VirtualAurea: perception through spatialized sound

J Sánchez, L Jorquera, E Muñoz and E Valenzuela

Department of Computer Science, University of Chile
Blanco Encalada 2120, Santiago, Chile

jsanchez@dcc.uchile.cl, ljorquer@c5.cl, elmunoz@dcc.uchile.cl, evalenzu@c5.cl

ABSTRACT

Blind children cognition is currently enhanced by using tactile-based methods such as the Braille system. The use of sound for cognitive purposes is not common in the blind educational community. Even though these facts, an increasingly number of studies are showing that blind children can develop cognitive structures through the interaction with spatial sound. Most studies use specific purpose software to help blind children to construct cognition and learning. A common characteristic of these studies is the absence of robust usability studies that inform of the adequacy and user satisfaction with this sound-based software. This paper introduces VirtualAurea, a flexible software editor for sound-based virtual environments designed "with" and "for" blind children to construct their cognition. It also presents a complete usability study to validate the editor with blind learners by using a kit of cognitive tasks. We also determined the impact of the interaction with this software editor on the development of basic cognitive functions, operations, auditory perception, laterality, comprehensive language, and tempo-spatial orientation mental skills of blind learners.

1. INTRODUCTION

A number of recent studies indicate that spatial structures can be developed and used by blind children through the interaction with virtual worlds through audio interfaces (Baldis, 2001; Lumbreras & Sánchez, 1998; MacCrindle & Symons, 2000; Mereu & Kazman, 1996; Sánchez, 2000a, 200b, 2001; Sánchez & Lumbreras, 1999; Stifelman, Arons & Schmandt, 2001; Sjostrom, 2001; Tan, 200; Winberg & Helltrom, 2000). Most of them are pilot studies with major emphasis on the development of computer applications rather than focused on blind user-centered design with fully usability testing to tailor virtual environments according to the interest, needs, and mental models of blind children.

In addition, most of these published applications are specific purpose software with limited flexibility to allow teachers, parents and the child to modify the structure and functionality of the software. There is no way that control, complexity, versatility, and flexibility can be determined by both software designer and end users.

Lumbreras and Sánchez (1998, 1999) developed a pioneer application that allows blind children to navigate and interact in a virtual world based on modeling the computer game Doom by using spatial sound interfaces. In this pilot study the researchers proved the potentiality of the exclusive use of audio to represent basic spatial cognitive structures depicted by main and secondary corridors. The navigation through sound-based virtual worlds triggered mental models of the spatial representations very close to the structure embedded in the software. A group of seven learners ages 8 to 11 experimented with the software during two months and attained concrete representations of the navigated virtual space including objects and entities through the use of Lego bricks. Authors concluded that it is possible to use spatial sound to stimulate blind learners to construct navigated spaces in an interactive virtual world. This initial experiment also informed about input and output interfaces concluding that devices such as the keyboard and ultrasonic joysticks adjust to the requirements of the blind children as well as external speakers instead of earphones.

This seminal study was the starting point of an entire research line about spatial sound and cognition in blind children. Even though the initial results were promissory a number of questions arose from this study. As a result, Sánchez & Lumbreras (2000) developed a full field study designed with a kit of cognitive tasks and a careful research planning by using the same sound-based computer application named AudioDoom (Sánchez & Lumbreras, 1998, 1999). They added a strong component of task concrete materials to support the representation of navigated spaces. The kit of cognitive tasks consisted of a set of learning activities with objectives and specific actions. This allowed focusing the research on the development of specific spatial skills. After working four months with forty blind children and following the stages of: 1. Interacting with AudioDoom, 2. Representing the navigated virtual environment, and 3. Constructing the models with sand tables, Styrofoam, Lego bricks, sand, clay, and corporal exercises, authors concluded that there are four levels of children´s cognitive attainment when interacting with spatial sound virtual environment: Entry, Exploration, Adaptation, and Appropriation. These levels imply an increasing cognitive complexity to model the cognitive tasks. Adaptation and appropriation are the highest level of spatial representation and mental modeling after using sound-based software and perform cognitive tasks. The whole research concluded by stating the viability of assisting the construction of mental structures through the use of spatial sound with a set of cognitive tasks. The sound *per se* does not have an impact or "effect" on the development of spatial cognitive structures. The cognitive tasks with didactic implementations are critical to obtain the expected results. The study also concluded that spatial mental images could be constructed with spatial sound and without visual cues. Images can be transferred through spatial sound that emerges from interactive virtual environments.

