5000$.

The last region is an exponentailly decreasing function with

$$fifthC = g - fourthC\left(\frac{k}{NT} - third\right); \tag{12}$$

$$fifthC_2 = \frac{fs/2 - fourth}{5}; \tag{13}$$

$$gain = 1 + (fifthC - 1)exp\left(-\frac{\frac{k}{NT} - fourth}{fifthC2}\right) \tag{14}$$

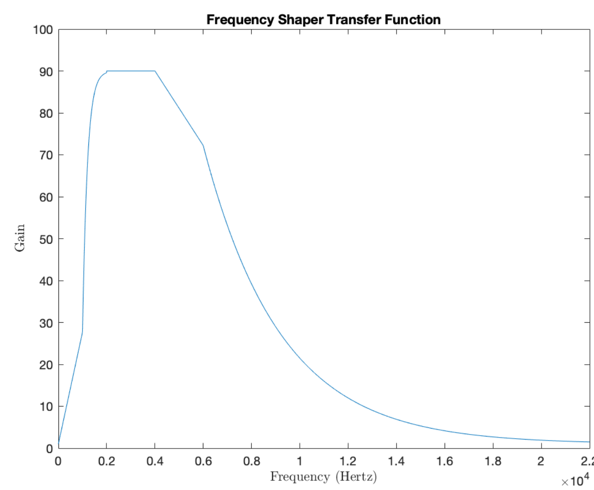

Fig. 1. Frequency Shaper transfer function

3) Amplitude Shaping: The processed signal will now pass through an amplitude adjustment filter. This acts like a compressor, checking the amplitude level in terms of output power of the signal and output level of each sample, comparing them with a higher and lower threshold level. If the signal level exceeds the upper threshold value P_{sat}, then it is reduced to P_{sat}. Vice versa, if the signal level is lower than the lower threshold value P_{low}, the signal is reduced to zero [17].

IV. Experiment

A. Stimuli

The experiment portion of this project involved selecting four musical tracks to be processed through the hearing loss and the hearing aid algorithms. The tracks are concert music movie soundtracks, and they are selected according to their reported success of eliciting happiness and sadness while 'avoiding familiar pieces to reduce potential biases due to memory effects' (Pereira et al. 2011, in [12]). This process led to 12 individual excerpts: four being unprocessed, four processed with hearing loss only (HL), and four with both hearing loss and hearing aid (HLHA). All excerpts were normalized and grouped according the their respective tracks, making for a total of four groups of three excerpts each. The excerpts were edited to last between 1-1:30 minutes each. The order both for the groups and the excerpts were randomized when presented to the participants. This was done in order to reduce biases when participants heard the processed tracks were played before the unprocessed excerpts and vice versa.

- Track 1: *Darcy's Letter - Dario Marianelli - Pride and Prejudice - Sad*
- Track 2: *Grand Hotel Fox Trot - Life is Beautiful - Happy*
- Track 3: *Papaya - Stelvio Cipriani - Happy*
- Track 4: *Song for Bob - Nick Cave & Warren Ellis - Sad*

B. Hypothesis

The leading question behind the experiment is: different levels of hearing loss affect musical enjoyment, but how does it affect the way we perceive happiness and sadness? The main hypothesis, described in II, is that both hearing loss and hearing aids simulations change listeners' emotional response to musical stimuli.

Additional expectations from the authors include: the participants were expected to rate valence and arousal values lower in the HL and HLHA compared to the unprocessed excerpts, because of them not being used to having hearing loss or using hearing aids, which could lead to boredom. Furthermore, their happiness and sadness ratings are expected to be reduced, potentially because their perceived emotion could shift towards boredom. Finally, their enjoyment was expected to be greatly reduced with the HL excerpts, and slightly less reduced with the HLHA excerpts. It is worth noting that participants might have 'difficulty with rating the valence of sad music' [11].

C. Participants

12 people participated in the experiment on different days (7 male, 5 female). None of the participants were compensated for their participation. The age of the participants ranged from 22 to 36 with a median of 25,5. The years of musical related activities of the participants ranged from 0 to 20 with a median of 11. Their usual music listening enjoyment, on a scale from 0 to 100, ranged from 55 to 100 with a median of 90. Most of the participants were students of the MSc in Sound and Music Computing course at Aalborg University. This contributed to a high number of participants with a background knowledge in audio and sound processing.

D. Experiment set up

The order of the tracks presented was randomized, as well as the internal order of the excerpts - in order to avoid potential biases. The reproduction method of choice was a pair of Sennheiser HD600 headphones, chosen for their open-back design, flat frequency response and to reduce the interference of room reflections. The experiments were conducted in relatively quiet environments, and after the initial volume has been set to a range that was comfortable to the listener, the volume level has been kept intact throughout the experiment. A similar playback volume was used for all participants.

Furthermore, an Empatica bracelet was used in order to measure skin conductivity response of the participants, with the intention of collecting data that could be correlated with their reported arousal ratings in the questionnaires. The results of the Empatica bracelet seemed too unpredictable to draw any sort of correlations reliably, so its use has been discarded for the rest of the experiment.

E. Procedure

The data collection method was a mixture between an adapted questionnaire on enjoyment of musical pieces with hearing aids (UCMLQHA) [18] and the VAHS-Model with ratings between 1-7 for each parameter. The participants were shown the questionnaires, which were filled partially by them and partially by the researchers. The questions were selected to collect basic background information from the participants (age, musical background, preferred music style or genre, etc.), as well as to get more subjective answers related to their perception of the individual tracks and excerpts. Further questions about their enjoyment of the pieces were included from the original UCMLQHA questionnaire for better comparison with previous research [5]. Participants were asked to rate the enjoyment of each excerpt on a scale from 0-100, where 0 is *didn't enjoy at all* and 100 is *greatly enjoyed* and impact of processing from 0-100 - where 0 is a *very negative* impact and 100 a *very positive* impact.

Furthermore, we used a dimensional model of emotion [11] in order to collect data about their emotional reactions to the excerpts. This is where the participants were asked to rate valence, arousal, happiness and sadness on 7-point scales for each of the 12 individual excerpts, as well as familiarity on a 4-point scale for each of the four tracks. Arousal ratings were asked as 'how intense is the emotion felt from listening to the excerpt from 1 (not at all) to 7 (very much so)?', and valence ratings 'how positively or negatively does the excerpt affect you, from 1 (very negatively) to 7 (very positively)?'.

The combination of those two models seemed to be an appropriate way to answer the following question: how do hearing loss and hearing aids affects people's emotional response to listening to music? Besides, they would allow for comparison between previous works that use the UCMLQHA questionnaire and the dimensional model of emotion. On this experiment, however, the main unusual characteristic is the participants' healthy hearing conditions, as well as the focus on emotional response. This might lead to different results if compared with people with actual hearing loss [5].

Familiarity ratings were also asked in the following manner: 'how well you know the excerpt, from 1(totally unfamiliar) to 4 (totally familiar)?'. In order to reduce biases - including potential ones related to the BRECVEMA model, such as evaluating conditioning and episodic memory [7] -, the participants familiar with specific tracks (familiarity rating 3 or 4) would have their ratings for all excerpts of that track invalidated.

