2D SPATIAL AUDIO IN A MOLECULAR NAVIGATOR/EDITOR FOR BLIND AND VISUALLY IMPAIRED USERS

*Ian Rodrigues, Ricardo Teixeira, Sofia Cavaco**

NOVA LINCS
Department of Computer Science
Faculty of Science and Technology
Universidade Nova de Lisboa
2829-516 Caparica, Portugal
in.rodrigues@campus.fct.unl.pt, scavaco@fct.unl.pt

Vasco D. B. Bonifácio

Centro de Química-Física Molecular and
Institute of Nanoscience and Nanotechnology
Instituto Superior Técnico
Universidade de Lisboa
Av. Rovisco Pais
1049-001 Lisboa, Portugal
vasco.bonifacio@tecnico.ulisboa.pt

*Daniela Peixoto, Yuri Binev, Florbela Pereira, Ana M. Lobo, João Aires-de-Sousa**

LAQV-REQUIMTE
Department of Chemistry
Faculty of Science and Technology
Universidade Nova de Lisboa
2829-516 Caparica, Portugal
{d.peixoto,y.binev, mf.pereira, aml, jas}@fct.unl.pt

ABSTRACT

In order to contribute to the access of blind and visually impaired (BVI) people to the study of chemistry, we are developing Navmol, an application that helps BVI chemistry students to interpret molecular structures. This application uses sound to transmit the information about the structure of the molecules. Navmol uses voice synthesis and describes the molecules using the clock polar type coordinates. In order to help the users to mentally conceptualize the molecular structure representations more easily, we propose to use 2D spatial audio. This way, the audio signal generated by the application gives the user the perception of sound originating from the directions of the bonds between the atoms in the molecules. The sound spatialization is obtained with head related transfer functions. The results of a usability study show that the combination of spatial audio with the description of the molecules using the clock reference system helps BVI users to understand the molecules' structure.

1. INTRODUCTION

BVI people have difficulty in achieving a higher education degree due to the lack of material adapted to their special needs. Moreover, the few students who reach higher education, are influenced to pursue a degree in humanities or music and to avoid science and engineering degrees. According to the 2011 census, there were 16 500 BVI people in Portugal, of which 42 000 were school-age people (5-29 years old) [1]. However, only 10 000 are referenced as students. In 2001 the percentage of BVI people in Portugal who completed a higher education degree was 0.9% while in the general population this number increases to 10.8% [2, 3]. In the European Union in 2002 the percentage of people without disabilities with 25-64 years of age who completed a university degree was 22%, but this number decreases to 10% when we consider people with disabilities [4]. The same study in Portugal concluded that 17% of the people without disabilities completed a university degree while only 1% of the people with disabilities completed a university degree.

The representation of molecular structures is a major challenge for the accessibility of blind students and professionals to chemistry and other science degrees that include chemistry courses. This is particularly true in organic chemistry that critically requires the interpretation and transmission of molecular structures - and these are commonly processed visually, by means of chemical drawings.

Only a few applications that tackle this issue have been proposed in the literature. The AsteriX-BVI web server resorts to tactile information, such as one using 3D printing techniques to construct tactile representations of molecules with Braille annotations [5], or another that functions as a molecular editor that uses 2D sketches and static Braille reports [6]. While this solution may help blind users to construct a mental representation of the molecules, it may not be accessible to everyone.

In order to contribute to the access of BVI people to higher education, we are developing tools adapted to their special needs that can be used to study chemistry courses. These consist of chemoinformatics strategies to teach chemistry to BVI users [7]. In particular, here we focus on the structure of the molecules. These strategies rely on computer representations of molecular structures not only as drawings, but as graphs, specifying the atoms in a molecule, their bond connections, and their 3D configuration [8].

Here we propose a solution based on spatial audio implemented with head related transfer functions (HRTFs) to transmit information about the structure of molecules to BVI students and chemists. This solution has the advantage of being able to reach more BVI users than previously proposed applications.

