Available online at www.sciencedirect.com
SCIENCE DIRECT®

Interacting with Computers 18 (2006) 853–868

Interacting with Computers
www.elsevier.com/locate/intcom

An overview of auditory display to assist comprehension of molecular information

Miguel Angel Garcia-Ruiz [a,*], Jorge Rafael Gutierrez-Pulido [a,1]

[a] *School of Telematics, University of Colima, Ave. Universidad 333, Colima 28040, Mexico*

Received 6 August 2005; received in revised form 7 December 2005; accepted 12 December 2005
Available online 3 March 2006

Abstract

This paper presents an overview of auditory display (the use of non-speech sounds to convey information) applied to the study of molecular properties in human–computer interfaces, particularly in virtual environments. Chemistry researchers and students have difficulty in analysing and comprehending molecular structure and bonding and other biomolecular characteristics. Research reports that non-speech sounds have been useful in identifying trends in gene sequences and molecular characteristics, which when used in virtual environments, can facilitate comprehension of complex relationships that are difficult to perceive through visualisation alone.

Keywords: Virtual reality; Multimodality; Auditory display; Non-speech sounds; Molecular structure; Bonding

1. Introduction

Chemistry students and researchers have difficulty in learning and understanding scientific concepts in molecular biology, biochemistry and related areas, due to the abstractness of their contents, as in the bonding of two molecules. The problems are basically due to molecular scale, three-dimensionality, and many factors that determine molecules to bond one another. Students particularly have difficulty in comprehending the bonding process and bond energy (Birk and Kurtz, 1999). In nature, three-dimensional

---

* Corresponding author. Tel./fax: +52 312 316 1093.
E-mail addresses: mgarcia@ucol.mx (M.A. Garcia-Ruiz), jrgp@ucol.mx (J.R. Gutierrez-Pulido).
[1] Tel./fax: +52 312 316 1075.

doi:10.1016/j.intcom.2005.12.001

molecular structure and bonding are present at the smallest microscopic scale (Rees and Sterenberg, 1984). Typically, small organic molecules, such as amino acids, have a size in the order of 0.3 nm and possess intricate shapes. Thus, the scale and shape of molecules (not only amino acids, but also other types of molecules as well) make their bonding more complicated to conceptualise.

Students have to discern that bonds between two molecules are composed of invisible and intangible electrostatic forces. Bonds can also vary in length, number and strength, amongst other properties (Rees and Sterenberg, 1984; Hart et al., 1999), not to mention the relationship between bond length and energy (Alcock, 1990). Moreover, traditional educational materials used to learn molecular structure and bonding present limitations. For many years, students have used molecular models made of either plastic or wood to show, to some degree, molecular conformation and shape. However, these models do suffer significant drawbacks. For example, these types of models can sometimes fall apart, their bond angles are often inaccurate, and students cannot freely rotate the bonds (Petersen, 1970). Additionally, it can sometimes take a long time to build large molecules that can occupy large spaces, complicating both storage and transportation.

Furthermore, many types of physical model sets cannot show bond order (the number of bonds between two atoms or molecules) if there is more than one bond involved. Another problem with the traditional tools of chemistry teaching can be found in chemistry textbooks. Students rely on illustrations and the accompanying textual explanation to understand molecular structure and bonding. Illustrations sometimes cannot convey the spatial characteristics of bonding, because they lack perspective and depth. As a result, students cannot rotate or zoom in the molecular representations in the illustrations, thus limiting the interaction between the students and learning materials (Rzepa and Whitaker, 1995). Furthermore, some books present illustrations of molecules that do not exploit colours to represent atoms and colour gradients to illustrate energy fields. Ranck (1997) commented:

> “Presenting dynamic, color representations as printed black-and-white figures are about as frustrating as presenting music in a sculpture studio. Although both statues and sonatas are art forms and although the sounds in a sculpture studio do have pitch, intensity, and harmonic structure, a hammer and chisel are poor substitutes for a piano.”

In the past few years, the analysis and study of molecular properties have been carried out using graphical representations in computer interfaces. More recently, however, computer technology has contributed to the more widespread use of virtual reality (a computer-generated three-dimensional space, called virtual environment, where a person can interact with it using more than one human sense). Along with graphical visualisations, molecular information has been represented through auditory properties. These audio–visual combinations (multimodality) have been proven to help comprehend molecular behaviour. What follows describes relevant research on auditory display applications in human–computer interfaces, in particular for the study of molecules.

2. Research on auditory display techniques in human–computer interaction

Auditory display is the application of non-speech sounds in human–computer interfaces to convey meaningful information (Kramer, 1994). Non-speech sound cues in auditory displays are non-verbal cues that convey information about events or objects in the computer interface. These can be made of digitally recorded or synthesized musical instruments, everyday sound effects, or electronically produced pure tones (Blattner et al., 1989; Gaver, 1986; Kramer, 1994). It is possible to classify non-speech sound cues according to their applications and sources mainly into three techniques: Earcons, auditory icons, and sonifications.

2.1. Earcons: abstract non-speech sounds that represent information

Blattner et al. (1992) defined an earcon as a combination of musical notes, called motives, or even a single one, with specific characteristics, such as changes in duration, tone/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. Earcons can be arranged into families to map groups of data and can inherit their particular properties.

2.2. Auditory icons: everyday sounds that represent information

Auditory icons are composed of everyday sounds, or sound effects, that should have a direct association with objects or events in a computer interface (Gaver, 1986). For example, a sound of splashing water played at different pitches (which is related to sound's frequency) can be associated with a computer simulation of a working water pump. The user should easily and rapidly identify this association because there is a semantic relationship between the sounds and the objects or actions.

