



UNIVERSITY
of
GLASGOW

Brown, L.M. and Brewster, S.A. and Ramloll, S.A. and Burton, R. and Riedel, B. (2003) Design guidelines for audio presentation of graphs and tables. In, Brazil, E. and Shinn-Cunningham, B., Eds. *9th International Conference on Auditory Display (ICAD)*, 6-9 July 2003, pages pp. 284-287, Boston, Massachusetts.

<http://eprints.gla.ac.uk/3196/>

DESIGN GUIDELINES FOR AUDIO PRESENTATION OF GRAPHS AND TABLES

*Lorna M. Brown, Stephen A. Brewster and
Ramesh Ramloll*

Glasgow Interactive Systems Group
Department of Computing Science
University of Glasgow G12 8RZ, UK
{lorna, stephen}@dcs.gla.ac.uk
www.multivis.org

Mike Burton and Beate Riedel

Department of Psychology
University of Glasgow
G12 8QQ, UK
{mike, beate}@psy.gla.ac.uk

ABSTRACT

Audio can be used to make visualisations accessible to blind and visually impaired people. The MultiVis Project has carried out research into suitable methods for presenting graphs and tables to blind people through the use of both speech and non-speech audio. This paper presents guidelines extracted from this research. These guidelines will enable designers to implement visualisation systems for blind and visually impaired users, and will provide a framework for researchers wishing to investigate the audio presentation of more complex visualisations.

1. INTRODUCTION

A major deprivation caused by blindness or visual impairment is the lack of access to information in the form of visualisations (e.g. graphs, tables, 3D plots). The methods currently available are somewhat limited. Tactile graphics, such as raised paper diagrams, are cumbersome to produce, and once rendered cannot be changed. In addition, they are limited in the amount of information they can contain without becoming cluttered [1]. Verbal descriptions of graphs (e.g. a description of a waveform in a physics talking book) can be confusing and difficult to understand. Speech readers can read out individual values from tables or graphs, but this can overload the short-term memory, and does not provide a means of gaining an overview of data.

The aim of the MultiVis Project is to find successful methods of presenting graphs and visualisations to blind and visually impaired people using the senses available to them, i.e. hearing (audio) and touch (haptics). This paper provides a summary of the audio design guidelines derived from the research carried out on this project. The paper contains two main sections, the first discussing the presentation of line graphs, and the second describing the presentation of tables. Each section is split into subsections describing the individual guidelines.

2. LINE GRAPHS

Line graph sonification has been the focus of the majority of the MultiVis work on audio presentation [2-5]. This section presents general guidelines extracted from our research into the sonification of line graphs, in addition to guidelines specific to the sonification of graphs containing two or three data series.

2.1. General Line Graph Sonification Guidelines

The first set of guidelines are general guidelines for the sonification of line graphs, and are applicable regardless of the number of data series present in the graph.

2.1.1. Map the y-axis of graphs to the pitch of musical notes

Research has shown that the pitch of musical notes can be used to represent data by mapping values to corresponding musical notes. Mansur [6] devised a method for line graph sonification, called *Sound-Graphs*, in which the y-axis of a graph is mapped to pitch, and the x-axis is mapped to time such that moving along the x-axis causes a musical note, representing the corresponding y-value, to be played. The higher the y-value, the higher the pitch (Figure 1). This sonification method is used by MultiVis for line graphs and tables.

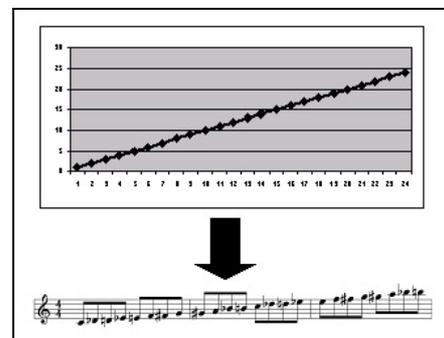


Figure 1: *As the y-value increases the pitch of the musical note gets higher*

2.1.2. Use musical sounds rather than pure sine waves

Musical sounds produced by a MIDI (Musical Instrument Digital interface) synthesiser or PC soundcard have a more complex spectrum than pure sine waves, making them easier to perceive, and more pleasant to listen to, than sine waves [7].

