

Scale

Setup, Conduction and Analysis of Listening experiments

User Documentation

Release: R16-0726 (third release)

July 3, 2016

Contents

1	Introduction	4
1.1	What is Scale?	4
1.2	Tests implemented	4
1.2.1	Adaptive	4
1.2.2	Double-blind Triple-stimulus with Hidden Reference (ABC-HR)	5
1.2.3	ABX	6
1.2.4	MUSHRA	6
1.2.5	SAQI	6
2	Type of stimuli available	7
2.1	.wav stimuli (WAV)	7
2.2	.wav combined with the SSR (BINAURAL_SSR)	8
2.3	Contact	8
3	Installing and starting the program	9
3.1	Running the program for the first time	9
3.2	Running an already installed version	10
3.3	Running an installed version from another computer	10
4	Running the program - Overview	11
5	Running the program - Create	12
5.1	General configuration window	12
5.2	Set Instructions Window	12
5.3	Procedure specific options	13
5.3.1	Adaptive	13
5.3.2	SAQI	14
5.3.3	ABC-HR	15
5.3.4	ABX and MUSHRA	16
5.4	Stimuli configuration	17
5.4.1	WAV	17
5.4.2	BINAURAL SSR	18
5.4.3	Order and number of stimuli	19
5.4.4	Finish	20
5.4.5	Files created	20
6	Running the program - Perform	21
6.1	Test preparation	21
6.2	Subject login	22
6.3	Test Introduction Window	22
6.4	Scenario Selection Window	23
6.5	Trial Window	24
6.5.1	Adaptive Methods	24
6.5.2	ABC-HR method	25
6.5.3	ABX method	26

6.5.4	MUSHRA Method	27
6.5.5	SAQI method	27
7	Running the program - Analyse	29
7.1	Analyse Window - Overview	29
7.2	Analyse Window - Adaptive, ABC-HR, ABX and MUHSRA	30
7.3	Analyse Window - SAQI	30
7.4	Exported Data	31
7.4.1	Common fields	31
7.4.2	Adaptive	31
7.4.3	ABC-HR	33
7.4.4	ABX	33
7.4.5	MUSHRA	33
7.4.6	SAQI	33
8	Closing Scale	35
9	The Sound Scape Renderer	36
9.1	Introduction	36
9.2	Sound Scape renderer and Scale	36
9.3	Installation	36
10	Jack Audio Connection Kit	37

1 Introduction

1.1 What is Scale?

Scale is a software tool that covers the full chain of setup, conduction and analysis of psychoacoustic experiments. It offers several testing procedures and the interaction between researcher or subjects and the software is done via a graphical user interface (GUI) and does not require any programming skills. Test setups or results can be easily ported from one instance of the program to another. Thus everything is portable and exchangeable between different computers and researchers. The first version of *Scale* was presented at the DAGA Conference in 2013 in Merano [1], Italy and the second version was presented at the DAGA Conference in 2015 in Nürnberg [2].

1.2 Tests implemented

The initial version of *Scale* included a selection of frequently used test procedures like simple or transformed staircase adaptive procedures and double blind triple-stimulus with hidden reference. A description of the latter can be found in [1]. In the second version three additional procedures, ABX, MUSHRA and SAQI, are implemented. In this third version, no more tests have been included however, the possibility of adding new procedures haven been made possible.

1.2.1 Adaptive

Adaptive procedures aim to find a threshold of detection in the psychometric function of a determined dimension of a sound. Stimuli are presented and varied in one dimension. The amount of variation is increased or reduced depending on the preceding subject's responses and on the respective adaptation method. In *Scale* the adaptation is made using different staircase methods like simple staircase (1up/1down) on the one hand, and some transformed staircase methods such as 1up/2down or 2up/1down as described in [3] on the other hand. The threshold estimation can either vary depending on the ending conditions of the trial (limited number of runs or reversals) or on how the average is calculated. In this version the threshold estimation using *QUEST* [4] has been also added.

The tests have to be combined with a paradigm. *Scale* provides different paradigms which can be divided into two groups: n-AFC (n alternative forced choice) paradigms and Yes/No paradigms. In the n-AFC paradigms n intervals (as used in [5]) are presented to the subject. When n is equal to 2 the subject has to decide in which of the two intervals a designated signal is present. When n is greater than 2 the subject has to decide in which of the n intervals the presented stimulus is different. The sample assignation to the intervals of the n-AFC paradigms is always automatically randomised. When using a Yes/No paradigm only one interval that includes one or more sounds is presented to the subject. The subject has to decide whether the signal occurs within the presented interval or not.

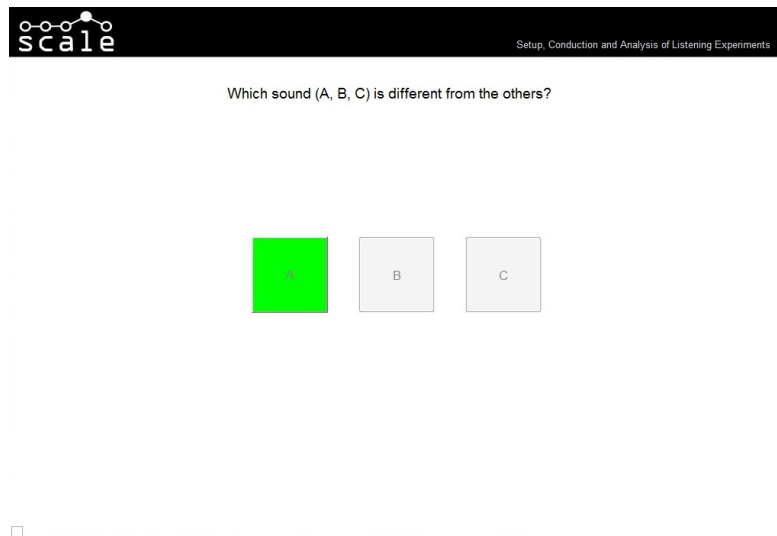


Figure 1: 3-AFC procedure trial window

1.2.2 Double-blind Triple-stimulus with Hidden Reference (ABC-HR)

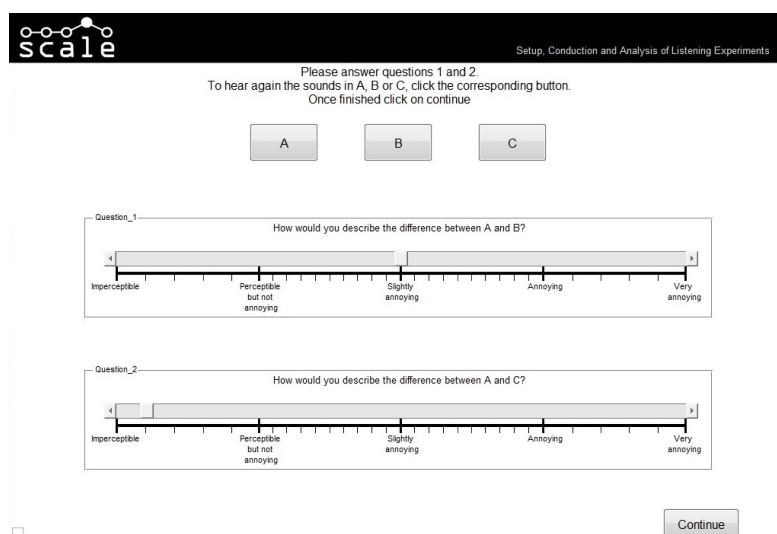


Figure 2: Double-blind triple-stimulus with hidden reference procedure trial window [6]

This procedure has become a standard in psychoacoustics and is used to assess small impairments between sound samples. In every trial, three stimuli are presented in three intervals ("A", "B" and "C"). The stimulus in "A" is presented as the known reference, the stimuli in "B" or "C" are randomly assigned, whereas one of them is a hidden reference and the other one is a sample which is varied in one determined dimension. After listening to the stimuli the subject is asked to assess the impairments between "A" and "B" and "A" and "C" using a rating scale. The rating is performed with a slider along a continuous scale with anchors. The number of grades in the scale as well as the text in the labels of every mark can be set. Thus all requirements to perform the test *"Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems"* described in the recommendations of the ITU [6] are met (see Figure 2).

1.2.3 ABX

The ABX test is a simplification of the *ABC-HR* test. In a trial of an *ABX* test, a subject is presented with two stimuli which are A and B, followed by a third one called X. After hearing A, B and X, the subject must select which of the stimuli in the intervals A or B is the same as in interval X. In this case, as opposite to the *ABC-HR*, no rating is done. In Scale, each scenario can contain many trials and for each trial, the position of the reference stimuli in A and B is randomized. Figure 3 shows a trial window of the test.

