Designing Auditory Cues to Enhance Spoken Mathematics for Visually Impaired Users

Emma Murphy
School of Computing
Dublin City University
Dublin 9, Ireland.
emma.murphy@computing.dcu.ie

Enda Bates
School of Computing
Dublin City University
Dublin 9, Ireland.
enda.bates@gmail.com

Dónal Fitzpatrick
School of Computing
Dublin City University
Dublin 9, Ireland.
dfitzpat@computing.dcu.ie

ABSTRACT

Visual mathematic notation provides a succinct and unambiguous description of the structure of mathematical formulae in a manner that is difficult to replicate through the linear channels of synthesized speech and Braille. It is proposed that the use of auditory cues can enhance accessibility to mathematical material and reduce common ambiguities encountered through spoken mathematics. However, the use of additional complex hierarchies of non-speech sounds to represent the structure and scope of equations may be cognitively demanding to process. This can detract from the users' understanding of the mathematical content. In this paper, a new system is presented, which uses a mixture of non-speech auditory cues, modified speech (spearcons) and binaural spatialization to disambiguate the structure of mathematical formulae. A design study, involving an online survey with 56 users, was undertaken to evaluate an existing set of auditory cues and to brainstorm alternative ideas and solutions from users before implementing modified designs and conducting a separate controlled evaluation. It is proposed that by involving a wide number of users in the creative design process, intuitive auditory cues will be implemented with the potential to enhance spoken mathematics for visually impaired users.

Categories and Subject Descriptors

G.4 [**Mathematical Software**]: User Interfaces; H.5.2 [**User Interfaces**]: auditory (non-speech) feedback, User-centred design, Voice I/O; K.4.2 [**Social Issues**] Handicapped persons/special needs

General Terms

Measurement, Design, Human Factors, Verification.

Keywords

Mathematics, Visually Impaired Users, Accessibility, Design Methods for User Interfaces, Non-Speech Sound, Spearcons.

ASSETS'10, October 25–27, 2010, Orlando, Florida, USA.

1. INTRODUCTION

Mathematics is a subject with considerable accessibility barriers for visually impaired students, academics and teachers. The spatial presentation of mathematical equations makes it difficult to render in a non-visual manner through the linear structures of Braille and synthetic speech. There are certain common problems that can cause ambiguity in spoken maths, whether communicated through human or synthesized speech. In algebra, for instance, fractions are a notable illustration of expressions where ambiguity can arise due to confusion of grouping. For example the spoken phrase "one over x plus 4" does not indicate whether the fraction incorporates "1 over x" as a single grouping or the entire phrase is part of the fraction.

In this paper a novel auditory math system is presented which uses a lexical structure in conjunction with a set of spatialized earcons and modified speech, to disambiguate the structure of mathematical formulae. A design survey is detailed from methods to results. The aim of the design survey is to firstly evaluate already designed auditory cues and modified speech parameters in the auditory math system and secondly to brainstorm alternative sound ideas with a wide sample of users.

2. ACCESSIBLE REPRESENTATIONS OF MATHEMATIC EQUATIONS

Karshmer et al. [13] highlight the two-dimensional nature of visual or printed mathematic equations, which is difficult to convey through both linear systems of synthesized speech and Braille. The spatial representation of equations can encode essential semantic information essential to understanding the mathematic construct. Previous studies have investigated the possibility of retaining this spatial information through non-visual interface solutions to enhance accessibility for visually impaired users. Solutions have included lexical structures [17], spatial displays (audio and haptic) [12], non-speech audio cues [22] active browsing [12, 22, 20] and prosody [10, 8].

Various studies have indicated that the use of non-speech audio can help improve access to non-visual user interfaces, by reducing the burden of information through speech alone [4, 16]. Brewster et al. [5] describe how musical parameters such as timbre, register pitch and duration can be manipulated to create earcons and convey information to the user through sound. Where earcons are sonic messages that have abstract musical qualities, auditory icons are based on sounds sampled from real-world environments. Auditory icons are based on the attributes of source events (such as size, material and force) rather than in terms of physical attributes of the sounds themselves [9].

While it is beyond the current scope of this paper to present a detailed literature review on previous approaches to creating an accessible platform to mathematics it is important to consider that there have been previous attempts to enhance spoken mathematics with non-speech sound in addition to prosody. Notably the Mathtalk system [7, 22] uses musical earcons to indicate structural delimiters (such as a pattern of 3 short ascending and descending notes to indicate the beginning and end of an expression respectively) and also provide an abstract overview of the entire equation [7, 22]. The significant drawback to this approach is that a complex and highly specific musical grammar must be learned prior to using the system. In addition, remembering and decoding musical patterns may be quite difficult for non-musicians and again, the additional cognitive effort required to decode each pattern could detract from the processing of the mathematical content. This is also symptomatic of a wider issue in the field of auditory display design. At present, most sound design methods for non-speech sounds in auditory interfaces are based on empirical knowledge. This often results in the creation of sounds derived from random selection or the personal preference of the sound designer. Such ad hoc approaches to interface design are often part of a simple implement and evaluate iterative cycle, without a proper design phase.