Three pilot tests were conducted. Initially we opted to use the severe simulation of the hearing loss algorithm, inspired by Mourgela's plugin implementation [1]. The reason behind this is that people who use hearing aids and have severe hearing loss have reported significantly lower music enjoyment than those with mild hearing loss [5].

However, the processing on the excerpts were reported to be too intense to allow for meaningful analysis when compared to the unprocessed ones, since they were unable to recognize most musical elements, aside from maybe rhythm, as reported in one of the questionnaires for the pilot tests. Furthermore, no direct correlations were found between Mourgela's severe hearing loss settings and 'real world' hearing loss. Therefore, the authors opted to use very mild settings (according to Mourgela's ratings) - for the frequency smearing portion of the hearing loss, as well as mild high frequency attenuation. It has proved to be enough to differentiate between processed and unprocessed tracks, and it made easier to distinguish between the intended processing differences between the HL and the HLHA excerpts.

On a side note, one of the pilot tests was executed grouping excerpts according to processing (unprocessed, HL, HLHA) instead if grouping them by track. The motivation behind that was to reduce the direct comparisons between different versions of the same track and focus instead on the effects of the different algorithms. The authors opted for the first procedure instead for the valid experiments, as it emphasized the comparison aspect between different forms of processing for the same tracks - as opposed to how different tracks fare with different kinds of processing.

F. Data Analysis

After the data has been collected the results where put into a CSV file and analysed using MATLAB. The process involved dividing data into individual groups representing tracks, kind of processing, and measured variable. The data is then represented in a form of *boxplots* in which values like median, $25th$ and $75th$ percentile, and outliers are represented. Although the questionnaire gives data such as happiness, sadness and valence, the primary variable of the analysis are valence and arousal, the reason of which is described in section II.

Next, the average value of the difference of the arousal and valence is calculated with an absolute mean error. The same process is repeated for enjoyment for each processed/unprocessed excerpt. Then a correlation is calculated between those variables to put the changes of emotional response in perspective of frameworks discussed in II.

V. Results

A. Impact of processing

Overall enjoyment ratings of HL and HLHA processed excerpts were generally very similar. For the first track (*Darcy's Letter - Dario Marianelli - Pride and Prejudice - Sad*), the impact of processing for enjoyment, 0 meaning really negatively and 100 really positively, ranged from 0 to 60 with a median of 30. For the second track (*Grand Hotel Fox Trot - Life is Beautiful - Happy*), the impact of processing for enjoyment ranged from 0 to 45 with a median of 30. For the third track (*Papaya - Stelvio Cipriani - Happy*), the impact of processing for enjoyment ranged from 0 to 60 with a median of 30. For the fourth track (*Song for Bob - Nick Cave & Warren Ellis - Sad*), the impact of processing for enjoyment ranged from 0 to 60 with a median of 30. Results are shown in Figure 2

B. Valence

For the first track the valence of emotions measured a median of 3 for HL (1A in Figure 3), 2.5 for HLHA (1B) and 5 for the unprocessed excerpt (1C), on a scale from 1 to 7. The average valence value of the difference between unprocessed 1C and HL 1A is 1.75 ± 1.50 while the average value of the difference between unprocessed 1C and HLHA 1B is 1.42 ± 1.15.

For the second track the valence of emotions measured a median of 3.5 for HL (2A in Figure 4), 3.5 for HLHA (2B) and 6 for the unprocessed excerpt (2C), on a scale from 1 to 7. The average valence value of the difference between unprocessed 2C and HL 2A is 1.75 ± 1.29 while the average

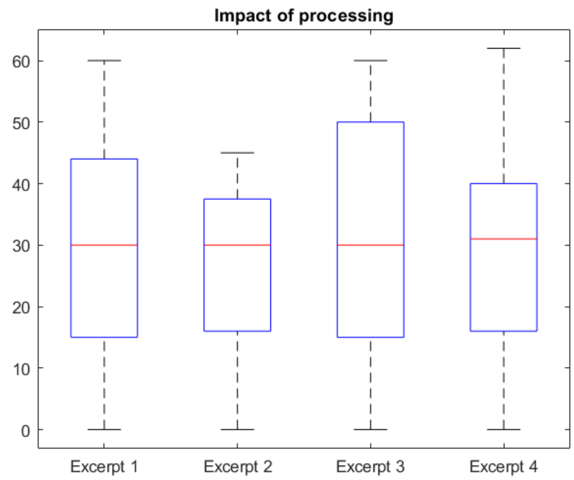

Fig. 2. Impact of processing for the four music excerpts

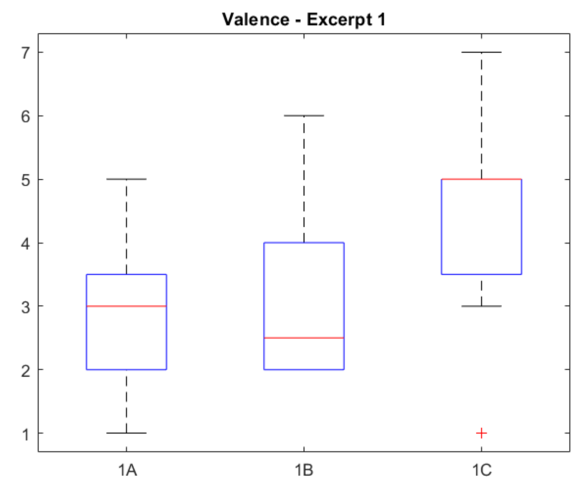

Fig. 3. Valence values for the first music track (*Darcy's Letter - Dario Marianelli - Pride and Prejudice - Sad*) with HL algorithm (1A), with HLHA (1B) and without processing (1C)

value of the difference between unprocessed 2C and HLHA 2B is 1.75 ± 1.17.

For the third track the valence of emotions measured a median of 4 for HL (3A in Figure 5), 3 for HLHA (3B) and 6 for the unprocessed excerpt (3C), on a scale from 1 to 7. The average valence value of the difference between unprocessed 3C and HL 3A is 2.17 ± 1.39 while the average value of the difference between unprocessed 3C and HLHA 3B is 2.58 ± 1.65.

For the fourth track the valence of emotions measured a median of 3 for HL (4A in Figure 10), 3 for HLHA (4B) and 5 for the unprocessed excerpt (4C), on a scale from 1 to 7. The average valence value of the difference between unprocessed 4C and HL 4A is 1.91 ± 1.40 while the average

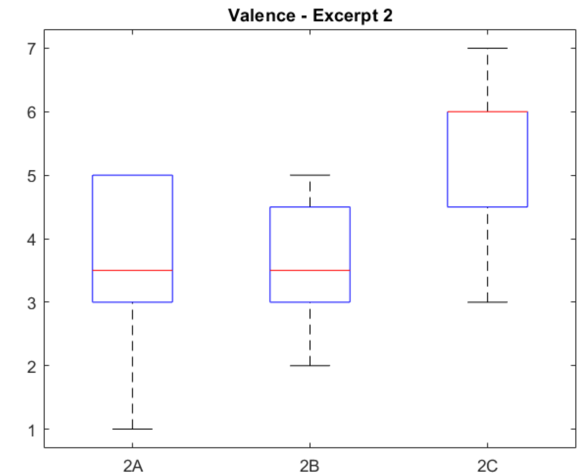

Fig. 4. Valence values for the second music track (*Grand Hotel Fox Trot - Life is Beautiful - Happy*) with HL algorithm (2A), with HLHA (2B) and without processing (2C)