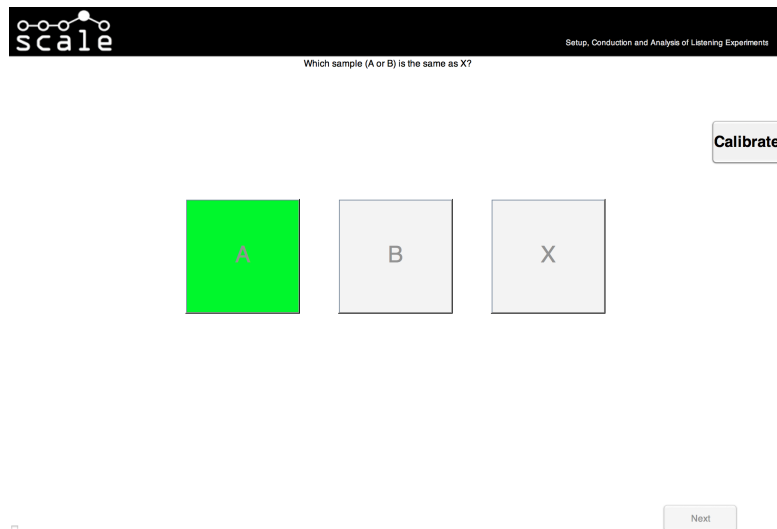


Figure 3: *Scale's* interface during an ABX test

1.2.4 MUSHRA

The aim of a multi-stimulus test with hidden reference and anchor (MUSHRA) test [7] is to rate global differences between several audio stimuli. All stimuli are presented in a single trial and have to be compared to a given reference. Each stimulus has a continuous scale (continuous quality scale) which goes from 0 (bad) to 100 (excellent), as shown in Figure 4. Reference and stimuli can be switched over instantly. The order of the stimuli is randomized and every trial has to include a hidden reference and an anchor.

1.2.5 SAQI

The spatial audio quality inventory (SAQI) test [8] has been specifically designed for the perceptual evaluation of virtual acoustic environments. In every trial a reference and a stimulus are presented together with 48 verbal descriptors of perceptual qualities that are assumed to be of practical relevance when comparing virtual auditory environments. Each descriptor comprises a rating scale with a pair of opposed adjectives in its scale ends. The subject's task is to compare the stimulus to a given or inner reference and give a rating for each perceptual quality.

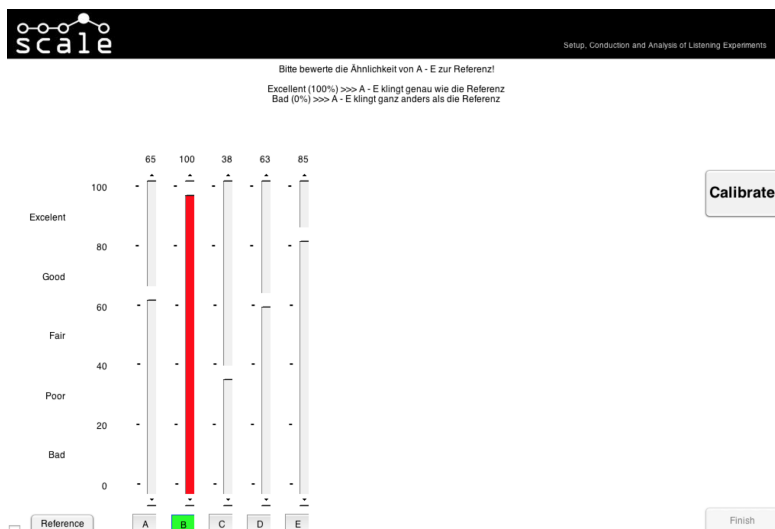


Figure 4: Scale's interface during a MUSHRA test

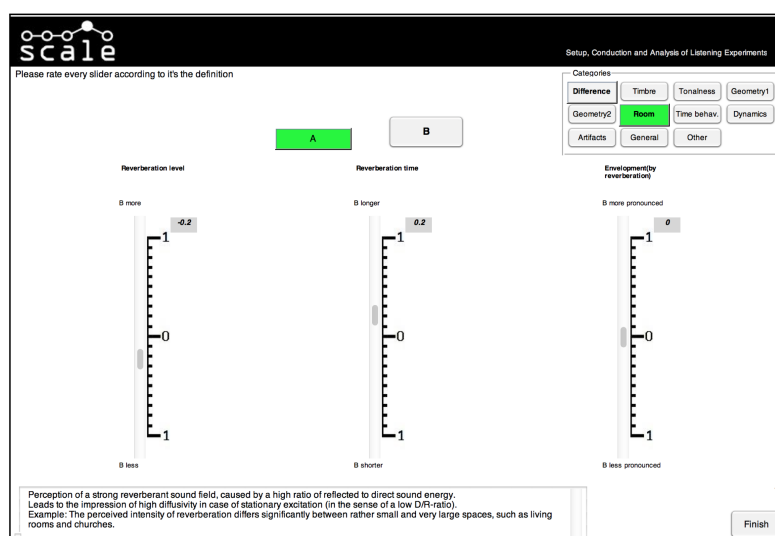


Figure 5: Scale's interface during a SAQI test

2 Type of stimuli available

In this version of Scale two different types of stimuli are available. Those are *.wav* stimuli and *.wav combined with SSR*. Other types of stimuli can be added by modification of the existing code files.

2.1 *.wav* stimuli (WAV)

The *.wav* stimuli are just audio files in the mentioned format which will be played by *Matlab* using the function `play` of the audioplayer objects.

2.2 .wav combined with the SSR (BINAURAL_SSR)

This type of stimuli are obtained by a combination of the rendering software Sound Scape Renderer (SSR) [9] and the .wav playback function of Matlab. As shown in Figure 6, *Scale* processes subject's inputs and operates the *SSR* using its network interface via TCP/IP protocol while the test is performed. The *SSR* runs in the background generating stimuli with the combination of the incoming audio signal, the tracker data and the corresponding head related impulse response (HRIR) or binaural room impulse response (BRIR) set.

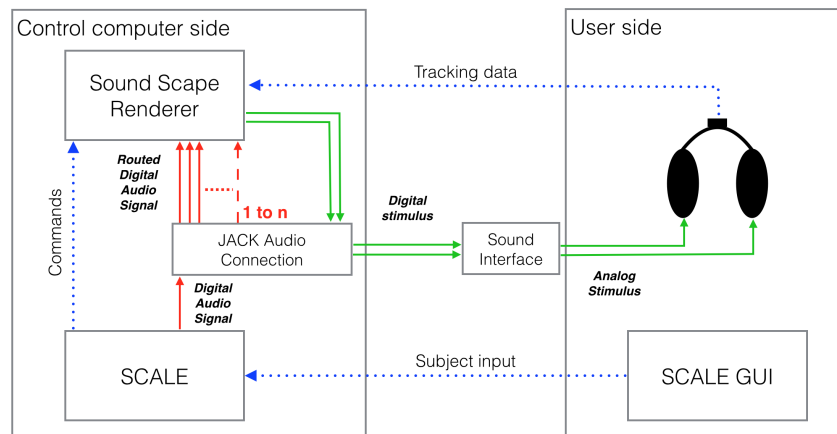


Figure 6: System architecture

2.3 Contact

For any further requests, technical support or bug reports feel free to contact:

- Arnau Vázquez-Giner, arnau.vazquez@gmail.com

For software updates and further information visit:

<http://www.audiogroup.web.th-koeln.de>

3 Installing and starting the program

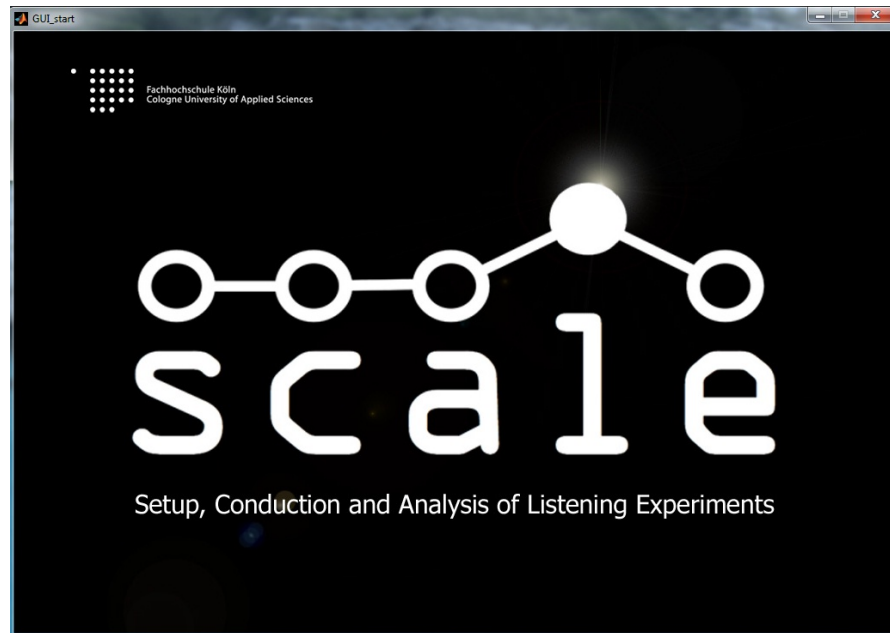


Figure 7: Program loading window

In this chapter, the instructions to start the program will be given.

3.1 Running the program for the first time

1. Download the item **Scale** from the following link:
<http://www.audiogroup.web.th-koeln.de>
2. Extract the file **Scale.zip** to a directory of your choice.
3. Once extracted open *Matlab* and set the current *Matlab* folder to the directory where you extracted the program (you should see a .m file with the name Scale and the folders Scale_code_files and Scale_user_folder).
4. Type **Scale** on the command line and press enter.
5. Scale will open and the window presented in Figure 7 should appear. After this point, you will not need to interact with *Matlab* anymore, but do not close it!

3.2 Running an already installed version

To run an installed version of Scale follow the steps mentioned in the last section starting at point 3.

3.3 Running an installed version from another computer

In case that you want to import your version of Scale, with your already configured test or results, just copy the folder Scale into the new computer. All the implemented tests and results will be automatically imported.

