

Sound Design and UX: The Dynamic Audio Application Guide

Luiz Roberto Carvalho^(✉) and Alice T. Cybis Pereira

Federal University of Santa Catarina, Florianópolis, Brazil
{semprecarvalho, acybis}@gmail.com

Abstract. Aiming to expose guidelines and procedures that lead to an efficient implementation of sound elements into interfaces, and given the need to develop a formal understanding of audio applications in interactive systems, this study presents the Dynamic Audio Application Guide. From the proposed methodology by Jesse James Garrett, the Dynamic Audio Application Guide was created, which aims to systematize and simplify the process of creation, production and implementation of sound into interactive systems. The guide is intended to highlight the important role played by sound in immersion and interaction processes.

Keywords: Sound design · Dynamic audio · Interactive sound · Hypermedia design

1 Audio for Interactive Devices: Sound is Information

The interactive systems development is a complex process, mainly evidenced by its interdisciplinary nature. To communicate, the interface needs to engage the user in its own dynamic potentialities. It is important to point that an interface is not only a visual stimulus, but a combination consisting of sound, image and hypertext. In 1996, [9] pointed that the use of technological resources in hypermedia was restricted to the visual modality. According to the author, little was invested in audio¹, an element that brings quality to content, facilitates the accessibility of information and makes them more attractive. Ten years after, according to [10], there has been no significant progress in relation to sound in interfaces. The author affirms that in the design market there is an overvaluation of visual communication, and because of its limitations, products and services can often present inconsistencies when other sensory properties are relevant, as in the case of sound and tactile information. While the tactile process and haptic² requirements are gradually beginning to develop in the interface field, sound and its cognitive functions remain in a largely unexplored territory.

¹ Audio comes from latin word *audio*, first person of the verb *audire*, that means listening. In a broader sense, refers to all wave phenomena occurring within the audible frequency range, i.e. between 20 Hz and 20 kHz. When in the acoustic field, audio waves is known as sound.

² The haptic adjective means “relating to the touch”, “synonymous of tactile,” and is derived from the Greek word *aptikó* that means “touch-sensitive”. It is the optical counterpart for visual, and sound for hearing.

The use of sound in interactive environments has not advanced at the same rate of the graphical elements, and consequently, it has been placed in low priority [1]. Reference [8] indicate that this over-emphasis on visual displays has constrained the development of interactive systems that are able to make better use of the auditory modality. Non-musical sounds have been accepted as technology sub-products, rather than being exploited by their intrinsic value. As a result, a world polluted acoustically has been experienced since the industrial revolution. However, there are already enough scientific knowledge and technology to start a way of *thinking about the sound* as one of the main dimensions of the environments that people live - whether physical or virtual. This means overcome the sound barrier as cultural noise and promote a targeted approach to the sound as *information* - a fact that already occurs with the visual elements.

2 DAAG: Dynamic Audio Application Guide

The game industry has proven that the knowledge of how audio works in an interactive environment is crucial for a proper engagement and immersion experience [4]. From this perspective, in order to meet the need to understand, identify and classify the existing modalities when using sounds in interactive environments, and contextualize their use, the Dynamic Audio Application Guide (DAAG) was developed. Starting from Garret’s User Experience Design approach [6], the guide was contemplated with features that intend to develop interfaces with a sound design committed to the user experience concepts. The guide intends to start a formal understanding of the principles that constitute the application of audio in interactive systems, starting from a better understanding of the sound production flow in these platforms (Fig. 1).

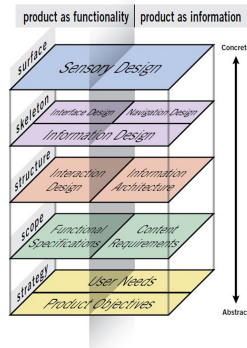


Fig. 1. Garret’s user experience design diagram [6]

The DAAG aims to bring the sound design concepts from the game industry, the user experience design approach, and put them together. Formed by five layers, the DAAG consists of a system of blocks of information that advances in steps aligned with the overall progress of the interface project. The guide is intended to be used in a

way to offer a unique approach to a proper management of sound elements in interface projects. From the exposure of guidelines and procedures to an effective implementation of sound features in interactive environments, the DAAG aims to systematize and simplify the process of creation, production and implementation of sounds in interfaces (Fig. 2).