One of original sides of this second study was the comparison made between the behavior of blind and sighted children when navigating the same audio-based virtual environment and solving cognitive tasks with concrete materials to represent the environment. As a result, sighted children reached only the exploration level and blind learners attained the highest level of cognitive appropriation. This implies that sighted children do not base their cognition in the aural perception, they heavily rely on the visual perception. Learning through sound is very poor in sighted children and perhaps these results can shed some lights to better stimulate cognition through the use of sound.

Diverse research lines emerge from this study. Perhaps the most direct is the necessity to have more versatile, flexible, and complete sound-based software to allow modifying the navigating scenario and the possibility to go back and shift in the software map. This called for a software environment that allows a wide diversity of cognitive tasks based on the work with the spatial sound elicited from the interaction with the computer. The possibility that the teacher or parent of the blind children can design software upon the framework of a generic editor was a challenge for the research.

The related literature present diverse experiences with auditory stimuli to mimic visual cues for blind children. Mereu y Kazman (1996) found that the use of 3D audio interfaces by a blind person could help to localize a certain point in the 3D space. This was performed in such a precise way as with a sighted person interacting with a graphical interface, even though it took more time for the blind person. They also found that in environments exclusively based on sound users with visual disabilities were more precise than sighted users.

Cooper y Taylor (1998) highlights the effect of 3D sound interfaces in virtual environments and applications for blind people. Lahav & Mioduser (2000) successfully developed a study on the premise that providing appropriate spatial information through compensatory channels can contribute to the spatial performance of blind people. They developed a multisensory virtual environment with the support of force feedback joysticks simulating real life spaces such as the school, the work and public buildings. Winberg & Hellstrom (2000) developed an audio version of the game Towers of Hanoi. The idea was to investigate the nature of the persistent presentation and what it may mean when implementing audio direct manipulation. The results showed that it was possible to implement audio direct manipulation by using the Towers of Hanoi as a model as well as the fact that having a set of objects, the nature of the continue presentation was not relevant as the interaction with the audio space.

McCrindle y Symons (2000) implemented a study in the same context of AudioDoom (Sánchez & Lumbreras, 1998, 1999) in order to probe that the standard characteristics of a traditional game such as Space Invader can be replicated by using audio 3D by using force feedback joysticks as input interfaces. They concluded that it is possible to produce a multimodal game played by blind and sighted people providing them a way to share the same game experience.

Baldis (2001) conducted an experiment to determine the effect of spatial audio on memory and perceived comprehension of conferences with nine participants, concluding that spatial audio improved memory and perceived comprehension and that participant preferred spatial audio instead of non-spatial audio. Stifelman, Arons and Schmandt (2001) developed AudioNotebook a tool that augments the user's activity by helping to find desired portions of audio. AudioNotebook is a device for taking notes and interacting with a speech recording as a new approach for navigation in the audio domain. This allows to free listeners to devote more attention to the talker, so they are not always required to take notes.

Loomis, Lippa, Klatzky & Golledge (2002) designed a study with blind and blindfolded sighted observers presenting them auditory stimuli specifying target locations by using 3D sound. They assessed the ability to mentally keep track of the target location without concurrent perceptual information about it (spatial updating), concluding that once an internal representation of a location has been determined, posterior-updating performance is independent of the representation mode.

Most of these studies called for a full usability study. Our experience with these sound-based computer applications has taught us that to map blind people mental models can only be possible if we test and retest with them the software and devices we built for them. This is the persistent paradox of developing software for blind children by sighted adults. Mapping the mental model of a child is a difficult task, mapping the mental models of blind children is more complex and an unexplored matter.