**Please note that if you are using the SSR stimuli type, you will need to adapt the .asd files*

4 Running the program - Overview

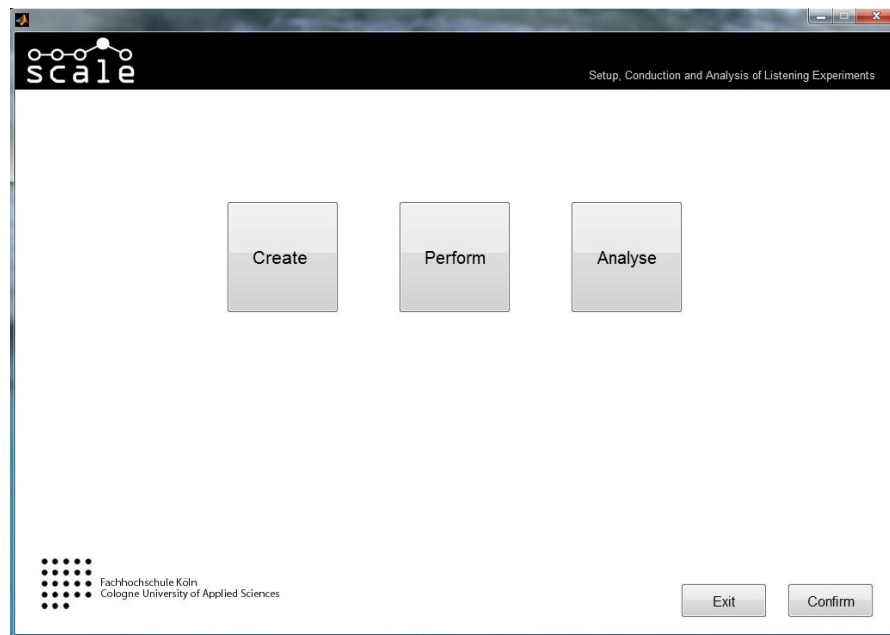


Figure 8: Main window

When the program has finished loading the **Main Program** window appears. This window appears also when the user at some point interrupts the program or decides to go back. The window contains three buttons (**Create**, **Perform** and **Analyse**) referring each to one of the basic features. When pressing any of the buttons, it will turn green and one of the actions described in the next lines will take place.

Create: Select this option to create a new test.

Perform: Select this option to conduct a test. A list with the available tests in the `scale_program_folder\tests` will appear. Select from the list the test to be performed and press **Continue**.

Analyse: A list with the available tests in the `scale_program_folder\tests` will appear. Select one test from the list and press the button **Continue**.

**If the test does not contain any result files, the action will be avoided!*

5 Running the program - Create

Results from two similar tests can be affected by several factors. On the one hand there are the external factors like the devices used along the playback chain and their settings or the characteristics of the experiment's location. On the other hand there are factors like e.g. the sequence of the presented stimuli or the information given to subjects in advance of or during the test. Taking the latter into account, different settings are adjustable in *Scale*. The settings are accessed in the creation module where different configuration options are available depending on which type of procedure is selected.

In this section the steps to follow during the creation of a test are described.

5.1 General configuration window

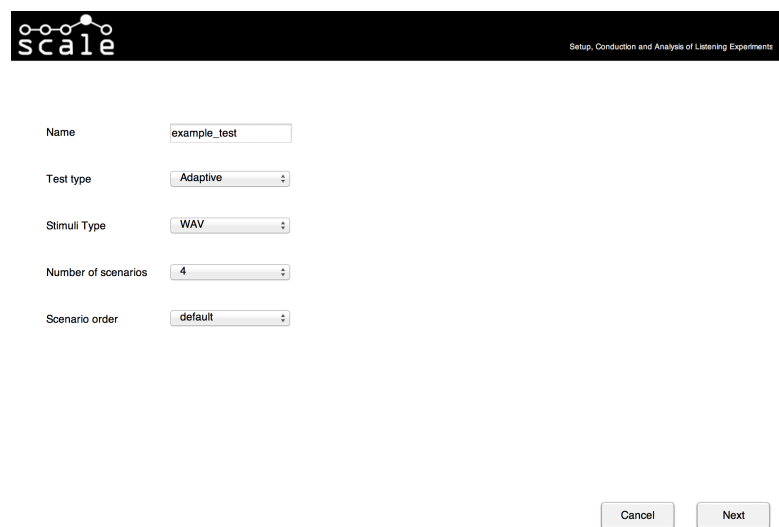


Figure 9: General configuration window

Figure 9 shows the general configuration window. This window is the first one to appear after selecting the option create in the main program window. Here the name of the test, its procedure, the stimuli type, the number of scenarios and the presentation order of them can be set. The order of the scenarios can be set to default or random. Default, means that scenario 1 will be performed in the first place, scenario 2 in the second place and scenario n in the nth place.

5.2 Set Instructions Window

Scale allows the researcher to write the instructions and commands that are shown to the subject during the test performance. The second window of the test configuration, presented in Figure 10, is used for setting those instructions. On the one side, the instructions and/or information that the subject will see at the beginning of their test is written on the panel **Instructions on test start** (upper panel). On the other side, the instructions that appear when the subject performs a run are written on the panel **Instructions on trial window** (lower panel). Although the instructions can be the same

for each scenario, the process must be done separately for each trial by clicking on the scenario buttons. To go faster, you can copy paste the text from one scenario to the other.

Figure 10: Set Instructions Window

5.3 Procedure specific options

In this window the specific configuration for each test type can be done. Depending on the procedure, the window will have different parameters. In some cases no specific parameters will be needed. Following, the window is presented and the options are explained separately for each procedure.

5.3.1 Adaptive

For the configuration of the adaptive tests, the configuration window will look like in Figure 11. The **general configuration panel** (upper panel) includes the parameters which are common for all adaptive tests. Those are, the number of repetitions, the paradigm and the adaption method.

Repetitions - Select the maximum number of times that a subject can perform a repetition in one run.

Paradigm - Select the paradigm (2AFC, 3AFC or Yes/No)

Adaptation method - Select the adaption method from one of the available (1up1down, 2up1down, 1up2down and QUEST).

Depending on the adaption method chosen, you will have to configure one of the panels below. For *1up1down*, *2up1down*, *1up2down* the panel to configure is the lower-left one (**staircase**). For *QUEST* it is the lower right one (**QUEST**).

Figure 11: Specific options for the adaptive tests

The options for the staircase methods are the following:

Number of runs - The trials end after a certain number of performed runs. If the option is left to 10000, this ending criteria will be ignored.

Number of reversals - The trials end after a certain number of reversals (changes from correct to false or false to correct subject decisions). If the option is left to 10000, this ending criteria will be ignored.

Activate half step width - If selected, next played stimuli after a correct/false decision will be the stimuli situated two positions above/below the preceding one. After the first reversal, the steps will be reduced to only one position.

The options for the *QUEST* test are: **tGuess**, **tGuessSD**, **pThreshold**, **stepSize**, **range**, **nTrials**, **beta**, **delta** and **gamma**. Those are described in the documentation of the *QUEST* Toolbox in [4].

5.3.2 SAQI

For the configuration of the *SAQI* tests, the configuration window will look like in Figure 12. The window includes a panel with general procedure options (lower panel) and the panels with the qualities that the researcher wants to get the ratings for.

Reference Type - In a *SAQI* test, the reference can be a stimuli (external) or a concept (internal). An internal reference example could be the sound of a guitar, so the subject would have to compare some stimulus with the idea of the sound of a guitar that he has in his memory instead of hearing a real guitar.

Figure 12: Specific options for the SAQI tests

Slider steps - In a *SAQI* test, every quality is evaluated using a slider. Here the number of steps of the sliders can be selected.

test language - Two languages are available for this type of test, those are English and German. The names of the qualities and the descriptions of them appearing during the test will be shown in the language selected.

To select a quality just click on the button next to its name. You can select all of them at once by clicking the **select all** button. You can also deselect qualities by clicking on their buttons again. Clicking on the button of a quality will also display an explanation of itself on the lower-right corner of the window.

5.3.3 ABC-HR

For the *ABC-HR* tests, the window will look like in Figure 13. Following each parameter, button and element in the window is explained.

Instructions - Here can be written the instructions that the subject will see right at the top of the question.

values... (*input text field*) - In those fields must be written the labels that the subject will see under the slider and the scale picture when performing the rating. The label in the left will correspond to the value 0 of the slider while the label in the right will correspond to the *max* value of the slider.

The screenshot shows the 'scale' software interface for configuring listening experiments. It features two identical panels for 'Question_1' and 'Question_2'. Each panel contains an 'Instructions for question' text area, five 'values' input fields, a 'Preview' button, a 'Scale picture' dropdown menu (currently showing 'aaa_no_scale.jpg'), a 'Values' input field (set to 5), and a 'Slider size' control with '+' and '-' buttons and a 'default' button. The interface is titled 'scale' and 'Setup, Conduction and Analysis of Listening Experiments'. At the bottom, there are 'Cancel' and 'Next' buttons.

Figure 13: Specific options for the ABC-HR tests

preview - A short preview of two seconds showing how will look like the question in a real test situation. This option is interesting to see if the text in the labels will be correctly placed in the real test.

scale picture (*list*) - This list let the researcher chose the image that will be displayed between the slider and the labels. When a picture is selected it is automatically displayed.

values (*list*) - Select the number of values that will appear in the rating scale. The number of fields will be automatically modified, and so their positions under the slider that will be always symmetrically distributed. The value in the left corresponds always to 0 while the value in the right, *max*, will correspond to the number selected in values, *n*, minus 1 ($max = n - 1$).

slider size (+/-) - Press this button to increment/decrement the width of the slider thumb.

slider size (default) - Press this button to set the width of the slider thumb to a default value depending on the number of values.