In contrast to previous research approaches this study aims to adopt a user-centred method in the design process in order to create sounds that are intuitive to the user without a need for complex pattern recognition, or the creation of a musical grammar. Through the user design survey presented in this paper it is intended that the operators and expression parts that will be sonified will be verified at an initial design stage. As part of this design study users can make choices between sounds, suggest alternatives and brainstorm possible metaphors or new sound ideas for mathematical concepts. Furthermore users will be consulted on which concepts are appropriate to be conveyed through non-speech sound.

3. AUDITORY MATH SYSTEM DESIGN

A novel system for spoken mathematics that utilises synthesised speech, prosody and non-speech sound is the focus of the design survey presented in this paper. The ideas and background to the techniques for the current design are described in more detail in [2]. The auditory math system utilises the Acapela speech synthesis[1] and all earcons and spearcons are spatially positioned using the IRRKLANG 3D[2] sound engine. Equations are synthesised through a mark up language where each operator is represented by a tag that triggers the relevant non-speech sound and binaural spatial position. The lexical structure that the spoken equations are based on is NemethMathSpeak [17].

3.1 Use of Binaural Cues in Spoken Mathematics

There have been previous attempts to directly map the visual spatial structure of printed mathematics to an equivalent spatialized audible structure to reduce the mental effort required by the user to process and solve the equation [12]. However, the spatial resolution of vision is much more accurate than audition and this suggests that accurately replicating the spatial layout of a printed equation with spatial audio is difficult to achieve. In fact, it has been found that the additional mental processing required to determine the spatial trajectory detracts from the processing of content [11]. The vertical layout of mathematical equations would appear to be well matched to a spatial audio presentation of an equation, however, auditory localization is particularly inaccurate in the vertical dimension and extremely difficult to synthesize, particularly with binaural spatialization techniques based on the head-related-transfer function, or HRTF [3, 6].

However spatial audio should not be dismissed as a redundant parameter to enhance spoken mathematics. Previous studies have illustrated that sounds produced from different spatial locations are easier to distinguish, which suggests that if additional sounds are added to the main speech signal, these should be produced from different spatial locations as in the system developed by Goose *et al* [11]. The externalization effect which generally occurs when binaural material is presented via headphones has also been found to be much less fatiguing than standard stereo, and this may also be of some benefit [3]. Therefore the binaural cues in this present system are intended to reinforce the lexical structure of spoken maths while at the same time relieveing some of the cognitive burden of speech processing. A left/right binaural spatial position has been implemented to convey the lexical open/close or begin/end.

3.2 Spearcons

Spearcons are created by time compressing a spoken phrase so that the resulting short sound is not entirely comprehensible as speech [23]. Spearcons are intended to be perceived somewhere between recognizable speech and abstract sounds such as earcons. It is proposed that spearcons are easier to learn and resulted in an increase in performance in interactive interface tasks such as menu browsing [23]. The major advantage of spearcons for this present system is that they can potentially function as either a descriptive lexical phrase or an abstract sound depending on its familiarity to the listener or its function as a structural delimiter.

3.3 Auditory Cues

The following provides a summary of the auditory cues implemented in the system:

3.3.1 Brackets

In this current system, brackets are indicated by a short beep-like earcon of 20ms, positioned to the left and with a rapid upward frequency glissando indicates an opening bracket, to the right and with a downward glissando indicates a closing bracket. To provide users with a design choice, an alternative bracket sound was also designed in contrast with the short beep earcon. This "noise" earcon consists of synthesised white noise which was dynamically filtered to become broadband for an opening bracket, and narrowband for closing. This sound was designed with a long attack and release to make it as unobtrusive as possible and had a duration of 400ms.

3.3.2 Fractions

Fractions are depicted using the spearcon "frac", played from the left to indicate the beginning of a fraction, and from the right to mark the end.

3.3.3 Superscripts/Subscripts

Superscripts are indicated using the spearcon "sup" and an EQ change which reduces bass frequencies and increases the treble.

---

[1] http://www.acapela-group.com

[2] http://www.ambiera.com/irrklang

This is also implemented using a left and right "sup" spearcon like the other spearcons. Subscripts are similarly indicated using the spearcon "sub" and an eq change that reduces the treble and increases the bass frequencies.