Our research team has explored spatial cognition but complex areas and spatial conflicts have been less explored with the use of sound. There is also a need for implementing tasks to stimulate tempo mental structures as well as working with small children to early support them in the development of tempo-spatial. This will help to level them constructively with the same cognitive knapsack of sighted children to face their intellectual development with similar conditions even though they can not decode visual stimuli that is embedded the surrounded cognition. Our study intends to consider some of these concepts.

The purpose of this study was to design and implement VirtualAurea and then to determine the impact on the development of basic cognitive functions, operations, and tempo-spatial skills of blind learners using interactive sound-based computer tools. We investigated a way to develop through the interaction with audio-based software editors, auditory perception (memory, consciousness, discrimination), laterality (concept of left and right with respect to the own body), comprehensive language and tempo-spatial orientation (front-back, between, alongside, before-after, slow-rapid).

2. RESEARCH MATERIALS AND METHODS

2.1 Materials

To implement the testing we used the following materials: four PCs Pentium III with quadraphonic cards and four speakers, VirtualAurea, colored geometric pictures with different sizes, sketches with the graphical representation of the software map, musical instruments (tambourine, pegs, Indian reed flute, oboe, guitars, etc.), panel boards to physically represent the walls and corridors of the software, and a digital camera.

2.2 Software

VirtualAurea is a software editor to navigate custom-made spatial structures through spatial sound. It is a flexible tool to create virtual scenarios. It allows children to perform diverse movements and opens the possibility to flexible place objects throughout the scenario. This allows having a better management of the level of difficulty of the tasks, allowing the creation of maps or spatial structures adequate to the cognitive level of the learner. The option to change constantly the scenarios allows keeping attention alive and the motivation of the child. VirtualAurea allows a high degree of freedom in the movements with a more intuitive and real simulation of the environment.

VirtualAurea was built upon a PC platform, operative system MS Windows 95, C++ and programming environment Visual C++ 5.0, using a libraries of Microsoft DirectX 5.0 multimedia routines. VirtualAurea is conformed of two independent programs: An editor that generates the files that define the maps, create, save and modify maps, and the game or software that opens the maps generating the visual and auditory representations. The editor has an interface and draw tool as displayed in Figures 1 and 2.

The editor gives total freedom to the user to create maps without limitations to displace elements. This allows that a map can be saved, modified or played anytime the user wants. The game mode interface is simple because the main information source for the blind child is delivered through auditory channels.

To move in the game we use three keystrokes: left shift (45º upon its own axis), right shift (45º upon its own axis) and to move forward. Due to the fact that in any standard keyboard the keystrokes "f" and "j" are marked with small dots, we chose the keystrokes "f", "j", and "k" to shift left, right and move forward.