2.3. Sonification: sound variations determined by scientific data

Sonification is defined by Kaper et al. (1999) as the "faithful rendition of data into sounds", where attributes of abstract sounds are parameterised by modifying their frequency, amplitude and duration to map data, this occurring in real time. The data to be 'sonified' is obtained from scientific experiments and analyses. The sounds used for scientific sonification are normally composed of synthesized musical and pure tones. Scientific sonification is the equivalent to scientific visualisation.

Exploratory research done by Sara Bly in early 1980s, about multivariate data analysis using sound characteristics of synthesized tones, pioneered the use of non-speech sounds in the computer interface (Bly, 1982a,b). During the 1980s and part of 1990s, however, past research about non-speech sounds in human–computer interaction was not as extensive as research on visualisation. One reason was because of the limited developments in audio-related technology, especially those developed in early personal computers (Shneiderman, 1998). With faster and more efficient developments in hardware and software for auditory applications, such as digital signal processing (DSP) and the musical instrument digital interface (MIDI) protocol, amongst others (Baggi, 1991; Burgess, 1992), the body of knowledge about auditory display is increasing. Published in 1989, the special issue about non-speech audio of the Human–Computer Interaction journal set a precedent of auditory display theory and practice in computer interfaces (Buxton, 1989). Nowadays, there are special interest groups devoted to non-speech sound research, being the International Community for Auditory Display (ICAD) one of the main forums. Many papers that capitalise useful auditory display applications, such as freeing visual information channel and providing redundant information to other sensory channels are available from ICAD website (www.icad.org).

Auditory display applications to enhance usability of computer interfaces, data analysis, and educational applications grounded that research. All of them are described in Salzman et al. (1996), Kaper et al. (1999), and Kramer (1994), and elsewhere in the seminal ICAD’92 proceedings book. The following reviews are relevant research about those applications. One of the first studies about usability enhancement of computer interfaces with non-speech sounds was carried out by Gaver (1986, 1989). He implemented and evaluated parameterised auditory icons that conveyed redundant feedback on user actions at the Macintosh operating system’s graphical environment. Gaver termed that sonified interface the SonicFinder. He claimed that auditory icons can be learned easily, and their parameterisation (changes in sounds’ pitch) conveyed extra information on user actions, such as the state of file copying. However, Gaver showed that it is hard to find suitable timbres (made of environmental sound effects) for developing auditory icons to be used in computer interfaces, because some of them could not accurately represent the meaning of their associated action. That could happen if the activity or object being represented by auditory icons is particularly abstract. Another problem found is that playing auditory icons for a long time can annoy people situated near the sonified interface.

Parameterised auditory icons have also been tested to assist collaborative processes. Gaver et al. (1991) developed a computer simulation of a bottling plant, called ARKola. In it, two teams of users attended two different stages of the bottling process by monitoring state processes through listening to up to 14 auditory icons that served as auditory feedback. The sounds that represented those states were played through digital MIDI-controller samples. The rhythm persons listened to indicate how fast the machinery was running, and the absence of sound represented a broken machine. In addition, a number of sound effects indicated other states. For example, a sound of crashing glass meant that some bottles were lost. Participants had to adjust parts of the machinery according to the auditory icons they heard to correct a problem. The main contribution of ARKola experiment was that participants were well aware of dynamic events at the interface by listening to parameterised auditory icons. However, those parameters conveyed little information about the magnitudes of those events. One possible reason was the similarity of timbres and pitches of the auditory icons, although they have been proven to be difficult to design. Moreover, Gaver et al. (1991) used many sounds in the ARKola simulator, causing confusion amongst users, who left some processes unattended because of the resulting cacophony of sounds.

Another approach of non-speech sounds applications in computer interfaces concerns using earcons (abstract musical tones that convey information about actions, events, or objects at a computer interface). It is important to note that the mappings betweenthe information and the earcons must be learned, because the sounds do not have a direct meaning or do not provide direct cues about the represented information (Blattner et al., 1989). Stephen Brewster et al. have carried out extensive studies about earcons applications in computer interfaces (Brewster, 1996, 1998a; Brewster et al., 1994), to name some, and has produced a set of guidelines for earcon creation and use (Brewster et al., 1995).

Brewster et al. (1994) used earcons that provide feedback about scrollbar manipulation, and do text searches and navigation tasks in a computer graphical user interface (GUI). A test with 12 experienced users with GUIs indicated that earcons improved search and navigation performance by up to 25%, although there were no significant differences in participants' errors. This experiment shows that earcons (which are composed of simple tones) can be helped to analyse occluded data and assist in performing visually demanding tasks at the computer interface.

Earcons associated with tool palettes were tested in a computer drawing program to enhance its usability (Brewster, 1998b). The problem with tool palettes is that users often do not remember the active palette tool. They often must focus their vision to see which one of the palette tools is currently active, which causes confusion and leads to the incorrect selection of the appropriate palette, thus contributing to drawing errors. Participants in the study made significantly fewer errors when selecting and using the palette tools accompanied by earcons. This study also looked for participant annoyance caused by the earcons they heard, although this was not significant between conditions.