With the intention to facilitate the integration of blind students in the study of chemistry, we are developing Navmol, a navigator and editor of molecular structures designed for BVI students and chemists that uses voice synthesis and other accessibility tools familiar to BVI users (text-to-Braille refreshable displays and the keyboard) [9]. Furthermore, it contains a graphical interface that replicates the same information that is heard so that sighted users can better communicate with BVI users, helping with their integration for example in a classroom (figure 1). Along with these features, we propose to use spatial audio to help BVI users to more easily understand the structure of the molecules.

The main novelty of this work is the use of spatial audio to transmit information about the structure of the molecules to BVI users. This work can have a real impact in the lives of people in a special needs group that is typically under-served, contributing to their access to higher education in science and engineering degrees.

Below we discuss the sound spatialization technique used in Navmol, and the results from a usability test performed with a group of chemists, a group of non-chemistry university students and a group of BVI volunteers. The results from all groups show that the users can successfully reconstruct the molecular structure of the molecules presented to them when Navmol describes the

Figure 1: Two BVI students using Navmol.

structure of the molecules with spatial audio. Moreover, the comments from BVI users also indicate that the sound spatialization improves Navmol, that is, with spatial audio it is easier and faster to grasp the structure of the molecules.

2. NAVMOL

Navmol[1] is a software application that provides BVI people with the possibility of navigating, interpreting and editing the representation of molecular structures in a plane (figure 2). For the communication of molecular structures to BVI users, Navmol uses the clock reference system, which is a polar coordinate system based on the representation of an analogue clock making an analogy between the hours of the clock and their corresponding directions. The clock reference system is widely used by BVI people in many situations, including during meals as a way to know where the food is located in the plate [10].

After selecting a certain atom in a molecule, the user gets information about the clock directions of all the neighbor atoms (that are bonded to the selected atom, except hydrogens which are omitted). For instance, the sentence “atom C 2 at 9 o’clock through a double bond" means that, in the loaded molecular representation (figure 2), there is a C 2 carbon atom to the left of the carbon atom C 1 that the user is currently selecting. Navmol uses a speech synthesizer to transmit this information.

In order to give the users more cues that help them to mentally conceptualize representations of the molecular structures, Navmol modifies the speech synthesizer signal with HRTFs so that the speech is spatialized. Provided the user wears headphones, the speech can be heard as if coming from different analogue clock directions.

Spatial audio has been used before in computer games to give a sense of immersion to the users. Other applications that use sound spatialization with HRTFs have been successfully used by BVI users. *Demor* [11] and *AudioSpaceStation* [12] are two such examples. These are games that are solely based on spatial audio sound to give BVI users the feeling of immersion in a virtual 3D world. The novelty of this work is to use spatial audio in an application whose purpose is to contribute to the integration of BVI users in chemistry courses.

[1]Navmol is available at `https://sourceforge.net/projects/navmol/files/Navmol-SoundSpacialization-beta.zip/download`.

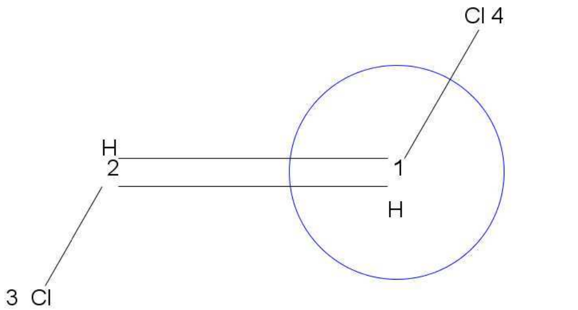


Figure 2: Navmol’s graphic representation of *trans*-1,2-dicloroethene. The circle marks the selected carbon atom C 1.

Navmol is implemented in Java and uses the Chemistry Development Kit (CDK) library and the FreeTTS text-to-voice synthesizer. For the spatial audio it uses the JOAL library [13].

