Ponencia presentada en la International Conference of New Technologies in Science Education (CINTEC), Aveiro, Portugal, 4-6 Julio de 2001.

USING NON-SPEECH SOUNDS TO CONVEY MOLECULAR PROPERTIES IN A VIRTUAL ENVIRONMENT

M. A. Garcia-Ruiz

School of Cognitive and Computing Sciences
University of Sussex, Brighton, BN19QH, England.
E-mail: miguelga@cogs.susx.ac.uk.

Abstract

This paper describes theoretical grounds and an ongoing research on the use of non-speech sounds (sound effects or musical tones that convey a message) intended for complementing visualization of properties and structure of biological molecules in a multimodal computer interface. This research suggests that a structured combination of everyday and abstract non-speech sounds is an effective way to associate and identify molecular properties in a virtual environment. A preliminary test was carried out to assess the memorization and usability of non-speech sounds played in a desktop virtual environment, where sounds represented molecular properties. Informal analysis of results shown high recognition rates when using a combination of sound effects and musical tones. Further tests and adaptations to the system are needed. The outcomes of this research are intended for application in molecular biology courses.

Keywords: Non-speech sounds, hybrid auditory interfaces, virtual environments, multimodal interfaces, molecular biology.

1. Introduction

The proper analysis and identification of physicochemical properties of biological molecules is an important issue for students of molecular biology and related courses. Peterson et al. [12] detected more than fifty misconceptions on molecular properties held by students from secondary education to graduate school. These researchers suggested that the misconceptions were due to inaccurate molecular representations.

Computer-aided learning (CAL) material used in chemistry usually has been used to convey molecular information through graphical user interfaces (GUIs). Recently, virtual reality techniques have been employed to display three-dimensional representations of molecules. One example of these techniques is the use of Virtual Reality Modeling Language (VRML) files that describe molecular structure and physicochemical properties, which are seen using Internet web browsers with special plug-ins installed. An example of this can be found in the web site http://www.webmolecules.com. A screen dump is shown in Fig. 1.

However, molecular visualizations using computer graphics occasionally lack on levels of detail, and in addition to this, some information is occluded. Likewise, the visual channel is sometimes overused in molecular visualization [9]. This research suggests an alternative to present visual-only molecular information by using non-speech sounds (musical tones or sound effects that can convey information about objects and actions in a computer interface [4]) to represent physicochemical properties of molecules, that otherwise are displayed along with molecular structure in a computer interface. According to Scaletti and Craig [14], data representations through non-speech sounds can overcome visual overload by supplementing and reinforcing visual information, as well as increasing the user's engagement in a computer interface.

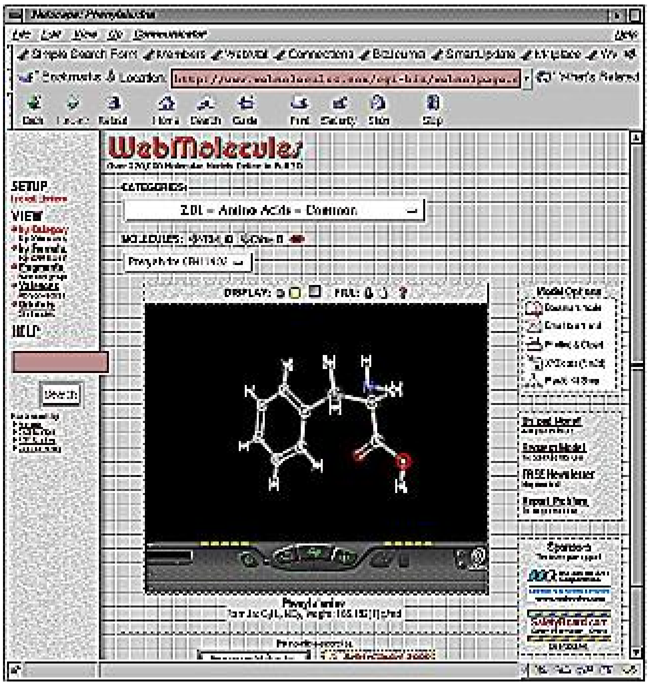

Fig. 1 A Web Page Showing a Molecule of Phenylalanine Using a VRML Plug-in.
(Copyright © 1997-2001, Molecular Arts Corporation)

What kind of sounds should be used in order to represent molecular information? As yet, sound effects associated with molecules and their properties (or sounds naturally generated by these properties, if they exist) are highly unusual in the field of molecular biology. This research proposes to use natural sound effects and abstract musical tones that resemble physicochemical properties of polarity and phobicity. For example, hydrophobicity, a property that literally means "fear of water", can be associated with a sound of a stream of water.