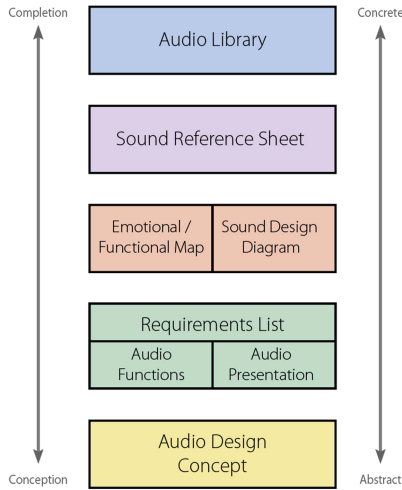


Fig. 2. The dynamic audio application guide diagram

The DAAG focuses on highlighting the important role played by sound in immersion and interaction processes. In interactive environments, the presence of dynamic audio³ intensifies the user immersion and empowers additional cognitive processes, driving the interaction experience into a different scenario. An experience surrounded by hypertext, image and sound is more complex and complete, since images, sounds, music, oral and written language, all these forms of expression are mixed in the same message. Once the interface reaches the users in distinct senses, their cognitive levels should be properly accentuated. In this context, the development of DAAG aims to explore efficient methods to make the sound design an essential part of interface projects. It is believed that the study results will open a debate of a yet low researched area, allowing its scientific basis to be useful in academic and professional practices, by providing a new tool to hypermedia designers.

2.1 The “Audio Design Concept” Layer

Every experience that man experiences is fundamentally perceived through the sense organs. In interactive systems, this involves consider which of the five senses (sight,

³ A more dense discussion about the dynamic audio can be found in [2].

hearing, touch, smell, and taste) are possible to apply in the interface and how this should be done. In this context, [7] warns that the use of sound is not something to be added as a late element in the project, given that it plays critical roles in important areas of the interface, such as the creation of *climates*, *atmospheres*, and to drive narrative (Fig. 3).



Fig. 3. The *audio design concept* layer

The first step in the DAAG workflow is to create an *Audio Design Concept* document. Create a list of possible sounds to the project is the first step in the interface sound design project. This involves a precise reading of user needs and product objectives, and based on such references, the audio design concept document should be made, pointing out the possibilities and limitations of using sounds. It is a simple text, in order to define a first communicative and aesthetics intention about what sound can bring to the interface, in an overall manner. During the rise of DAAG layers, this document will gradually become the *Sound Reference Sheet* (see in item 2.4), which will conduct the creation and application of the sounds in the interface. Therefore, this first document should be reassessed and reorganized as the interface project gradually progresses.

It is crucial to mention that in the initial phase of the interface project, even counting only with preliminary information about user needs and product objectives, many decisions can be made to begin the sound design work, thus ensuring that the audio will play an important role throughout the project. In fact, the audio design document has its importance since it enables designers to start thinking about sound at the beginning of the interface project.

2.2 The “Requirements List” Layer

Significant improvements in the sound design field are given when the sound is analyzed in a broad perspective with respect to their content *characteristics*, *form* and *function*. This DAAG layer has an objective to promote a targeted approach to audio as information, in a way that becomes possible to use the sound stimuli to transmit messages systematically. Acoustic signals must be explored in order to maximize the communicative effects of the interface (Fig. 4).

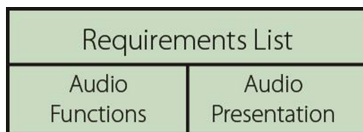


Fig. 4. The *requirements list* layer

When identifying all types of formats (text, audio, video, image) associated with a specific content of the interface, it is possible to determine what will be needed for its production [6]. In this way, the same content block can be presented in different formats and also different content can be displayed using a single media (text only, for example). However, when the range of presentation formats are expanded, the user can decide which kind of content to interact with. Since users have distinct cognitive models, the parallel use of texts, images, videos and sounds offers more pronounced interaction possibilities. This is a solution aligned with the principles of user experience design.