Brewster and Pengelly (1998) have proposed the use of non-speech sounds to complement haptic and visual information in multimodal virtual environments, to assist users in learning abstract concepts and skills. They devised a desktop virtual environment for visually impaired people to perceive graphic bar charts through haptic and sound modalities, although the paper showed no conclusive results about their tests. However, this paper highlights that timbre and spatial location properties of non-speech sounds could be effective for information discrimination in a virtual environment, and it shows the importance of using qualitative and quantitative data collection methods to evaluate multimodality in virtual environments. Similarly, a simulator for veterinary training carried out by Crossan et al. (2000), where non-speech sounds provided state information and warning feedback about a multimodal virtual environment for training veterinarians to analyse horse ovaries. Ovary palpation is a very delicate skill that veterinarians have to learn. Participants did the palpations to virtual ovaries using a commercial force-feedback input device with six degrees of freedom, and heard non-speech sounds with its pitch mapped onto the force applied with the force-feedback device, meaning the higher the pitch, the higher the force applied to the virtual ovary. Very high pitch could mean a great danger done to the ovary.

Few experimental studies about educational auditory display have been reported to date. One of them is an investigation done by Upson (2001). This researcher tested sonifications of math graphs with six sighted students with ages ranging from between 7 and 13. Qualitative evaluations revealed that students were able to identify data trends from the sonifications, and most of them commented that the use of sounds was motivating and fun. Interestingly, Upson (2001) also found that most teachers who were presented math sonifications were reluctant to include their use in math classes, although Upson did not explain why that happened. Similarly, an ongoing study was proposed by Bonebright et al. (2001), which includes testing of mono and stereo sonifications of graphs, as well as rhythm changes. Other studies about educational non-speech sounds are reviewed in a further section of this paper. What we can see from those studies is that the sounds increased students' motivation and engagement, fostering learning, perhaps as a consequence of short-term novelty. Learning with non-speech sounds also agrees with past research about general applications of multimedia technology to assist learning, in the sense that conveying information through multiple media can be engaging and motivating, provided that they have meaningful contents and adequate interaction (Kozma, 1991; Aldrich et al., 1998). In addition, a particular characteristic about combining auditory and visual modalities in a multimodal interface is that presenting information through both modalities allows relief of learner's cognitive load (Mousavi et al., 1995), which in turn should facilitate learning of abstract or complex information, as described in Chandler and Sweller's Cognitive Load Theory (Chandler and Sweller, 1991).

3. Past research regarding HCI and auditory display applications in chemistry

Project GROPE (Brooks et al., 1990) investigates the use of haptic and visual modalities for understanding and performing protein docking in an immersive virtual environment. Brooks et al. represented molecular force fields as force feedback using an adapted six degrees-of-freedom haptic manipulator, originally used for manipulation of toxic material at Argonne Laboratories in the US. The researchers found in experiments done with seven Computer Science students that performance increased about twofold when students carried out a molecular docking task with the haptic device. The multimodal system has been tested with professional chemists, which reported feeling immersed in the docking experience and gaining new insights about how new molecules dock and react with one another. Brooks et al. (1990) pointed out advantages of haptic display over docking visualisation. A future implementation of this project plans to use auditory feedback along with haptic feedback for performing molecular docking, although no auditory implementations are reported to date. The haptic device and the virtual environment, as well as smaller desktop versions, have been tested a number of times since the project began in late 1960s.

One of the first research projects that used sounds to analyse chemical data was carried out by Yeung (1980). This researcher mapped seven sound characteristics (pitch, loudness, damping, direction, duration, repetition and rest) to seven chemical variables used in analytical chemistry. Four professional chemists took part in an experiment to test those mappings. With only a short training session, three chemists achieved a recognition rate of 90%. After a second training session, two chemists obtained a rate of 98%, and one chemist, who had previous musical experience, could recognise 100% of chemical patterns. Although Yeung claimed that using this auditory method over data visualisation offered advantages, he did not explain in detail which were those advantages. In addition, Yeung did not explain that loudness might be annoying at high levels. Moreover, Yeung did not mention that loudness is not appropriate for chores requiring accurate data discrimination. Loudness resolution (the variations of sound intensity, measured in decibels) can be difficult to perceive (Brewster et al., 1995).

Bly (1982a,b) carried out similar experiments to Yeung's. Bly mapped multivariate data onto seven sound characteristics of a note: pitch, volume, duration, the fundamental wave shape, the attack envelope, and the addition of a fifth and ninth harmonics, and tested the mappings with a group of participants. Her experimental results showed that participants improved their ability to discriminate between two data sets when they used the audio mappings. This indicates that sound characteristics different to the ones used by Yeung in his experiments, like the attack and adding harmonics, are also useful for representing and discriminating multivariate data.

Past research has made remarkable attempts to map sounds onto DNA sequences to analyse scientific data and for artistic purposes. This idea has been suggested by Douglas Hofstadter in his book Gödel, Escher, Bach (Hofstadter, 1980), where the author makes analogies about how music can be related to protein sequences to help identify their patterns, and to decompose information about complex molecules into smaller pieces for easier identification. Hofstadter wrote:

> "Imagine the mRNA to be like a long piece of magnetic recording tape, and the ribosome to be like a tape recorder. As the tape passes through the playing head of the recorder, it is 'read' and converted into music, or other sounds… When a 'tape' of mRNA passes through the 'playing head' of a ribosome, the 'notes' produced are amino acids and the pieces of music they make up are proteins… Music is not a mere linear sequence of notes. Our minds perceive pieces of music on a level far higher than that. We chunk notes into phrases, phrases into melodies, melodies into movements, and movements into full pieces. Similarly, proteins only make sense when they act as chunked units."