5.3.4 ABX and MUSHRA

For both *ABX* and *MUSHRA* tests, no specific options need to be configured. The window will look in both cases like in Figure 14.

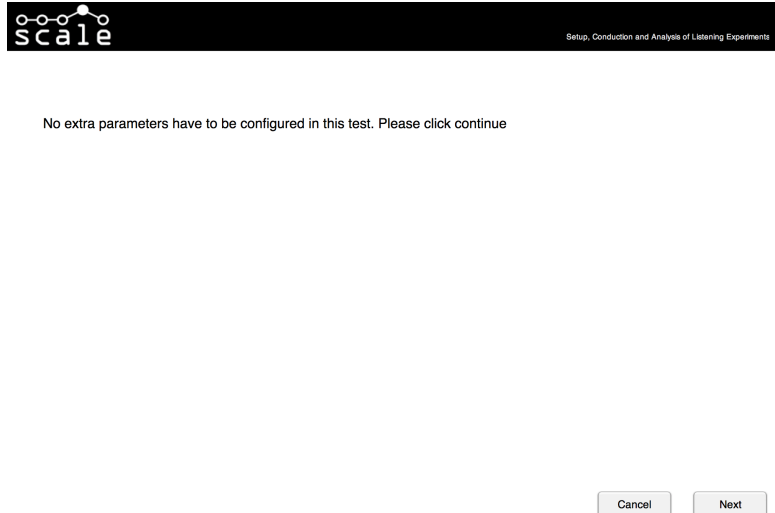


Figure 14: Specific options for the MUSHRA and ABX tests configuration

5.4 Stimuli configuration

In this version of *Scale* two types of stimuli are available, those are simple *.wav* samples and *.wav* samples processed by the SSR. The process of adding the stimuli is different for each type of stimuli, for this reason both procedures are explained separately. At the end of the section, some hints about the number of stimuli to be added in each type of test are given.

5.4.1 WAV

Figure 15 shows the window where the *.wav* samples are added. The first step is to add the samples into the program folder so they are selectable for the test. This is done by pressing the button **add** located inside the **Available samples** panel (right). When the button is clicked a dialog window opens and the user can select the samples from the directory. Samples can be selected one by one or several at the same time. This action can be performed as many times as the user wants. Once the samples are loaded they will appear in the list of the same panel.

The panel samples contains the slots where the samples can be added to the test. To add or delete slots the buttons **erase last slot** or **add new slot** at the lower right corner of the panel can be used. The lists of samples inside the slots, correspond to the list on the panel of available samples. For each slot we will select one sample. This process must be made for each scenario. To change the scenario, use the scenario buttons on the lower-left panel **Scenario**. More information about the samples order in each type of test is given at the end of this section.

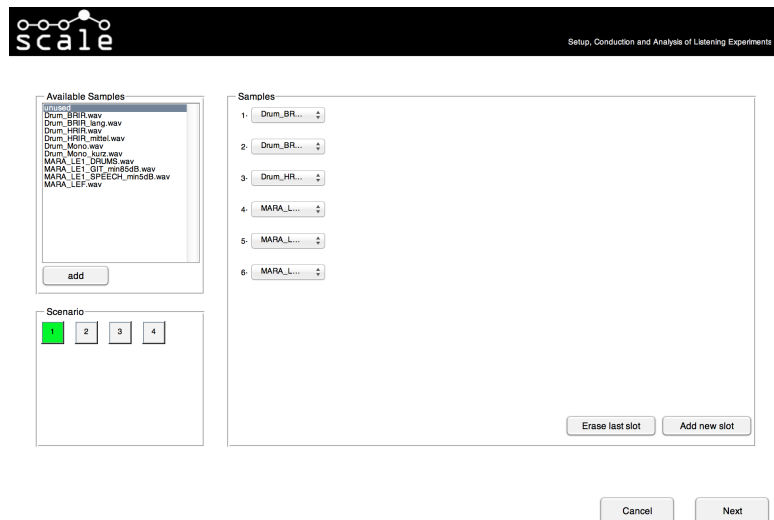


Figure 15: .wav stimuli setting window

5.4.2 BINAURAL SSR

BINAURAL SSR refers to the type of stimuli that use the *SSR* renderer. Two parts are involved in the addition of the stimuli in this case. The first part is the setting of the filters and the second the setting of the .wav file.

Figure 16 shows the window where the filters used for the stimuli are added. We will add one filter for each stimulus. On the right of the window we can select from the options available, the sample rate of the filter files and also the tracker that will be used. On the main panel (left) the user will add the filters on the slots. Each slot corresponds to one stimuli. The variable **volume** refers to the volume that the source will have in the .asd file and **x** and **y**, the position on the scene. Since the *SSR* will work only in *BRIR* mode, **x** and **y** will not affect the sound at all. To select the filters, we must click the button **select**.

After clicking, a dialog window will open where the folder containing the *BRIR* files can be selected. Once selected, the files will be selectable on each slot. To add or delete slots, the buttons **Erase last sample** and **Add sample** can be used. More information about the samples order in each type of test is given at the end of this section.

The second part of the stimuli configuration consists in adding the .wav files. Only one .wav file will be added for each trial or scenario.

After adding the filters, and .asd file (scene file used by the *SSR*) is created. This file contains the paths to the filters, for this reason, the filters should not be moved from their folders after a test is created. If the test needs to be played in another computer, the .asd file should be modified and the new paths should be added by hand. The .asd files are placed in *Scale/Scale_user_folder/Tests/<nameOfTheTest>*.

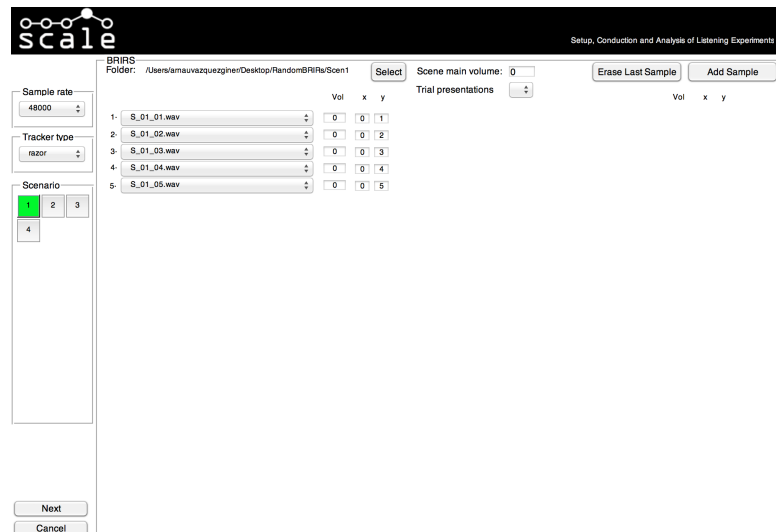


Figure 16: Setting of the filters for the BINAURAL SSR stimuli type

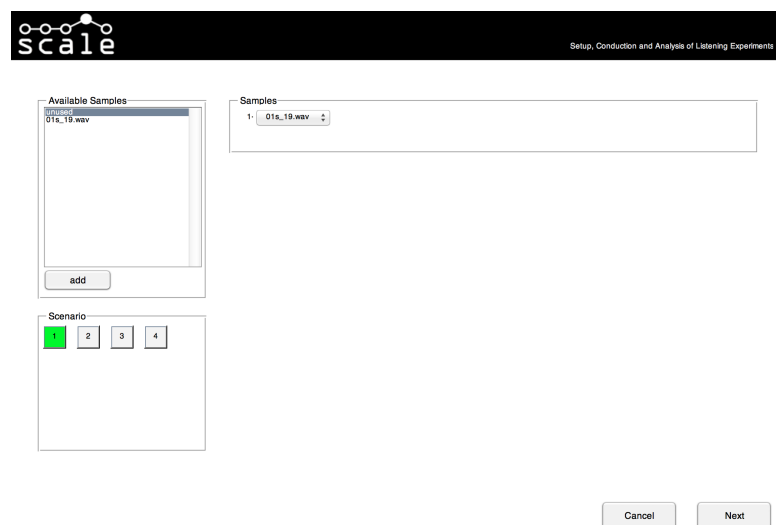


Figure 17: Setting of the .wav files for the BINAURAL SSR stimuli type

5.4.3 Order and number of stimuli

Adaptive - An adaptive test can have a maximal of 50 stimuli. The first stimuli on the first slot will correspond to the reference stimuli, the second will correspond to sample most similar to the reference and so on. The last slot will correspond to the sample with the biggest difference to the stimuli.

SAQI - If internal reference was selected, only one stimuli will be added for each trial. If external reference was selected, the first slot will contain the reference and the second the stimulus to be compared.

ABC-HR - In ABC-HR tests, each run compares two different stimuli. The slots will represent those pairs. Slot 1 and 2 will be the first pair, slot 3 and 4 the second, 5 and 6 the third and so on. The presentation of the pairs will be randomized differently for each subject, however, in the results analysis they will be presented in the order they were added.

ABX - In ABX tests, each run compares two different stimuli. The slots will represent those pairs. Slot 1 and 2 will be the first pair, slot 3 and 4 the second, 5 and 6 the third and so on. The presentation of the pairs will be randomized differently for each subject, however, in the results analysis they will be presented in the order they were added.

MUSHRA - In a MUSHRA test, a maximum of 15 stimuli can be presented in a single trial. The reference will be placed in the first slot and the rest of the samples in the remaining slots. The reference will be played when clicking on the reference button during the trial, but will also be presented in one of the sliders like the rest of the stimuli. The presentation of the stimuli will be randomized differently for each subject, however, in the results analysis they will be presented in the order they were added.

