# Collaborative Cross-Modal Interfaces


N. Bryan-Kinns
Queen Mary University of London
Interactional Sound and Music Group
School of Electronic Engineering and Computer Science
London
E1 4NS
nickbk@eecs.qmul.ac.uk

T. Stockman
Queen Mary University of London
IMC Research Group
School of Electronic Engineering and Computer Science
London
E1 4NS
tonys@eecs.qmul.ac.uk

O. Metatla
Queen Mary University of London
Interactional Sound and Music Group
School of Electronic Engineering and Computer Science
London
E1 4NS
oussama@eecs.qmul.ac.uk



## ABSTRACT
Diagrams form a key part of how we collaborate in the workplace and the home. However, the visual nature of diagrams makes them difficult to use and manipulate when people have differing perceptual resources available to them, for example, due to visual impairment, or reduced access to computing and network bandwidth. In this paper we describe the potential for cross-modal interfaces to support collaboration with diagrams when participants use a varied mix of perceptual channels. Our position is that developing tools for cross-modal collaboration will increase social and workplace inclusion.



## Categories and Subject Descriptors
H.1.2 [User/Machine Systems]: *Human factors*; K.4.2 [Social Issues]: *Assistive technologies for persons with disabilities;* H.5.3 [Group and Organization Interfaces]: *Collaborative computing, Computer-supported cooperative work.*

## General Terms
Design, Human Factors.

## Keywords
Collaboration, cross-modal, multimodal, accessibility, inclusion.


## 1. INTRODUCTION
Diagrams are a key form of representation used in all manner of collaborations from creating music to designing jet engines. Indeed, diagrammatic representations have often become common standards for expressing specialised aspects of a particular discipline, e.g. meteorologists use weather maps, architects use floor plans, and computer scientists make extensive use of node-and-arrow diagrams. Diagrams are particularly important in the software engineering domain where they are used to capture and structure solutions to complex problems e.g. Unified Modelling Language (UML) and Entity-Relationship (ER) diagrams are used throughout the design process by commercial software engineering teams to help convey and refine designs. However, there is currently no practical way for visually impaired co-workers to view, let alone edit, such diagrams. Teammates typically resort to using expensive and time-consuming static representations (e.g. using heat reactive paper to print diagrams using raised lines), or reverse engineering the diagrams (e.g. coding diagrams using C++ and passing them to sighted team mates to render as UML). Lack of access to these types of diagrams also potentially forms a barrier to progression for visually impaired programmers into jobs such as Software Engineer or Systems Analyst, where the ability to produce and use such diagrams forms an even more central part of the job role. This is a major barrier to workplace collaboration that contributes to the exclusion and disengagement of visually impaired individuals. Indeed, the RNIB estimate that 66% of blind and partially sighted people are currently unemployed [8].



*Digital Futures'10*, October 11-12, 2010, Nottingham, UK.


## 2. THE CHALLENGE: CROSS-MODALITY
The research challenge we are exploring is: how to design support for collaboration where participants have differing access to modalities, which we refer to as *cross-modal* collaboration. Addressing this challenge in the workplace has the potential to significantly improve the working lives and inclusion of perceptually impaired workers. In the teaching and learning environment, methods of cross-modal collaboration will increase social inclusion, providing a valuable resource that will tap into potential that would otherwise remain dormant.

Current technological developments mean that it is increasingly feasible to support cross-modal collaboration in a range of devices and environments, yet there are no practical examples of such systems, e.g. Apple's iPhone provides touch, visual, and speech interaction, but there is no coherent way of collaborating using differing sets of modalities. There is significant potential impact as, in addition to increasing social and workspace inclusion, the substantially ageing workforce, and elderly people in general, will also benefit from the outputs of this research which will support continued independent living and working. Indeed, the amount of time that we can expect to live with a limiting illness or disability is increasing as our life expectancies rise, and over eighty year olds are currently the fastest growing population group in the UK [6]. Supporting lifelong health and wellbeing is a key Research Councils UK priority area which we will contribute to by providing mechanisms by which people can continue to collaborate with others as they acquire sensory impairments in old age such as remaining active in the workplace, engaging in cross-modal social networks, and informal collaborations.