A scientific application of sounds to analyse DNA sequences was carried out by Hayashi and Munakata (1984), who mapped musical tones within a range of a fifth onto each of the four DNA sequence bases following their thermal stability. These researchers tested this method on them, and found that with a short training session before the test they could easily recognise large DNA sequences, three times larger than those analysed with conventional methods. Hayashi and Munakata (1984) also pointed out that the mappings produced motivating and pleasant musical arrangements, although this could be due to novelty effect. This emotional capability of sounds can be used to enhance a motivational learning experience about molecular bonding. As Roussos et al. (1999) found out, motivation plays an important role in the learning of scientific concepts in virtual environments (as well as in other learning environments).

Another example of auditory display use for DNA analysis is reported by Ohno and Ohno (1986). In this paper, these researchers discussed a method that used musical tones to highlight pattern repetitions in DNA base sequences, which they called as the "principle of repetitious recurrence". The researchers pointed out that these repetitions frequently occur in DNA base sequences and their derivatives, especially in genes. Ohno and Ohno (1986) also explain that the sound mappings of base sequences are remarkably melodic, a sound property that can be further exploited in auditory display research. Their rule for assigning musical notes to each of the DNA bases is based on "the assignment of two consecutive positions each in the octave scale to four bases in the ascending order of A, G, T, and C". The researchers applied that rule to encode notes of a segment of Frederic Chopin's Nocturne Opus 55 number 1 onto sequences of DNA bases. Ohno and Ohno (1986) pointed out that the musical repetitions of the nocturne closely resembles the patterns of base repetitions of an enzyme molecule found in mice. However, the researchers did not mention in the paper why or how they chose the nocturne for doing the comparisons with base repetitions.

King and Angus (1996) developed an algorithm to translate DNA base sequences and molecular properties onto musical notes. They used musical mappings of sequences of the four DNA bases to form the musical top line, and seven amino acid characteristics (which defined those bases) to form the bass line, all within a C major scale. Their algorithm produced more harmonious sounds and was more open to analysis than methods from previous research, such as Hayashi and Munakata (1984) and Ohno and Ohno (1986). The researchers argued that their algorithm offers advantage over visualisation of multivariate data, and pointed out that mapping base sequences and molecular properties onto colours could yield to visual blending, making data identification tiresome.

An application of sound associations with DNA sequences to unravel complex patterns in biochemistry has been developed for edutainment purposes. The Exploratorium Museum of San Francisco in the US exhibits an audiovisual representation of DNA sequences called 'Musical Mutants' (Carlson et al., 1994; Musical Mutants, 2002), where notes from different musical instruments (to ease data discrimination) were mapped onto visual representations of mutations of amino acid sequences, pertaining to different organisms, to show their mutation effects.

It is possible to use multiple sound characteristics (apart from using only simple pitch variations, as we saw it in previous research on DNA sonification) to represent and analyse molecular properties. PROMUSE (Hansen et al., 1999) is an audiovisual computer program for perception and analysis of protein structural alignments (the superposition of two or more molecules in 3D). Protein molecules were visualised using the Rasmol molecular modeller, and jazz quartets (composed of bass line, drums, rhythmic accompaniment, and a lead instrument for the melody) were used to map molecular parameters of secondary structure, polarity, exposure, and goodness of fit. The jazz arrangements used different musical instruments to help data discrimination. Hansen et al. evaluated PROMUSE by conducting an experiment for discriminating the alignment parameters with three conditions: Audio, visual and audio/visual. The researchers found in their experimental results that auditory data presentation scored higher than audiovisual or visual presentations. In addition, these researchers pointed out that the auditory mapping of PROMUSE was an effective way for data disambiguation of cluttered molecular visualisations. As well as musical timbres, Hansen et al. used audio panning effects to play the jazz sounds components at different stereo positions to ease data discrimination. It could be possible to enhance data discrimination by presenting the sounds in 3D played according to the spatial positions of the molecules being analysed. This can give a better idea about the molecular alignments in 3D.

Kaper et al. (1999) developed a computer system called Digital Instrument for Additive Sound Synthesis, or DIASS, along with the M4CAVE program, which plays spatial sound in virtual environments. Both programs were used to explore and analyse complex and multivariate scientific data of molecular bonding. Kaper et al. conducted a preliminary experiment on scientific sonification in a virtual environment using DIASS and M4CAVE. A computational chemist analysed the binding of a carbon atom to a protonated thiophene molecule, before and after binding happened, by listening to sonification of different energy levels arranged in space in a mesh of static points that formed a $128 \times 128 \times 128$ cubic lattice. The researchers assigned a sound to each point, which mapped sound frequency (pitch) to horizontal coordinates, amplitude (loudness) to vertical coordinates, and arbitrarily assigned time to the spatial coordinates. Thus, loudness was an indicator of the energy distribution before and after the binding. Although the authors mentioned no conclusive results in their paper, they reported that the sonification was helpful to identify subtle variations of data about the energy levels of the binding. In addition, the experiment served to test the technical capabilities of DIASS and M4CAVE. The researchers agreed that sonification can bring out periodic data over time, and it offers valuable redundant feedback that ease data exploration. They pointed out in the paper that scientific sonification has not been widely used because it seems that vision is more dominant than sound (Kramer, 1994), and because of a lack of proper scientific sonification tools (either hardware or software based).

