AUDIO SPATIALISATION STRATEGIES FOR MULTITASKING DURING TELECONFERENCES.

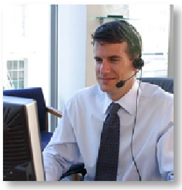
Stuart N. Wrigley[1], Simon Tucker[2], Guy J. Brown[1] and Steve Whittaker[2].

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

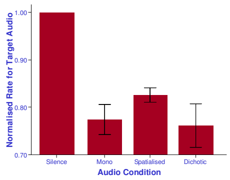
{s.wrigley, g.brown}@dcs.shef.ac.uk, {s.tucker, s.whittaker}@sheffield.ac.uk

Introduction.

It is increasingly common for workers to multitask during virtual meetings. For example, responding to email whilst listening to the meeting.



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.



Our previous study[2] established that the use of spatialised audio significantly increased multitasking efficiency compared to traditional audio presentation.

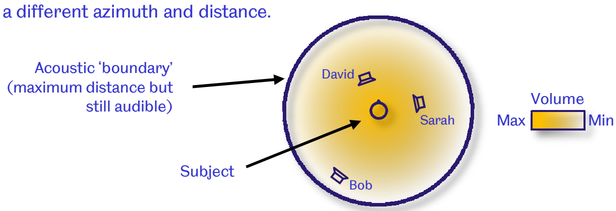
Here, we extend our previous work by giving the subjects full control over the locations and distance of the participants.

We are interested in:

- the positioning of the target talkers relative to the interfering talkers;
- subjects' strategies for reducing the influence of the interfering talkers;
- what effect these strategies have on their multitasking performance.

Audio techniques.

Acoustic virtual reality environment created by spatialising each participant to be at a different azimuth and distance.



- Word-level transcripts were used to remove crosstalk from each channel.
- Each channel was amplitude normalised to ensure the RMS values of the speech portions were equal.
- Channels were positioned using OpenAL (Open Audio Library) audio API[3].

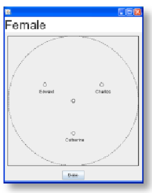
Experimental design.

Multitasking scenario: subject monitors meeting audio for topic of interest (keyword) to occur while text processing (e-spotting).

- Each meeting contained 3 participants; a mix of genders.
- Each presentation had 2 phases.

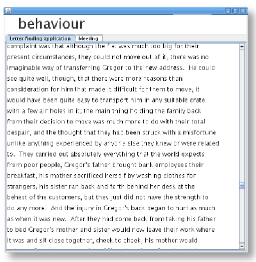
Phase 1 – participant positioning

- Subject told name or gender of target participant (single or dual target).
- 30 seconds to arrange participants.
- Continuous speech recordings for each participant played concurrently allowing subject to hear, in realtime, effect of moving any of the three talkers.
- In half of trials, distance was fixed.



Phase 2 – multitasking

- Subject monitors meeting for keyword.
- Subject finds as many occurrences of the letter 'e' as possible from a section of text and clicks on them using the mouse.
- 60 second scenarios each using a different section of text.
- Time and location of each letter click logged.
- When keyword heard, subjects clicked a button on the interface.



Subjects and procedures.

- 15 native English speaking subjects were used (8 male and 7 female).
- Subjects sat in a single walled sound-attenuating booth (IAC 402-A Audiometric Booth). Audio was presented to a pair of Sennheiser HD 25 SP headphones.

Stimuli.

Audio

- Taken from a number of meetings within the AMI corpus[4].

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

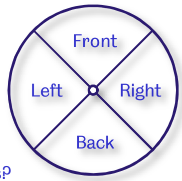
- E-spotting text extracted from *The Metamorphosis* by Franz Kafka
- Each presentation used a different, randomly selected, portion.

Evaluation.

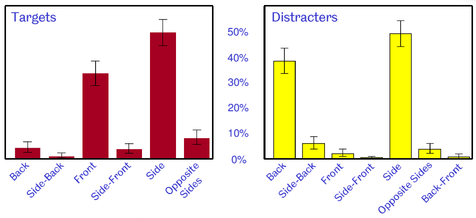
Performance metric was the number of e's spotted per second. Computed for:

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

To investigate which arrangements were employed by subjects, we split the acoustic space into four regions: front, back, left and right.



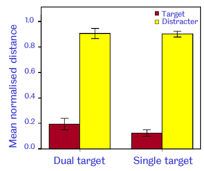
1. Which locations are preferred for targets and distracters?



- Most common location for target talkers was at the side.
- Left side favoured for targets (73% of trials).
- Right side favoured for distracters (66% of trials).
- Target location did not influence e-spotting rate.
- 'Opposite positioning' was most common strategy.

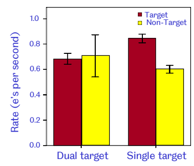
2. Did subjects show consistency when using the distance cue?

- Acoustic space normalised: subject at distance 0, acoustic boundary is 1.
- Targets generally placed close to the subject.
- Distracters placed far away.
- Multiple distracters placed at similar distances.
- Multiple targets placed at similar distances.
- No effect on e-spotting rate.



3. How does the number of targets influence multitasking performance?

- Multitasking performance during target portions higher when listening to a single target.
- Number of targets / distracters has less effect on multitasking during distracter portions.



Conclusions.

Listeners generally employ consistent and effective strategies to maximise their multitasking performance. Common to place target and distracter talkers opposite each other. However, some strategies were not necessarily optimal for multitasking: subjects preferred to separate dual target talkers but multitasked more efficiently when targets were closer in acoustic space

Target talkers generally placed to one side consistent with natural two-talker interactions[5].

Subjects made full use of the ability to alter the distance of talkers in the acoustic space: target talker(s) were consistently moved closer to the subject while distracter talker(s) were moved further away. Contributes to positive user experience rather than increased multitasking performance.

Outcomes raise important points for the design of future teleconference presentation approaches.

[1] Bregman, A.S. (1990). *Auditory Scene Analysis. The Perceptual Organization of Sound*. MIT Press.
[2] Wrigley, Tucker, Brown and Whittaker (2008). The influence of audio presentation style on multitasking during teleconferences. *Interspeech 2008*,
[3] http://www.openal.org
[4] http://corpus.amiproject.org
[5] Ciolek and Kendon (1980), Environment and the spatial arrangement of conversational encounters, *Sociological Inquiry*, **50**(3-4), pp. 237–271.

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