At this stage of DAAG, a *Requirements List* must be formatted. In this document, the sounds that will be in the interface should be classified according to their *Audio Presentation* possibilities, being divided in *dialogue*, *background music* and *sound effects*. This implies that designers will consider the sound feature (through the human voice, musical compositions and specific sound effects) as means of transmitting information relating to interface's content blocks. These sound elements should also comply with specific functions in the interface, so that their *Audio Functions* are fully met. To fulfill these specifications, the sound element shall corroborate to specific functional aspects, exercising *structural*, *narrative*, *immersive*, *aesthetic*, *kinetics* functions, as pointed out by (quote). Through the Requirements List layer, it is intended to define which kind of content will be transmitted through sounds, what is the nature of these sounds, and how these sounds will contribute to achieve the interface goals.

2.3 The “Maps and Diagrams” Layer

The quality of a sound can be considered high if an acoustic event is perceived as an information carrier and processed in a way that an specific meaning is extracted from it. From this perspective, sound designers are information architects, since they amplify the meaning of the messages (Fig. 5).

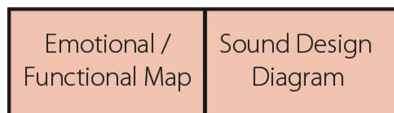


Fig. 5. The *maps and diagrams* layer

In the *Maps and Diagrams* layer the sound is considered in the context of the environment's dynamic interactions, more specifically, in the interactions that occur between different interfaces and connect large blocks of content, i.e., the *macro interactions*. At this stage, the contents belonging to the Audio Requirements' List (containing elements such as narration, dialogue, music, and sound effects) start to be added in the architecture diagram, also called interface's *hypermap*⁴, and according to

⁴ The hypermaps are produced from an organization of the interface information and become the navigation system array. The hypermap aims to show graphically, in the more systemic possible way, the information architecture and the various gateways and inter-relations thru content links [5].

[6], called *visual vocabulary*. As the other kinds of contents (images, texts) are being distributed to form the interface architecture, sounds should also follow this pattern. With the architecture diagram, it becomes possible to plan the audio behavior by setting up, for example, which audio tracks will be used in a group of graphical interfaces without the need for sound interruption in screen changes, and which audio tracks will be used as a transition zone between one interface and another.

This chart containing the interface ramifications allows defining, for example, about the use of music and ambience tracks as a connection between two distinct interfaces, or that indicate the opening or closing of a specific content block. The absence of sound - silence - can also transmit information to the user, for example, that some specific task was completed and the user can already get out of a certain area, progressing by the available links. A sound interruption may indicate a change in the interface direction, and the continuous use of music in disparate interfaces can help to signal the same theme of a specific content (Fig. 6).

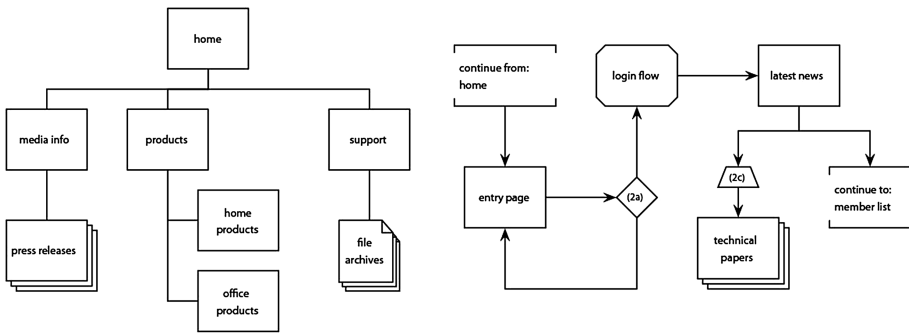


Fig. 6. Garret's visual vocabulary [6]

In this context, starting from the architecture diagram, a *sound architecture diagram* should be designed and conceived as an *interface feature*⁵. A feature usually indicates which areas of the hypermedia possess attributes that should be further implemented by the programming team, so, the sound architecture diagram is a feature. The end result should be the *sitemap*, which unifies the hypermap, the sound architecture diagram, and the whole structure of the interface, showing the dynamic spaces and connections between its content blocks. At this project step, the sounds begin to take a closer characteristic of how they will be in the final interface. After properly allocated on specific areas of the *sitemap*, it becomes possible to visualize how sounds will work between the interface sections and nodes. The creation of these references allows the sounds to behave according to the proposed dynamic for the project.

⁵ In interactive environments, the *feature* concept is related to the characteristics and functions of the interface, when analyzed as software. It is commonly used to describe systems and programs that require some kind of programming language.