2. Research Background

Non-speech sounds at the computer interface can be defined as non-verbal auditory cues that convey information about events or objects. These can be made of synthesized musical instruments, everyday sound effects, or electronically produced tones. Non-speech sounds can function either as alarms and warning systems, status and monitoring indicators, or encoded messages and data [4]. They can be classified into three branches: Earcons, auditory icons, and sonification.

Earcons: Abstract Non-speech messages that represent data. The application of abstract non-speech sounds in computer interfaces was initially studied by Blattner et al. [2]. These researchers defined an earcon as a cluster of combined musical notes, or even a single one, with specific characteristics, such as changes in duration, timbre and loudness. Usually their timbre is made of sounds of wind, stringed or percussion instruments. Earcons are associated with either objects or actions presented in a computer interface. Because earcons make abstract associations with data, users must learn them in an initial training process.

Auditory Icons: Everyday sounds that represent data. Auditory icons are composed of everyday sounds, also called sound effects, that have a direct association with objects or events in a computer interface. For example, a sound of splashing water (that is, an auditory icon) played at different pitch (which is related to the sound's frequency) can be associated with a computer simulation of a working water pump. This association should be easily and rapidly learned by the user because there is a semantical relationship between the sounds and the objects or actions [6].

Sonification: Sound variations determined by scientific data. Sonification is defined by Kaper et al. [10] as the "faithful rendition of data into sounds", where abstract sounds' variables are parameterized by modifying their frequency, amplitude and duration to map data, often happening in real time. The data to be *sonified* is usually obtained from scientific experiments. Sounds used for sonification are normally composed of synthesized tones. Scientific sonification is the equivalent to scientific visualization.

2.1 Some Benefits of Using Non-speech Sounds

Kramer [11] pointed out that non-speech sounds can increase selective attention and discrimination of visual information. This can be achieved through the so-called "cocktail party effect", which also is described by Arons [1]. Moreover, Kramer suggested that sounds used in conjunction with other modalities can enhance learning by encouraging "fresh interpretations not suggested by traditional representation techniques". Finally, Kramer reported that non-speech sounds are useful to overcome visual constraints by increasing its data dimensionality.

2.2 Some Drawbacks of Using Non-speech Sounds

These are some difficulties that may appear when using non-speech sounds, which are described by Kramer [11]: Sounds played for a long time can be annoying and distracting if they are not well designed. Other negative factors are that some sounds have low spatial resolution, not all people have "absolute pitch" (the ability to recognize a specific musical note) and some people cannot easily discern a specific sound (or changes in sound variables) from a polyphonic interface.

2.3 Non-speech Sounds in Virtual Environments

Three-dimensional sounds (also called spatial sounds) played in virtual environments have been used and tested for providing navigational cues. As well, spatial sounds can convey informative feedback on interactions, and can be used for representing complementary or redundant information along with visualization. According to Wenzel [16], non-speech sounds played at different spatial positions in the computer interface can enhance discrimination between two or more sound sources. Also, this researcher reported that spatial audio cues can be detected faster than visual ones.

According to Dede, et al. [5], auditory cues can enhance attention, engagement and usability of multimodal virtual environments. These researchers created a virtual environment system called 'Newton world', intended for users to learn Newtonian mechanics. This system provides the user with haptic, auditory and visual cues. Dede et al. tested 'Newton world' with secondary school students. Students rated the system easier to use with auditory and haptic conditions than visual condition alone. Also, these researchers reported increased students' attention and engagement, which was positively reflected in the posttests.

The reason why spatial sounds properties have been added to this research is that this kind of sounds could ease the identification of amino acids placed at different spatial positions in proteins. Further tests are needed to assess this idea.

3. Related Work

One of the first successful applications of non-speech sounds was in representation and identification of data patterns in analytical chemistry, reported by Yeung [15]. He associated seven chemical variables to seven sound variables (pitch, loudness, damping, direction, duration, repetition, and rest). With only one training session, his subjects (professional chemists) achieved a recognition rate of 90% in the test. After a second training session, subjects could achieve an accuracy of 98%. This researcher also pointed out the advantages of this method over data visualization.

