The Dynamic Audio
Direct-Interactive and Indirect-Adaptive Sounds

L.R. Carvalho, A.T.C. Pereira
Universidade Federal de Santa Catarina/EGR-CCE, Florianópolis, Brazil

Abstract—**With the evolution of Information and Communication Technologies (ICTs), the society gradually became audiovisual - the picture and sound have become part of a communication module which the individuals are becoming more adept, the Human-Computer Interface (HCI). Recognizing the HCI as an integrating element of media and visual, sound and tactile metaphors, this study will demonstrate investigations that contextualize the role of sound into interactive environments by proposing a better approach for the term 'interactive sound'. Commonly, audio elements belonging to an interactive system can react to a user´s input, but there's another group of sounds that can be started by parameters that aren't directly manipulated by the user. These two different situations represent two different types of sounds into interactive systems. This study aims to expose a better approach for these types of sounds, proposing its classification into direct-interactive and indirect-adaptative sounds, as well as pointing out its meanings and applications.**

Index Terms—**Adaptative Audio, Dynamic Audio, Interactive Audio, Interface Design' Sound Design.**

I. INTRODUCTION

Reference [4] points that "interface is the surface to access and exchange information". Reference [8] says that the usage of sound, tridimensional representations and animations is growing in order to improve the appeal and the possibilities of content presentation in the interfaces, as well to better attend the cognitive characteristics of the users. In order to make a user undergo to an optimized execution of a task, exceptional conditions can be presented to attract their attention. On the definition of [6] "The attention is a concentration of mental activity. It allows the user to select a perceptual channel, deciding which information should be prioritized in determined contexts".

Reference [7] points out that "the attention consists on the process of selecting one thing to concentrate on a certain moment among the variety of things available possibilities". The author emphasizes that it involves auditory and visual senses, since it allows users to focus on relevant information for what they´re doing.

Into an interactive environment, it's important to point out that the graphic interface will influence upon the use of the auditory resources. Both must be congruent, that is, projected as complements of each other. Reference [5] stands out that to use the auditory resource we must be cautious with the number of audio objects that will compose the interface, since its excess can cause a negative impact and even noise pollution. The same author says that on his works there was an enlargement on the usage

of sounds into interactive systems and their majority is because of dynamic interactions on the environment.

For an adequate development of an interactive project, the management of interactive auditory objects has some primary points that should be observed, as:

- The object that will use the auditory resources must be clearly defined;
- The interface must be projected to allow the creation and be able for possible alterations on the interactive acoustic environments.

Reference [8] explains that audio can be used into interactive systems according to three different ways: narration or dialogue, ambient sound or music, and as sound effects. In narration, the human speech intends to be informative, presenting explanations when necessary. The background music defines the mood and the rhythms of the presentation and it's linked to emotional interpretations that can stimulate reactions on the user. The sound effects are brief and have a function to stand out some point of the message, focusing and enlarging its impact. The lack of sound – the silence – can also transmit some sort of information. A pause can indicate a change into the narrative script, as well as the use of continuous music in different scenes can help to point out the maintenance of a particular theme. On the cognitive point of view, the auditory channel is an excellent way to transmit information and give feedback to the user. While the background music can be used to evocate emotions and define an environment; a specific sonorous effect can be used to transmit specific alert information [3].

II. DYNAMIC AUDIO: DIRECT-INTERACTIVE AND INDIRECT-ADAPTIVE SOUNDS

The applications of sound in interactive platforms differ from the commonly used in music and movies, which are essentially linear. In games, the most important element of interactivity and that offers meaning to the terms, according to Richard Rouse (*apud* [3]) is the non linearity, because "without the non linearity the developers of games would be developing films". The term non-linear refers to the fact that the user has choices, and each choice will result on the construction of a different narrative. It is the primary distinction that separates the interactive environments of traditional linear applications, as cinema and television, which the succession of events happens in a fix and unchangeable way.