In addition with the *sound architecture diagram* definitions, an emotional/functional map should be created. An *emotional map* is focused in the sense that the interface has an inner narrative, and is capable of enabling different emotional reactions from its users during its navigation. A *functional map* is useful in interfaces that has a task managing behavior (e-mails readers, governments spreadsheets), to graphically indicate the amount of tasks that the users should accomplish during navigation. The emotional/functional maps are useful since they systematize the interface according to acts, chapters or segments, and define the upcoming events that the users should face according to the interface's intensity patterns, or task completion levels. Thus, it becomes possible to identify which interface areas require a more precise cognitive support to help users to achieve their objectives (Fig. 7).

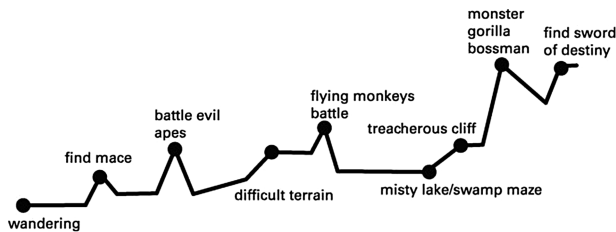


Fig. 7. Collins' emotion map, pointing tension's intensity patterns [3]

Successful interfaces are those which users can immediately identify the relevant material in its arrangement [6]. Emotional and functional maps indicate important points, helping to decide, for example, which interface area should be more emphatically projected. Through the systematic use of sound in the critical points of the map, it is possible to infer deeper meanings to the interface events. This happens to the extent that users can realize that there is a set of messages, actions and specific tasks that are in evidence and require more attention.

As the project advances thru layers, new design approaches can arise, suggesting that new sounds can be added, and the ones that were previously planned can be removed. It is vital that sounds meet the overall structure of the project, and this implies to reassess previously made decisions, and take new directions at any time.

2.4 The "Sound Reference Sheet" Layer

The sound architecture diagram offers a broad, systematic view of the interface's sound design. In this DAAG layer, a detailed document called *sound reference sheet* will show how this broader vision will be fulfilled individually, on each interface, in every interaction node. It's a document that defines the behavior of each audio element in the interface, that is, how the sounds will be integrated into the interactive environment. To this end, some questions should be readily raised. Sounds are merely reactive or will be triggered by a specific user action? The sound will transmit some kind of perceptible information, alerting, indicating something on a specific region of the interface? These

and other questions about the nature of the interaction that sound will perform should be made at this project stage (Fig. 8).



Fig. 8. The *sound reference sheet* layer

The sound reference sheet is a document that makes possible to organize and create the interface soundtrack. It is a document that synthesizes the interface's sound design management, and has a major importance to the acquisition (recording, editing and mixing) and implementation of the sounds in the interface. The sound reference sheet allows an overview of the sounds that will be integrated in the interface, and also provides a possibility to reassess the previously made decisions, in order to refine them. This document will guide the programming staff, providing a detailed script to inserting sounds in the interactive environment. To fully comply with these premises, the sound reference sheet must be formulated with two main items, namely: (a) soundtrack description; and (b) interaction rules.

The *soundtrack description* is a detailed description of what should be the sound itself. Since sounds were already sorted in categories such as voice, ambience music and sound effects, it is possible to schedule the recording sessions. It is recommended to choose a short name and increasing numbers for every sound that will be used in order to control its production and properly organize them in the interface's audio library.

Interaction Rules. The second parameter to be determined is about how the sounds will interact with the user. This means defining under what circumstances the sound is triggered, by being reactive, dynamic, interactive or adaptive, that is, what is the audio behavior with the user. In other words, the interaction rules defines how a sound will start and finish, in which place of the interface, and under what circumstances it will happen (by clicking on a visual element, according to an execution of a task, in a timing triggered situation).

A critical aspect to control dynamic sounds is the definition of interaction key-points. The interaction rules needs starting and stopping points (play-ins and play-outs) by selecting variables (runtime, difficulty of task, performance, user status) that will qualify them in the programming platform. These points occur in periods of great importance, usually when users have to make relevant decisions. Getting a sound response according to a specific interface behavior is a fundamental matter for the sound design project. The interaction rules must be clearly defined and explicitly detailed, since it will be a guide to insert the sounds in the interactive platform.