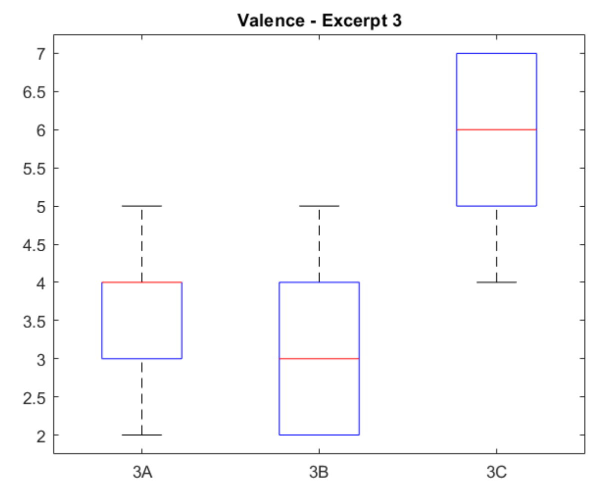

Fig. 5. Valence values for the third music track (*Papaya - Stelvio Cipriani - Happy*) with HL algorithm (3A), with HLHA (3B) and without processing (3C)

value of the difference between unprocessed 4C and HLHA 4B is 1.92 ± 1.26.

C. Arousal

For the first track the arousal of emotions measured a median of 2.5 for HL (1A in Figure 7), 3 for HLHA (1B) and 5.5 for the unprocessed excerpt (1C), on a scale from 1 to 7. The average arousal value of the difference between unprocessed 1C and HL 1A is 2.25 ± 1.96 while the average value of the difference between unprocessed 1C and HLHA 1B is 1.83 ± 1.83.

For the second track the arousal of emotions measured a

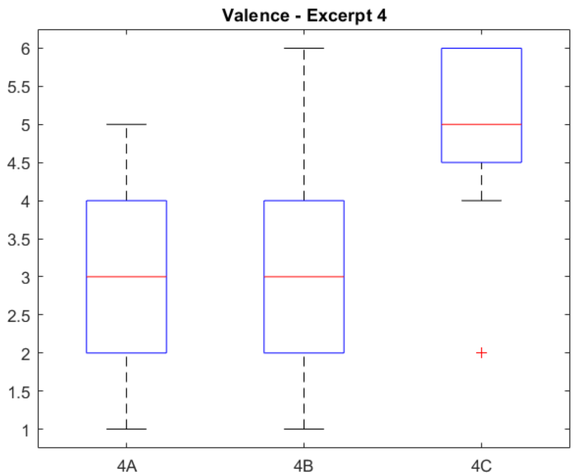

Fig. 6. Valence values for the fourht music track (*Song for Bob - Nick Cave & Warren Ellis - Sad*) with HL algorithm (4A), with HLHA (4B) and without processing (4C)

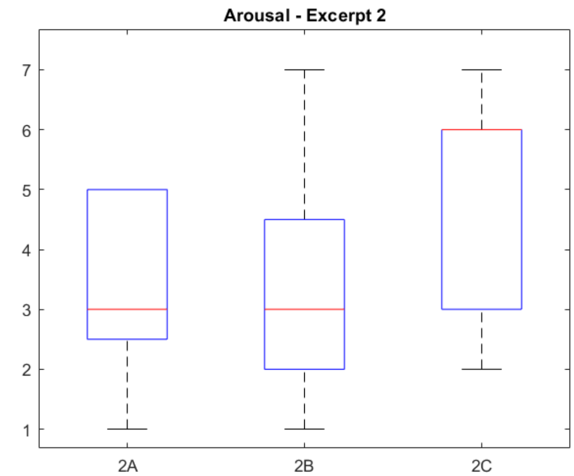

Fig. 8. Arousal values for the second music track (*Grand Hotel Fox Trot - Life is Beautiful - Happy*) with HL algorithm (2A), with HLHA (2B) and without processing (2C)

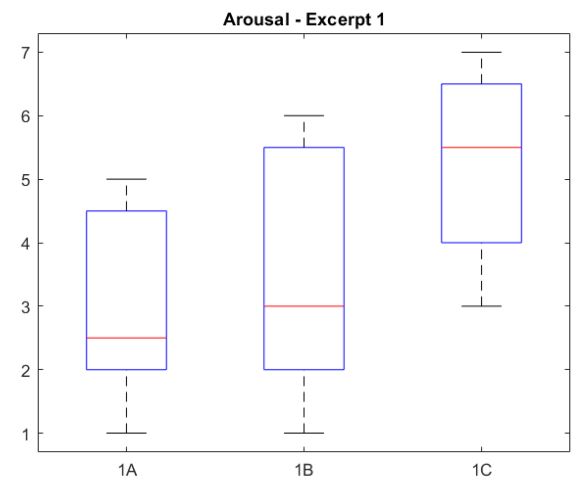

Fig. 7. Arousal values for the first music track (*Darcy's Letter - Dario Marianelli - Pride and Prejudice - Sad*) with HL algorithm (1A), with HLHA (1B) and without processing (1C)

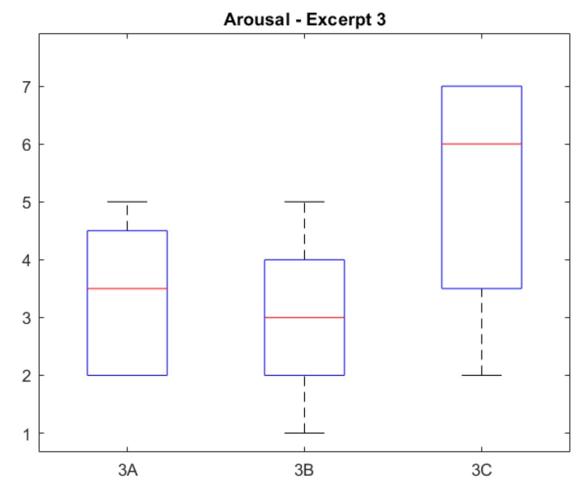

Fig. 9. Arousal values for the third music track (*Papaya - Stelvio Cipriani - Happy*) with HL algorithm (3A), with HLHA (3B) and without processing (3C)

median of 3 for HL (2A in Figure 8), 3 for HLHA (2B) and 6 for the unprocessed excerpt (2C), on a scale from 1 to 7. The average arousal value of the difference between unprocessed 2C and HL 2A is 1.42 ± 1.08 while the average value of the difference between unprocessed 2C and HLHA 2B is 1.50 ± 1.75.

For the third track the arousal of emotions measured a median of 3.5 for HL (3A in Figure 9), 3 for HLHA (3B) and 6 for the unprocessed excerpt (3C), on a scale from 1 to 7. The average arousal value of the difference between unprocessed 3C and HL 3A is 1.83 ± 1.05 while the average value of the difference between unprocessed 3C and HLHA 3B is 2.25 ± 1.25.