In the last years, with the digital devices development, the dynamic behavior of sound offered a significant improvement above the interaction and immersion user´s level. The transitions between graphic spaces and sound

assets become more integrated, allowing the system to offer immediate responses to users. According to [1], interactive audio is any type of sound produced by an audio system that is programmed to generate a real-time response to an input stimulus, on the form a predetermined sonorous expression. This audio system is composed by an "interactive audio engine[41]" that receives users' commands (input stimuli), select the data (interactive data directory), and then send it to an interactive audio output directory. This output will then send an electric impulse thru the speakers, which will result in the generation of sound waves (Figure 1).

Not all of the audio systems that react to an input stimulus can be defined as interactive. Reference [1] affirms that an interactive system allows changes on the input command to modify the auditory behavior, while a reactive system simply statically reproduces audio events without any kind of response related to the user's stimuli. This interactive system can be classified in two categories: as direct audio, the user controls the audio responses consciously; as indirect audio, the user is controlling some other parameter that affects the audio behavior. According to [2], the main characteristics of interactive audio are:

- Improves the user´s experience;
- Inspires the user´s involvement;
- Creates a unique personality for the products;
- Allows users to make new kind of activities;
- Creates an participative experience;
- It is potentially cheaper to implement;

Among the possible platforms for interactive audio's application, there are virtual environments, games, web applications, music players, e-books, softwares, smartphones, vehicles, household appliances, toys and others. Since digital devices are in constant growth and development, possibilities for interactive audio´s applications get larger.

In discussions about audio in interactive environments, the term *interactive* is sometimes used to exchange or amplify the meaning of the terms *reactive* and *adaptive*. In fact, the interactive audio refers to sound events resulting from the user's direct interaction, and the adaptive audio, however, is a type of sound that reacts to the interface's state and its status, reacting to distinctive parameters. To avoid that the name *interactive* can result into a technical ambiguity, [3] suggests the term *dynamic audio* to define sound events occurring on the interface, which is divided and classified into *interactive* and *adaptive sounds*.

Reference [3] defines the dynamic audio as a mutable audio, a wide concept that embraces interactive and adaptive sounds. It is the audio that reacts to environmental changes and to the user's response. The interactive audio is defined by a sound event that reacts to a direct user's signal, as a sound emitted when a button is pressed. In a different way, the adaptive audio reacts to an interface's status, reacting to pre-established parameters that aren't directly controlled by the user. On the adaptive audio, the emitted sounds are not plainly determined by the user's action, that is, it involves other variables that the user has no direct control of it. On the game Super Mario Bross (Nintendo Inc.) the rhythm of the music increases when

---

[1] A group of software or hardware of algorithms that process interactive data based on entrance stimuli to process an audio exit.
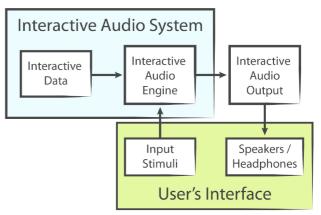

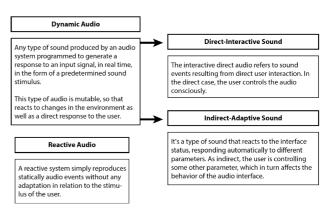Figure 1.   Interactive sound structure (adapted from [1])

Figure 2.   Terminological proposition of dynamic and reactive audio, direct-interactive and indirect-adaptive sounds.

the stage level timer is about to end, to warn the user that he should get hurry to finish his task.

All types of sounds that are put into an interactive environment, as music, dialogue and sound effects, can be classified according to [1] as *direct* or *indirect interactive* audio, or according to Collins (2008), as *interactive* and *adaptive dynamic* audio, and according to both, as *reactive* audio. Since these concepts agree with one another, and aiming to minimize the possibility of ambiguity, the categorization of the terms direct-interactive sound and indirect-adaptive sound are assumed, both being part to the dynamic audio group, as proposed on Figure 2.