5.4.4 Finish

Once the stimuli have been added, the test configuration is finished and the user just needs to click on the button **next** on the window presented in Figure 18. If the user cancels the configuration of the test or closes *Scale* before this point, all the files created during the steps mentioned in this section will be deleted.

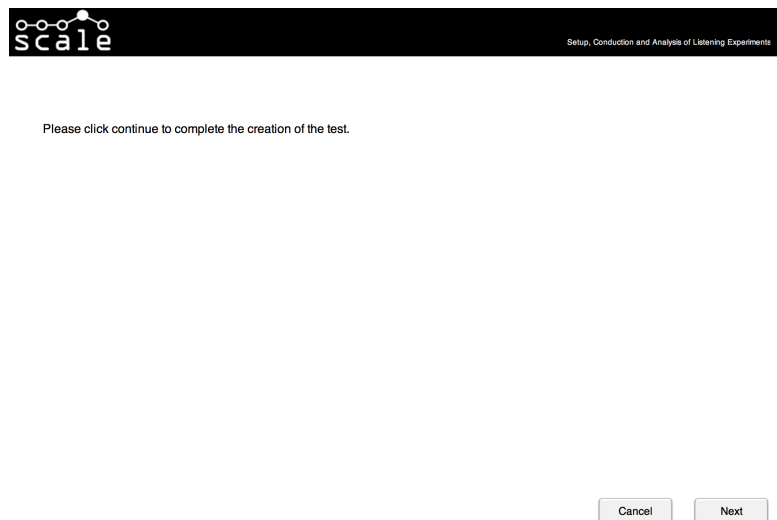


Figure 18: Completion of test creation

5.4.5 Files created

Once a test is configured, a folder with the name of the test is created inside *Scale/Scale_user_folder/Tests*. This folder contains the struct **test_info** which contains the information of the test, it also contains the *.wav* samples that will be played, the *.asd* files in case of *BINAURAL_SSR* stimuli type and a folder called *results* containing the files corresponding to the results. One result file is created each time a test is performed.

6 Running the program - Perform

In this section the process of performing a test is explained step by step. Specific information about how the stimuli are selected and presented and which are the tasks of the subjects will be given for every type of test.

6.1 Test preparation

The first window to appear is dedicated to prepare the system for the test. If you are using *.wav* stimuli type, nothing will be done at this point.

If you are using the *SSR*, Figure 19 will open. On this window, the processes of turning on the *SSR* and achieving a connection are done. The user can also click on the option **SSR Visible** if he wants to see the *SSR* user's interface during the test. Sometimes it is useful to leave it on to control that everything is working. In case no connection is not achieved, try several times before restarting the program. Please note that the *JACK* interface should be turned on before *Matlab* is started. In case it was not you will get an error message. The first time that *Scale* is used on a new system, this window will ask the sudo password, this is needed during the *SSR* starting process. The password will be saved and it will not be asked again.

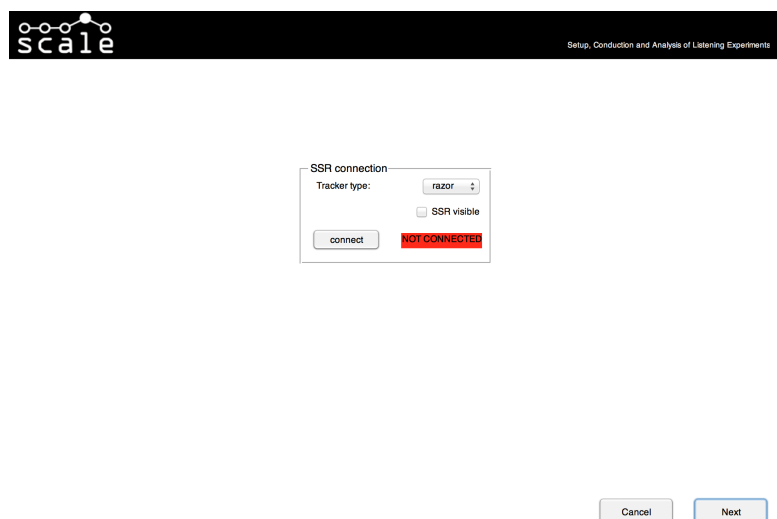


Figure 19: System preparation window

6.2 Subject login

In order to perform a test, the subject must be logged in, to log in, introduce the user data in the required fields on the window shown in Figure 20. If the data is correct, the subject will be logged in and the test can be started by pressing the button **next**. In case that the subject did already started the test, a new button will appear on the lower-left corner called **restart**. If **restart** is pressed, the results of that subject will be deleted and the test will be restarted. Please note that only a registered subject can perform a test. To register a subject, use the button **Register subject** on the main program window shown at the beginning of the manual (Figure 8).

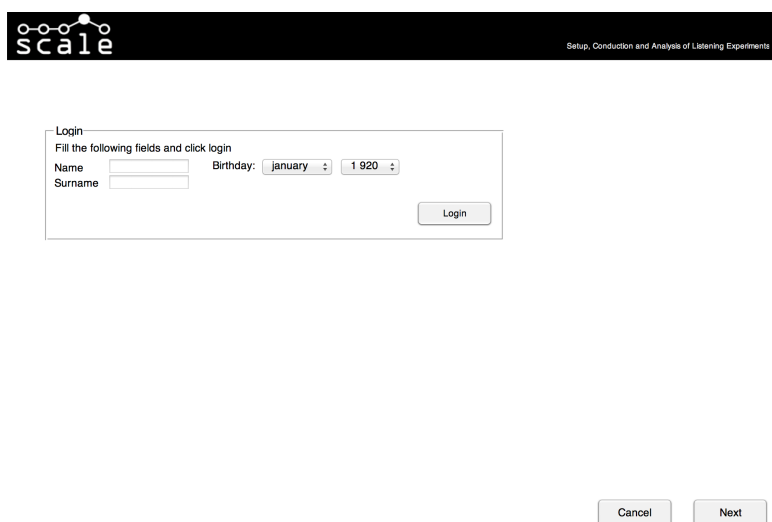


Figure 20: Subject login window

6.3 Test Introduction Window

Figure 21 shows the **Introduction** window of the test. As an example of how the introduced text look like, the text written in the field general instructions of Figure 10 is shown. In this window it would be useful a description of how the subject must interact with the **Scenario Selection** window, see Figure 22, due to the fact that the scenario order can be different from one test to another.

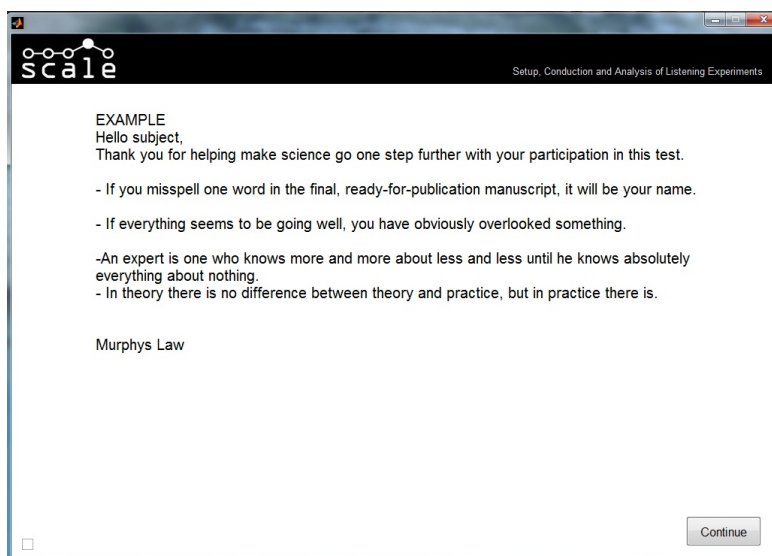


Figure 21: Test Introduction window

6.4 Scenario Selection Window

Figure 22 represents the **Scenario Selection** window. Subjects interact with this window at the beginning of the test and every time they complete a scenario.

In this example the test consists in six scenarios; tests with a different number of scenarios will make the window look different because a specific distribution of the buttons in the window is performed in every case. The text written in the window explains to the subject the meaning of the different buttons and what they have to do with them. To perform a scenario, the subject must click **next**. The scenario that he is going to perform is the one that has its button number painted in green. Buttons corresponding to completed scenarios are unenabled and the word "completed" is written in them. When all scenarios are completed, button **Next** will change its text to **Finish**. The selection of the next scenario is done automatically by *Scale*.

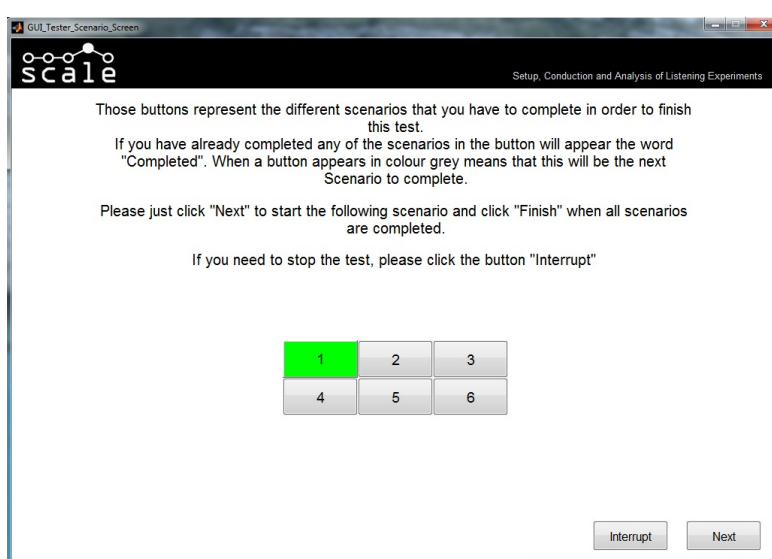


Figure 22: Scenario Selection window

The button **Interrupt** stops the test. The information of the results until this point will be saved and the test will be able to be continued later. It is only possible to interrupt a test in the pause between scenarios. If a test is interrupted during the performance of an scenario, the information will be saved only until the last completed scenario.