2.1.3. Only use MIDI notes within the range 35-100

Sonification systems can convert data to sound by mapping data values to MIDI notes. The range of available MIDI notes is from 0-127. However, on a standard PC soundcard only notes

in the range 35-100 should be used. Some informal testing found that notes outside this range are difficult to perceive and differentiate on a SoundBlaster Live soundcard, and are inaudible on some other soundcards and synthesizers. This range is appropriate when using the General MIDI Acoustic Grand Piano instrument (MIDI Instrument 0) and the General MIDI Trumpet (MIDI Instrument 56) [5]. Many other instruments have a more limited range. In order to provide as much detail as possible in a sonified graph it is beneficial to be able to represent many discrete points. Therefore, instruments with a wide pitch range should be used.

The MIDI note number can be calculated from a given value by using Equation 1, where $note_{min}$ is the minimum MIDI note (in this case 35), $note_{max}$ is the maximum midi note (in this case 100), val_{min} is the minimum data value in the graph or table being sonified, val_{max} is the maximum data value in the graph or table being sonified, and value is the data value which is to be converted into a MIDI note. The result of this equation is rounded to the nearest integer.

$$MIDINote = note_{min} + \frac{(value - val_{min})}{(val_{max} - val_{min})} * (note_{max} - note_{min}) \quad (1)$$

2.1.4. Present graphs at a speed that does not impair comprehensibility.

The graph sonification software developed on the MultiVis project allows users to explore graphs at their own pace. The numeric keypad is used to navigate around the graph, one data point at a time, and enables users to control presentation speed and change direction at will.

In addition, the software allows users to hear an overview of a graph [7]. The overview is accessed by holding down the key that is used to move through the graph. In the overview mode, if a graph is played too fast it will be difficult for the user to interpret. Conversely, if it is played too slowly, the user may have difficulty remembering the shape of the graph. Pilot testing showed that a delay of 50-70 milliseconds between successive tones allows graphs to be presented at a speed that does not impair comprehensibility [3].

2.2. Sonifying Line Graphs Containing Two Data Series

Sighted people are able to access graphs containing many data series, and identify the relationships between them. It was therefore important to see if it was possible to sonify multiple data series so that blind and visually impaired people could also access these graphs.

Two experiments were carried out with blind users, in which participants were asked to explore graphs containing two data series in order to find intersection points, global maxima and global minima. Two modes of exploring the graphs were compared—*parallel* mode and *serial* mode [4]. In the parallel mode users listen to both data series simultaneously (with one data series panned to each ear through headphones), such that at any point on the x-axis they will hear the y-value of one data series in their left ear and the y-value of the other data series in their right ear. In the serial mode, users hear one data series at a time. Thus, at a given x-coordinate, they will hear a single note representing the y-value in one ear and can switch to the other ear to hear the corresponding y-value for the second series, allowing comparisons to be made.

In the first of these experiments, blind participants completed a set of tasks in each mode, and the accuracy and

speed with which the tasks were completed were compared [4]. In the second experiment participants were able to switch freely between modes during a task, in order that any preference for one mode for a particular task type could be identified.

A pilot study was also carried out with sighted users to discover whether, in addition to finding specific points, users could build up a mental model of the graphs they were listening to when they were presented using the parallel mode. This was tested by asking the users to draw the graphs after listening to them, and calculating the percentage of key features of each graph which were drawn correctly. This experiment also compared the results of using the same instrument (piano) to represent both series and using different instruments (piano and trumpet) to represent each series. The guidelines which follow are the result of these three experiments.