A.   The specificity of Dynamic Audio

Koji Kondo (2007, apud [3]), musical compositor of the Super Mario game series, describes four components of the dynamic music, typical of interactive environments:

- The ability to create music that change according to user's interaction;
- The ability to develop a multilayered production, by creating distinctive themes for the same composition;
- The ability to add surprise elements, enlarging the interactivity;
- The ability to add musical elements with specific characteristics that matches to the interface's condition;

According to the author, dynamic music should evidence the participative character of the interface and indicate changes of its condition. These objectives can be achieved by the use of rhythmic changing, instrument and

voice additions, and even by altering the whole music according to distinctive stages of the interface, with possibility of adding or varying the reproduction of the sonorous sentences. The dynamic music must react or interact according to the narrative and fulfill the user's expectations.

Reference [8] point that a useful distinction of dynamic sounds is to classify them between *familiar sounds* - that are called *auditory icons* - and *abstract sounds*, called *earcons*. Auditory icons (as the sound of a door being opened or a ball jumping) help to reinforce the visual metaphors of the interface. Sonorous icons represent the kinds of sounds that aren't known by the user, and their meanings must be learned. Other category of sound includes the *cartoonified sounds*, which exaggerate the aspects of familiar sounds. Cartoonified sounds must be considered as belonging to the abstract sounds group because they normally don't have any relation with the sounds that are noticed on a physical environment, because of their characteristics of exaggeration and increasing proportions. Figure 3 is a scheme of these concepts.

The Interactive Audio Special Interest Group (IASG) proposes a series of functions attributed to dynamic audio, affirming the existence of a considerable superposition between these categories and pointing that they shouldn't be considered mutually excluding. These functions are divided in semiotics, emotional (intimately linked to the semiotics function), structural, narrative, immersive, esthetical and kinetic.

The semiotics function of dynamic audio seek to transmit an emotional meaning, aiming to guide the user's attention to identify objectives in a way that it is possible to diminish the learning curve, creating a positive level of inclination with the interface. Sounds symbols help to identify objectives and focus on the user's perception in determined objects. In many games, for example, the presence of enemies starts a tense music, and when the user finds benefic elements, as coins and heart shaped life, these end up having the same good sonorous suggestions. In other words, sound symbols are used many times to help users to identify other elements of the interface. These symbols can induce suggestions of humor and feelings when incorporated in the interactive environment, in a way that it is possible to make the interface more comprehensible.

A crucial semiotics role that the dynamic sound performs is its preparatory and anticipatory function. In games, anticipate action is a critical element, especially on adventure titles. Sounds without clear visual indication can incentive the user to look to the direction of a sound. The usage of sounds to add a behavioral and cognitive tendency is as important as the preparatory elements of dynamic audio. They change the user's perspective about the interface.

The emotional functions are intimately related to the semiotics. Here, a distinction must be done between transmitting a message through sounds and its condition of humor induction: the mood changes according to what the user feels, while emotional aspects simply transfer information. The user can understand that a sound exposes sadness without being sad. Considerable quantities of sounds on the interfaces have emotional effects that can enlarge or diminish the degree of difficulty for the execution of a task, as the case of the increasing rhythm of a
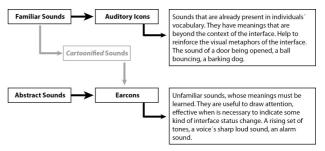

Figure 3. Division of sounds on the interface: familiar, abstract and cartoonified (adapted from [8]).

composition while a task is being executed. In this way, sound can control or manipulate the user's emotions, guiding responses through to the interface.

Structural functions of dynamic audio are linked to the act of creating, reinforcing or masking the interface structure to indicate changes on the narrative and situate the user. As in the movies, music and sounds are used frequently in interfaces to reinforce or improve its continuity structure. A significant example of the utilization of structural functions of dynamic audio is on the game *Vib Ribbon* (SCEI, 1999) in which the music can literally direct the narrative of the content. The game allows users to insert their preferable music to be used as a reference to map the difficulty levels. The system scans the audio signal characteristics, and then executes two obstacle courses for each song (one easy and other hard). The narrative structure of the interface can vary according to the chosen song. Although this is a very singular case, the utilization of songs and audio tracks for the creation of interface structures is a resource that has a meaningful potential. However, the dynamic audio is used with more frequency to improve the general structure of the interface. The inclusion of sonorous clues intermediating two interfaces acts as a linking element for a gradual and continuous transition of the contents. A brief silence can also inform the user that the suggested time for the execution of a task can be over, or that something different is going to happen, indicating possible changes of the interface condition.