3.3.4 Nested Layers

Nested layers have been implemented by adjusting the spatial position and pitch of the beep-like bracket earcons. Nested layers are also indicated by adjusting the speaking rate by 6% as defined on a prosody model in [8].

4. DESIGN STUDY

Stevens et al. [21] identified a problem with evaluation in the development of assistive technology for visually impaired users. It is difficult to apply a controlled evaluation paradigm when there is a great deal of variability between people. This is particularly true when testing systems designed for visually impaired users. The inherent variability is large, due to factors such as length of time since the onset of blindness, education, technical skills and other disabilities. It can be difficult to evaluate research that is truly innovative as the system may not have an alternative to perform a comparison in a controlled evaluation. It is also difficult to test on sufficient numbers of visually impaired users to attain statistical reliability (ibid.). As numbers of blind people in communities are relatively small, there can be difficulties in assembling a group of participants for visually impaired users. It was considered that by creating an online survey at least the issue concerning a small sample size highlighted by Steven's et al. would be addressed. Furthermore this design study supports the ideas proposed in [19] that involving users early in the creative design process will create a more coherent and cohesive design solution.

4.1 Design Study Research Aims:

1. To capture qualitative feedback on the current auditory math system design from a wide number of users both blind and sighted through online dissemination.
2. To brainstorm alternative creative sound design ideas to enhance spoken mathematics by introducing users to design concepts in a task-driven context.
3. To investigate an appropriate way to evaluate auditory cues without encountering the barriers and ambiguities to maths equations through screen readers.

4.2 Participant details

56 users (22 Female, 34 Male) took part in the online survey, including 35 sighted users and 21 visually impaired screen reader users between the ages of 18 and 64 (AV:38; S.D.:12). Participants received the survey through university contacts in Ireland, Finland and the United States. 10 of the 21 visually impaired users had a congenital visual impairment and the remaining 11 users had lost their sighted later in life. All visually impaired users completed the survey using a screen reader.

As the survey was disseminated through personal contacts and mailing lists in the areas of mathematics, sound and accessibility, the majority of participants described their occupation as students, researchers or lecturers in the fields of Mathematics, Computer Science or Music Technology or employees in the field of disability services or accessibility consultancy. Participants were also asked how long it had been since the last exam or study they had undertaken in mathematics. 12 participants had taken a maths exam within the last year or were currently researching or teaching in the area. The average length of time for users since the last exam or research in mathematics was 10 years (S.D.: 9 years). 41 of the 56 participants reported undertaking musical training in the past or playing a musical instrument.

4.3 Structure of Sound Design questionnaire

While the sound design questionnaire is based on a mixed methods design, the main intention was not to evaluate users on quantitative performance data, but rather to evaluate firstly if aspects of the already designed cues are intuitive and also to brainstorm some creative sound design ideas from participants. Training is an obvious issue that needs to be addressed in any interface and particularly one with abstract feedback cues. For the purposes of this design study, it was considered that if users were trained that a certain cue represents a particular mathematic concept they may be less likely to suggest an alternative. Also it is difficult to devise a controlled training set up through remote testing without information on users concentration and timing. This will be implemented in a controlled evaluation setting. Alternative sound ideas suggested by users will be considered in the next phase of design of the system and will be tested in a more controlled setting to evaluate users preference and performance of sound cues.