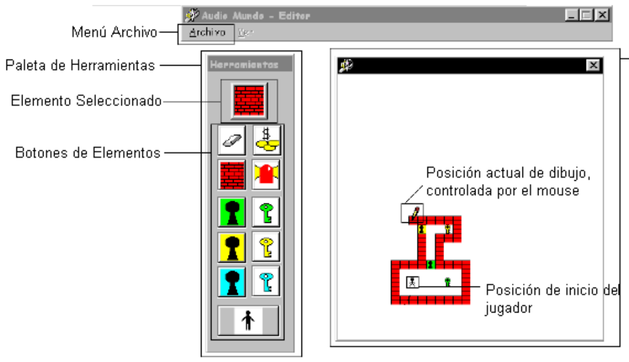

Figure 1. *Graphic interface of VirtualAurea.*

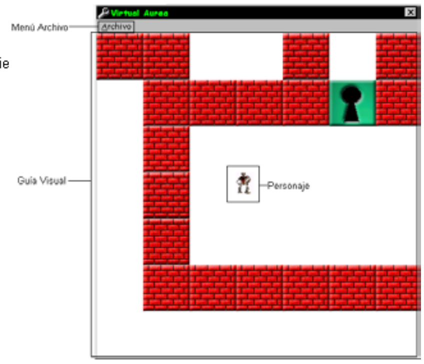

Figure 2. *Graphic interface of one game generated by VirtualAurea.*

VirtualAurea is a generic platform upon which we can construct specific applications. For this study we built an application called "The Music Castle". The idea behind the software is that while the child picks pieces of a song in the correct order s/he can get higher score and can be able to listen to the song. The level of difficulty can vary when different pieces of diverse songs are introduced. In such a case the child should be able to discriminate what pieces correspond to the song that is looking for. To work the notions of spatiality each object in the map emits a sound that is placed in a determined position and location within the auditory spectrum. Thus the child can exercise concepts by searching objects that are to the left, right, front or back, according to the location of the sound emitted. The notions of temporality are exercised through the recognition of the corresponding position of a musical piece. This allows to reinforce concepts such as position (first, second, third, fourth) and numbers (one, two, three, four).

The Music Castle is based on VirtualAurea then it is conformed of two modes: the editor of structures and the game that reproduces auditory and visually the structures generated by the editor. The editor (see figures 3 & 4) is a program that allows creating diverse structures of The Music Castle by the teacher.

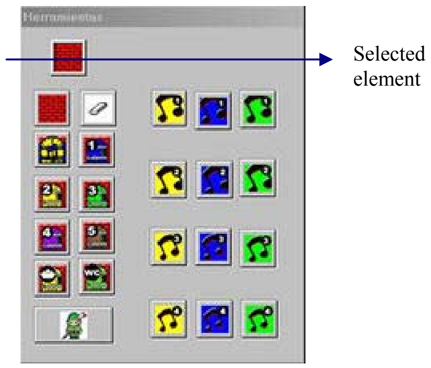

Figure 3. *Graphic interface.*

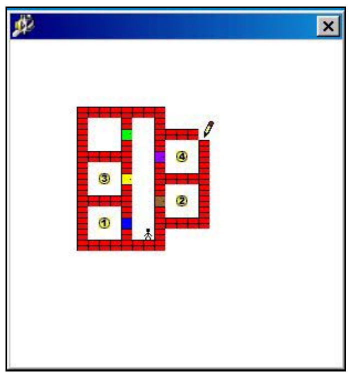

Figure 4. *Draw area.*

The main goal of The Musical Castle is to help children to identify the corresponding pieces of a song. To do this they have to recognize the object's sound placed in the navigating spatial structure and identify the position. Once the children find a piece of the song the melody is listened and they choose a fragment of the song. While they use this application the teacher guides their performance by reinforcing the concepts of laterality, temporality and language.

To get a spatial position the game emits sounds coming from different positions of the space. For example, if an object is behind the subject controlled by children, the sound comes from the back speakers as it is placed in the back. If the object is placed to the left the sounds come from the speakers that are placed to the left of the child. To move toward an object the children have to shift until they perceive the sound of the object as it comes from the front through the front speakers. Once the child places the subject in this way he later can move toward the object with an intention in mind.

2.3 Methods

The study was held during the year 2001 in a Chilean school for blind children. The school is conformed of children that are totally blind and others with residual vision. A multidisciplinary team of professionals assists these children. The school has enough infrastructures to lodge learners in a Monday to Friday internship system. Children come from economically poor and socially deprived families. The research was implemented in the computer laboratory.

The research was implemented during six months with nine children with visual deficits ages 6 to 12, seven boys and two girls. Seven children have residual vision and two were totally blind. Three of them have also diverse intellectual deficits, four have slow mental performance and two were tested intellectually normal. Each week they worked during three sessions of 30 minutes with VirtualAurea and 25 minutes solving cognitive tasks

We used an intact sample of 12 learners, three of them left during the study for family problems. Then we systematically worked with nine cases. These children were evaluated each year by a multidisciplinary team. They evaluated both the visual capacity and IQ by using the standardized test Wisc- r, Wechsler reviewed scale for blind children. Using a national test battery for verbal exploration, BEVTA to evaluate psycholinguistic processes. We also used two specific tests of this battery, TAVI and 3-S. TAVI evaluates the auditory receptive of the oral language, immediate verbal reception, oral comprehension of simple paragraphs, short time attention to verbal stimuli and verbalism of a answer. 3-S evaluates verbal abstraction, sameness relationships and vocabulary recognition.