For the fourth track the arousal of emotions measured a median of 2 for HL (4A in Figure 10), 3.5 for HLHA (4B) and 6 for the unprocessed excerpt (4C), on a scale from 1 to 7. The average arousal value of the difference between unprocessed 4C and HL 4A is 2.83 ± 1.39 while the average value of the difference between unprocessed 4C and HLHA 4B is 2.25 ± 1.46.

D. Enjoyment

Based on a scale described in IV-E, the average enjoyment value of the difference between unprocessed 1C and HL 1A is 47.25 ± 22.08 while the average value of the difference between unprocessed 1C and HLHA 1B is 43.25 ± 22.79.

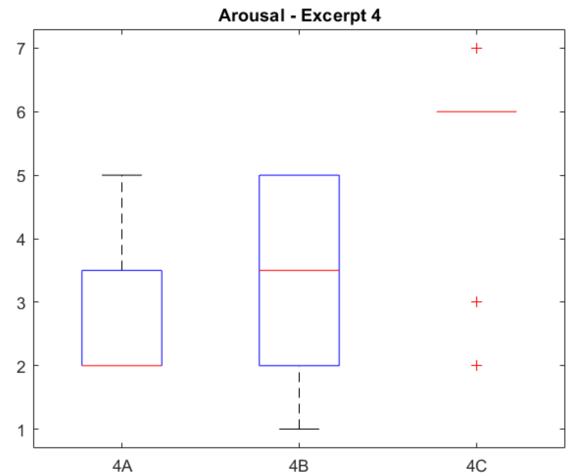

Fig. 10. Arousal values for the fourth music track (*Song for Bob - Nick Cave & Warren Ellis - Sad*) with HL algorithm (4A), with HLHA (4B) and without processing (4C)

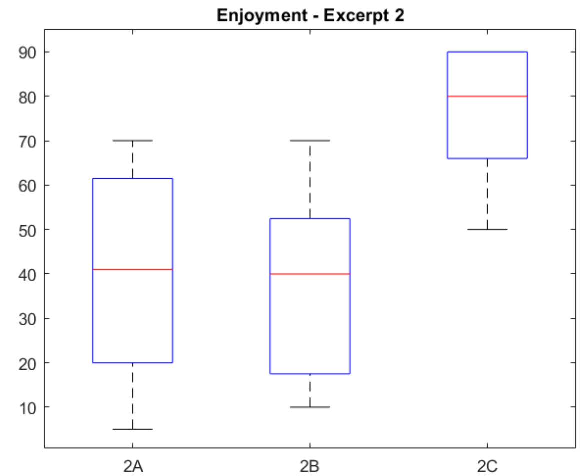

Fig. 12. Enjoyment values for the first music track (*Grand Hotel Fox Trot - Life is Beautiful - Happy*) with HL algorithm (2A), with HLHA (2B) and without processing (2C)

The average enjoyment value of the difference between unprocessed 2C and HL 2A is 47.25 ± 27.28 while the average value of the difference between unprocessed 2C and HLHA 2B is 39.08 ± 22.08. The average enjoyment value of the difference between unprocessed 3C and HL 3A is 37.00 ± 22.50 while the average value of the difference between unprocessed 3C and HLHA 3B is 49.58 ± 23.82. The average enjoyment value of the difference between unprocessed 4C and HL 4A is 45.75 ± 20.21 while the average value of the difference between unprocessed 4C and HLHA 4B is 46.00 ± 22.50.

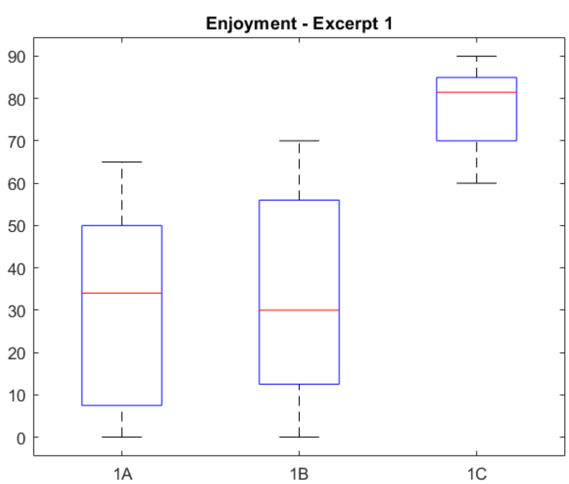

Fig. 11. Enjoyment values for the first music track (*Darcy's Letter - Dario Marianelli - Pride and Prejudice - Sad*) with HL algorithm (1A), with HLHA (1B) and without processing (1C)

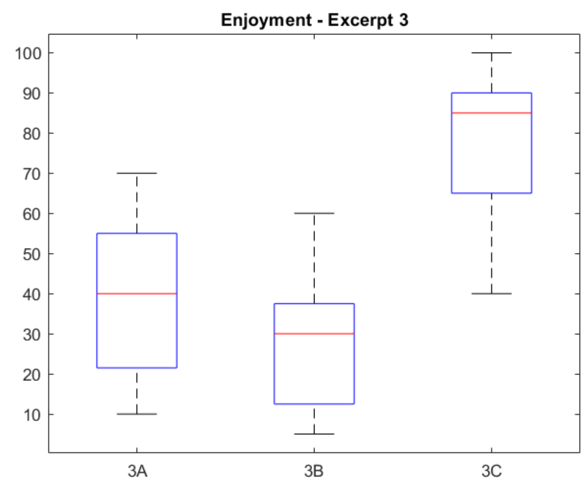

Fig. 13. Enjoyment values for the first music track (*Papaya - Stelvio Cipriani - Happy*) with HL algorithm (3A), with HLHA (3B) and without processing (3C)

E. Other data

Similar decreases were observed in all other sets of measured data (happiness, sadness) related to HL (A) and HLHA (B), which will be shown in the Appendix A.

The data of the Empatica bracelet has not been used because it did not seem to provide any useful data.

F. Correlation Analysis

To help characterize the change of emotional response in therms of the causation, a correlation between arousal and enjoyment is calculated and results shown in Table I. The

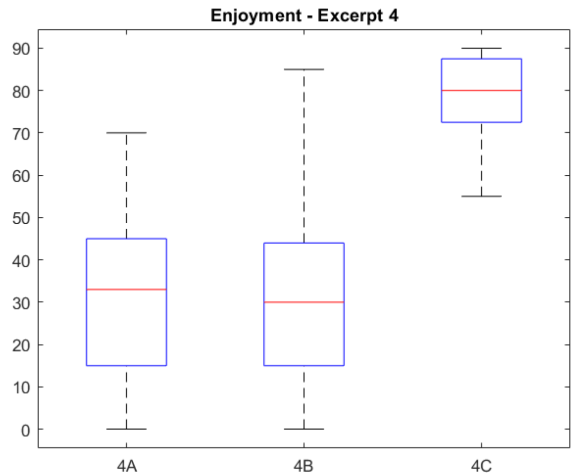

Fig. 14. Enjoyment values for the first music track (*Song for Bob - Nick Cave & Warren Ellis - Sad*) with HL algorithm (4A), with HLHA (4B) and without processing (4C)

Table II shows mean value and standard deviation of changes of arousal. Tables showing the information for valence and enjoyment can be found in Appendix A.

