EFFECT OF SOUND SPATIALISATION ON MULTITASKING IN REMOTE MEETINGS.

Stuart N. Wrigley[1], Simon Tucker[2], Guy J. Brown[1] and Steve Whittaker[2].

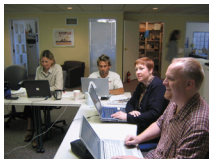
[1]Department of Computer Science and [2]Department of Information Studies, University of Sheffield, UK.

{s.wrigley, g.brown}@dcs.shef.ac.uk, {s.tucker, s.whittaker}@sheffield.ac.uk

The University Of Sheffield.

Introduction.

Workers are experiencing a significant rise in the number of meetings that they are expected to attend – a growing proportion of which are virtual. However, they are still expected to achieve their productivity goals as normal.



This pressure means it is now common for participants to multitask during virtual meetings. For example, responding to email whilst listening to the meeting.

Virtual meeting participants are more susceptible to confusion due to lack of non-verbal communication cues.

It is important that the technology used to present the meeting to the participant does so in a manner that allows them to multitask with greatest efficiency. In this study we look at different audio playback techniques.

Motivation

How do listeners process sound environments?

- Auditory Scene Analysis[1] to create 'streams' (mental representation of sounds)
- Attend to single stream, but monitor multiple talkers?
- Create multiple streams first, then attend to single talker stream?
- Does cueing aid the attentional selection process (and hence multitasking ability)?

Audio techniques.

This study examines three different techniques for presenting the audio from a virtual meeting to the listener: mono, dichotic and spatialised.

Mono

- Audio signals from each talker are mixed in equal proportions
- Equivalent to standard teleconference approaches

Dichotic

- Audio signals from one or more talkers presented to the left ear
- Audio signals from the remaining talkers presented to the right ear

Spatialised

- Simulates a full 3D sound environment
- Each talker's audio signals can be placed at any position around the listener's head.

Experimental design.

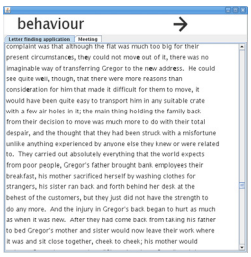
Experiment 1: mono and spatialised speech equally disruptive to the subject's multitasking performance?

Experiment 2: any benefit in cueing the subject about the direction from which the keyword would be said?

Experiment 3: any benefit in cueing the subject about the gender of the participant who utters it?

Subjects sat at a PC performing a task which involved finding as many occurrences of the letter 'e' as possible from a section of text and clicking on them using the mouse.

Some (but not all) scenarios were also accompanied by an audio playback of a meeting.



Common to all experiments:

- 60 second scenarios each using a different section of text
- Time and location of each letter click logged. Allowed computation of e-spotting rate (e's per second).
- When audio present, subject asked to listen for a keyword *in addition* to performing the e-finding task.
- When keyword heard, subjects clicked a button on the interface.

Subjects and procedures.

- 12 native English speaking subjects were used (6 male and 6 female).
- Subjects sat in a single walled sound-attenuating booth (IAC 402-A Audiometric Booth). Audio was presented to a pair of Sennheiser HD250 linear II headphones.

Stimuli.

Audio

- Taken from a number of meetings within the AMI corpus[2]
- Word-level transcripts were used to remove crosstalk from each channel
- Channels were upsampled from the original 16 kHz to 48 kHz to ensure sufficient spatial resolution when spatialised
- Each channel was amplitude normalised to ensure the RMS values of the speech portions were equal
- Channels were spatialised using head related transfer functions (HRTFs)

Keyword

- High TF*IDF score
- Occurred after 20 to 50 seconds into clip (clip length 60 seconds)
- Start times were evenly distributed between these two limits.

Text

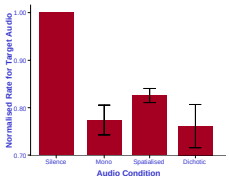
- E-spotting text extracted from *The Metamorphosis* by Franz Kafka
- Each presentation used a different, randomly selected, portion

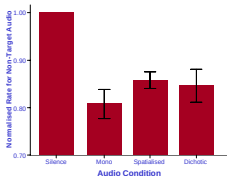
Evaluation.

Performance metric was the number of e's spotted per second. Computed for:

- audio portions containing relevant cue (target)
- audio portions not containing relevant cue (non-target)

1. Which audio presentation technique allows best multitasking performance?

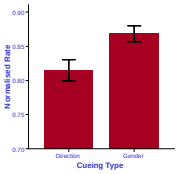




- Performance best with no audio
- When listening to target audio, performance with spatialised audio best
- Performance suffers in target audio portions in dichotic condition

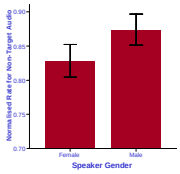
2. Does cueing by gender or direction influence multitasking performance differently?

- Significant difference in performance
- Subjects performed better when cued to gender
- No user preference for cueing type



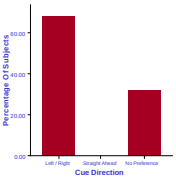
3. Does cueing a particular gender influence multitasking performance?

- Performance better when listening to male talkers
- No subject preference stated for either gender
- No effect of subject gender on performance or preference



4. Does cueing a particular direction influence multitasking performance?

- No effect of direction on performance
- No interaction between direction and cueing type
- Users didn't state a preference for straight ahead



Conclusions.

When cued only by keyword, spatialised audio provided no improvement over mono.

However, significant differences were observed in more realistic scenarios, e.g.,

- increased multitasking efficiency when listening to spatialised audio.
- suggests that extracting a spatially distinct stream and subsequently attending to it involves a lower cognitive load than simply attending to a single stream containing multiple talkers.

Future experiments will allow subjects to position the participants in a virtual auditory space including a distance metaphor by allowing subjects to place participants of less relevance to their interests further away (i.e., lower amplitude and increased reverberation).

[1] Bregman, A.S. (1990). *Auditory Scene Analysis. The Perceptual Organization of Sound.* MIT Press.
[2] http://corpus.amiproject.org/

Funded by AMIDA: an EU 6th FWP IST Integrated Project (project FP6-0033812).

AMI Consortium