Another example of this application is a program called 'PROMUSE', developed by Hansen et al. [7], which is a visualization/sonification computer program intended for analyzing protein structural alignments. This program allows the user to associate visualization of protein structures and other information with musical compositions based on Jazz. Hansen et al. reported in their experimental results that this auditory mapping is an effective way for data disambiguation among protein alignments.

Hayashi and Munakata [8] developed and tested a method to analyze DNA sequences by mapping distinctive musical tones into each of the four DNA sequence elements. They found that after a short training session, they could recognize and memorize large sequences, three times larger than those analyzed with conventional methods.

4. Pilot Study

An informal study was carried out to test the memorization of the molecular properties using non-speech sounds. The usability of the non-speech sounds played in the computer system was also tested.

4.1 Participants

Eight graduate students (5 men and three women) took part in the study, which were recruited in a voluntary basis. Only one subject had previous knowledge about the molecules used in the study and previous musical training (the ability to read music and play a musical instrument).

4.2 Material

4.2.1 Molecules Used in This Study

The molecules were six amino acids (the building blocks of proteins), taken from a total of twenty. Amino acids can be classified according to their phobicity and charge properties. The chosen molecules are examples of each of the charge and phobicity values. The amino acids standard nomenclature (letters) was used. The amino acids and their properties can be seen in Table 1. For more information about the amino acids see [13].

Amino acid	***Designed Letter***	***Phobicity***	***Charge***
Arginine	R	Hydrophilic	Positive
Aspartic acid	D	Hydrophilic	Negative
Asparagine	N	Hydrophilic	Uneven
Valine	V	High hydrophobicity	No charge
Cysteine	C	Low hydrophobicity	No charge
Glycine	G	Hydrophilic	No charge (neutral)

Table 1 Amino Acids and Their Properties.

4.2.2 Description of the Virtual Environment System

A software browser/editor for virtual environments called 'DIVE' (Distributed Interactive Virtual Environments), version 3.3X was used for the study, which was developed by the Swedish Institute of Computer Science. The three-dimensional molecular structure and extra information was converted into virtual reality modeling language (VRML) format file from Protein Databank (PDB) files, which were downloaded from the Winthrop University's Department of Chemistry, Physics and Geology web site (used with permission): http://bohr.winthrop.edu/vrml/vrml.html. The conversion was made using program Babel, version 1.6 and program pdb2vrml, developed at the Department of Chemistry, University of Darmstadt. A Silicon Graphics O2 workstation was used. The molecules and their information on properties were visualized on the workstation's screen, and sounds were heard through headphones Sennheiser, model HD 570.

4.2.3 Characteristics of the Non-speech Sounds Used in This Study

The sound sources used in this study were arbitrarily chosen by the author from a sounds library on musical instruments called 'Prosonus' (TM). Library samples were provided with the Silicon Graphics O2 workstation. The chosen sounds were composed of wind, stringed and percussion instruments, and they were stored with Sun/Next format, sampled at approximately 44 kHz with 16 bits depth. Their duration was no longer than two seconds. Earcons and auditory icons were edited using program 'GoldWave' for Windows (TM).

Each hydrophilic amino acid had an earcon attached, which was composed of three tones. Each earcon represented charge and phobicity (See Fig. 2). The earcons were made of distinctive musical timbres. The sounds used for making earcons are listed in Table 2. The earcons were designed according to the guidelines developed by Brewster et al. [3].

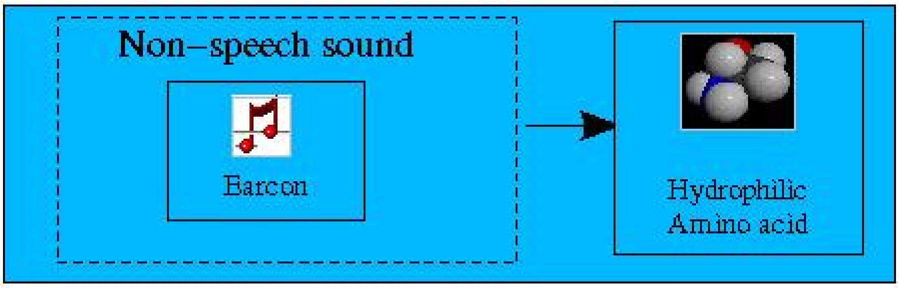

Fig. 2 A Non-speech Sound Associated With a Hydrophilic Amino Acid.