TABLE I
CORRELATION OF AROUSAL AND ENJOYMENT OF HEARING LOSS SIMULATION (PROCESSING A) AND HEARING AID WITH HEARING LOSS SIMULATION (PROCESSING B)

	Processing A	Processing B
Correlation Excerpt 1	$R = 0.67$	$R = 0.64$
Correlation Excerpt 2	$R = 0.65$	$R = 0.66$
Correlation Excerpt 3	$R = 0.83$	$R = 0.86$
Correlation Excerpt 4	$R = 0.47$	$R = 0.38$

TABLE II
MEAN ($\bar{x}$) AND STANDARD DEVIATION(σ) OF DIFFERENCE OF AROUSAL FOR HEARING LOSS SIMULATION (PROCESSING A) AND HEARING LOSS WITH HEARING AIDS SIMULATION (PROCESSING B)

	Processing A	Processing B
Excerpt 1	$\bar{x} = 2.25$	$\bar{x} = 1.83$
	$\sigma = 2.45$	$\sigma = 2.32$
Excerpt 2	$\bar{x} = 1.41$	$\bar{x} = 1.50$
	$\sigma = 1.37$	$\sigma = 2.32$
Excerpt 3	$\bar{x} = 1.83$	$\bar{x} = 2.25$
	$\sigma = 1.53$	$\sigma = 1.60$
Excerpt 4	$\bar{x} = 2.83$	$\bar{x} = 2.25$
	$\sigma = 1.70$	$\sigma = 1.91$

VI. DISCUSSION

In the project the simulations of hearing loss and hearing aid where developed based on Mourgela et al. 2016 [1] and other existing solutions with utilization of a wavelets-based denoising algorithm, and applied to process four excerpts. The questionnaire then was constructed to gather information about arousal, valence, happiness and sadness, as well as enjoyment of each of descriptive information about processed/unprocessed excerpts.

A. Data

From the data analyzed in Section V-A and shown in Figure 2 we can then assert that the process algorithm impacted negatively on the overall enjoyment of the proposed tracks, being the average median of this value 30, on a scale from 0 to 100, as shown in Figure 2. This validates our expectations in regards to enjoyment, as stated in section IV-B. Average of differences measured in V-C between xC - xA and xC - xB showed that overall, arousal was higher in all the unprocessed tracks (C), while the values for HL (A) and HLHA (B) were similar, as reported in Figures 7, 8, 9 and 10 in Section V-C. Similarly, Section V-B shows that valence values were higher in all unprocessed tracks (C) than in HL (A) and HLHA (B) as reported in Figures 3, 4, 5 and 6. This once again confirms our expectations of changes in emotional response. Based on the measurements reported in Section V-D and in Figures 11, 12, 13 and 14 we can state that the overall Enjoyment decreased significantly and similarly in the HL (A) and HLHA (B) listening situations compared to the unprocessed excerpts (C). When looking at the correlations regarding arousal and enjoyment it can be concluded that the relation between those parameters is strong reaching up to $R = 0.83$ in case of third excerpt. This may suggest, from BRECVEMA taxonomy perspective, that the emotional response measured as arousal is largely altered by aesthetic judgement, that might be linked with enjoyment factor, changes between the versions of the excerpts. It is also worth mentioning that there were no significant differences between ratings of happy tracks (*Grand Hotel Foxtrot* and *Papaya*) and sad tracks (*Darcy's Letter* and *Song for Bob*).

Summarising the results of experiments it can be stated that the data supports the hypothesis of a decrease in emotional response (measured mainly through arousal and valence) when musical excerpts are processed with simulations of hearing aid and hearing loss. The effect of individual processing, however, is not clear since the difference of arousal and valence vary between the excerpts, but the difference between them are not large (median value of said parameters are close to each other, though always significantly lower then those of not processed versions). This invalidates the expectation that the experiment would show slightly higher results in the HLHA in relationship to the HL excerpts, and will be addressed in the following sections. But first, objective and subjective answers from the questionnaires will be analyzed.

B. Questionnaire

In section V, we can visualize the ratings of the 12 participants. The questionnaire included questions about the *best* and *worst sounding* excerpts of each track, where participants could express their opinions and elaborate on it. Generally, the results were according to expectations for the best excerpt:

the unprocessed one was chosen most of the time. The less predictable results were the excerpts declared the worst ones.

For track 1 (Darcy's Letter, sad), 8 participants opted for excerpt A (HL) as the worst one, 3 for B (HLHA) and 1 did not know. This track consists mostly of a solo cello melody and soft piano accompaniment. Therefore, the sound of the recording is warmer, with emphasis on the lower mids. Further analysis, such as brightness analysis from the MIR Toolbox could provide a deeper understanding of the frequency content of each excerpt. Those have not been made because of time restrictions. That said, we can assume that an already dark track would suffer from further frequency attenuation and low pass filtering - and benefit from the extra brightness from the frequency shaping of the HA simulation without making it sound harsh.

For track 2 (Grand Hotel Foxtrot, happy), 8 participants opted for excerpt B and 4 for excerpt A. This could be because the main melodies of this excerpt are located in instruments in higher registers. Those can become more negatively affected by the high frequency smearing, making them sound more distorted and overall harsher. This can also be confirmed by some of the participants' answers. Looi et al found that there is a stronger aversion to brighter instruments for actual hearing aid users as well, so this correlates positively with the results of our questionnaires [5]. The mean ratings for enjoyment (40) also indicate that this was the track that suffered the least from both HL and HLHA algorithms.

For track 3 (Papaya, happy), only one participant opted for excerpt A as the worst one; 10 others opted for B and 1 opted for both. In this case, the overall brighter mix could have influenced the subjects' choices, as well as the bright instrumentation (guitar, trumpet, strings). A lot of the negative artifacts can be clearly noticed on excerpt A, but they get further amplified in a negative way in excerpt B. Participants reported the amount of distortion and harshness on this excerpt more than in any other.

For track 4 (Song for Bob, sad), 7 participants chose excerpt A as the worst one, 4 chose excerpt B and 1 did not know.

C. Other considerations

In evaluating other relevant aspects of this paper, we should take into account the characteristics of the selected music tracks. All excerpts were normalized. This means that hearing loss vs. hearing aid had roughly the same perceived volume, which is obviously not the case in real life. That said, experiments were organized in that manner, because we did not find a straightforward way of replicating the volume changes - coming from e.g. the massive drop in high frequencies expected with hearing loss - with participants with healthy hearing. Even though hearing loss was simulated using other algorithms and block diagrams as guidelines, there were no references found for precise levels of high frequency loss, smearing or any of the parameters. The ones used as base were Mourgela's [1]. Participants had healthy hearing. This means that they don't get the time to get used to hearing aids and they don't have hearing loss as their previous standard, as other studies do [5]. Those abrupt changes probably influence the results drastically, and this needs to be taken into considerations when comparisons are made between our results and Looi's. Also the genre and instrumentation of the selected tracks is a relevant aspect, as they were mostly classical music, film soundtracks with a significant amount of string instruments. Those are instruments characterized by a complex spectrum of frequencies, as well as very different textures, despite the fact of having a, homogeneous acoustic effect [19] [20].