We used this test because language structure is related not only with temporal sequences in a grammatical structure of a phrase, rather it also involves abstract thought and the capacity to understand directions. The interactive virtual environment used in this study has a great deal of information given in oral language that has to be comprehended and structured coherently.

The progress of children was evaluated through the use of cognitive tasks. We developed three tasks to work with VirtualAurea. They were designed to help children to 1. Integrate tempo-spatial references perceived as sequential realities, 2. Demonstrate the construction of knowledge about the virtual space of the software, and 3. Evidence the knowledge/mapping of the virtual space. The idea with this kit was to provide a tool to the mediator to assist the use of VirtualAurea by blind children. It also provides diverse activities to complement the use of VirtualAurea. Finally, it is a way to exercise tempo-spatial skills necessary for reading and writing and pre-cane techniques.

In addition, besides to cognitive tasks and tests, we documented our results by using diverse methods such as checklists, anecdotal notes, interview, recording, filming, and taking digital photographs.

3. RESULTS

Our main results can be analized around the cognitive tasks designed.

Cognitive task 1. **To relate tempo/spatial concepts and the mapping of the sound-based virtual space**. Blind learners integrated tempo-spatial concepts and mapped the virtual space through the interaction with the game by solving progressive complex tasks to put in order elements of the software. Thus they progressively attained the exploration stage. All of them attained the goal for this task. They interacted with VirtualAurea very easily (see Figures 5, 6, 7, 8 and 9).

Cognitive task 2. **To evaluate temporal concepts related to the rythm in the performance of actions (fast/slow).** Learners were able to evidence progress in their movements, orienting through acoustic stimuli. This requires a certain level of concentration on the activity. They can discriminate rhythm and sound in different velocities and reached the level of adaptation. They again represented very precisely the structure of the virtual environment (see Figure 10).

Cognitive task 3. **To develop laterality and spatial concepts**. They had to recognize their ownn space to orient themselves and demonstrate the spatial knowledge of VirtualAurea. Learners recognize the space for orientation and demonstrate the spatial knowledge of the program. They applied tasks related to tempo-spatial notions by using their bodies and moving physically through the represented environment. They not only can interact with the virtual environment but can also move through the real environment that

VirtualAurea represented. They are able to relate doors with the music instrument that represent them and also to order in a chronological time as they appear in the software (see Figures 11 and 12).

3.1 A sequence of the experience illustrated with pictures

Figure 5. *The learner interacts with the game receiving information through the four speakers.*

Figure 6. *The learner develops the Concept of chronological ordering.*

Figure 7. *The learner orders puppets and dolls representing the chronological time.*

Figure 8. *The learner orders concrete elements in a board.*

Figure 9. *The learner orders illustrations in a chronologically.*

Figure 10. *The learner works in a sketch of the virtual environment reflecting the mental image of the structure.*

Figure 11. *The learner performs laterality tasks*.

Figure 12. *The learner performs spatial tasks by using his own body*.

4. DISCUSSION

The main goal of this study was to determine the impact on the development of basic cognitive functions, operations, and tempo-spatial skills of blind learners, of using interactive sound-based computer tools. Our results evidence that sound not only supports the child´s development of spatial representation skills as in past studies already discussed but also facilitates the development of basic cognitive functions and temporal skills. This extends the scope of cognitive processes that can be improved with adequate use of sound-based virtual environments.

Qualitative and quantitative results indicate that when we use VirtualAurea + spatialized sound interfaces + cased-based cognitive methodology including tasks and tests to measure space and time representations of navigated virtual worlds, blind learners can develop tempo-spatial cognitive structures, diversify and make a deeper development by transferring them to everyday tasks. Spatial sound delivered through the interaction with computers by navigating and interacting with objects and virtual world entities can be a powerful medium to represent information and knowledge by blind learners. This can be critical with early learners with urgent needs to minimize access deficit that separate and divide them from the cognitive experiences of sighted learners. Once more we highlight the potentiality of sound as powerful tool to enhance learning and cognition in blind children.