The sounds that represented hydrophobic (which literally means 'fear of water') amino acids were composed of a hybrid of an earcon and an auditory icon. The latter was made of a sound of a stream of water (see Fig. 3). This sound effect is supposedly intended to facilitate its association with a hydrophobic molecule.

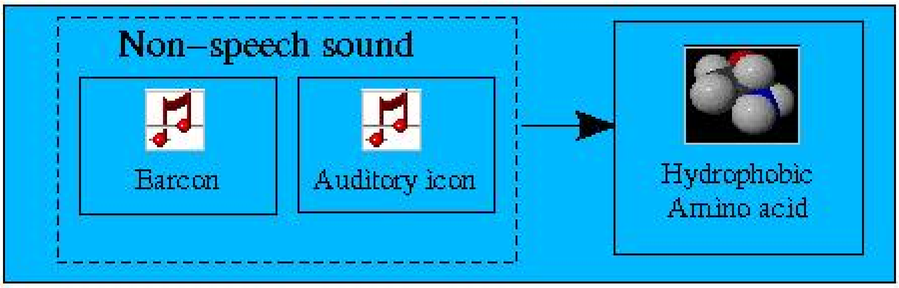

Fig. 3 A Non-speech Sound Associated with a Hydrophobic Amino Acid.

As suggested by Blattner et al. [2], to facilitate the learning of associations between non-speech sounds and objects at the computer interface, molecules, earcons and auditory icons used in this study were arranged into *families*, according to the charge and phobicity. Each family had a distinctive musical timbre, which is determined by a musical instrument. The families are described in Table 2.

Families:	***Musical instruments for making the sounds:***
Positively charged (Hydrophilic):	Wind instrument:

Families:	***Musical instruments for making the sounds:***
Arginine (R)	Pan flute
Negatively charged (hydrophilic): Aspartic acid (D)	Stringed instrument: Acoustic grand piano
Uneven charge (hydrophilic): Asparagine (N)	Stringed instrument: Violin
No charge (Hydrophobics): Valine (V) (high hydrophobicity) Cysteine (C)(low hydrophobicity)	Percussion instruments: Xylophone and sound of strong stream of water Tambourine and sound of weak stream of water
No charge (Hydrophilic and neutral): Glycine (G)	Percussion instrument: Low bongo

Table 2 Classification of Amino Acids and Their Respective Sounds Associations.

These are the associations of charge with earcons' tones:

* *Positive* charge: Earcons had sequential tones with *increasing* pitches (low-medium-high tones sequence).

* *Negative* charge: Earcons had sequential tones with *decreasing* pitches (high-medium-low tones sequence)

* *Uneven* charge: Earcons had sequential tones with *non-scaled* pitches (high-low-high tones sequence)

* *No charge:* Earcons had sequential tones with *similar* pitches.

These are the associations of phobicity with auditory icons:

* *Hydrophobic* ('fear of water'): The musical tones (that represented charge) were played along with a sound effect of running water. If the stream of water sounded *weak*, that meant the hydrophobic property was *low*. If the stream of water sounded *strong*, that meant the hydrophobic property was *high*.

Hydrophilic ('water loving'): There were no sounds of water associated. Sounds had musical tones only.

The earcons associated with the hydrophobics consisted of two musical tones with same pitch.

The frequencies of the tones and sound effects can be seen in Table 3. They were calculated using program 'Sox', version 12.17.1.

Associated property	***Musical instrument or sound effect***	***Rough frequency (in Hz)***
Positive charge	Pan flute	437 572

Associated property	***Musical instrument or sound effect***	***Rough frequency (in Hz)***
		632
Negative charge	Acoustic grand piano	1157 974 810
Uneven charge	Violin	815 678 781
High hydrophobic (no charge)	Xylophone Strong stream of water	3595 3595 3269
Low hydrophobic (no charge)	Tambourine Weak stream of water	9590 9590 1458
Neutral (no charge)	Low bongo	1022 1022 1022

Table 3 Rough Frequencies of the Musical Tones And Sound Effects.