These considerations led the authors toward another important aspect of this research, that is the existing problem of reproducing broad spectral information and complex textures, such as music. The use of hearing aid amplification for music listening therefore presents unique challenges, as amplification requirements depend on a wide variety of variables including acoustics, instrumentation, sound source, level of musical training, hearing loss configuration and the specific hearing-aid technology used. Although new hearing aid technologies have somewhat reduced problems of music enjoyment experienced by hearing impaired people, audiologists are aware that a significant amount of patients may have difficulties with listening to music and may require extra attention to minimize those problems. Issues identified here include poor sound quality, difficulties hearing words in songs, distortion, and feedback [21]. The results in [22] also show that hearing aids can be problematic for music for two-thirds of hearing-aids users. As hearing aids are not expressly designed to amplify music, the diverse range of experiences reported is perhaps to be expected. In this cited study, audiologists corroborated this, the majority reporting that hearing aids were *"sometimes useful, sometimes not"*. Hearing-impaired listeners perform poorly compared to normally hearing listeners on music perceptual tasks such as pitch discrimination, melodic intonation,and identifying instruments due to threshold elevation, reduced frequency selectivity, loudness recruitment, and anomalies in pitch perception [4]. As also presented in [17], it is clear that the algorithms implemented in this paper for hearing loss and hearing aid simulations perform significantly better with speech as input source.

Furthermore, it should also be taken into consideration that the definition of arousal in the questionnaires could probably have been clearer and more directly related to its reported meaning. This has proven to be a challenge, as the definition of the term as 'activation-deactivation' or 'alertness-sleepiness' seemed clear enough in the references, but not so clear as to how it could be asked to participants of an experiment [13] [14] [11] [12]. It is possible that a more clearly defined question would lead to different arousal ratings.

Another aspect to discuss is the environment in which experiment took place. The tests were not performed in a fully controlled environment; in fact it was reported external noise, as well as a close proximity of the participants to the researchers, allowing the former to see the facial expressions and reactions of the latter as the experiment was conducted. Consequently, unknown participants' conditions might have been relevant were while taking the survey and if they might have influenced their answers.

Lastly, most of the participants came from similar academic backgrounds (mostly from Msc in Sound and Music Computing course at Aalborg University). This uniformity of sample contributed to common and similar knowledge in sound processing, audio technology and music interest and, together with the limited number of participants taking part in the experiment, contributed to a narrow range of data variety, even though the data acquisition methods were designed in a way to account for a low sample size. Furthermore, the questionnaire included objective, as well as subjective answers, in order to collect as much data.

VII. Conclusions

The goal of the research described in this paper was to measure and analyse emotional responses of hearing loss and hearing aid. Gathered data from questionnaire showed significant decrease in arousal and valence of perceived emotions as well as enjoyment of presented excerpts. The experiment utilized a simulation of certain aspects of hearing loss and hearing aids and was used on normal hearing subjects. This allowed to observe changes in emotional response from purely signal processing perspectives. A similar study could be performed on hearing impaired people to compare the results and describe the accuracy of the prediction.

Furthermore, the simulation utilizes only one setting of hearing aid. As discussed in VI-C, hearing aid proved not sufficient to allow comfortable music listening experience for patients. Additional research could involve different solutions and algorithms that are designed for this specific function.

Although data gathered from the questionnaires was consistent, further analysis using psychophysiological measurements could be performed to objectify the reported subjective changes of emotions (test like skin conductivity, AAR or fMRI).

References

[1] A. Mourgela, T. R. Agus, and J. D. Reiss, "Investigation of a real-time hearing loss simulation for use in audio production," 2016. [Online]. Available: https://code.soundsoftware.ac.uk/hg/hearing-loss-

[2] R. Dhawan and P. Mahalakshmi, "Digital filtering in hearing aid system for the hearing impaired," *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*, pp. 1494–1497, 11 2016. [Online]. Available: https://research.vit.ac.in/publication/digital-filtering-in-hearing-aid-system

[3] J. G. Desloge, C. M. Reed, L. D. Braida, Z. D. Perez, and L. A. Delhorne, "Auditory-filter characteristics for listeners with real and simulated hearing impairment," *Trends in Amplification*, vol. 16, pp. 19–39, 2012.

[4] M. J. Kirchberger and F. A. Russo, "Development of the adaptive music perception test," *Ear and Hearing*, vol. 36, pp. 217–228, 7 2015.

[5] V. Looi, K. Rutledge, and T. Prvan, "Music Appreciation of Adult Hearing Aid Users and the Impact of Different Levels of Hearing Loss," *Ear and Hearing*, vol. 40, no. 3, pp. 529–544, may 2019.

[6] P. N. Juslin, "Handbook of Music and Emotion: Theory, Research, Applications," apr 2010.

[7] ——, "From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions," *Physics of Life Reviews*, vol. 10, no. 3, pp. 235–266, 2013.

[8] "The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise." *undefined*, vol. 130, no. 1, pp. 334–349, jul 2011.

[9] C. Lorenzi, G. Gilbert, H. Carn, S. Garnier, and B. C. Moore, "Speech perception problems of the hearing impaired reflect inability to use temporal fine structure," *Proceedings of the National Academy of Sciences*, vol. 103, no. 49, pp. 18 866–18 869, dec 2006. [Online]. Available: https://www.pnas.org/content/103/49/18866 https://www.pnas.org/content/103/49/18866.abstract

[10] J. M. Vaisberg, E. A. Macpherson, and S. D. Scollie, "Understanding hearing aid sound quality for music-listening," Tech. Rep., 2019. [Online]. Available: https://ir.lib.uwo.ca/etdhttps://ir.lib.uwo.ca/etd/6145

[11] T. Eerola and J. K. Vuoskoski, "A comparison of the discrete and dimensional models of emotion in music," *Psychology of Music*, vol. 39, pp. 18–49, 2011.

[12] L. Taruffi, C. Pehrs, S. Skouras, and S. Koelsch, "Supplementary information for effects of sad and happy music on mind-wandering and the default mode network," 2017.

[13] J. Posner, J. A. Russell, and B. S. Peterson, "The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology," *Development and Psychopathology*, vol. 17, pp. 715–734, 7 2005.

[14] J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, pp. 1161–1178, 12 1980.

[15] U. Schimmack and A. Grob, "Dimensional models of core aect: A quantitative comparison by means of structural equation modeling," *European Journal of Personality*, 2000.

[16] W. Chunli, Z. Chunlei, and Z. pengtu, "Denoising algorithm based on wavelet adaptive threshold," *Physics Procedia*, vol. 24, pp. 678–685, 2012.

[17] B. Saha, S. Khan, C. Shahnaz, S. A. Fattah, M. T. Islam, and A. I. Khan, "Configurable Digital Hearing Aid System with Reduction of Noise for Speech Enhancement Using Spectral Subtraction Method and Frequency Dependent Amplification," *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, vol. 2018-Octob, no. September, pp. 735–740, 2019.

[18] K. L. Rutledge, "A music listening questionnaire for hearing aid users," 2009.

[19] F. Leccese, G. Salvadori, G. Bernardini, and P. Bernardini, "The bowed string instruments: acoustic characterization of unique pieces from the italian lutherie," *IOP Conference Series: Materials Science and Engineering*, vol. 364, p. 012022, 6 2018.

[20] T. D. Rossing, "The science of string instruments," *The Science of String Instruments*, pp. 197–208, 2010.