6.5 Trial Window

Each type of test uses a different trial window, following each procedure is individually explained together with screenshots.

6.5.1 Adaptive Methods

In the adaptive methods stimuli that differ in a certain parameter are presented. The subject is aimed to find that difference. The decision can vary depending on the adaptation method and on the paradigm; *Scale* offers three paradigms which are **Yes/No**, **Two Alternative Forced Choice (2AFC)** and **Three Alternative Forced Choice (3AFC)**. In a Yes/No paradigm the subjects are asked directly if they could find that mentioned difference after hearing a single interval. In alternative forced choice paradigms they must find the difference indirectly comparing two or more intervals. In all cases, when the subject give a positive answer, next run should present a situation where the difference is lesser. This process should be performed until the point where the subject reaches the border between noticeable and unnoticeable. This point corresponds to the psychophysical threshold. Depending on the adaption method, the stimuli used and other specific settings, more or less runs will be needed to reach that threshold. Figure 23 presents an example of the *Yes/No* and *3-AFC* paradigms.



Figure 23: Adaptive Method Run windows

The sample to play in each run will depend on the answer given and on the adaptation method. Following, a list explaining the different adaption methods is given. For a better understanding it will be assumed that the last presented stimuli was the one corresponding to sample n .

- **oneUp1Down** - The first sample to be compared to the reference will be sample max , that means the last sample of the samples vector. Reference will be always sample 1. If the answer is correct, next sample will be $n-1$. If the answer is wrong next sample will be sample $n-1$.

- **oneUp2Down** - The first sample to be compared to the reference will be sample *max*, that means the last sample of the samples vector. Reference will be always sample 1. If both current and previous answer are correct, next sample will be $n-1$. If current answer is correct but previous answer is wrong, next sample will be n . If current answer is wrong next sample will be $n+1$.
- **twoUp1Down** - The first sample to be compared to the reference will be sample *max*, that means the last sample of the samples vector. Reference will be always sample 1. If current and previous answer are wrong, next sample will be $n+1$. If current answer is wrong but previous answer is correct, next sample will be n . If current answer is correct next sample will be $n-1$.
- **QUEST** - Reference will be sample 1. The samples to play will be calculated using the *QUEST* algorithm which varies in function of the parameters given by the user, explained in [4].

Every time the subject gives an answer (*positive/correct or negative/wrong*) that is different from the previous, a **reversal** occur. For a non *QUEST* test, there are three ways to finish a scenario and two of them can be set by the researcher in the **ending criterion** options:

- **Number of runs** - The scenario will end after a certain number of runs.
- **Number of reversals** - The scenario will end after a certain number of reversals occur.
- **Sample index above the limits** - The scenario will end when the samples having indexes 0 or $n+1$ are required. This means that the subject has not only achieved the maximum or the minimum, but surpassed it. This would mean that the threshold is out of range.

6.5.2 ABC-HR method

ABC-HR follow the "double blind principle". Those tests are aimed to compare two samples: the original sample, that works as a reference, and a modified sample. To make this comparison, the subject hears three intervals (*A*, *B* and *C*). Interval *A* contains always the reference. Intervals *B* and *C* contain one the reference and the other one the modified. Nor the researcher or the subject will know which one is each. This assignation is always randomly made for every run.

Subject hears first interval *A* followed by interval *B* and then interval *A* followed by interval *C*. After hearing them for the first time, he can listen them again an unlimited number of times just clicking buttons **A**, **B** or **C**. After hearing the intervals, the two questions must be answered. First question should be "How would you evaluate the difference between samples A and B?" and second question, "How would you evaluate the difference between samples A and C?". The answer is performed positioning the slider thumb in a certain place in the horizontal scroll. After the user has clicked at least once in every slider the button *Continue* will set to enable and once pressed the run end. The recommendation *ITU-R BS.1116 for the evaluation of small impairment in coded*

Figure 24: ABC-HR Method Run window

audio signals [6] uses this method under certain conditions. Figure 24 shows an example of a run in an ABC-HR test using the mentioned recommendation.

6.5.3 ABX method

The ABX test is a simplification of the *ABC-HR* test. In a trial of an *ABX* test, a subject is presented with two intervals, which are A and B, followed by a third one called X. After hearing A, B and X, the subject must select which of the stimuli in A and B resembles to X and click on it. In this case, as opposite to the *ABC-HR*, no rating is done. In *Scale*, each scenario can contain many runs and for each run, the position of the reference stimuli in A and B is randomized. Figure 3 shows a trial window of the test.

Figure 25: ABX Method Run window

6.5.4 MUSHRA Method

This procedure is aimed to compare multiple samples with medium or big impairments between them. The advantage of this method is that it enables to compare multiple samples at the same time having also a reference and an anchor. The reference is presented two times, one as a known reference and the other as a hidden reference. The anchor must be also hidden. The rating is performed with vertical sliders placed on the top of every stimuli button.

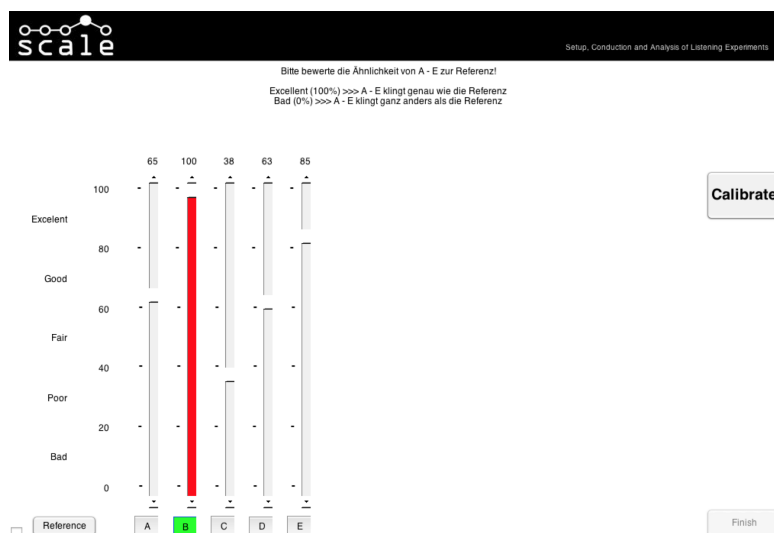


Figure 26: MUSHRA Method Run Window

Figure 26 shows a run window of a **MUSHRA** test, the buttons containing the letters A to O represent the different stimuli and the button *Reference* the reference stimuli. When the subject clicks on one of the buttons, the stimuli contained in that letter will be played. At the right side of the screen there are some options related to the playback such as *Pause*, *Resume*, *Stop* and *Loop*. Also a slider shows the playback instant. On the left side of the window the rating scale is presented and also the rating grades. All sliders have anchors on their side every ten points and at the top a text field show the position of the slider thumb. After the user has clicked at least once in every slider the button *Continue* will set to enable and once pressed the run end.

6.5.5 SAQI method

The spatial audio quality inventory (*SAQI*) [8] test has been specifically designed for the perceptual evaluation of virtual acoustic environments. In every trial a reference and a stimulus are presented together with 48 verbal descriptors of perceptual qualities that are assumed to be of practical relevance when comparing virtual auditory environments. Each descriptor comprises a rating scale with a pair of opposed adjectives in its scale ends. The task of the subject is to compare the stimulus to a given or inner reference and give a rating for each perceptual quality.

Figure 27 shows a run window of a *SAQI* test. The test window may differ depending on which qualities were selected by the researcher for evaluation. On the upper-right

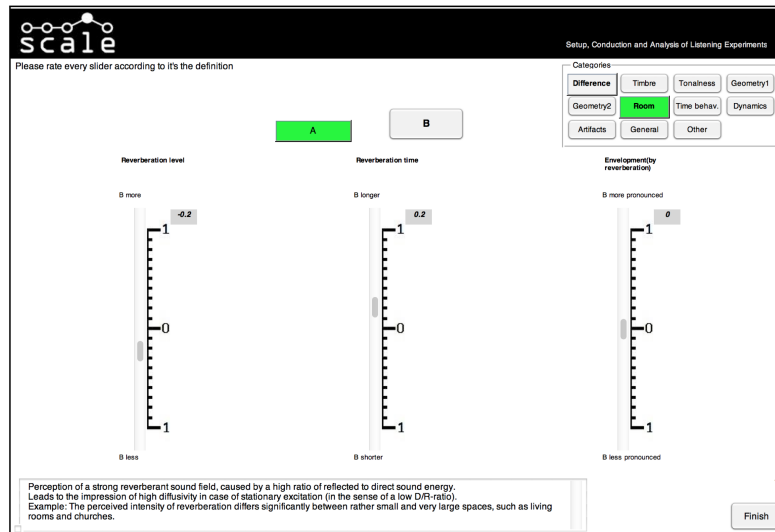


Figure 27: SAQI Method Run window

corner, the panel **Categories** contains a button for each quality category. When pressing on a button, the test window will show all the sliders corresponding to the qualities of the selected category. When the test starts, the category **Difference**, which contains a single slider, is shown. At this point, the other buttons are blocked. This is made on purpose because if no difference is found, all the qualities will be automatically set to 0 and the run will be finished by pressing the button **Finish** (bottom-right corner). In case the slider of the Difference quality is moved, the rest of the Category panel buttons will become active.