4. The use of non-speech sounds and multimodal interfaces in chemical education

Non-speech sounds have been used in chemical education to help visually impaired students manipulate measurement instruments and perceive their output data in conventional chemistry laboratory classes. Lunney and Morrison (1981) devised and implemented a method using electronic transducers and converters to produce musical sounds with pitch variations that represented non-linear and one-dimensional output data of pH electrodes, visible and infrared spectrophotometers, and a piston burette. They used a voltage-to-frequency converter connected to a pH multimeter to play tones representing pH peaks. The researchers also used electronic text-to-speech translators to tell readouts of exact values of some laboratory instruments. The objective of Lunney and Morrison's method was to give non-sighted students the opportunity to operate laboratory instruments, thus, they learn by doing and enhance their self-confidence and independence, which otherwise the students had to rely on a sighted assistant to do most of the work in chemistry experiments. The researchers also explain in detail how they used sounds in previous implementations in Lunney and Morrison (1990), where also pointed out that using sounds that closely resemble human speech, for example, some vowel sounds, helped visually impaired students to identify patterns in chemical data. In addition, using pitch changes for data discrimination of chemical information, such as extrema in a data array, could yield an accuracy of 0.1%. Lunney and Morrison's instrumental adaptations are a compelling example about non-speech sounds uses in education. Lunney continued improving and adapting speech and non-speech sound uses to help visually impaired students operate and perceive laboratory data by using an enhanced acquisition system to obtain data from potentiometer titrations, UV measurements, infrared spectroscopy, and gas and liquid chromatography. Lunney (1994) explains that one-dimensional chemical data can be converted through a computer soundcard into musical sounds with rising and falling pitches. In addition, software-based text-to-speech conversions tell through the soundcard exact data values. Their technical implementations are depicted in Lunney (1994).

A laboratory instrument similar to the ones developed by Lunney and Morrison is a simple electronic device that measured electrochemical conductivity of substances, designed and tested by Berenato and Maynard (1997). This device emitted audio tones that changed pitch according to electrochemical resistance of the substance being measured. The mapping consisted of the inverse relationship between resistance and tone frequency. That is, the higher the resistance, the lower the tone's pitch. The researchers found in their experiments done with students (they did not mention whether the students were sighted or not) that this device was an 'attention getter' and students had fun using it. This provides another example of motivating non-speech sounds, although it shows that novelty effect was present. Berenato and Maynard also argued that by using tone pitch changes, their device could serve to explain a chemical phenomenon, for example, covalent bonding. Both Lunney and Morrison's and Berenato and Maynard's developments depict the potential use of the auditory characteristic of pitch to convey chemical information, which otherwise could be difficult to perceive if that information is presented through another sensory channel.

Another application of non-speech sounds to perceive and learn chemical information is described by Miner and Della Villa (1997), who carried out a three-day class activity where their students created musical compositions that mapped notes onto DNA base sequences. The students' main activity was to reverse transcribe 10 given amino acid sequences into DNA base sequences, and assigning each base to a musical note. Miner and Della Villa (1997) applied cooperative learning strategies in their project. Students were arranged into groups, and those who had musical and computer skills took leadership roles in each group, which helped other students in their group with the computerised music edition. The students used mappings of pitch to bases that are similar to the base mappings done by King and Angus (1996), where the musical note A represented Adenine, note G represented Guanine, note C represented Cytosine, and the exception was Thymine, who was represented by note E. Thus, the mapped sounds produced a musical alphabet that was easy to remember. After the translations and sound mappings were done, the researchers asked students with musical skills to change the sounds' rhythms and octaves to make them more melodic. After doing that, the students used music edition software to create and print musical scores of their DNA songs. After the 3-day class activities finished, each group of students gave a musical performance by playing their DNA compositions through musical instruments or a synthesizer. The students used instruments like a harp, full orchestra, and a saxophone to play the DNA songs. Miner and Della Villa (1997) observed mixed students' reactions to DNA mappings. Some of the students found the sounds repetitive. Others said the sounds sounded melodic, and some students liked them very much. A student commented about the DNA songs: "it may not be a tune you can hum or remember easily, but when you realize where it came from—Wow! It's the music of life!". Indeed, auditory display applications in science courses can cause impact in students. During and after the activities, students were assessed on three issues: Ability to reverse-translate amino acid to DNA sequences, research on the role of protein function in cells, and DNA music presentation. Miner and Della Villa (1997) found that their students were successful with DNA reverse translations; they worked methodically, consistently and with positive interdependence. In their paper the researchers explained an interesting example of students' constructive activities about designing and playing their own non-speech sounds to represent and understand molecular information.

In her doctoral research, Chris Byrne (Byrne, 1996) investigated the effectiveness of immersion and interaction in a fully immersive virtual environment for learning concepts of atomic structure. Byrne carried out an experiment with high-school students where performed tasks in four treatments: Non-immersive and non-interactive (video), non-immersive and interactive (a 2D program running on a Macintosh computer), and a highly immersive—highly interactive, where students watched a fully immersive virtual environment through a head-mounted display (HMD), and interacted with the virtual environment with a 3D mouse. Although Byrne found no significant differences in learning with the virtual environment treatment, she found that interactivity was more important than immersion. Byrne used auditory feedback in the form of auditory icons in the virtual environment. When participants put virtual atoms in a correct atomic orbital, they heard a laughing sound that indicated successful task completion. If participants made an incorrect action, like spinning an electron in a wrong way, a belching sound was heard. Even though Byrne made no analyses of participants' reactions or performance about the sounds heard, she considered that adding auditory feedback could help participants in the task of atom building in the virtual environment.