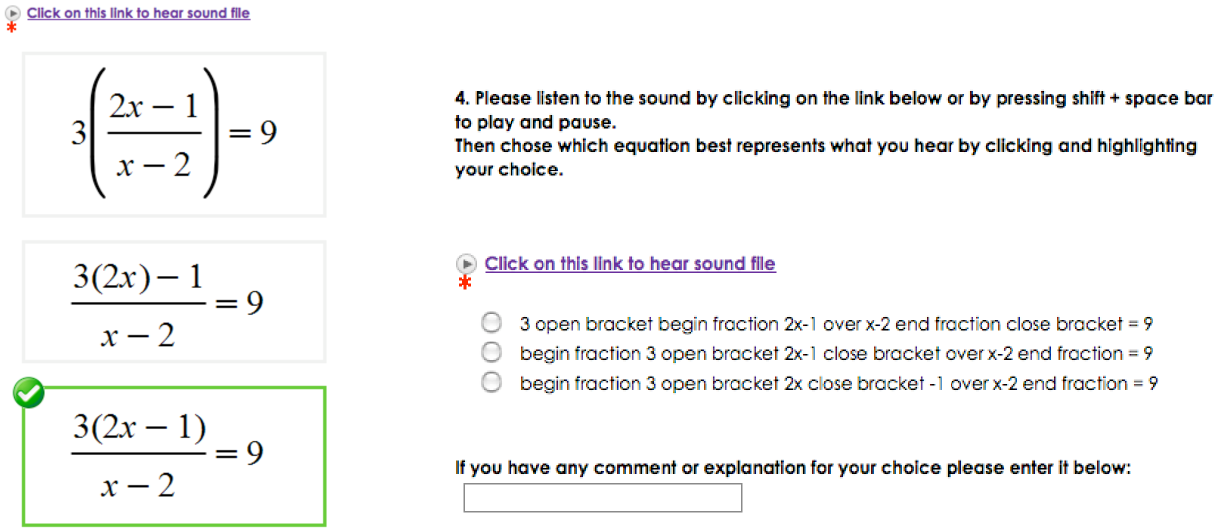
Although the system is intended for visually impaired users, sighted users were also invited to participate in the design survey in order to widen the sample group. It was considered important to have a design input from sighted users who work with visual maths equations, to try to capture their mental representations of the spatial attributes of maths equations. It is appreciated that sonifying visual representations is not always the most effective approach to auditory interface design. However it is also important to create designs that are inclusive and allow the potential for collaboration rather than creating assistive technology solutions in isolation. There are existing approaches that attempt to bridge the gaps between Braille math representations, mark-up languages and visual representations to promote collaboration [1, 14]. However it is intended that the present design will be inclusive to all users by creating auditory cues that are intuitive and therefore accessible to both sighted and visually impaired users.

4.4 AuditoryMath Design Survey

The survey was designed using surveygizmo[3] as it had been cited [18] as an accessible survey tool for visually impaired users. A number of web media-players were evaluated with a visually impaired screen reader user before choosing an audio player to embed in the survey. Many Flash based media players were inaccessible within the survey pages as it was difficult to navigate in and out of the player using keystrokes. The most accessible solution was the yahoo media player [http://mediaplayer.yahoo.com], as this player floats on the web page and can be navigated using keyboard commands without the user having to "tab" into the player area or have a virtual cursor in the vicinity.

---

[3] www.surveygizmo.com



Click on this link to hear sound file

*

$$3\left(\frac{2x-1}{x-2}\right)=9$$

$$\frac{3(2x)-1}{x-2}=9$$

$$\frac{3(2x-1)}{x-2}=9$$

4. Please listen to the sound by clicking on the link below or by pressing shift + space bar to play and pause.
Then chose which equation best represents what you hear by clicking and highlighting your choice.

Click on this link to hear sound file

*

○ 3 open bracket begin fraction 2x-1 over x-2 end fraction close bracket = 9
○ begin fraction 3 open bracket 2x-1 close bracket over x-2 end fraction = 9
○ begin fraction 3 open bracket 2x close bracket -1 over x-2 end fraction = 9

If you have any comment or explanation for your choice please enter it below:

Figure 1. Comparison of multiple-choice selection for sighted users and screen reader users (order multiple choice answers are randomised). While sighted users were presented with image choices of the equation, screen reader users were presented with an equivalent verbose lexical text.

4.4.1 Protocol

In the introductory section of the questionnaire, users were asked to listen to a speech file containing the word left which was panned left and the word right, panned right. Users were asked to make sure that their headphones were placed on the correct ear and also to enter the make and model of their headphones. This was intended to reduce the possibility of users completing the survey without headphones or through speakers, which would have made the binaural cues redundant.

Most screens in the survey followed the format of an audio file containing a whole or partial equation followed by multiple choice equations or text boxes for open-ended questions. All audio files contained 500ms of silence at beginning of the spoken equation so as not to conflict with the screen reader speaking the audio file link. Users were only able to navigate forwards through the survey.

4.4.2 Section 1: Spearcon Recognition

Currently in the system, lexical cues such as fraction, subscript and superscript are depicted with a using shortened spearcons, “frac”, “sup”, “sub”.

One of the aims of this section was to test word recognition or longer phrases at different speeds. Another purpose for this section was to introduce users to the lexical language of spoken mathematics, and some of the concepts and operators involved in the design process.

Users were presented with the following phrases:

“Begin/End Fraction, Begin/End Subscript, Begin/End Superscript.”

While these spoken phrases are longer than the phrases used in the current design, it was considered a fairer task to present users with fuller phrases when presented in isolation. The phrases were created using the Acapela speech synthesiser used in the current system design. The spoken phrases were speeded up to 50% and 70% of their original durations. Participants were asked to type in the words that they heard in a text box. All audio clips were randomised for each user within this section.

4.4.3 Section 2: Full Equations with non-speech sound

Designing this task for users to understand a spoken equation highlighted the barriers to mathematics that screen reader users currently face.