3. SPATIALIZED SOUND

When a sound is emitted, it suffers many modifications before it reaches our ears. Some of these modifications are caused by our own body, namely by the head, torso and pinnae. In particular, when sound travels through the pinnae it is reflected in different ways depending on the orientation of the head relative to the sound source. In other words, a series of reflected waves are generated due to the shape of the pinnae. The phase differences between the direct and reflected waves depend on the sound source’s location [14]. As a result, the intensity of the sound’s frequency components change, that is, the pinnae have a directional filtering effect. These spectral differences as well as the delays of the reflected waves are some of the cues used by the brain to perform sound localization.

Synthesized spatial audio can be generated by changing the right and left channel signal to simulate the pinnae filtering effect. This can be performed with HRTFs, which are frequency fil-

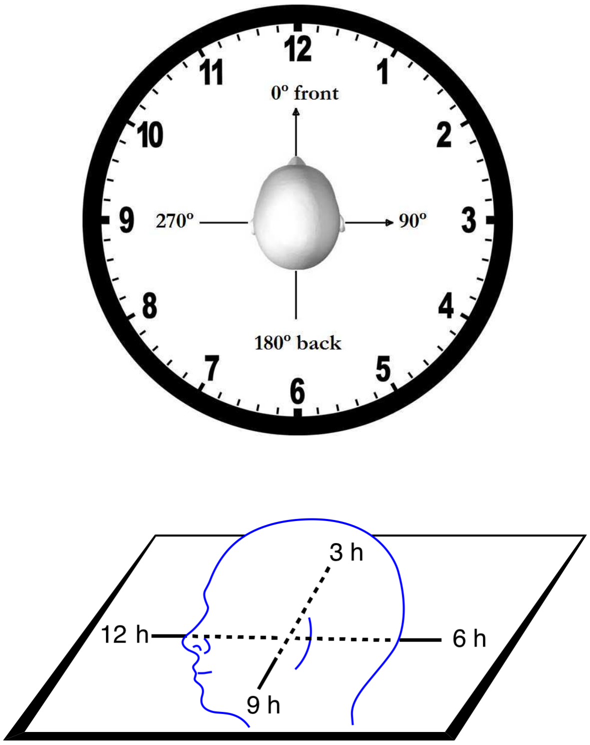


Figure 3: Representation of the head (facing 0° in the azimuth) and the direction of sound according to an analog clock. From above (top) and from a side perspective (bottom).

ters that change the signals according to the effects that the head, torso and pinnae have on the sounds. These filters depend on the source's azimuth, elevation and distance relative to the listener [15].

Navmol modifies the voice synthesizer signal with HRTFs. Instead of generating a mono audio signal with the description of the atom's neighbors, Navmol generates two signals (for the left and right channel) that when heard through headphones give the perception of spatial audio. Here we are using 2D sound, with the sound's direction varying in the plane perpendicular to the head, cutting across the nose and ears, and with 0° ahead of the listener as shown in figure 3.

As an example, let us consider the molecule in figure 2, which has a planar structure, with the center of all its atoms occupying positions in the same plane. When the users are positioned at atom C 1, Navmol indicates that the atom has two neighbors: "atom Cl 4 at 1 o'clock through a single bond" and "atom C 2 at 9 o'clock through a double bond" (carbon atoms are represented only by the number). If the users wear headphones, they will be immersed in a virtual scene in which they are centered at the atom with the circle (C 1 atom) and will hear those sentences as if coming from 30° and 270° in the horizontal plane, respectively.

4. HEAD RELATED TRANSFER FUNCTIONS SETS

HRTFs can be obtained by recording sounds with microphones inside the ears of dummy heads with pinnae. Different pinnae cause slightly different effects on the sounds as their filtering effect depends on their actual shape. Thus, HRTFs depend on the pinnae used to do the recordings.