4.3 Experimental Design

The study consisted of two parts. The first one was for training purposes. In this part, all the subjects listened to the same stimuli at different volume levels and stereophonic positions, due to the spatial positions of the molecules in the virtual environment. The second part was for assessing recognition of the molecular properties, where subjects listened to the sounds only (this time at same volume level) and gave information about the molecular properties. Each subject listened to the six sounds in different order. A latin square array determined the order of presentation.

4.4 Procedure

In part one, subjects were given a piece of paper with instructions on how to carry out the training. The instructions also contained information on the association of molecular properties with the sounds. Subjects watched the six molecules, which were previously placed in a random order (a screen dump can be seen in Fig. 4), and read the textual information on the molecular properties shown on screen. That textual information was placed in front of each molecule. Thus, subjects could read the properties information on screen and in the instructions paper. Subjects had to click using the mouse on each of the molecule to listen to its associated sound, as much and as long as they wanted. This training part lasted about five minutes on average. There was no specific order for hearing the sounds. The subjects were not allowed to move the molecules around the virtual environment.

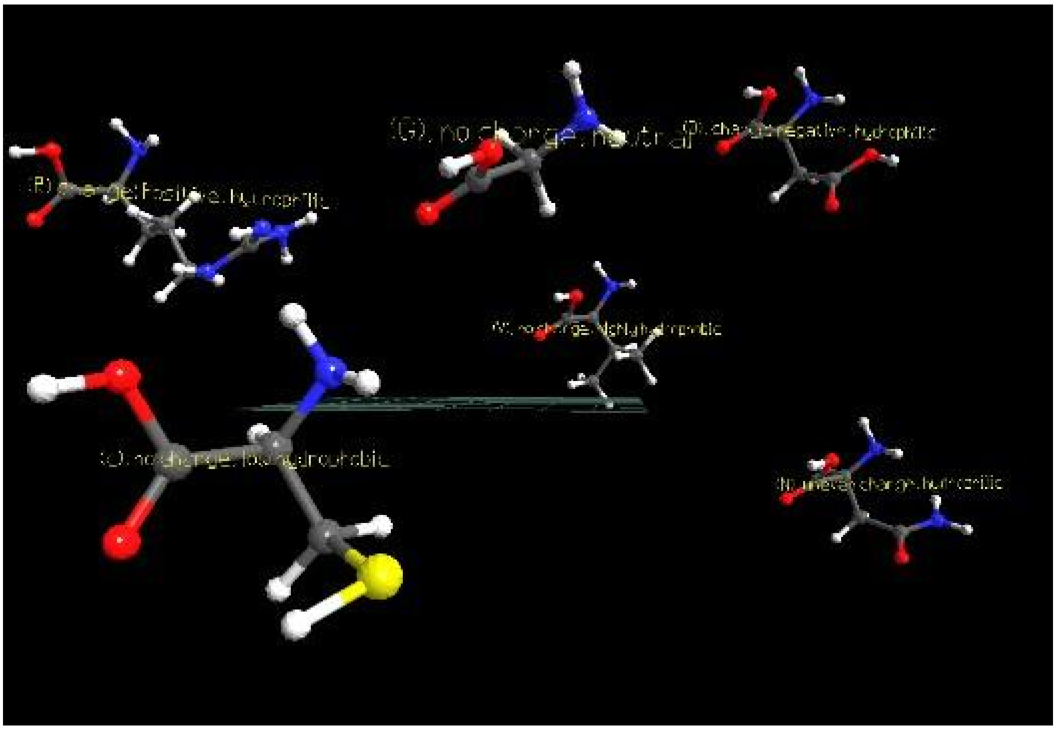

Fig. 4 Amino Acids Displayed in the Virtual Environment.

Part two was for evaluating the subjects' recognition of molecular properties by listening to the sounds only. Once the subjects thought they were ready in the training part, the screen went blank and each non-speech sound was played twice, with one second in between. All sounds were played with same volume and as monophonic. Subjects had to tell the name, the charge and the phobicity (and whether the latter was high or low hydrophobic) of the molecule, according to each of the non-speech sound they just heard. A mark was assigned for each correct answer. Responses were scored with 0 points for each incorrect answer and 1 for each right one.

4. 5 Results and Discussion

Informal analysis of the results showed that subjects recognized well the property of phobicity (whether the molecule was hydrophobic or hydrophilic) with 81% of correct answers, as it is shown in Fig. 5. Most of the subjects reported that it was easy for them to identify phobicity because of the absence or presence of auditory icons in the associated non-speech sounds. The results of the only subject with previous musical training and knowledge on the molecular information were excluded when the analysis was made.