An open-ended question was considered to evaluate whether users understood the non-speech sound elements within full equation examples. However interpreting users descriptive text would result in the same ambiguities that pervade spoken mathematics. In order to overcome this, multiple-choice solutions were presented to users with the correct equation and two distracter equations.

This section was the only part of the questionnaire that was different for visually impaired and sighted users. All users were presented with an audio file containing the spoken equation with non-speech sound, sighted users were presented with a choice of visual equations and visually impaired users were presented with the equivalent verbose lexical version based on Nemeth Mathspeak [17]. An example of one of the questions in this section is illustrated in figure 1.[4] This decision to present alternative sections was not ideal in creating a fully accessible survey design. However it was considered that sighted users may have understood the design context when presented with the familiar visual modality. All multiple-choice answers were randomised for each question and all questions were randomised within the section.

---

[4] A sound example of this equation is also available at http://mcg.computing.dcu.ie/amexamples

4.4.4 Section 3: Partial Equations with Spatial Attributes

In order to assess whether users found the spatial distribution intuitive to represent open/close or begin/end users were presented with the test sound file followed by three multiple-choice solutions. For example users were presented with a sound file containing the spoken letter x followed by the beep earcon positioned left. This was intended to represent x(. Users were then asked to choose between the following options; x(, x) or (x. This section contained six questions based on the two bracket designs (beep and white noise sweeps), and the spatially distributed spearcon "frac". All multiple choice answers were randomised for each question and all questions were randomised within the section.

4.4.5 Section 4: Qualitative Sound Design

While the survey is based on a mixed methods design the emphasis on this study is on the qualitative responses from users. The quantitative questions were intended to illustrate the system to users by undertaking the task so that they could assess the design choices available and answer the open-ended questions in the final section. Every question from each section also contained an open-ended comment box for users to express points of confusion or suggestions.

This final qualitative design section was intended to create some alternative sound ideas. Users were asked to evaluate the representations of brackets, fractions, subscripts, superscripts and square roots, that they had been presented with. Also they were asked to suggest alternatives for these representations or provide further comments through open-ended questions.

4.5 Results

4.5.1 Section 1: Spearcon Recognition.

Average accuracy rates for all participants were significantly higher ($t(110)=1.98$, $p<0.001$) for the slower spearcon speed (70% of original speech file). Figure 2 illustrates average accuracy results by user group. While visually impaired users had higher average accuracy rates overall, the differences between user groups were not significant. T-test analysis of accuracy rates at the slower speed (70% of original) revealed a significant difference between visually impaired and sighted users ($t(54)=2.0$, $p=0.02$). This is not surprising as all visually impaired participants were screen reader users and therefore would be accustomed to working with fast rates of speech. It also confirms previous research in the area [15]**,** which revealed that blind users have higher levels of comprehension of synthetic speech particularly at ultra fast speeds.

In terms of individual words, at the faster speed (50% or original), the words "Begin" and "Fraction" had high recognition rates. The word end was confused with "in" or "and" and "superscript" was confused with "pseudo-script". Superscript and subscript were also confused.

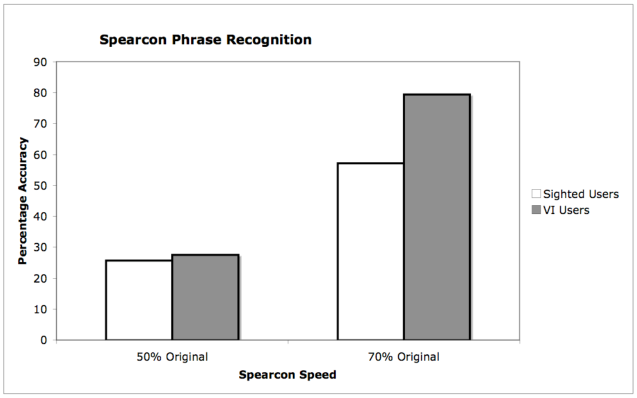


Figure 2: Graph illustrating percentage accuracy of recognition for full phrases i.e. Begin Fraction, End Superscript according to user groups.

4.5.2 Results Section 2: Full Equations with non-speech sound

The average rate of accuracy for all users for the 7 questions in this section was 67% (S.D: 25). T-test analysis revealed significantly higher average results for sighted users than visually impaired users ($t(38)=2.0$, $p<001$).This makes sense as visually impaired users had a more tedious task in having to process verbose lexical multiple-choice equations. Two visually impaired participants reported survey fatigue in completing this section. This highlights the difficulty and tedious nature of spoken mathematics perceived through a screen reader.

This task also highlighted some interesting issues in learning effects for the sounds under evaluation. A number of users commented that while they were confused at the beginning of this section on completing the first few questions, they understood that left right binaural presentation of cues indicated opening and closing of lexical operators.