2.2.1. Use Stereo Panning to separate data series

Stereo panning can be used to help users to distinguish between two data series. In order to sonify line graphs containing two data series, each data series is sonified as described above. One data series is panned hard left (MIDI panning value=0) and the other is panned hard right (MIDI panning value=127), thus the two data streams are perceptually separated [4]. Stereo panning is used to separate simultaneously playing earcons [8] (short structured audio messages used to communicate information in computer interfaces), and the same technique is effective for separating simultaneously playing data series [5]. Deutsch [9] and Bregman [10] have suggested that separating simultaneously playing streams by spatial location alone may not be sufficient to allow users to differentiate between them. Our results showed, however, that users were able to draw each line, with over 80% accuracy, after listening to the graph [5].

2.2.2. Use the Parallel Mode rather than Serial Mode

The parallel mode enables users to complete intersection location tasks significantly faster than the serial mode. In addition, users demonstrated a preference for the parallel mode for these tasks, choosing to spend a significantly greater percentage of time in this mode than in the serial mode when looking for intersections [4]. No significant differences in performance were found between the parallel and serial mode when users were asked to locate global maxima or minima, or in the workload [11] experienced during any tasks. Since users have expressed a preference for the parallel mode graph sonification systems should default to this mode, but should allow users the option to switch to serial mode.

2.2.3. Allow users to choose between instrument combinations

The pilot study [5] found that users could draw graphs with 80% accuracy on average after listening to them. No significant differences were found between the two presentation modes – i.e. when the same instrument (piano) was used to represent both lines, or when different instruments (piano and trumpet) were used to represent each line. Due to the small number of subjects, this is not surprising. However, there were indications that using the same instrument for each line might be better for identifying features of the interactions between the two lines (e.g. intersection points, relative y-positions of maxima and minima). However, the majority of participants expressed a preference for the use of two different instruments and the

results suggested that this might be more successful for differentiating between lines and recognising features of the individual lines (e.g. maxima, minima, curve-shape). Since both instrument combinations are successful, and each has its own benefits, it is recommended that graph sonification systems feature both modes, and allow users to choose between them based on their preference and current task [5].

2.3. Sonifying Graphs Containing Three Data Series

An experiment was carried out with sighted users to discover whether it is possible to interpret sonified graphs when they contain three data series [3]. It also aimed to discover whether the parallel mode or serial mode was more effective for carrying out intersection location, and shape recognition tasks.

In order to sonify a graph containing three data series, a spatialised audio environment can be created. Each sonified series is placed at one corner of an equilateral triangle with the user in the center (Figure 2) and played to the user through headphones. In the parallel mode all three series play simultaneously, and the user can move around the triangle in order to focus on a single line, two lines at once, or all three. In the serial mode users can only hear one series at a time but can switch between them, using the keypad, to make comparisons. The following guidelines were derived from this study.

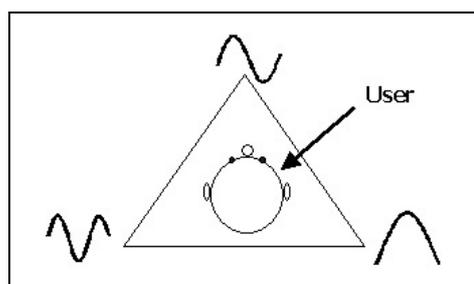


Figure 2: One series is positioned at each corner of an equilateral triangle, with the user in the centre

2.3.1. Provide opportunities for smoothly changing the salience of data series

Users should be allowed to change their spatial relationship with the 3D sources, by moving around within the triangle. This allows them to focus on the lines in which they are interested. In addition, allowing the user to move closer or further from sources can resolve ambiguities between whether sources are behind or in front, which are common in 3D audio presentation.

2.3.2. Allow users to use both Parallel and Serial Modes

The results showed that people were able to locate intersection points significantly more accurately in the parallel mode, and that the workload experienced when carrying out these tasks was significantly lower in the parallel mode. In addition, there were no significant differences in participants' abilities to match the sonified graphs to their visual counterparts in either condition, showing that the parallel mode does not interfere with participants' abilities to recognise and recall the shape of the individual graphs.