[21] M. R. Leek, M. R. Molis, L. R. Kubli, and J. B. Tufts, "Enjoyment of music by elderly hearing-impaired listeners," *Journal of the American Academy of Audiology*, vol. 19, pp. 519–526, 8 2008. [Online]. Available: http://www.thieme-connect.de/DOI/DOI?10.3766/jaaa.19.6.7

[22] A. Greasley, H. Crook, and R. Fulford, "Music listening and hearing aids: perspectives from audiologists and their patients," *https://doi.org/10.1080/14992027.2020.1762126*, vol. 59, pp. 694–706, 9 2020. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/14992027.2020.1762126

APPENDIX

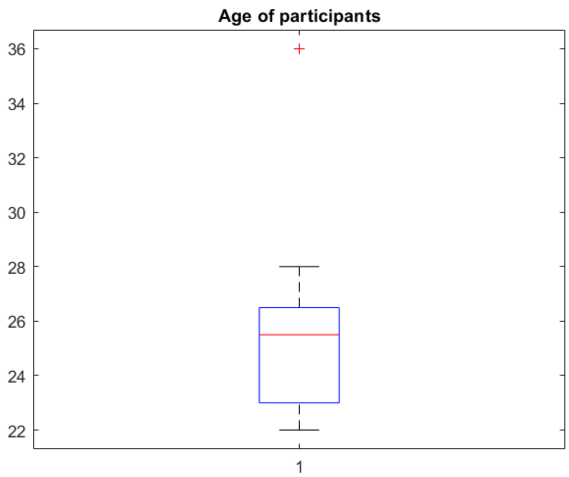

Fig. 15. Age of participants involved in the experimentation

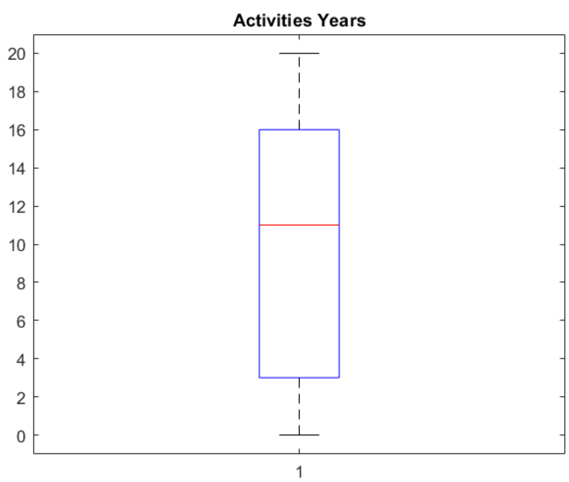

Fig. 16. Years of musical activity of the participants

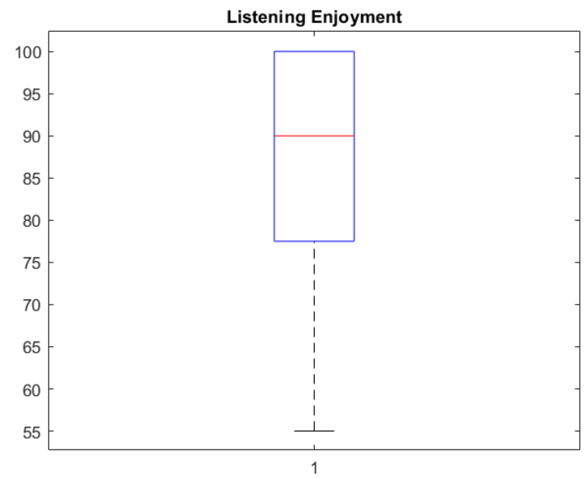

Fig. 17. Participants' usual enjoyment of music listening

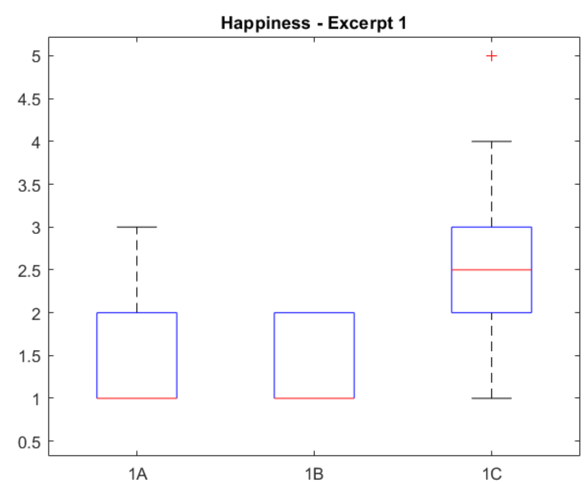

Fig. 18. Happiness values for the first music excerpt (*Darcy's Letter - Dario Marianelli - Pride and Prejudice - Sad*) with HL algorithm (1A), with HLHA (1B) and without processing (1C)

TABLE III

MEAN ($\bar{x}$) AND STANDARD DEVIATION (σ) OF DIFFERENCE OF VALENCE FOR HEARING LOSS SIMULATION (PROCESSING A) AND HEARING LOSS WITH HEARING AIDS SIMULATION (PROCESSING B)

	Processing A	Processing B
Excerpt 1	$\bar{x} = 1.70$ $\sigma = 1.86$	$\bar{x} = 1.41$ $\sigma = 1.44$
Excerpt 2	$\bar{x} = 1.67$ $\sigma = 1.72$	$\bar{x} = 1.75$ $\sigma = 1.60$
Excerpt 3	$\bar{x} = 2.17$ $\sigma = 1.75$	$\bar{x} = 2.58$ $\sigma = 1.88$
Excerpt 4	$\bar{x} = 1.91$ $\sigma = 1.62$	$\bar{x} = 1.91$ $\sigma = 1.56$

TABLE IV

MEAN ($\bar{x}$) AND STANDARD DEVIATION (σ) OF DIFFERENCE OF ENJOYMENT FOR HEARING LOSS SIMULATION (PROCESSING A) AND HEARING LOSS WITH HEARING AIDS SIMULATION (PROCESSING B)

	Processing A	Processing B
Excerpt 1	$\sigma = 28.52$ $\bar{x} = 47.25$	$\sigma = 27.80$ $\bar{x} = 43.25$
Excerpt 2	$\sigma = 28.33$ $\bar{x} = 35.58$	$\sigma = 26.74$ $\bar{x} = 39.08$
Excerpt 3	$\sigma = 29.85$ $\bar{x} = 37.00$	$\sigma = 28.24$ $\bar{x} = 49.58$
Excerpt 4	$\sigma = 24.58$ $\bar{x} = 45.75$	$\sigma = 29.68$ $\bar{x} = 46.00$

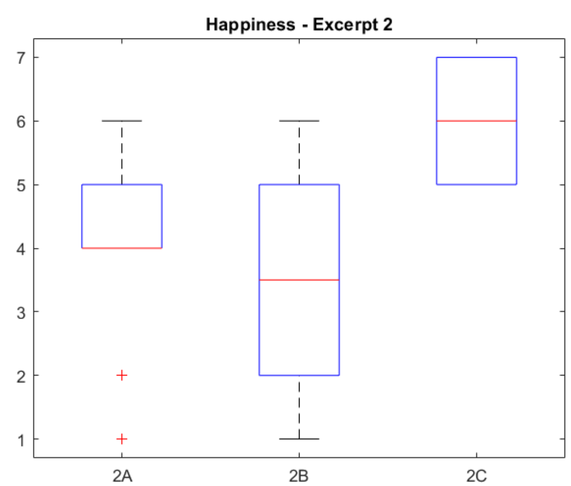

Fig. 19. Happiness values for the second music excerpt (*Grand Hotel Fox Trot - Life is Beautiful - Happy*) with HL algorithm (2A), with HLHA (2B) and without processing (2C)