As a result, when sound is modified with a given set of HRTFs, not everyone perceives the spatialization in exactly the same manner. Therefore, the set of HRTFs that should be used depends on the listener. It should be the one that approximates better the filtering of the listener's pinnae.

When users start using Navmol, they need to choose the set of HRTFs that works best for them. To make this task easy, Navmol has an accompanying application, the *HRTFs-test-app*, that helps users choose the set of HRTFs. When the users run this application, they will hear the sentence "atom at X o'clock" several times and coming from random directions. The users must introduce the direction from where they think the sound is coming using the clock reference system and then the application calculates the error. The users can run the application with all available sets of HRTFs and choose the one for which they obtained the lowest error.

Many sets of HRTFs exist. We used 53 sets: one set from KEMAR [16], one from CIAIR [17] and 51 from IRCAM [18]. The sets were compiled from the initial HRTFs wave files to the *mhr* format, so as to be in accordance with the OpenAl-Soft implementation of the HRTFs done by JOAL, while also drastically reducing the size of the HRTF measurements.

While Navmol's installation package includes all the 53 sets of HRTFs, it is not convenient for the users to test all the sets. In order to reduce the number of suggested sets to the user, we tried all 53 sets with five volunteers and chose the five sets that worked better for them all, that is, the five best sets for all volunteers. Then we ran a small test with 10 volunteers and these five sets to determine if it was necessary to include more than one HRTFs set in *HRTFs-test-app* sets' suggestion. This procedure is described in section **The HRTFs sets test**.

Following the results of this procedure, Navmol's package suggests the following sets of HRTFs: KEMAR, CIAIR, IRC05, IRC25 and IRC44. The users can run the HRTFs-test-app with these five sets and choose the one that works better for them. Afterwards, they can install Navmol with that set of HRTFs and start using it with spatial audio. (If the users do not feel satisfied with these sets, they can run the application with any other set.)

4.1. The HRTFs sets test

The goal of this informal test was to reduce the number of HRTFs sets suggested to the users, so as to make the task of choosing a set of HRTFs easy and not tiring. All the tests described here were made using the same Asus laptop running a 64-bit version of Windows 8 and a Sennheiser HD 202 headset.

We tried the 53 sets of HRTFs in random order with five volunteers. Each volunteer heard the sentence "atom at t o'clock", with t varying from 1 to 12. The sentence was spatialized so as to be heard from the corresponding direction. For instance, if the sentence was "atom at 3 o'clock", the sound's direction was 90° (figure 3).

The volunteers heard this sentence in a clockwise direction twice. The sound would start at 0° in front of the user (12 o'clock)

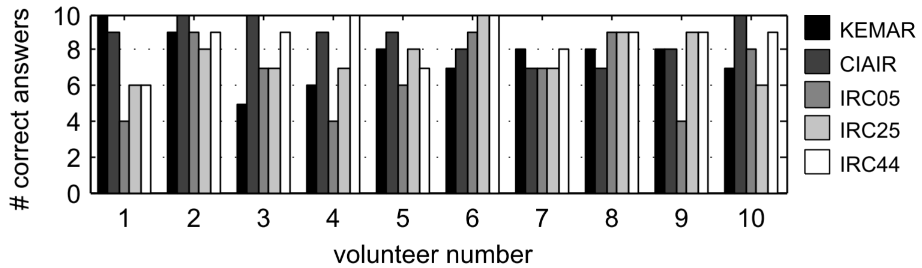


Figure 4: Number of correct answers for each HRTFs set.

and would then proceed for each of the distinct directions, 1 o'clock, 2 o'clock, etc. all the way to 12 o'clock again. The volunteers would then classify the sound spatialization from 1 to 4, with 1 being "I could not determine at all the sound's directions" and 4 being "I could hear all the directions perfectly".

Afterwards, the volunteers heard the sentence again but coming from 10 random directions and we asked them to identify the direction of the sound (in terms of hours in an analog clock). The correct answers were counted for each of the 53 sets of HRTFs.