In many cases, audio signals can help to situate the user on the interface's narrative context. When listening to distinctive musical compositions in different areas, the user is capable of identifying his whereabouts through the response of sounds. The audio is commonly used to locate the user on the plot, anchoring the user in terms of where, when and what is happening, as well as serving as an anticipation element for what is yet to come. The dialogue can also serve as a big event on the narrative, and it is used to reveal the attribution of objectives and specific tasks of the application. Non-verbal sounds can also reveal details about environment and objects through the usage of ambient moods that are particularly useful to create empathy and familiarization, making the interfaces more immersive.

The immersive function develops a critical role on the dynamic audio. It deals directly with the suspension of interface's incredulity, adding realism through the creation of an illusion of the *real*, which is indispensable for the user's immersion. The illusion of being immerse in a virtual environment is really reinforced by audio. Besides integrating the user on the narrative of the interface, the dynamic audio can also be used to make the sounds of the

application direct focus on the user in a way that he won´t be distracted by sonorous stimuli produced by the environment around him; they can be noise, sounds, voices. Dynamic audio can help masking these external sounds, as it progressively enlarges the user's concentration and focus on the interface task.

The kinetic functions connect a sensorial stimulus of audio to a specific motor response of the individuals. Some interfaces are projected so that the users can directly interact with sonorous stimuli and physical moves, like in the game *Dance Dance Revolution* (Konami, 2000). Kinetic dynamic audio works as the main motivator factor for the execution of movements because is the primary element that confirms (or rejects) the correct execution of a task on the interface's context.

The esthetical functions of the dynamic audio deal with the creation of identity and contextual references of the interface in which the sound is used to create beauty, generating acceptance and familiarity. The esthetical function offers the possibility to create moods that supply clues about the characteristics of the interface. Slow and smooth introductory songs, for example, normally indicate that the interface has a light rhythm of tasks. Musical compositions that are more accelerated normally indicate action and dynamic. Certain kinds of music adapt well to certain types of interfaces, and knowing that different types of narratives have distinct interactivity requirements, sound elements should also follow these patterns.

However, there is a series of variables that difficult the insertion of dynamic audio. The duration of an interface condition is a complex element to quantify, having in mind that each user manipulates the interface according to his familiarity and knowledge. The user experience will also influence the interactivity with sound elements. Sometimes, sounds are not sufficiently relevant to the interface's context, and become repetitive and boring, generating a listener's fatigue. The concept of *listener's fatigue* must be treated with caution; some interfaces are projected for consecutive use and repetitive sounds can be tiring, especially if the user spends a lot of time on a particularly specific area of the interface.

To solve this difficulty, some games started to incorporate timings for the audio clues, in a way that if the user stays in a determined environment, the music won't repeat endlessly; instead, it just stops being played. Marty O'Donnell (apud [3]) argues about the game *Halo*, that has a command called *I'm bored* that – if the user haven't completed a determined objective and 5 minutes are gone –the song just disappears, in fadeout[2].

According to [8], since the origin of the desktop's interfaces, a series of sounds was used to indicate tasks, warnings, or even to point out the conclusion of an action, as the sound that is emitted when an archive is put on the trash can for elimination into an operational system. The effect for most of the users is a satisfactory confirmation of actions; on the other hand, after a few hours the sounds can become a distraction instead of a contribution, especially where there are many machines and many users on the same area. Some applications ring a bell or the sound of a tune when an error occurs. This alarm can be useful if the user could lose the mistake but it can also be embarrassing if other people are on the same area. Sound de-

signers must find a way between drawing attention to a problem and avoid embarrassment to the user. Considering the ample and distinctive range of experience and temperament of the users, the most appropriate solution is offer to user some control over the sounds, making the dynamic audio approach consonant with the principles of *user´s experience design*.