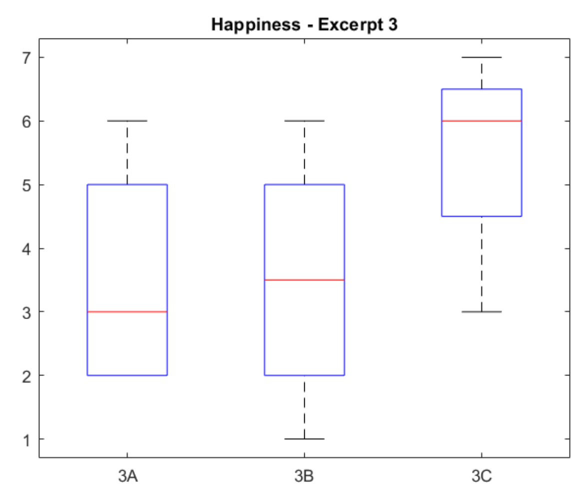

Fig. 20. Happiness values for the third music excerpt (*Papaya - Stelvio Cipriani - Happy*) with HL algorithm (3A), with HLHA (3B) and without processing (3C)

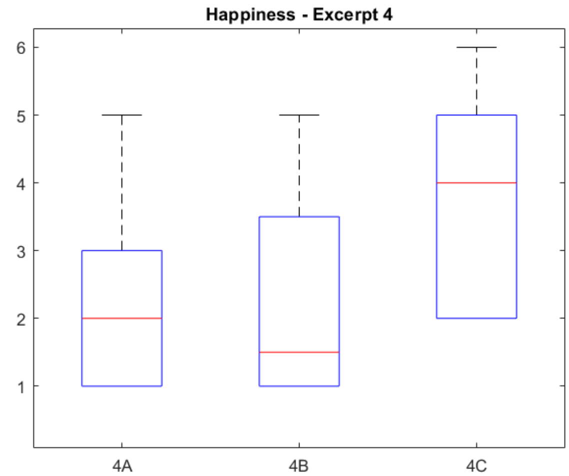

Fig. 21. Happiness values for the fourht music excerpt (*Song for Bob - Nick Cave & Warren Ellis - Sad*) with HL algorithm (4A), with HLHA (4B) and without processing (4C)

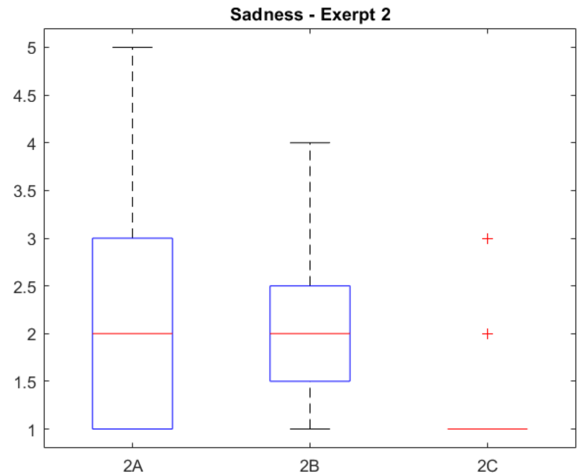

Fig. 23. Sadness values for the second music excerpt (*Grand Hotel Fox Trot - Life is Beautiful - Happy*) with HL algorithm (2A), with HLHA (2B) and without processing (2C)

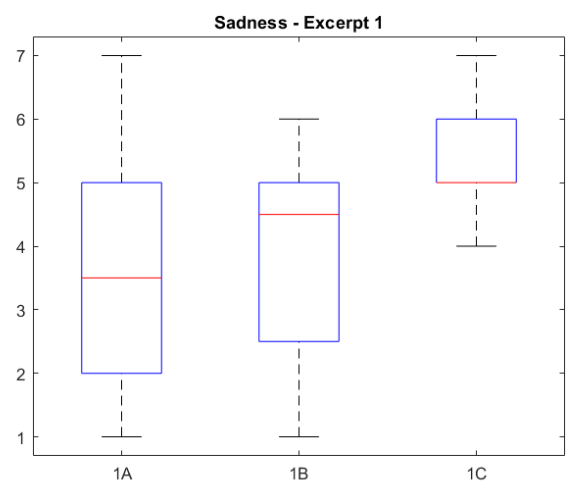

Fig. 22. Sadness values for the first music excerpt (*Darcy's Letter - Dario Marianelli - Pride and Prejudice - Sad*) with HL algorithm (1A), with HLHA (1B) and without processing (1C)

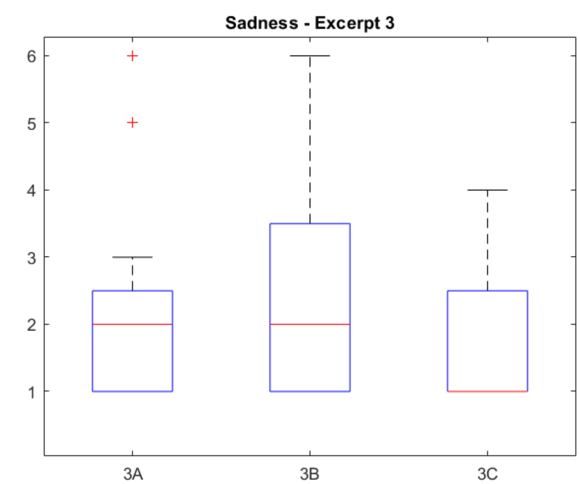

Fig. 24. Sadness values for the third music excerpt (*Papaya - Stelvio Cipriani - Happy*) with HL algorithm (3A), with HLHA (3B) and without processing (3C)

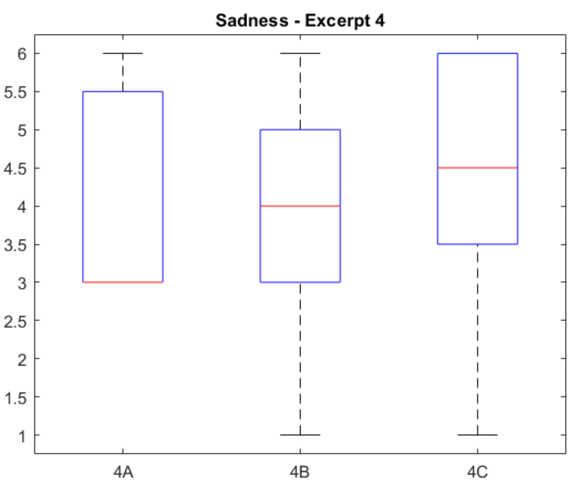

Fig. 25. Sadness values for the fourth music excerpt (*Song for Bob - Nick Cave & Warren Ellis - Sad*) with HL algorithm (4A), with HLHA (4B) and without processing (4C)