A common way to address style questions about audio tracks for the interface can be done through the creation of temporary tracks. Temporary tracks should be set in preliminary positions by defining the basic parameters from which a creative team can lean on. The use of provisional tracks is also useful to test its incorporation in the platform, making possible to check if the proposed interaction rules are working properly.

Through the *sound reference sheet*, the programming team can exactly understand the behavior of the audio tracks in the interface, and how these sounds should be inserted. The sound reference sheet is a document that summarizes a group of relevant information that make possible to implement sounds in the interface, by exposing the nature of their behavior in the navigation context. In short, the sound reference sheet brings together in one document, all decisions made about the interface's sound design.

2.5 The “Audio Library” Layer

After finishing the sound reference sheet, then the *audio library* of the interface can be created. At this point, the stages of acquisition and processing of the sound files are done, making them ready for subsequent application by the programming team. This process takes place in three stages: production, post-production and implementation (Fig. 9).



Fig. 9. The *audio library* layer

The production stage involves capturing the sounds or acquiring them in sound libraries. The use of sound libraries is very common, and often the picked sounds are manipulated to achieve specific acoustic effects. However, to record custom sounds, a studio will be needed, or even in a recording field, outdoors, proper equipment should be set, and professional assistance should be considered. Technical aspects of audio, such as sample rate, resolution and other considerations to ensure that the audio files will have a compatible technical quality must also be defined. After that all the sounds listed in the sound reference sheet were captured and manipulated to achieve their aesthetic goals, these sounds will go thru a mixing stage for its subsequent implementation in the interactive platform.

In interactive applications, the post-production stage typically involves some degree of mixing. It is a procedure that aims to balance and adjust the various sound sources to present them in a clear and interesting way. It is in the mix stage that volume, panorama, equalization and other acoustic effects are added in order to establish a harmonic relationship between the sounds, so they can work together in the interface. Mixing is required since it considers the mutual relation of the sounds, and ensures that no overlapping between their frequency bands will occur. At this point, the design team must clarify which group of sounds will be emphasized and which ones will have less relevance in the overall context of the interface, so these concepts could be incorporated to the mix process.

Finally, the implementation should be considered, including the available tools and technologies. With the finished audio files and in conjunction with the sound reference sheet, the programming team can then add the sounds accurately in the interactive application.

3 Preliminary Use Cases

A preliminary test of the DAAG was performed in class, in order to provide students an articulate understanding between sound design and interface design in a conceptual, technical and practical manner. The guide was used in the academic discipline Project 6 (EGR 7140, Bachelor in Design, Federal University of Santa Catarina, Florianópolis, Brazil), that utilizes the User Experience Design approach from [6], and was held in the second semester of 2012. As the user experience layers were introduced in class, the sound recommendations from DAAG were progressively presented to students. In order to bring a better understanding of the guide, practical examples of the use of sound in interactive environments were shown to the students. Students enrolled in the course were divided into teams, and each team had to develop an interactive environment in the form of a website or an application, in a real scenario, with clients. The following is a brief description of two of the six works that were developed during the discipline.

3.1 The “SAT” Interface

The “SAT” project aimed to redesign the website of the Tax Administration System, from the Department of Finance of the State of Santa Catarina. This portal has a header area, with an initial menu that leads users to its applications. The user audience is composed of internal users (SAT officials), tax accountants, taxpayers, accountants, municipality officials, and the State Attorney General. The main objective of the working team was to reshape the interface in order to reduce the graphical content to its essentials, contributing to an easy-to-use operation.

In relation to the intended objectives, one of the requirements was “to call the user’s attention to unexpected occurrences in the system”, which means making the user confirms actions and take care about losing possible data during navigation. To achieve that, two content requirements were selected: to build informative messages (indicating error and warning) and dialog boxes to confirm actions (sending and losing data). To make it more perceptible to users, audio alerts were added to these boxes. Since the boxes only appear when something wrong happens, the sounds will not be considered repetitive. Since it is an interface with highly functional premises, the use of sounds in the “SAT” project were limited to these ones, but they were rated with high importance by the working team, that were chasing simple but effective solutions.