Only 21% of the subjects could identify the molecules' names (letters). Perhaps it happened because all of the letters were consonants, and there was not a possible association between names and visual patterns of the letters.

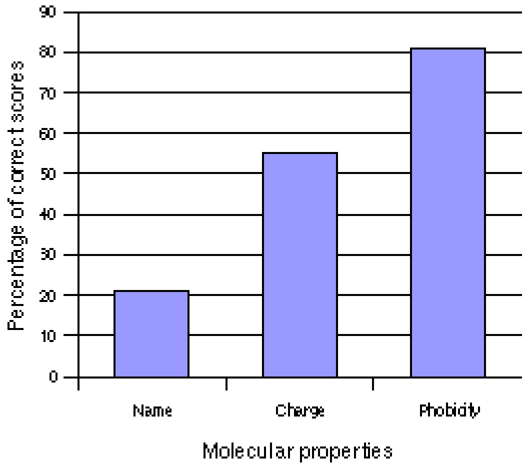

Fig. 5 Scores of Recognized Molecular Properties.

Recognition of charge using earcons showed a correct score of 55%. The reason of this score is perhaps due to faults in the earcons design [3], especially earcons that represented positive charges. This earcon did not have a large difference in its tones' pitch.

As it can be seen in Fig. 6, 71% of the subjects correctly recognized the properties of the molecules with negative and uneven charges, but 57% of the subjects recognized positive charges. Similarly, subjects scored lower when they tried to identify high or low hydrophobicity. 57% of the subjects correctly answered when the sounds represented high hydrophobicity (the ones with the auditory icon of a strong stream of water). This could be because of faults on non-speech sounds design (especially in its recorded volume level). 71% of the subjects identified the property of low hydrophobicity. This reflects that careful design of auditory icons is needed [6].

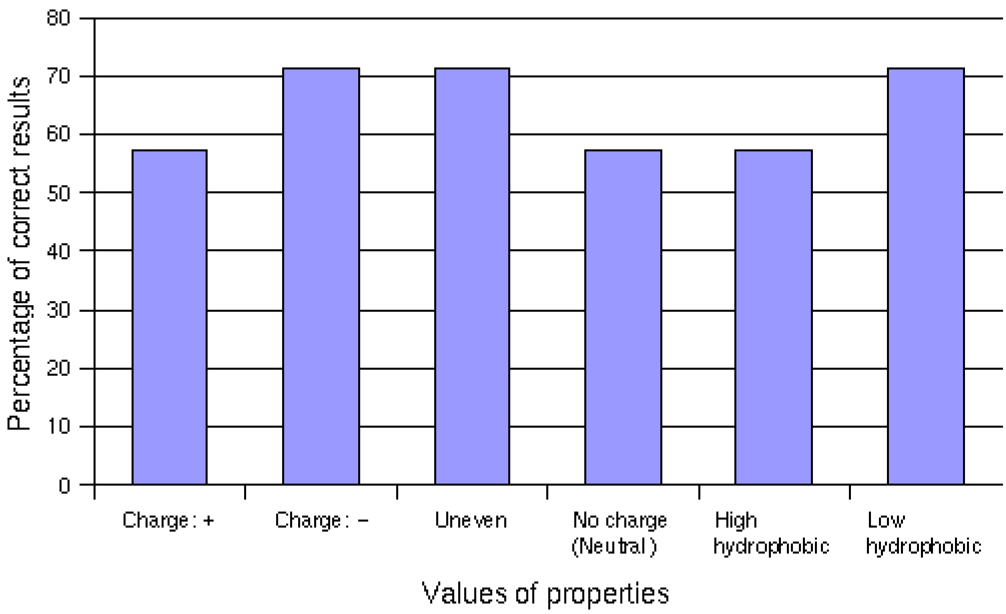

Fig. 6 Scores of Recognized Values of Molecular Properties.

Remarkable results were obtained from the only subject with previous knowledge on molecular biology and on musical training. As it was expected, this subject could easily identify all the molecular properties from all the six molecules, and could identify 66% of the letters. This suggests that users with previous knowledge on the molecular properties could benefit of further elaborated non-speech sounds for use in complex molecular visualizations. Non-speech sounds could represent more molecular properties. More studies are needed to assess this suggestion. It is important to mention that according to Brewster et al. [3], previous musical training does not have a relevant effect in recognizing earcons.