Since the parallel mode is more successful, graph sonification systems should default to this mode. However, if a user wishes to focus on a single data set for a specific task it is

useful for them to be able to use the serial mode, so this should be provided as an option.

3. PRESENTING TABLES USING AUDIO

3.1. Combining Speech and Non-speech Audio

An experiment was carried out with blind users to compare speech output alone with a combination of speech and non-speech audio for the presentation of tables, which contained only numeric values [7]. Speech is widely used in computers, in speech readers such as JAWS, and can be used to read out individual values in tables. While speech is excellent for presenting accurate or detailed information, as it can provide exact values and figures, it is not good for presenting overviews, due to the limits of short-term memory.

In one condition, users were only able to access speech output. In the other condition, non-speech audio was used to sonify the table but users could press a key at any time to access the speech output. The tables were sonified using the method used for line graphs. Each datum in the table was mapped to a musical note, with the value of the datum represented by the pitch of the note. In both conditions, the numeric keypad was used to navigate around the table. The guidelines extracted from this study were as follows.

3.1.1. Include non-speech audio in tables

The results of this experiment showed that the inclusion of non-speech audio lead to significantly faster, and more accurate, completion of tasks. In addition, the workload, which was measured using NASA TLX scales [11], was significantly lowered when non-speech sounds were available.

3.1.2. Use Stereo Panning to indicate location in a table

Stereo panning through headphones can be used to indicate the location of a cell within a table. The location of the sound, from left to right, can be used to indicate the row or column in which the cell is located. If a user is browsing by column, the location of the sound will indicate the column number, whereas if he/she is browsing by rows, the location of the sound will indicate the row number. For example, if the user is browsing by columns and the current cell is in Column 1, the user will hear a sound in their left ear, while if the current cell is in the middle column they will perceive the sound to be above their head.

This was not formally evaluated but pilot studies showed that the majority of participants were pleased with the panning, and none found the mapping confusing [7]. This provides a low budget method for introducing spatial cues.

3.1.3. Inform users when they reach a table boundary

Users should be informed when they reach the boundary of a table or a graph. Upon reaching the boundary a non-speech sound (such as a percussive sound), which is distinct from the sounds heard in the table, should be played to indicate this. In addition they should not be allowed to cross these boundaries. If they attempt to do so the boundary sound cue should be repeated to indicate that they are still at the boundary and cannot cross it. Pilot testing showed that allowing users to navigate out-with the table boundaries was confusing.

3.1.4. Provide access to location information at all times

It was found that users can easily become lost in a large table presented in audio so they should be able to access information about their location (row and column numbers) at any time. However, unnecessary information is distracting and overloads short-term memory so users should be allowed to choose when they want to access their location information rather than being forced to hear it at all times.

3.1.5. Interrupt speech when task is halted

If a user presses another navigation key while any speech information is being read out, the speech should be stopped immediately. Pilot studies found that it was frustrating for users to have to wait for speech to finish before moving on [7].

3.1.6. Provide table navigation shortcuts

Pilot testing showed that navigation shortcuts are extremely helpful, so users should be able to jump to the start or end points of rows or columns at any time.

3.2. Imposing Structure on Speech Output

A series of six experiments [unpublished] investigated whether imposing auditory structure on verbally presented tables of numbers could improve memory load. The aim of the experiments was to increase the short-term memory capacity for digits presented verbally. Typically, people remember only about seven items, and this holds whether the items are numbers, letters or words [12]. Chunking of independent units into meaningful larger units can be used to increase the capacity of memory [13] allowing people to store seven larger "chunks", rather than single items, of information. This property of memory has often been exploited when attempting to improve people's short-term memory. The hypothesis of these experiments was that improvements in memory span can be made if verbally presented material from a table of numbers is introduced in a way that encourages chunking of the information into larger than single-digit units.