3.2 The “Gym” Interface

This group performed the development of a fitness gym’s website. Giving the proximity of its physical place to the campus, the target audience is the university community, representing more than 50 % of the clients. Since a large group of customers are young, the website should be a key-factor to create a virtual space that can be accessible outside of the physical space of the gym. To be part of experiences that customers can live outside of the physical environment of the gym, the working group decided to insert a playlist in the website’s front-page, as mentioned in the audio design concept document:

The gym's website could contain a playlist and make it available to anyone accessing it, regardless of location and devices (desktops, smartphones, tablets). The songs could be heard since there is a connection to the internet. These presented songs could be pre-selected by the academy team, containing songs that are played in the fitness classes, or even the user could ride your own playlist with the songs that are played during his training in gym. An online radio could also be created and, by on-line surveys, the radio content should fit with the users' taste and style.

With this proposal the students created a competitive advantage, enabling customers the access of the songs that are featured in the gym, so they can listen to it when performing outdoors activities, like walking, jogging, cycling. Looking upside down, the solution was quite simple and nothing really outstanding was needed to make these decisions feasible. But looking from a user's perspective, it was possible to bring a new degree of experience that was previous limited to a physical space, and now exceeded those barriers. That's exactly the DAAG's goal: to start a new approach when thinking about sound, interface and user experience design.

4 Final Considerations

Audio for interactive environments should be considered a collaborative process. The programmer cannot implement the sound files without the composition, and the sounds and music, in turn, depends largely on how it will be applied. One should take into account a constant dialogue between these two aspects: the conception and its application in the interactive platform.

As this is a relatively new area in the academic field, it is not yet sufficiently able to develop strong theories without the basic and substantial empirical research, which will evaluate the practice of audio production in interactive environments. The fact that the studies in dynamic audio are a recent effort means that much empirical evidence have not been sufficiently researched, and the available content is still scattered.

The Dynamic Audio Application Guide cast a new approach on the use of sound in interactive systems. Principles, foundations and assumptions of dynamic audio should be included in interface design methodologies, so the sound stimulus may be used with clearly defined goals: as an information carrier.

It is believed that the study results will open a debate of a yet low researched area, allowing its scientific basis to be useful in academic and professional practices, by providing a new tool to interface designers. The Dynamic Audio Application Guide aims to explore efficient methods to make the sound design an essential part of interface projects.

References

1. BAR-B-Q, Project. Group Report: Group Report: Providing a High Level of Mixing Aesthetics in Interactive Audio and Games. The Thirteenth Annual Interactive Music Conference PROJECT BAR-B-Q 2008, San Antonio, USA, December 2008. <http://www.projectbarbq.com/bbq08/bbq08r8.htm>. Accessed 15 Oct 2014
2. Carvalho, L.R., Cybis Pereira, A.T.: Interface design and dynamic audio. In: Kurosu, M. (ed.) HCI 2014, Part II. LNCS, vol. 8511, pp. 523–531. Springer, Heidelberg (2014). doi: http://link.springer.com/chapter/10.1007%2F978-3-319-07230-2_50

3. Collins, K.: *Game Sound: an Introduction to the History, Theory, and Practice of Video Game Music and Sound Design*, p. 92. MIT Press, Massachusetts (2008)
4. Collins, K.: *Making Gamers Cry: Mirror Neurons and Embodied Interaction with Game Sound*. AudioMostly 2011, September 7–9, 2011, Coimbra, Portugal (2011)
5. Darras, B.: *Design cognitivo e design participativo*. In: *Hipermídia: desafios da atualidade*/Vânia Ribas Ulbricht, Alice Theresinha Cybis Pereira (orgs.). Pandion, Florianópolis (2009)
6. Garrett, J.J.: *The Elements of User Experience: User Centered Design for the Web and Beyond*, 2nd edn, pp. 29–103. New Riders, Berkeley (2011)
7. Mckee, H.: *Sound matters: Notes toward the analysis and design of sound in multimodal webtexts*. *Comput. Compos.* **23**(3), 335–354 (2006). [s.l.]: Elsevier/ScienceDirect
8. Rocchesso, D., et al.: *Sonic interaction design: sound, information and experience*. In: *Conference on Human Factors in Computing Systems: CHI 2008 Extended Abstracts on Human Factors in Computing Systems*. ACM Association for Computing Machinery, New York (2008)
9. Stuart, R.: *Design of Virtual Environments*. McGraw-Hill, New York (1996)
10. Susini, P., et al.: *Closing the Loop of Sound Evaluation and Design*. In: *2nd ISCA/DEGA Tutorial & Research Workshop on Perceptual Quality of Systems*. Berlin, Deutschland, set. 2006