Although for the training part the non-speech sounds were situated at different spatial positions in the virtual environment, the spatial sound capabilities of the virtual environment was not assessed. It was not assessed whether spatial positions of the non-speech sounds affected the sounds' recognition.

5 Future Work

It is necessary to modify and re-test the non-speech sounds that scored low, by using different timbres and pitches. Also, an experiment with proper statistical analysis is needed to compare the effectiveness of audio over visual modality of molecular properties, which should include people with previous knowledge on those properties.

6 Conclusion

The purpose of this paper was to describe a research on the use of non-speech sounds that represent molecular properties in a virtual environment. An early test was carried out to assess the association of hybrid sounds (earcons and auditory icons) with molecular properties. Informal results indicated high recognition rates when using these hybrids of sounds. Further adaptations and tests are needed to confirm the early results. This paper also suggested the application of this multimodal interface research in molecular biology learning.

7. References

[1] Barry Arons, "A review of the cocktail party effect", *Journal of the American Voice I/O Society*, 12, 1992, 35-50.

[2] Meera M. Blattner, Denise A. Sumikawa, Robert M. Greenberg "Earcons and icons: Their structure and common design principles", *Human Computer Interaction*, 4(1), 1989, 11-44.

[3] Stephen A. Brewster, Peter C. Wright and Alistair D. N. Edwards, "Experimentally derived guidelines for the creation of earcons", *In Adjunct Proceedings of HCI'95, Huddersfield, UK*, 1995.

[4] William Buxton, "Introduction to this special issue on nonspeech audio", *Human-Computer Interaction*, 4, 1989, 1-9.

[5] Chris Dede, Marilyn C. Salzman, R. Bowen Loftin, Katy Ash, "Using virtual reality technology to convey abstract scientific concepts", to appear in M. J. Jacobson & R. B. Kozma (Eds.), *Learning the Sciences of the 21st Century: Research, Design, and Implementing Advanced Technology Learning Environments,* Lawrence Erlbaum, Hillsdale, in press.

[6] William W. Gaver, "Auditory icons: Using sound in computer interfaces", *Human-computer interaction*, 2, 1986, 167-177.

[7] Marc Hansen, Eric Charp, Doanna Meads, Alex Pang, "PROMUSE: A prototype for multi-media visualization of protein structural alignments", *Pacific Symposium on Biocomputing*, 4, 1999, 368-379.

[8] Kenshi Hayashi, Nobuo Munakata, "Basically musical", *Nature*, 312, 1984, 96.

[9] Wolf-D. Ihlenfeldt, "Virtual reality in chemistry", *Journal of Molecular Modeling,* 3, 1997, 386 - 402.

[10] Hans G. Kaper, Sever Tipei, Elizabeth Wiebel, "Data sonification and sound visualization", *Computing in Science and Engineering*, 1(4), 1999, 48-58.

[11] Gregory Kramer, "An introduction to auditory display", In G. Kramer (Ed.), *Auditory display, sonification, audification and auditory interfaces. Proceedings of the First International Conference on Auditory Display*, Addison-Wesley, Reading, 1994, 1-77.

[12] Raymond F. Peterson, David F. Treagust, Patrick Garnett, "Development and application of a diagnostic instrument to evaluate grade-11 and 12 students' concepts of covalent bonding and structure following a course of instruction", *Journal of Research in Science Teaching*, 26(4), 1989, 301-314.

[13] Anthony R. Rees, Michael J.E. Sterenberg, "*From cells to atoms/an illustrated introduction to molecular biology"*, Blackwell, Oxford, 1984, 12-13.

[14] Carla Scaletti, Alan B. Craig, "Using sound to extract meaning from complex data" *In Proceedings of SPIE, Extracting meaning from complex data: Processing, display, interaction II*, 1459, 1991.

[15] Edward S. Yeung, "Pattern recognition by audio representation of multivariate analytical data", *Analytical Chemistry*, 52, 1980, 1120-1123.

[16] Elizabeth M. Wenzel, "Spatial sound and sonification", *In G. Kramer (Ed.), Auditory display, sonification, audification and auditory interfaces. Proceedings of the First International Conference on Auditory Display*, Addison-Wesley, Reading, 1994, 127-150.