There was no significant difference in users performance in terms of the presence of the lexical operator "sup". Therefore users were able to rely on the EQ changes in the voice to denote the duration of a superscript. A comment from a number of users regarding the implementation of the EQ modification is that the volume was significantly reduced. While a slight amplitude reduction for superscript text reflects the smaller print of a subscript, sounds may need to be amplified after the EQ process to make sure that they are intelligible. In the example of a nested bracket in this section, users correctly identified the pitch differences to represent the layering of brackets in the comment field.

4.5.3 Results Section 3: Partial Equations with Spatial Attributes

The average accuracy rate for partial equations was 61% (S.D.: 28). There was no significant difference between visually impaired and sighted responses in terms of accuracy. It is interesting to analyse averaged results for the 2 types of bracket sounds presented and the spatially distributed frac spearcon (illustrated in figure 3).

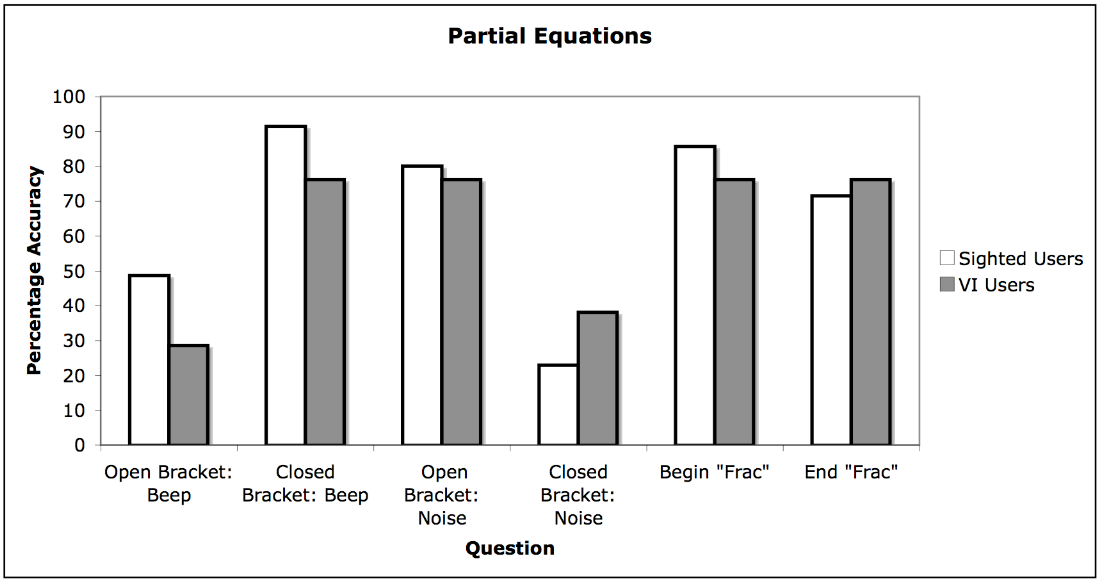


Figure 3: Graph illustrating average accuracy for visually impaired and sighted users for each auditory cue tested in section 3.

The spatially distributed (left for open and right to depict closed) spearcon "frac" received the highest results in terms of accuracy from all users. There were less errors for the beep glissando bracket sound to represent closing whereas the opposite was true for the noise sweep sound. This received higher accuracy for the "open" version.

An explanation for this distribution of results is that users found the left/right binaural cues intuitive to depict open and close or begin and end. This is illustrated by the high accuracy rates for localizing the spearcon "frac". In the non-speech versions, perhaps the glissando or noise filter sweep upwards and downwards added confusion. It is difficult to conclude from this study whether the sweep/glissando effect is redundant or perhaps that users need to be trained on this. This will be tested further in a controlled experimental set-up. It may be that the left right spatial distribution will be sufficient to depict the lexical operators without the addition of a pitch cue.

4.5.4 Results Section 4: Qualitative Sound Design

When presented with both sound design choices for brackets, 32 out of 56 users (57%) stated a preference for the beep sound with glissando. The most common reason for this preference was the fact that the beeps stood out from the texture of spoken sound. Users also described this sound as "clear" and "short".

Conversely, the 14 users that preferred the noise sweep sound explained their preference as a result of the sound having a softer attack and therefore more aesthetically pleasing to listen to. Users also reported that the "whoosh" (filter sweep) of the noise-like bracket was a stronger indicator of open or closed in addition to the binaural cue. Users also reported that this sound was easier to spatialize than the beep earcon. Although this was not supported by user performance in the results of sections 3, it needs be explored further in a controlled setting with some user training. Most users referred to this sound as a "swoosh" or a "whoosh" and two users claimed that it sounded like a pencil drawing the brackets. One user suggested that the noise-like earcon should be speeded up so that it became, short and noticeable yet unobtrusive.