Sciencespace (Salzman et al., 1996) is an ongoing multidisciplinary research project that evaluates fully immersive and multimodal virtual environments for learning physicochemical properties of matter. In it, researchers assess aspects like learning by doing, remediation of misconceptions, motivation, engagement, immersion, multisensory cues, and multiple frames of reference. The project is divided into three modules, each containing a different virtual world: Laws of motion (Newtonworld), electrostatic fields (Maxwellworld), and molecular bonding (Paulingworld), and are intended for high school students. In the latter module, explained in Loftin et al. (1998) and Su and Loftin (2001), a student visualises a virtual environment with small and large virtual molecules through a head-mounted display (HMD), and steers the molecules using a 3Ball, an input device that allows molecule translations and rotations in 3D. The 3Ball can also serve to select options from the interface menu. A central computer tracks HMD and 3Ball movements in 3D. The students' tasks are to bind two molecules according to certain docking parameters, such as molecules' specific 3D position and orientation. Another task is to observe and analyse different molecular representations, such as van der Waals' spheres and ball-and-stick, which are selectable from a menu. Paulingworld uses simple fixed tones as feedback in menu selection and molecule placements. Paulingworld has recently been modified to be shared by two or more students through the Internet, to collaborate in analysing the molecules for its understanding. The Paulingworld researchers conducted usability tests, depicted in Su and Loftin (2001), where they found that users felt comfortable and in control of the 3D environment and input devices. In addition, Paulingworld seems usable for collaborative applications.

The work of Garcia-Ruiz (2001, 2002) describes the use of earcons to represent data of amino acids in a virtual environment, otherwise difficult to perceive if they were visualised alone. This research showed that non-speech sounds are useful for conveying molecular properties in combination of visualisation of molecular structure. In addition, the same authors also propose in Garcia-Ruiz et al. (2004) the use of earcons to identify each of the four bases that compose a DNA molecule, all this presented in a virtual environment.

5. Discussion

Although most of the research reviewed in this paper does not mention that they used auditory display applications, they implicitly used sonification, auditory icon, and earcon techniques to represent data and interface events as auditory feedback. Perhaps, some of the researchers who did the studies were not aware of the field of auditory display.

Having read all the positive results from the reviewed projects and applications of this paper, the question arises, why are auditory display techniques not widely and extensively used today? We believe there are some significant reasons, including:

- Studies, guidelines and applications related to the field of auditory display have only been objects of study over the last 10–15 years. Consequently, educational applications of auditory display are even more recent, requiring more formal study. Although auditory display has been applied for learning mathematics and physics, among other subject areas, it is hoped that future applications of earcons, auditory icons and sonification in areas other than chemistry will be developed in the near future, based on the research projects described here and elsewhere.
- Auditory display has not been formally integrated into the curricula and taught within the curriculum at most undergraduate and graduate computer science programs around the world. A search on the Internet shows that there are few cases where the topic of auditory display has been integrated into human–computer interaction, information systems and related courses. This has greatly affected the study, development and general awareness of auditory representation of information, which perhaps explains why some of the research projects and applications described in this paper do not discuss the use of auditory display, especially the early ones. Currently, there is sufficient body of knowledge to adequately support the integration of auditory display into computer science courses.
- Some auditory display techniques are difficult to implement, which is especially true in the case of auditory icons. As they are everyday sounds used at the computer interface, it is not easy to find a suitable sound effect to map to an action or object at the interface (Gaver, 1986). When developing an auditory interface, one option is to let many people choose this mapping. Moreover, careful auditory icons design is needed to develop useful and pleasant auditory interfaces.
- To our knowledge, there are no standard (or at least commonly used) programming libraries for developing auditory icons, earcons and sonification. Furthermore, there are no standard libraries of sampled or synthesized sounds specifically made to develop auditory icons, earcons or sounds to be used in sonification.
- Some people argue that western culture is visually dominant, and this has slowed down the widespread use of auditory display (Forrester, 2000). However, although

Table 1
Putative benefits and limitations of modalities

Form of modality	Putative benefits	Limitations
Physical model set	Able to manipulate relations	Not good with large data sets
	Suitable for visually impaired students	Offers limited feedback about molecular/atomic properties
	Easy to manipulate	Offers no feedback cues on dynamic molecular properties/events
	Strong direct manipulation	
	Tangible thus concrete structural model	
Visualisation alone	Colour and graphic cues help discriminate information	Partially blind and blind colour students may experience problems
	Visual information is persistent	Visual occlusion of objects and processes can make tasks difficult to do and perceive
	Useful for finding information patterns	
Visual + sound	Provides synesthetic experience	Mismatch of modalities can cause cognitive dissonance
	Sound working as redundant information to visual enhances feedback, engagement and learning	Sounds can be annoying if badly designed
	Some tasks are easier to perform than visualisation alone	
	Useful for finding information patterns	
	Complementary visual/auditory information allows efficient sensory correlations	
	Auditory and visual modalities useful in virtual environments	
	Intermodal correlations	
	Sounds can drive attention on particular visualisation tasks	

> the learning style of western cultures may be predominantly visual, this does not exclude developing or using other learning styles or incorporating other senses in learning.

Table 1 summarises what we consider some of the putative benefits and limitations about each modality and their combinations, and their comparison with a physical model set, which is generally used in chemistry classes and research. We can see from research explained in this paper and in Schomaker et al. (1995) that we can achieve the best of each modality by combining them in a multimodal interface.

6. Conclusion

In this paper we have reviewed research about applications of auditory display techniques in human–computer interfaces, particularly about non-speech sounds and multimodal interfaces in chemistry, for research and educational applications, making emphasizing multimodal virtual environments for learning abstract chemical concepts.