As expected, even though all HRTFs sets had positive scores, different volunteers gave different scores to each HRTFs set. The five best distinct HRTFs for all volunteers were chosen. These were the CIAIR, KEMAR, IRC05, IRC25 and IRC44.

Only five volunteers participated in this procedure because the tasks described took many hours for each person. Afterwards we performed another test with 10 volunteers but using only the five HRTF sets mentioned above. Again, the order of the HRTFs sets was random.

For this second test, we used the sentence "atom at X o'clock" (without replacing X by 1, 2,. . . , 12). This test started with a training phase during which the volunteers heard that sentence in clockwise order (for two complete turns) and the direction of the sound varying accordingly.

The main task of the test was performed after the training period. The volunteers heard the sentence 10 times (from random directions) and were asked to identify the direction of the sound. Since humans have a localization accuracy that can diverge up to a few degrees from the correct direction, we considered correct all answers that were displaced by less than 30° to either side of the correct direction (for instance, if the correct direction was 3 o'clock, we considered 2, 3 and 4 o'clock correct and all other answers incorrect) [19].

Figure 4 shows the number of correct answers for each of the five HRTFs set. As it can be observed, while there is at least one set of HRTFs for each subject with more than eight correct answers, as expected the set with the highest number of correct answers varies from subject to subject.

Following these results, Navmol's installation package includes many sets of HRTFs so that the users can use the set they feel the most comfortable with, without having the need of searching and compiling the measurements by themselves.

5. RESULTS

In order to assess if users can understand the molecules' structure when Navmol uses sound spatialization, we ran a usability test in which the volunteers had to interpret two molecules using Navmol. Due to the difficulty in having a high number of BVI volunteers to test Navmol, before running the test with BVI people, we ran it with sighted volunteers. This way we could guarantee that the expected outcome was obtained and there were no corrections or modifications to the software that had to be made before running the test with the BVI volunteers.

Before the actual test started, the participants tried the five HRTFs sets to determine which worked better for them. They heard the sentence "atom at X o'clock" from 10 random directions (where X was not replaced by the time) and were asked to estimate the sound's direction. The HRTFs set with a minimal number of wrong answers (using the same tolerance as above) was chosen and the rest of the test was performed with that set.

After the HRTFs set was chosen, there was a training phase. During training the participants heard the same sentence again but where X o'clock was substituted by the matching direction (1 to 12 o'clock). They heard the sentence for all 12 directions in clockwise order and for two full circles. Then they heard the same sentence again from 10 random directions but without information on the actual direction (*i.e.*, X was not replaced by the corresponding time in hours). They were asked to estimate the sound's direction and received feedback on their answers, that is, the training application spoke back the correct direction of the spatialized sound.

After the training period, the main task of the test started. This consisted of exploring two molecules with Navmol (figure 5) and reproducing their structure, which for sighted users consisted of drawing them in paper. Before starting the task, the participants were shown a molecule in Navmol's window and a demonstration on how to explore it. They tried Navmol by themselves so that they could familiarize with the controls, interface and the synthesized