There were some creative responses from users to create alternative sound ideas for brackets. Six users suggested a metaphor of a door opening and closing to indicate the beginning and end of parenthesis. Two users referred to the Microsoft Windows system sound of a USB device disconnecting (beeps rising or falling) to denote the sound of brackets opening or closing.

When asked to suggest design ideas for ways that a fraction could be conveyed in spoken mathematics, a number of users suggested a spoken cue; either "fraction" for beginners or "frac" as users became acquainted with the system. There were also some creative sound suggestions to depict a fraction. Three users suggested sounds or metaphors relating to cutting, for example the "sound of broken glass" or a "laser cutting" or "scissors" sound.

Alternative superscript design included ideas related to the already implemented higher pitched voice for superscript and low pitched for a subscript. Users also suggested using a spoken lexical cue, there were some alternatives to "sup" and "sub". "Super" was suggested as a shortened version of superscript to avoid confusion with "sub". "Power" or "to the power of" was also suggested as a spoken alternative convey a superscript. While "power" could be a lexical replacement of "sup" or "super" "to the power of" will cause the same ambiguities as simple using "over" for a fraction because there is no indicator of an end to the superscript grouping. Perhaps an alternative could be to use "to the power of" in conjunction with the EQ modified voice as in the original design.

5. CONCLUSION AND FUTURE WORK

The quantitative results discussed above form an interesting evaluation of the current auditory cues that have been implemented. The results on spearcon recognition generated feedback concerning what words might be chosen as lexical spearcons. For example " Subscript/Superscript" are too close to be correctly identified as rapid speech. Rather some of the suggested alternatives in the final design section may be more effective such as “power” and “subscript”. Also the confusion of the word “end” could be changed to “end of” for clarity. The second section in the design survey containing full equation expression choices raised issues regarding the perceptual testing of auditory mathematics with a screen reader. While multiple-choice option answers with spoken lexical operators provide unambiguous detail they are tedious and fatiguing for a screen reader user. An alternative idea, for future evaluations, is to create a task where users need to perform a practical analysis or computation of the spoken equation, which would illustrate that they have understood the structure and auditory cues.

The partial presentation of equations worked well for both sighted and visually impaired users as there was minimal processing of text/speech options. This may be an effective method to evaluate the binaural cues in a controlled setting, including a more detailed evaluation of nested layers. As discussed previously the quantitative evaluation data collected in this survey are far from complete due to omission of training and the lack of control in the online dissemination of the design survey. However this data does provide a basis for designing a more controlled evaluation of auditory cues. Furthermore the quantitative sections in the survey created a means for users to understand the context of using sound in spoken equations and allowed users to offer informed suggestions in the final qualitative section of the questionnaire.

The qualitative responses from the survey presented in this paper have generated many new ideas for design for the auditory math system. It is intended that the results of the final evaluation of design choices will be implemented and the system will be tested using a controlled comparison evaluation against a speech-based system for spoken maths.

The process of creating and disseminating an online sound design survey that is accessible to screen reader users was in itself an interesting undertaking. Conducting pilot evaluations with screen reader users to determine the clearest survey question layout, as well as the search for an accessible web player were lengthy processes. This highlighted the fact that many popular web tools remain inaccessible to visually impaired screen reader users. The survey design process also emphasised that many tasks often taken for granted by sighted users, can become an arduous procedure for a screen reader user when technology is not accessible.

While the concepts explored in this paper are intended to inform a system design for visually impaired users, a usable auditory math system would be a valuable tool for sighted mathematicians in mobile contexts or contexts where users are described as having situation blindness. It is proposed that by implementing the auditory design ideas presented above, it is possible to enhance spoken mathematics with sound in an intuitive and unambiguous manner, taking advantage of the specific strengths and capabilities of audition.

6. ACKNOWLEDGMENTS

This research is supported by Science Foundation Ireland (SFI: www.sfi.ie). We would also like to thank all survey participants for their time, enthusiasm and creative responses.

7. REFERENCES

[1] Archambault, D., Stöger, B., Batusic, M., Fahrengruber, C., and Miesenberger, K. 2007. A software model to support collaborative mathematical work between braille and sighted users. *Proceedings of ASSETS 2007 9th International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 115–122, Tempe, AZ, USA.

[2] Bates, E., Fitzpatrick, D. 2010. Spoken Mathematics using Prosody, Earcons and Spearcons, *Proceedings of ICCHP, the 12th International Conference on Computers Helping People with Special Needs*, July 14-16, Vienna, Austria.