As we can see, promising research has been carried out respecting auditory display applications in multimodal interfaces for chemical education. The research reviewed here indicates that auditory display applications in multimodal interfaces for chemical education have potential for future research. However, auditory display needs to be more widely disseminated and promoted in related venues, conferences and academic events, such as human–computer interaction and others, by showing its contribution and its potential. In addition, there are very few auditory display-related forums like those organised by the International Community for Auditory Display (ICAD) and the International Computer Music Association (ICMA). Auditory display is a relatively young branch of computer science that needs greater support from younger computer scientists that receive instruction about this important topic at schools.

References

Alcock, N.W., 1990. Bonding and Structure. Ellis Horwood, New York.

Aldrich, F., Rogers, Y., Scaife, M., 1998. Getting to grips with 'interactivity': helping teachers assess the educational value of CD-ROMs. British Journal of Educational Technology 29 (4), 321–332.

Baggi, D.L., 1991. Computer-generated music. IEEE Computer 24 (7), 6–9.

Berenato, G., Maynard, D.F., 1997. Audio conductivity device. Journal of Chemical Education 74 (4), 415.

Birk, J.P., Kurtz, M.J., 1999. Effect of experience on retention and elimination of misconceptions about molecular structure and bonding. Journal of Chemical Education 76 (1), 128.

Blattner, M.M., Sumikawa, D.A., Greenberg, R.M., 1989. Earcons and icons: their structure and common design principles. Human Computer Interaction 4 (1), 11–44.

Blattner, M.M., Greenberg, R.M., Kamegai, M., 1992. Listening to turbulence: an example of scientific audialization. In: Multimedia Interface Design. ACM Press/Addison-Wesley, New York/Reading, MA, pp. 87–102.

Bly, S.A., 1982a. Presenting information in sound. In: Proceedings of the First Major Conference on Human factors in Computers Systems. National Bureau of Standards, Gaithersburg, MD.

Bly, S.A., 1982b. Sound and Computer Information Presentation. PhD thesis, University of California, Davis, also Available as Technical Report UCRL-53282, Lawrence Livermore National Laboratory.

Bonebright, T.L., Nees, M.A., Connerley, T.T., McCain, G.R., 2001. Testing the effectiveness of sonified graphs for education: a programmatic research project. In: Hiipakka, J., Zacharov, N., Takala, T. (Eds.), Proceedings of the 2001 International Conference on Auditory Display. Helsinki University of Technology, Espoo, Finland, pp. 62–66.

Brewster, S.A., 1996. A sonically enhanced interface toolkit. In: Proceedings of ICAD'96, Xerox PARC. Addison-Wesley, pp. 47–50.

Brewster, S.A., 1998a. Using earcons to improve the usability of a graphics package. In: Proceedings of BCS HCI'98. Springer, London, UK, pp. 287–302.

Brewster, S.A., 1998b. Using earcons to improve the usability of tool palettes. In: Summary Proceedings of ACM CHI'98. ACM Press, Los Angeles, CA, pp. 297–298.

Brewster, S.A., Pengelly, H. 1998. Visual impairment, virtual reality and visualisation. In: Proceedings of the First International Workshop on Usability Evaluation for Virtual Environments. British Computer Society, Leicester, UK, pp. 24–28.

Brewster, S.A., Wright, P.C., Edwards, A.D.N., 1994. The design and evaluation of an auditory-enhanced scrollbar. In: Adelson, B., Dumais, S., Olson, J. (Eds.), Proceedings of CHI'94. ACM Press/Addison-Wesley, Boston, MA, pp. 173–179.

Brewster, S.A., Wright, P.C., Edwards, A.D.N., 1995. Experimentally derived guidelines for the creation of earcons. In: Adjunct Proceedings of HCI'95. Springer, Huddersfield, UK.

Brooks, F.P., Ouh-Young, M., Battert, J.J., Kilpatrick, P.J., 1990. Project GROPE-Haptic displays for scientific visualization. Computer Graphics 24 (4), 177–185.

Burgess, D.A., 1992. Techniques for low cost spatial audio. In: Proceedings of the Fifth Annual ACM Symposium on User Interface Software and Technology. ACM Press, Monterey, CA, pp. 53–59.

Buxton, W., 1989. Introduction to this special issue on nonspeech audio. Human–Computer Interaction 1 (4), 1–9.

Byrne, C.M., 1996. Water on tap: the use of virtual reality as an educational tool. PhD Thesis, Department of Industrial Engineering, University of Washington.

Carlson, C., Comer, R., Fontes, R., et al., 1994. The human genome: science and the social consequences. In: DOE Human Genome Program Contractor-Grantee Workshop IV, US Department of Energy.

Chandler, P., Sweller, J., 1991. Cognitive load theory and the format of instruction. Cognition and Instruction 8, 293–332.

Crossan, A., Brewster, S., Reid, S., Mellor, D., 2000. Multimodal feedback cues to aid veterinary training simulations. In: Proceedings of the First Workshop on Haptic Human–Computer Interaction, pp. 45–49.

Forrester, M.A., 2000. Auditory perception and sound as event: theorising sound imagery in psychology, electronic publication. Sound Journal, www.kent.ac.uk/sdfva/sound-journal/forrester001.html

Garcia-Ruiz, M.A., 2001. Using non-speech sounds to convey molecular properties in a virtual environment. In: Proceeedings of the International Conference of New Technologies in Science Education (CINTEC). University of Aveiro, Aveiro, Portugal.

Garcia-Ruiz, M.A., 2002. Binding virtual molecules sounds good!: exploring a novel way to convey molecular bonding to students. In: Proceedings of E-Learn 2002. Association for the Advancement of Computing in Education, Montreal, Canada.