5. REFERENCES

J. Baldis (2001). Effects of spatial audio on memory, comprehension, and preference during desktop conferences. *Proceeding of the ACM CHI ´01*, Vol 3, 1, pp. 166-173.

L. Bravo & A. Pinto (1995). *Pruebas psicopedagógicas de lenguaje y lectoescritura para evaluar dificultades del aprendizaje escolar*. Documento no publicado.

M. Cooper & Taylor (1998). Ambisonic sound in virtual environments and applications for the blind people. *Proceedings of the Second European Conference on Disability, Virtual Reality, and Associated Technologies*, pp. 113-118.

O. Lahav & D. Mioduser (2000). Multisensory virtual environment for supporting blind persons' acquisition of spatial cognitive mapping, orientation, and mobility skills. *Proceedings of the Third International Conference on Disability, Virtual Reality and Associated Technologies*, pp. 53-58.

J. Loomis, Y. Lippa, R. Klatzky & R. Golledge (2002). Spatial updating of locations specified by 3-D sound and spatial language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(2), 335-345.

M. Lumbreras & J. Sánchez (1998). 3D aural interactive hyperstories for blind children. *International Journal of Virtual Reality 4(1), 20-28*.

M. Lumbreras & J. Sánchez (1998). 3D Aural interactive hyperstories for blind children. *The 2nd European Conference on Disability, Virtual Reality, and Associated Technologies*. Skovde, Sweden.

R. McCrindle & D. Symons (2000). Audio space invaders. *Proceedings of the Third International Conference on Disability, Virtual Reality and Associated Technologies*, pp. 59-65.

S. Mereu & R. Kazman (1996). Audio enhanced 3D interfaces for visualli impaired users. *Proceedings of CHI ' 96*, ACM Press.

J. Sánchez (2001). Interactive virtual acoustic environments for blind children. *Proceedings of ACM CHI ´2001, pp. 23-25.* Seattle, Washington, April 2-5.

J. Sánchez & M. Lumbreras (1999). Virtual Environment Interaction through 3D Audio by Blind Children . *Journal of Cyberpsychology and Behavior*, *CP&B* 2(2), pp. 101-111..

J. Sánchez (2000). Usability and cognitive impact of the interaction with 3D virtual interactive acoustic environments by blind children. *Proceedings of the 3rd International Conference on Disability, Virtual Reality and Associated Technologies*, pp. 67-73. September 23-25, 2000, Alghero, Sardinia, Italy.

J. Sánchez (2000). 3D interactive games for blind children. *Proceedings of Technology and Persons with Disabilities, CSUN 2000.* Los Angeles, USA.

J. Sánchez (2000). Usability and cognitive impact of the interaction with 3D virtual interactive acoustic environment by blind children. *Medicine Meets Virtual Reality: 7*, Newport Beach, California, USA.

J. Sánchez (1999). Interactive 3D sound hyperstories for blind children. *Proceedings of ACM-CHI 99*, pp. 318-325.

J. Sánchez & M. Lumbreras (1999). Virtual environment interaction through 3D audio for blind children. *Medicine Meets Virtual Reality: 7*, San Francisco, USA.

L. Stifelman, B. Arons & Ch. Schmandt (2001). The Audio Notebook. *Proceeding of the ACM CHI ´01*, Vol 3, 1, pp. 182-189.

C. Sjostrom (2001). Using haptics in computer interfaces for blind people. *Proceeding of the ACM CHI ´01*, Vol 3, 1, pp. 245-246.

H. Tan (2000). Haptic interfaces. *Communications of the ACM*, 43(3), 40-41.

F. Winberg & S. Helltrom (2000). The quest for auditory manipulation: the sonified Towers of Hanoi. *Proceedings of the Third International Conference on Disability, Virtual Reality and Associated Technologies*, pp. 75-81.