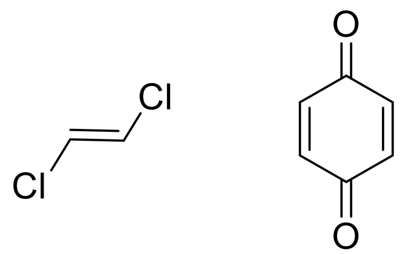
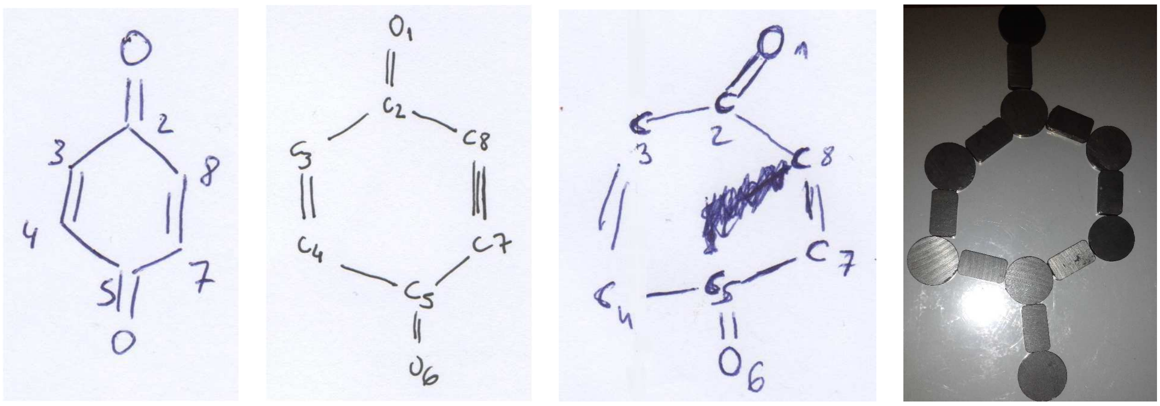


Figure 5: Molecules used in the first test.



(a) (b) (c) (d)

Figure 6: Answers from the usability test. (a)-(c) Answers from three sighted participants. (d) Answer from a blind participant.

voice. After they felt confident on how to use Navmol, the task started. For this, the participants had no access to the screen. They could only use the keyboard to explore the molecule and hear its information.

The point of this task was to see if the participants could understand and reproduce the complete molecules structure from just the spatialized sound cues. All the sentences with the molecule name or constitution were omitted and only the atom bonds were left intact but with the omission of the direction of the bond. For instance, if the sentence for Navmol's normal version was "atom O 2 at 6 o'clock through a double bond" the sentence in the task would only say "atom O 2 at X o'clock through a double bond". This task had two distinct molecules that were given to the participants in a random order.

5.1. Results from sighted volunteers

We ran the test with 10 chemist volunteers and 15 non-chemist volunteers. The chemist volunteers had ages between 22 and 56 (8 women and 2 men, 6 PhD, 3 MSc and 1 BSc). Two of them reported having some hearing problems. The non-chemist volunteers were university students with ages between 18 and 27 and all with normal hearing. There were 10 men and 5 women. They all wore headphones to perform the test and used an external keyboard (so that they did not face the laptop screen when Navmol was running).

It was observed that using only the sound spatialization cues, the participants were able to navigate and understand the molecules of this study. All participants were able to draw the two molecules with only minor errors: on average there were 0.5 errors out of 12 possibilities (12 atoms), being that the chemist participants had an average of 0.9 errors and the other group had an average of 0.3 errors out of the possible 12. The typical errors were a missing bond in a molecule or the representation of a bond not being in the exact correct direction. Figure 6.a-c shows three examples of typical answers from this test (all representing the same molecule). The leftmost and center answers are completely correct. The rightmost answer is also correct but it has only some small orientation errors: the top oxygen atom should have been drawn right on top of the C 2 atom, like in the other two answers, and the C 4 atom has also a small orientation error.

After having drawn the complete molecules or most of the atoms and bonds, most of the chemist volunteers recognized the molecules in question, being that some of them reported that they used this knowledge to finalize drawing them. While this knowledge could have helped them to perform the task, the answers from the non-chemist volunteers show that people who do not have this type of knowledge can also draw the molecules. Whilst the non-chemist volunteers did not recognize the specific molecules of the test (as expected since they did not have the same chemistry background as the first group), they could grasp their structures correctly. This shows that Navmol's sound spatialization can successfully transmit the molecular structure information to users.