Garcia-Ruiz, M.A., Bustos-Mendoza, C., de la O, L.G., Andrade-Arechiga, M., Santos-Virgen, M., Acosta-Diaz, R., 2004. Exploring multimodal virtual environments for learning biochemistry concepts. In: World Conference on Educational Multimedia, Hypermedia and Telecommunications. AACE, Lugano, Switzerland, pp. 2143–2147.

Gaver, W.W., 1986. Auditory icons: using sound in computer interfaces. Human–computer Interaction 2, 167–177.

Gaver, W.W., 1989. The SonicFinder: an interface that uses auditory icons. Human–computer Interaction 4 (1), 67–94.

Gaver, W.W., Smith, R.B., O'Shea, T., 1991. Effective sound in complex systems: the ARKola simulation. In: Robertson, S., Olson, G., Olson, J. (Eds.), Human Factors in Computing Systems CHI'91 Conference Proceedings. Addison-Wesley, Reading, MA, pp. 85–90.

Hansen, M.D., Charp, E., Lodha, S., Meads, D., Pang, A., 1999. PROMUSE: a system for multi-media data presentation of protein structural alignments. In: Proceedings of the Pacific Symposium on Biocomputing.

Hart, H., Craine, L.E., Hart, D.H., 1999. Organic Chemistry, A Short Course, 10th ed. Houghton Mifflin, Boston, MA.

Hayashi, K., Munakata, N., 1984. Basically musical. Nature 310, 96.

Hofstadter, D.R., 1980. Godel, Escher, Bach: an Eternal Golden Braid. Vintage Books, New York.

Kaper, H.G., Tipei, S., Wiebel, E., 1999. Data sonification and sound visualization. Computing in Science and Engineering 1 (4), 48–58.

King, R.D., Angus, C.G., 1996. PM—protein music. Computer Applications in Biosciences 12 (3), 251–252.

Kozma, R.B., 1991. Learning with media. Review of Educational Research 61 (2), 179–211.

Kramer, G., 1994. An introduction to auditory display. In: Kramer, G. (Ed.), Auditory Display. Proceedings of the International Conference of Auditory Display'92. Addison-Wesley, Reading, MA, pp. 1–77.

Loftin, R.B., Pettitt, B.M., Su, S., 1998. PaulingWorld: an immersive environment for collaborative exploration of molecular structures and interactions. In: Proceedings of the 17th Nordic Internet Conference Nordunet'98, University of Tromso, Norway.

Lunney, D., 1994. Development of a data acquisition and data analysis system for visually impaired chemistry students. Journal of Chemical Education 71, 308.

Lunney, D., Morrison, R.C., 1981. High technology laboratory aids for visually handicapped chemistry students. Journal of Chemical Education 58 (3), 228–231.

Lunney, D., Morrison, R.C., 1990. Auditory presentation of experimental data. In: Proceedings of the SPIE Extracting Meaning from Complex Data: Processing, Display, Interaction, vol. 1259, pp. 140–146.

Miner, C., Della Villa, P., 1997. DNA music. The Science Teacher 64 (5), 19–21.

Mousavi, S., Low, R., Sweller, J., 1995. Reducing cognitive load by mixing auditory and visual presentation modes. Journal of Educational Psychology 87 (2), 319–334.

Musical mutants, 2002. The Exploratorium, 3601 Lyon Street, San Francisco, California, 94123, Available at: www.exploratorium.edu/genepool/exhibits.html

Ohno, S., Ohno, M., 1986. The all pervasive principle of repetitious recurrence governs not only coding sequence construction but also human endeavor in musical composition. Immunogenetics 24, 71–78.

Petersen, Q.R., 1970. Some reflections on the use and abuse of molecular models. Journal of Chemical Education 47 (1), 24–29.

Ranck, J.P., 1997. Visualization for chemists. In: Swift, M.L., Zielinski, T.J. (Eds.), Using Computers in Chemistry and Chemical Education. American Chemical Society, Washington, DC, pp. 227–240 (Chapter 13).

Rees, A.R., Sterenberg, M.J., 1984. From cells to atoms. An Illustrated Introduction to Molecular Biology. Blackwell, Oxford, UK.

Roussos, M., Johnson, A., Moher, T., Leigh, J., Vasilakis, C., Barnes, C., 1999. Learning and building together in an immersive virtual world. Presence 8, 3.

Rzepa, H., Whitaker, B., 1995. Chemists surf to active sights. Times Higher Education Supplement 13.

Salzman, M., Dede, C., Loftin, B., 1996. Science space: virtual realities for learning complex and abstract scientific concepts. In: Proceedings of IEEE Virtual Reality Annual International Symposium. IEEE Press, New York, pp. 246–253.

Schomaker, L., Munch, S., Hartung, K., 1995. A taxonomy of multimodal interaction in the human information processing system. A Report of ESPRIT Project 8579 MIAMI, Technical Report, Nijmegen Institute of Cognition and Information, The Netherlands.

Shneiderman, B., 1998. Designing the user interface. Strategies for Effective Human–Computer Interaction, third ed., Addison-Wesley, Reading, MA.

Su, S., Loftin, R.B., 2001. A shared virtual environment for exploring and designing molecules. Communications of ACM 44 (12), 57–58.

Upson, R., 2001. Sonifications as mathematics teaching tools. In: Hiipakka, J., Zacharov, N., Takala, T. (Eds.), Proceedings of the 2001 International Community of Auditory Display. Helsinki University of Technology, Espoo, Finland, pp. 217–221.

Yeung, E.S., 1980. Pattern recognition by audio representation of multivariate analytical data. Analytical Chemistry 52, 1120–1123.