In the experiments participants were read numbers forming a matrix, and then immediately asked to recall items, in their correct position. The size of the matrices ranged from 2x2 to 4x4. The changes introduced in order to improve memory performance were as follows: changing the pitch of the voice reading the values for each row or column; using a different voice (i.e., a different speaker) for each row; changing the pitch dependent on the value of the data in a cell. No advantage was found for any of these manipulations over a simple unchanging pitch condition. The following guidelines were extracted.

3.2.1. Do not use different pitches for speech output

Changing the pitch of speech output voice in order to chunk verbally presented items does not improve memory performance. Equally, memory performance is not enhanced by using different pitches corresponding to different value ranges to impose structure on single items of information.

3.2.2. Use one voice for speech output

Memory performance is not improved when different speech output voices are used to organize verbal information into

larger units auditorily. One voice should be used to present tabular data verbally. This should not be manipulated by imposing auditory structure on the information with a view to improving memory performance.

4. CONCLUSIONS

This paper has presented a set of guidelines for the audio presentation of graphs and tables. These guidelines could be used by researchers as a basis for the sonification of more complex visualisations, or by designers of audio presentation systems for blind or visually impaired users.

5. REFERENCES

- [1] B. Challis and A. D. N. Edwards, "Design Principles for Tactile Interaction", in *Haptic Human-Computer Interaction (LNCS 2058)*, S. A. Brewster and R. Murray-Smith, Eds. Berlin, Germany: Springer, 2001.
- [2] R. Ramloll, W. Yu, S. A. Brewster, B. Riedel, A. M. Burton, and G. Dimigen, "Constructing Sonified Haptic Line Graphs for the Blind Student: First Steps", In *Proceedings of ACM Assets 2000*, Arlington, VA, pp. 17-25, 2000.
- [3] R. Ramloll and S. A. Brewster, "An Environment for Studying the Impact of Spatialising Sonified Graphs on Data Comprehension", In *Proceedings of Information Visualisation 2002*, London, UK, pp. 167-174, 2002.
- [4] L. Brown, S. Brewster, R. Ramloll, B. Riedel, and W. Yu, "Browsing Modes for Exploring Sonified Line Graphs", In *Vol. II Proceedings of HCI 2002*, pp. 6-9, 2002.
- [5] L. M. Brown and S. A. Brewster, "Drawing By Ear: Interpreting Sonified Line Graphs", In *Proceedings of ICAD 2003*, University of Boston, MA, 2003.
- [6] D. L. Mansur, *Graphs In Sound: A Numerical Data Analysis Method for the Blind*, M.Sc. Thesis, Department of Computing Science, University of California, 1975
- [7] R. Ramloll, S. A. Brewster, W. Yu, and B. Riedel, "Using Non-speech Sounds to Improve Access to 2D Tabular Numerical Information for Visually Impaired Users", In *Proceedings of IHM-HCI 2001*, Lille, France, pp. 515-530, 2001.
- [8] S. A. Brewster, P. C. Wright, and A. D. N. Edwards, "Parallel earcons: Reducing the length of audio messages", *International Journal of Human-Computer Studies*, vol. 43, pp. 153-175, 1995.
- [9] D. Deutsch, "Grouping Mechanisms in Music", in *The Psychology of Music, Cognition and Perception*, D. Deutsch, Ed. London: Academic Press Inc., pp. 99-134, 1982.
- [10] A. S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*, 1st. ed: Massachusetts Institute of Technology Press, 1994.
- [11] S. Hart and L. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research", in *Human Mental Workload*, P. Hancock and N. Meshkati, Eds. North Holland B.V., Amsterdam, pp. 139-183, 1988.
- [12] G. A. Miller, "The magical number seven plus or minus two", *Psychological Review*, vol. 63, pp. 81-97, 1956.
- [13] H. Simon, "How big is a chunk?" *Science*, vol. 183, pp. 482-488, 1974.