When clicking on a slider, a description of the quality corresponding to it will be shown at the bottom of the window. In order to switch from reference to tested sample click on the buttons **A** or **B**. The stimuli are looped, so the subject hears the sound continuously, if the sound needs to be stopped, this is done by clicking on the active sample button **A** or **B**, in green. If the test uses an internal reference, then only one button is present.

7 Running the program - Analyse

Results obtained in tests are analysed at the researcher's discretion. Hence, the data can be represented in many different ways. In *Scale* some of the most common methods of representation have been implemented in order to facilitate this process. The analysis module offers a graphical representation of the collected data and results can be exported to MATLAB Structs (*.mat*) and MATLAB Figures (*.fig*).

7.1 Analyse Window - Overview

The analyse of the results is made through the window represented in Figure 28. This window can vary on the test procedure, but most of the settings are common for all of them.



Figure 28: Example of the analyse window

The common options, Figure 29, include the **Results** panel and the **Scenario** panel. In the results panel, the available results sets from a certain test are presented as a list. To select a result set just click on it and then on the + button, you can add more results by repeating the process or select **all** of them by clicking **select all**. To not include a single result set, select it on the **Selected** panel and click the - button, to deselect all result sets click on **Clean all**.

In the panel **Scenario** the scenarios to be seen can be selected. Each scenario has a color assigned and this color will correspond to the color of the line representing the data of it in the graphic.

The panel **Save Data** contains two buttons. If the user click on **plot** the graphic can be saved as a *Matlab* figure, if **all data** is clicked, the results will be exported in a *Matlab* struct file which can be used outside *Scale*.

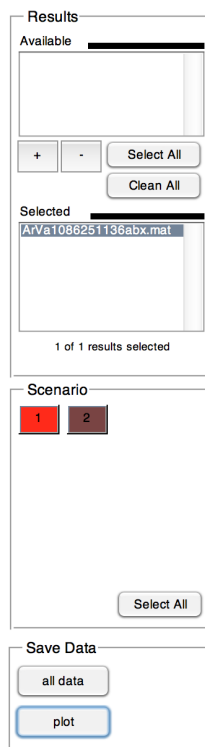


Figure 29: Analyse window common options

7.2 Analyse Window - Adaptive, ABC-HR, ABX and MUHSRA

The window to observe the results is common for the Adaptive, ABC-HR, ABX and MUHSRA procedures, see Figure 30. In the panel **options** three options can be selected.

Single - For every subject and for every scenario a line will be drawn with his obtained value/s.

Average - The values obtained on the all the result sets are averaged and a single value/s are presented.

Combined - Both information shown in **Single** and **Average** options is plotted.

7.3 Analyse Window - SAQI

The **options** panel include several buttons, each one corresponding to a category of qualities, and a list of the visualization possibilities (single, average and combined), see Figure 31. When observing the categories, please take in account that only 4 categories can be plotted simultaneously.

Single - For every subject and for every scenario a line will be drawn with his obtained value/s.

Average - the values obtained on the all the results sets are averaged and a single value/s are presented.

Combined - Both information shown in **Single** and **Average** options is plotted.



Figure 30: Analyse window for Adaptive, ABC-HR, ABX and MUHSRA procedures

7.4 Exported Data

In this section the exportation of the data will be explained. The results data is exported in a single struct called **data_subs**. This struct has several fields which some of them are common for all type of tests. The common data will be now explained and the test specific results will be later explained specifically for each test type.

7.4.1 Common fields

Figure 32 shows all the fields of the results struct. The fields inside the red frame, are common for all type of tests, each field is a cell array. The cell arrays have as many positions as result sets. In the case of the figure, there are two sets of results, for this reason, the length of the cell arrays is 2. The not common part of the results will be saved in the fields *scen<number of scene>*. The contain or the data class of this *scene* structs will be following explained separately for each test.

7.4.2 Adaptive

For the adaptive tests, every trial includes two different values. The first one is the threshold value *Scen<trial_number>_tValue* which in this case is the value of the last stimuli played and then a vector with all the stimuli played in the trial saved in *Scen<trial_number>_runValues*. This vector has a fixed length of 100, and if the trial was performed in less than 100 runs, the extra positions will be filled with *NaN*. In the figure, the struct contains only one set of results, if it contained more sets of results, *Scen<trial_number>_tValue* and *Scen<trial_number>_runValues* would be a vector and a vector of vectors with a length equal to the number of results. This is shown in Figure 33.

**If the staircase method QUEST is used, two more fields are added which are the *Scen<trial_number>_st_dev* containing the standard deviation in each trial and the *Scen<trial_number>_QuestStruct* containing the QUEST information struct in each trial. This is shown in Figure 34.*

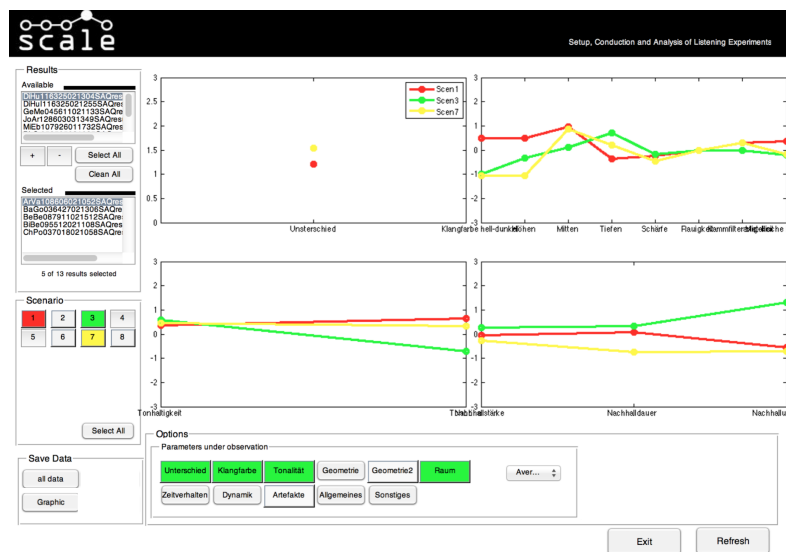


Figure 31: Analyse window common options for SAQI procedure

Variable Editor – data_subs			
Stack: Base			
Select data to plot			
Field	Value	Min	Max
subjectID	<2x1 cell>		
resultsCode	<2x1 cell>		
Surname	<2x1 cell>		
Name	<2x1 cell>		
Gender	<2x1 cell>		
Country	<2x1 cell>		
Experience	<2x1 cell>		
Audio_Knowledge	<2x1 cell>		
Hearing_Problems	<2x1 cell>		
Scen1	<2x7 double>	NaN	NaN
Scen2	<2x8 double>	NaN	NaN

Figure 32: Exported results struct

Audio_Knowledge	<1x1 cell>		
Hearing_Problems	<1x1 cell>		
Scen1_tValue	11	11	
Scen2_tValue	10	10	
Scen1_runValues	<1x100 double>	NaN	NaN
Scen2_runValues	<1x100 double>	NaN	NaN

Figure 33: Exported results struct

Scen1_tValue	0	0	0
Scen2_tValue	10	10	10
Scen1_runValues	<1x100 double>	NaN	NaN
Scen2_runValues	<1x100 double>	NaN	NaN
Scen1_stDevValue	0	0	0
Scen2_stDevValue	0.7921	0.7921	0.7921
Scen1_QuestStruct	<1x1 struct>		
Scen2_QuestStruct	<1x1 struct>		

Figure 34: Exported results struct

7.4.3 ABC-HR

In an *ABC-HR* test every trial includes a single vector of values called *Scen<trial_number>*. Those values correspond each to the comparison between a pair of stimuli. For example, the first position of the vector *Scen1* will show the comparison between stimuli added in slot 1 and stimuli in slot 2. The second position, will show the comparison between stimuli in slot 3 and slot 4 and so on. Please note that during the test, the presentation of the stimuli pairs is randomized but here is ordered again. The value itself can have a range from *-nValues* to *+nValues*. The negativity or positivity of the value depends on the answer of the subject. For example if C was the same as A and the subject rated "very different" to the question asking the similarity between A and C, the value would be equal to *-nValues*.

7.4.4 ABX

In an *ABX* test every trial includes a single vector of values called *Scen<trial_number>*. Those values correspond each to the comparison between a pair of stimuli. For example, the first position of the vector *Scen1* will show the comparison between stimuli added in slot 1 and stimuli in slot 2. The second position, will show the comparison between stimuli in slot 3 and slot 4 and so on. Please note that during the test, the presentation of the stimuli pairs is randomized but here is ordered again. The value itself can be 1 or 0. If it is 1 it means that the subject guessed correctly (the reference was pointed correctly to A or B), if it is 0 it means that the subject guessed wrong.

7.4.5 MUSHRA

In a *MUSHRA* test every trial includes a single vector of values called *Scen<trial_number>*. The vector has as many positions as stimuli presented in the trial. The order of the ratings is not related to the order that the stimuli were presented in the trial window because they are randomised for every subject. The order of the results corresponds to the order that the stimuli were added.