[3] Begault, D. R. & Erbe, T. R.: 1993. Multi-channel spatial auditory display for speech communications. *Audio Engineering Society 95th Convention*, Preprint No. 3707.

[4] Brewster SA. Using non-speech sound to overcome information overload. *Displays 1997*; 17: 179-189.

[5] Brewster, S.A., Wright, P.C. & Edwards, A.D.N. 1995. Experimentally derived guidelines for the creation of earcons. *Adjunct Proceedings of HCI'95*, Huddersfield, UK.

[6] Crispien, K. and Petrie, H. 1993. Providing Access to Graphical-Based User Interfaces for Blind People: Using Multimedia System Based on Spatial Audio Representation. 95th AES Convention, *J. Audio Eng. Soc, (Abstracts),* vol. 41, pp. 1060.

[7] Edwards, A. D. N. 1997. Using sounds to convey complex information. in A. Schick and M. Klatte (ed.), *Contributions to Psychological Acoustics: Results of the Seventh Oldenburg Symposium on Psychological Acoustics*, Oldenburg, pp. 341-358.

[8] Fitzpatrick, D. 2006. Mathematics: how and what to speak. In *Proceedings of ICCHP*, *International Conference on Computers Helping People with Special Needs,* Springer Verlag, pp. 1199-1206 .

[9] Gaver, W. W. (1986). Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction,* 2(2), 167-177.

[10] Gellenbeck, E., Stefik, A. 2009. Evaluating Prosodic Cues as a Means to Disambiguate Algebraic Expressions: An Empirical Study, *Proceedings of ASSETS 2009, the 11th International ACM SIGACCESS Conference on Computers and Accessibility* October 26-28, Pittsburgh, PA, USA.

[11] Goose, S. and Möller, C. 1999. A 3D Audio Only Interactive Web Browser: Using Spatialization to Convey Hypermedia Document Structure. In *Proceedings of the ACM International Conference on Multimedia*, pp. 363–371, Orlando, USA.

[12] Harling, P. A., Stevens, R. and Edwards, A. 1995. Mathgrasp: The design of an algebra manipulation tool for visually disabled mathematicians using spatial-sound and manual gestures. HCI Group, University of York, UK. Available: ftp://ftp.cs.york.ac.uk/pub/alistair/publications/ps/mathgrasp.ps

[13] Karshmer, A.I., Farsi, D. 2007. Access to Mathematics by Blind Students: A Global Problem, In *Proceedings of the WMSCI 2007 Conference*, Orlando, Florida.

[14] Karshmer, A.I., Gupta, G., Pontelli, E., Miesenberger, K., Ammalai, N., Gopal, D., Batusic, M., Stöger, B., Palmer, B., Guo, H-F. 2004. UMA: a system for universal mathematics accessibility, *Proceedings of the 6th international ACM SIGACCESS Conference on Computers and Accessibility*, October 18-20, Atlanta, GA, USA.

[15] Moos, A. and Trouvain, J. 2007. Comprehension of ultra-fast speech – blind vs. "normally hearing" persons. *Proceedings of the 16th International Congress of Phonetic Sciences*, pages 677-680.

[16] Mynatt, E. 1997. Transforming graphical interfaces into auditory interfaces for blind users. *Human-Computer Interaction,* 12(1), 7-45.

[17] Nemeth, A.: Mathspeak. Available: http://people.rit.edu/easi/easisem/talkmath.htm

[18] Ohio State University, Web Accessibility Initiative, A Survey of Survey Tools, Available: http://wac.osu.edu/workshops/survey_of_surveys/

[19] Pirhonen, A., Murphy, E. 2008. Designing for the Unexpected: The role of creative group work for emerging interaction design paradigms. *Journal of Visual Communication: Special Issue on Haptic Interactions*, 7(3) 331-344.

[20] Raman, T. V. 1994. *Audio Systems for Technical Reading*. PhD Thesis, Department of Computer Science, Cornell University, Ny, USA.

[21] Stevens, R. D. & Edwards, A. D. N. 1996. An approach to the evaluation of assistive technology. In *Proceedings of the 1996 Annual ACM Conference on Assistive Technologies*; April 11-12; Vancouver, British Colombia, Canada.

[22] Stevens, R. D., Principles for the Design of Auditory Interfaces to Present Complex Information to Blind Computer Users. 1996. PhD Thesis, University of York, UK.

[23] Walker, B., Nance, A. and Lindsay, J.: Spearcons. 2006. Speech-based Earcons Improve Navigation Performance in Auditory Menus. In *Proceedings of the 12th International Conference on Auditory Display*, London, UK, June 20-23.