III. FINAL CONSIDERATIONS

There are many examples of a growing tendency pointing to dynamic audio beyond the traditional computer desktop applications: videogames, smartphones, tablets, domestic devices, vehicle systems. The most advanced dynamic audio system that exists nowadays is found on videogames platforms. As games become more sophisticated, instances of audio that are reproduced in response to a user's stimuli are also becoming more intelligent. That said, the development of dynamic audio into any kind of interactive environment requires tracking innovations that are found in video games.

Interactive spaces offer a wide range of possibilities to intersect different modalities of language in a never before offered way. On the digital environment, there is no construction of meaning only by a unique semiotics system. It deals with systems that allow navigation through distinct groups of information in a multi linear way, involving many integrated language modalities - as verbal, image, sound, animation, the use of colors, typography and other resources to produce meanings. From this point of view, it´s important not to privilege a certain type of language upon another.

Since the effects that each element involved into an interactive production is noticed, the interface development process become more conscious, contributing for the elaboration of better virtual spaces. It´s important to figure out the meanings that each element can produce, considering its relation and integration as a whole into the interactive system. The dynamic audio intensifies the processes of immersion and cognitive processing of the user, making the interaction experience more evolving. This kind of experience, enclosed by image and sound, is more complex and complete because it reaches the user in distinctive senses through a unique communication object: the interface. Background songs, music, oral and writing language, images, animations and texts; all of these forms of expression – languages – get mixed on the same message. There is a tendency in direction of the interactivity that was already documented in many areas. Consumers – especially the young – demand dynamic activities in place of passive devices. Dynamic audio has the potential to feed this demand, but if not well applied can suffocate innovation and disappoint this new and important public. Establishing a solid structure for dynamic audio can guarantee that the next generation of applications based in audio will produce results that are closer to the human cognitive model, in a way that the information presented on a interactive context will better accommodate the content´s absorption structures. This way, when systematizing the dynamic audio structures, the sounds used in the interfaces will be made in a more efficient way, and the applications would be plainly developed.

REFERENCES

[1]   BAR-B-Q, Project. Group Report: What is Interactive Audio? And What Should It Be? The Eighth Annual Interactive Music Confer-

---

[2] Technical term used to indicate gradual diminishing of a sound until it becomes inaudible.

ence PROJECT BAR-B-Q 2003, San Antonio, USA, section 5, dez., 2003. Available at: http://www.projectbarbq.com/bbq03/ bbq03r5.htm. [Accessed 20 November 2011].

[2] 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, dec., 2008. Available at: http://www.projectbarbq.com/bbq08/bbq08r8.htm. [Accessed 20 November 2011].

[3] COLLINS, Karen. Game Sound: an introduction to the history, theory, and practice of video game music and sound design. Massachusetts, MIT Press, 2008.

[4] CAVALCANTE, Marianne Carvalho Bezerra. Mapeamento e Produção de sentido: os links no hipertexto. In: Hipertexto e Gêneros Digitais: novas formas de construção de sentido. Luis Antônio Marcuschi, Antonio Carlos Xavier (orgs.) - 3. ed., São Paulo: Cortez, 2010.

[5] MENZIES, Dylan. 2002. Scene Management for Modelled Audio Objects in Interactive Worlds. Proceedings of the 8th International Conference on Auditory Display, July 2nd ~ 5th, 2002.Advanced Telecommunications Research Institute (ATR), Kyoto, Japan, 2002.

[6] PADOVANI, Stephania; MOURA, Dinara. Navegação em Hipermídia. Rio de Janeiro : Moderna. 2008.

[7] PREECE, Jennifer. Design da Interação. Porto Alegre: Bookman. 2005.

[8] SHNEIDERMAN, Ben; PLAISANT, Catherine. Designing the User Interface: Strategies for Effective Human-Computer Interaction. Addison Wesley, 5a. ed., mar., 2009.

AUTHORS

L. R. Carvalho is with Universidade Federal de Santa Catarina, Florianópolis, SC, Brasil (e-mail: semprecarvalho@gmail.com).

A.T.C. Pereira is with Universidade Federal de Santa Catarina, Florianópolis, SC, Brasil (e-mail: semprecarvalho@gmail.com).