7.4.6 SAQI

In a *SAQI* test every trial includes a single vector of values called *Scen<trial_number>*. The vector has 48 positions and the rating for each quality is written in this vector. Since the user can select qualities to not be asked, the positions of the not asked qualities will not be taken in account. Figure 35 show the correspondence between the vector positions and the quality names.

-
- | | |
|-----------------------------|----------------------------------|
| 1 Difference | 25 Envelopment(by reverberation) |
| 2 Tone color bright-dark | 26 Pre-echoes |
| 3 High-frequency tone color | 27 Post-echoes |
| 4 Mid-frequency tone color | 28 Temporal disintegration |
| 5 Low-frequency tone color | 29 Crispness |
| 6 Sharpness | 30 Speed |
| 7 Roughness | 31 Sequence of events |
| 8 Comb filter coloration | 32 Responsiveness |
| 9 Metallic tone color | 33 Loudness |
| 10 Tonalness | 34 Dynamic range |
| 11 Pitch | 35 Dynamic compression effects |
| 12 Doppler effect | 36 Pitched artifact |
| 13 Horizontal direction | 37 Impulsive artifact |
| 14 Vertical direction | 38 Noise-like Artifact |
| 15 Front-back position | 39 Alien source |
| 16 Distance | 40 Ghost source |
| 17 Depth | 41 Distortion |
| 18 Width | 42 Tactile vibration |
| 19 Height | 43 Clarity |
| 20 Externalisation | 44 Speech intelligibility |
| 21 Localizability | 45 Naturalness |
| 22 Spatial disintegration | 46 Presence |
| 23 Reverberation level | 47 Degree-of-liking |
| 24 Reverberation time | 48 Other |

Figure 35: List of SAQI qualities

8 Closing Scale

The correct way to close *Scale* is by clicking **Exit** button in the **Main program** window. The **Main program** window is accessed when a task is completed, for example when a test is created, or when some actions are cancelled, for example when a test is interrupted. In order to avoid errors and to provide a good functionality the **close** button in the upper-right corner of the windows is under normal conditions deactivated; it would be a problem for example when a subject accidentally closed a test. For this reason a little "trick" has been implemented. In the bottom-left corner there is a little input text field. When the number 22 is written in this field and the close button is pressed the program can be closed.

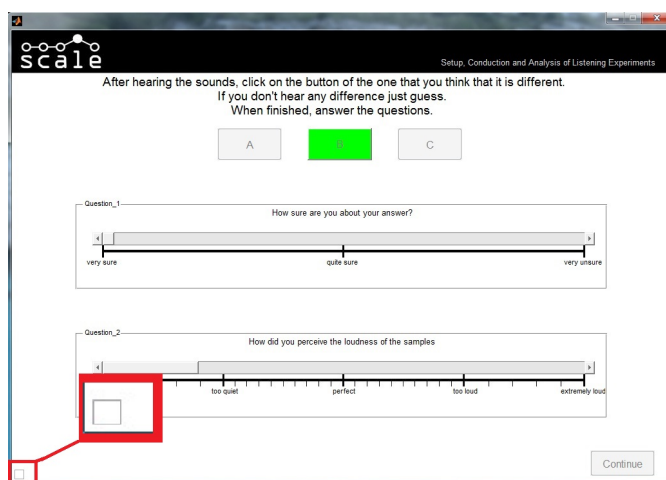


Figure 36: Closing input text field

9 The Sound Scape Renderer

9.1 Introduction

The *SoundScape Renderer (SSR)* [9] is a tool for realtime spatial audio reproduction, which includes several rendering methods. It is free open source software running on *UNIX*-based systems and using the *JACK* [10] audio framework. The *SSR* can be controlled either using a graphical user interface (GUI) or a TCP/IP network interface. Sound source positions and other attributes represented in a spatial audio scene, as shown in Figure 37, can be imported or exported using audio scene description files (.asd) [11].

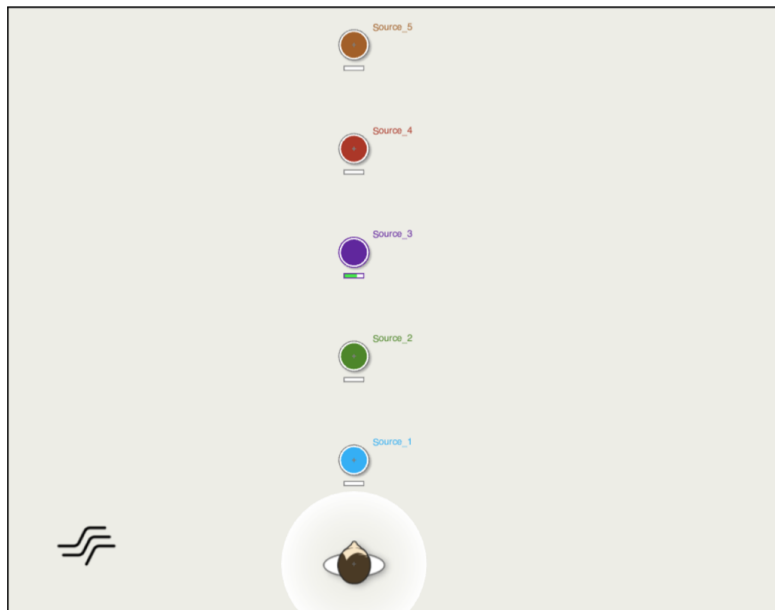


Figure 37: SSR user interface

9.2 Sound Scape renderer and Scale

As shown in Figure 38, *Scale* processes subject's inputs and operates the *SSR* sending *XML* messages through its network interface via *TCP/IP* protocol while the test is performed. The *SSR* runs in the background generating stimuli with the combination of the incoming audio signal, the tracker data and the corresponding head related impulse response (HRIR) or binaural room impulse response (BRIR) set. The extension *Scale-SSR* uses the BRIR mode because it allows the use of arbitrary impulse response datasets in the same scene.

9.3 Installation

The *SSR* needs to be installed in the system in order to be used. Please follow the next steps to complete the installation:

1. Click the link user manual on the site <http://www.spatialaudio.net/>.

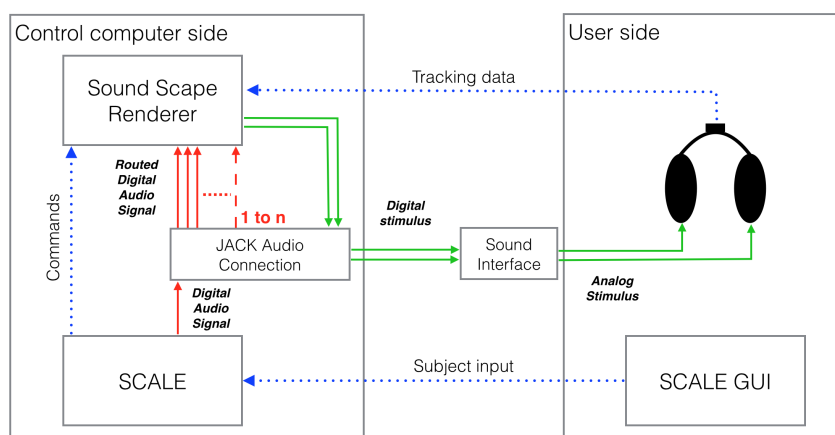


Figure 38: SSR System architecture

2. In the user manual there are the instructions about downloading and installing the SSR.
3. The Scale-SSR extension has been developed using the version 0.4.2 of the SSR. Please note that while older versions should not have any problem, earlier versions could.

10 Jack Audio Connection Kit

The SSR uses the Jack Audio Connection Kit to make the routing of the sound and for this reason it needs to be installed on the system. Remember that the Jack audio connection must be running before you start Matlab. If it is not the case, it will not be detected. To install the software go to the site <http://www.jackaudio.org/downloads/> and download the Jack1 for OSX version 0.124.1. or greater.

References

- [1] A. Vazquez, “Scale, a software tool for listening experiments,” *Proceedings of the DAGA*, 2013.
- [2] A. Vazquez, “Scale, conducting psychoacoustic experiments with dynamic binaural synthesis,” *Proceedings of the DAGA*, 2015.
- [3] H. Levitt, *The City University of New York*. PhD thesis, University of Wollongong, 1970.
- [4] A. Watson and D. Pelli, “Quest: a bayesian adaptive psychometric method,” *Perceptive Psychophysics*, vol. 33(2), pp. 113–120, 1983.
- [5] E. Zwicker and H. Fastl, *Psychoacoustics. Facts and Models*. Springer, 2nd ed., 1999.
- [6] I.T.U., “Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems,” *ITU-R BS*, no. 1116-1, 1997.
- [7] I.T.U., “Method for the subjective assessment of intermediate quality level of coding systems,” *ITU-R*, vol. BS, pp. 1534–2, 2014.
- [8] A. Lindau, V. Lepa, *et al.*, “A spatial audio quality inventory (saqi),” *Acta Acustica united with Acustica*, vol. 100, pp. 984–994, 2014.
- [9] M. Geier, J. Ahrens, and J. Spors, “The soundscape renderer: A unified spatial audio reproduction framework for arbitrary rendering methods,” *Proceedings of the 124th Convention of the AES*, no. 7330, 2008.
- [10] P. Davis *et al.*, “Jack audio connection kit.” <http://jackaudio.org/>, October 2013. Version 1.9.
- [11] M. Geier, J. Ahrens, and S. Spors, “Asdf: ein xml format zur beschreibung von virtuellen 3d-audioszenen,” *Proceedings of the DAGA*, 2008.