As reported above, the chemist volunteers had a slight higher error average than non-chemists. Some of the people from the chemist group recognized the molecules and used this knowledge to help them perform the task. It appears that this can have led them to make more "mistakes" than the people in the other group because reproducing a molecular structure from a chemist's point of view is not necessarily reproducing a drawing. Such a molecular structure is chemically defined by the elements of the atoms, their connectivity and the bond orders, independently of the orientation of the molecule.

5.2. Results from BVI volunteers

In spite of our efforts, we only managed to have four BVI volunteers for this test: two blind male participants and two low vision female participants. The blind participants were 54 and 67 years old. The youngest was blind since he was 2.5 years old, and the oldest was blind from birth. The low vision participants were 22 and 46 years old. The youngest was a university student, who was visually impaired since she was 7 years old, and the oldest had low vision from birth. The three older participants had a bachelor's degree.

The same protocol as described above was used but, due to the vision impairments of the volunteers, some adjustments were made. The instructions were printed both in Braille (for blind volunteers) and with size 16 font (for low vision volunteers). We also read the instructions to them if they preferred it this way.

While the low vision participants were able to draw the molecules, the blind participants used another technique. We provided the

Figure 7: Metal board and magnets for answers from blind participants.

blind volunteers with a set of magnets that they could arrange in a metal board to replicate the molecules' structure (figure 7). There was a set of circular magnets (representing atoms) and rectangular magnets (representing connections between atoms).

All four BVI volunteers correctly represented the two molecules without any errors. Figure 6.d shows the answer from one of the blind volunteers to the same molecule as that of figure 6.a-c. This answer is correct and, as it can be observed, represents the same structure as the drawings in figure 6.a-c. This shows that Navmol's sound spatialization is appropriate to transmit the molecular structure information to BVI users.

Having said that, there is still another question about the sound spatialization that we have not mentioned: is Navmol better suited to transmit the molecular structure information to BVI users with or without spatial audio? While the users can choose to use Navmol with or without spatial audio, according to the youngest low vision volunteer it is easier to interpret the molecules' structure with spatial audio. This volunteer had participated in a previous Navmol test with no spatial audio [9]. As such, she said that the sound spatialization helped very much with the mental conceptualization of the molecules. She referred that with the previous Navmol version, the users had to wait until the end of the sentence that described the atom ("atom ... at ... o'clock...") to know where the atom was positioned. Now, with the sound spatialization, she could identify the direction where the atom was located right from the beginning of the sentence. She mentioned that this helped her immensely to build the correct mental representation of the structure of the molecules.

6. CONCLUSIONS

While spatialized sound has been used before in computer games, here we propose to use it in an application that can make a difference in the lives of BVI students. In order to help these special needs users to understand the molecular structures, we are developing Navmol, a molecular navigator/editor for BVI users that describes molecular structures with synthesized voice. As a contribution to help these users to more easily conceive a mental representation of the structures, we propose to use spatial audio. Thus, we modified the speech synthesizer's signal with HRTFs such that when users hear the molecule description they will immerse in a virtual scene where they stand on the atoms of the molecule and hear the information coming from different angles of the azimuth, which depend on the directions of the bonds in the molecule.

As observed in the usability test described above, the introduction of spatialized sound to Navmol is a positive addition to its usability, being that by itself, the sound spatialization is enough for the user to realize the correct and complete perception of planar molecular structures. A BVI volunteer who had experimented Navmol with no sound spatialization observed that when the spatial audio is used it is easier to understand the molecules' structure.

The main novelty of this work is the use of spatial audio in conjunction with a polar coordinate system based on the representation of an analogue clock to transmit information about the structure of molecules to BVI people. As a result, this application can have impact in the lives of real people who are typically under-served.

7. ACKNOWLEDGMENTS

We thank all the volunteers who participated in the tests. Special thanks to the team from ACAPO and from Biblioteca Nacional de Portugal, namely to Carlos Ferreira and Maria Aldegundes, for all the help with the tests with BVI subjects.