## 3. MULTIMODALITY
Multimodal interaction is a term used to describe the integration of multiple user input or system output modalities which has the

potential of promoting richer and more efficient interactive experiences [7]. Multimodality could also overcome the constraints that hinder information flow between a system and its users when limited sensory channels are available to convey information. Currently, the visual modality is the predominant modality in human-computer interaction, but audio and haptics are increasingly being used to augment, or in some cases substitute graphical displays. Auditory display is the use of speech and non-speech sounds to convey information [3]. They are typically used to present information to visually impaired people, to draw attention to activity outside of the field of view, or to provide additional information in situations where the eyes are occupied or there is limited screen space. To date, auditory interfaces have been successfully employed in a variety of areas including monitoring applications for complex environments, such as operating rooms and aircraft flight decks, improving accessibility to visually represented information, and supporting data exploration through sonification. Haptic displays on the other hand, are interfaces that convey information through cutaneous or kinesthetic sensation. They allow visually represented objects to be augmented with rich physical properties, such as mass and textures, and can be used to simulate most physical sensations that can be mathematically represented, such as gravitational fields [2]. This is usually achieved by using vibrating or robotic devices to convey haptic sensations, allowing a user to perform physical manipulations like pulling, pushing and feeling objects. Research has produced a variety of techniques for conveying information through haptic feedback. Tactons for instance, are a form of structured tactile signal that can be used to convey abstract messages non-visually and are equivalent to visual icons and audio earcons [1].

However, despite the progress in multimodal research, accessibility research has also shown that simply developing an accessible display does not guarantee that collaboration will necessarily be supported, and moreover, the social interaction between the participants is as crucial as the ability to access and manipulate the shared content [9]. Furthermore, a key drawback of existing collaborative systems is that they do not explicitly support cross-modal interaction where participants have access to differing modalities. This reduces the transformational potential of these systems by systematically excluding collaborators with differing perceptual needs.

# 4. OPPORTUNITY: CROSS-MODALITY

We have developed and tested a prototype research tool in Java which allows users to inspect and edit node-and-link diagrams in a structured manner using a cross-modal combination of audio and visual interaction (see [5] for a description). To date we have conducted laboratory based testing of the approach which has demonstrated its utility for sighted and visually impaired users editing object-oriented modelling diagrams [4,5]. Whilst we have demonstrated that our research idea is feasible, there are a number of scalability issues which we intend to address in order to turn our research idea into a solution with a clear route to adoption. Firstly, the testing has so far been undertaken in laboratory conditions which shows that the forms of interaction are feasible, but does not demonstrate its practical and effective use in the wild. We will address this by demonstrating the utility of the approach in the real world of the workplace with real design problems to be solved. Second, our current research prototype is proprietary which reduces possible routes to exploitation. We will address this by developing the software further into an open source system which can be integrated with other open source diagram editing tools, thus significantly expanding the potential user and developer community. Third, the tool currently only supports a small number of forms of diagrams (UML and ER) which limits its potential uptake. We will develop the system into a modular tool which allows for different diagram forms to be 'plugged in' thereby expanding the user base. Fourth, our tool currently primarily supports audio; again, we will develop the tool in a modular manner to allow different modalities to be used which increases the potential user-base. We aim to achieve the above by coordinating with our project partners to recruit and work closely with software engineering design teams involving sighted and visually impaired collaborators. We will introduce the cross-modal tool developed as part of this project in the teams' design process and observe its impact on their collaborations. Such observations, in addition to our interactions with the teams through interviews and focused group sessions will provide insights into how to improve the tool and, hence, the design and support for cross-modal collaboration in the workplace.

# 5. SUMMARY

Cross-modal interfaces have the potential to significantly improve social and workplace inclusion by supporting a range of perceptual modalities in collaborative settings. Our work will deploy and test a collaborative cross-modal interface to demonstrate the transformational potential of this approach.

# 6. ACKNOWLEDGMENTS

Funded by the RCUK Digital Economy Programme Research in the Wild call (EP/H042865/1). We would like to thank our partners: the RNIB and BCAB for their contribution and support.

# 7. REFERENCES

[1] S. Brewster & L. M. Brown. Tactons: structured tactile messages for non-visual information display. In *Proc. AUIC'04*, 15–23, 2004. Australian Computer Society.

[2] P. Kortum. *HCI Beyond the GUI: Design for Haptic, Speech, Olfactory, and Other Nontraditional Interfaces*. Morgan Kaufmann, 2008.

[3] G. Kramer. *An introduction to auditory display*. SFI studies in the sciences of complexity. Addison Wesley Longman, 1992.

[4] O. Metatla, N. Bryan-Kinns, & T. Stockman. Using hierarchies to support non-visual access to relational diagrams. In *Proc. of BCS HCI'07*, 215–225, 2007. BCS.

[5] O. Metatla, N. Bryan-Kinns, & T. Stockman. Constructing relational diagrams in audio: the multiple perspective hierarchical approach. In *Proc. of Assets'08*, 97–104, 2008. ACM.

[6] Office for National Statistics, 2009.

[7] S. Oviatt. Ten myths of multimodal interaction. *Communications of the ACM*, 42(11):74–81, 1999.

[8] RNIB, *Looking forward to 2014 – RNIB's strategy to end the isolation of sight loss*, RNIB, 2009.

[9] F. Winberg & J. Bowers. Assembling the senses: towards the design of cooperative interfaces for visually impaired users. In *Proceedings of CSCW '04*, 332–341, 2004.