This work was funded by Fundação Calouste Gulbenkian under project "Ciência para todos: STEREO+", and Fundação para a Ciência e a Tecnologia through LAQV, REQUIMTE: UID/QUI/50006/2013, and NOVA-LINCS: PEest/UID/CEC/04516/2013.

8. REFERENCES

[1] *INE Resultados Definitivos Censos 2011*, Instituto Nacional de Estatística, Lisboa, Portugal, 2012.

[2] *SNR Inquérito Nacional às Incapacidades, Deficiências e Desvantagens - Resultados Globais*, Secretariado Nacional de Reabilitação, Caderno n.8., Retrieved Sep/2008 from `http://www.inr.pt/content/1/117/informacao-estatistica`.

[3] *INE Resultados Definitivos Censos 2001*, Instituto Nacional de Estatística, Lisboa, Portugal, 2002.

[4] "APPLICA & CESEP & ALPHAMETRICS, men and women with disabilities in the EU: Statistical analysis of the LFS AD HOC module and the EU-SILC, final report," 2007.

[5] V. Lounnas, H. Wedler, T. Newman, G. Schaftenaar, J. Harrison, G. Nepomuceno, R. Pemberton, D. Tantillo, and G. Vriend, "Visually impaired researchers get their hands on quantum chemistry: application to a computational study on the isomerization of a sterol," *J. Comput. Aid. Mol. Des.*, vol. 28, no. 11, pp. 1057–1067, 2014.

[6] V. Lounnas, H. Wedler, T. Newman, J. Black, and G. Vriend, "Blind and Visually Impaired Students Can Perform Computer-Aided Molecular Design with an Assistive Molecular Fabricator," *Lecture Notes in Computer Science*, vol. 9043, pp. 18–29, 2015.

[7] "MOLinsight web portal," Retrieved Apr/2017 from `http://molinsight.net`.

[8] F. Pereira, J. Aires-de Sousa, V.D.B. Bonifácio, P. Mata, and A. M. Lobo, "MOLinsight: A Web Portal for the Processing of Molecular Structures by Blind Students," *Journal of Chemical Education*, vol. 88, pp. 361–362, 2011.

[9] R. Fartaria, F. Pereira, V.D.B. Bonifácio, P. Mata, J. Aires-de Sousa, and A. Lobo, "NavMol 2.0 - A Molecular Structure Navigator/Editor for Blind and Visually Impaired Users," *European Journal of Organic Chemistry*, vol. 8, pp. 1415–1419, 2013.

[10] L.C. Widerberg and R. Kaarlela, *Techniques for eating, a guide for blind persons*, Western Michigan University, 1981.

[11] "Demor, location based 3D audiogame," Retrieved Apr/2017 from `http://blendid.nl/index8803.html`.

[12] D. Simões and S. Cavaco, "An orientation game with 3D spatialized audio for visually impaired children," in *Proceedings of Advances in Computer Entertainment Technology Conference*, 2014.

[13] "Java Bindings for the OpenAL API," Retrieved Apr/2017 from `https://jogamp.org/joal/www/`.

[14] S. Handel, *Listening: An Introduction to the Perception of Auditory Events*, The MIT Press, 1989.

[15] J. Blauert, *Spatial Hearing*, The MIT Press, 1983.

[16] B. Gardner and K. Martin, "HRTF measurements of a KEMAR dummy-head microphone," Tech. Rep., MIT Media Lab Perceptual Computing, 1994.

[17] T. Nishino, S. Kajita, K. Takeda, and F. Itakura, "Interpolation of head related transfer functions of azimuth and elevation," *J. Acoust. Soc. Jpn.*, vol. 57, no. 11, pp. 685–692, 2001.

[18] "Listen HRTF Database," Retrieved Apr/2017 from `http://recherche.ircam.fr/equipes/salles/listen/index.html`.

[19] D. Wang and Eds. Brown, G. J., *Computational auditory scene analysis: principles, algorithms and applications*, IEEE Press/Wiley-Interscience, 2006.