Intelligibility vs comprehension: understanding quality of accessible next-generation audio broadcast

Shirley, BG and Ward, L

<table>
<thead>
<tr>
<th>Title</th>
<th>Intelligibility vs comprehension: understanding quality of accessible next-generation audio broadcast</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors</td>
<td>Shirley, BG and Ward, L</td>
</tr>
<tr>
<td>Type</td>
<td>Conference or Workshop Item</td>
</tr>
<tr>
<td>URL</td>
<td>This version is available at: <a href="http://usir.salford.ac.uk/id/eprint/47780/">http://usir.salford.ac.uk/id/eprint/47780/</a></td>
</tr>
<tr>
<td>Published Date</td>
<td>2018</td>
</tr>
</tbody>
</table>

USIR is a digital collection of the research output of the University of Salford. Where copyright permits, full text material held in the repository is made freely available online and can be read, downloaded and copied for non-commercial private study or research purposes. Please check the manuscript for any further copyright restrictions.

For more information, including our policy and submission procedure, please contact the Repository Team at: usir@salford.ac.uk.
Accessible Audio Research

Intelligibility vs Comprehension
Understanding Quality of Accessible Next-generation Audio Broadcast

Dr. Ben Shirley
@bengshirley

Lauren Ward
@thepengineer

University of Salford
MANCHESTER
Current Access Needs

11 Million People
have hearing impairment in the UK in 2015 [1]

Current Access Needs

11 Million People
have hearing impairment in the UK in 2015 [1]

87% struggle to understand speech on TV [2]

Current Access Needs

11 Million People
have hearing impairment in the UK in 2015 [1]

Key Issues identified:

- 47% balance between audio objects
- 18% accents and clarity of speech
- 18% recording & reproduction problems
- 17% other

87% struggle to understand speech on TV [2]

Current Access Services

Standardised services (UK figures[3])

Subtitles (*100%), signing (*5%), audio description (*10%) – Ofcom mandated

VoD now covered under Digital Economy Act (2017)

*Maximum based on viewer share with some exemptions

Guidance only

On speech levels and speech clarity[4]

---

Channel-based approaches

Clean Audio

Made use of speech being (mostly) in centre channel of 5.1 broadcast [5, 6]


Channel-based approaches

Clean Audio

Made use of speech being (mostly) in centre channel of 5.1 broadcast [5, 6]

Centre speaker improves intelligibility compared with phantom centre [7]
Channel-based approaches

Clean Audio

Made use of speech being (mostly) in centre channel of 5.1 broadcast [5, 6]

Centre speaker improves intelligibility compared with phantom centre [7]

Channel-based approaches

Clean Audio

Similar approach adopted by HBB4ALL, exploiting HBBTV 2.0 specification\[8\]

Clean audio algorithm using IRT ‘centre cut’ approach

\[8\] D4.4 – Pilot-B Evaluations and recommendations, in HBB4ALL, Connected TV Accessibility. 2016.
Channel-based approaches

Speech Enhancement

Early Work: based on blind source separation

- Shown to reduce listening effort but not intelligibility [9]


Channel-based approaches

Speech Enhancement

**Early Work:** based on blind source separation

- Shown to reduce listening effort but not intelligibility [9]

**Current State:** Use in making small adjustments to level or position of objects in original content

- Ongoing work determining suitable objective measures of quality for this [10]

Object-based Audio

New opportunity for accessible TV

• ‘Next-generation’ Object-based audio

• Personalisation for accessibility
What is object-based audio?
Channel based audio (now)

- dialogue
- music
- reverb
- noise
- foot steps
- broadcast
Object-based audio

- Channel-based audio (now)
- Dialogue
- Fan noise
- Music
- Footsteps
- Reverb
- Noise
Object-based audio: An Analogy

Traditional Broadcasting

Shirley & Ward – Intelligibility vs. Comprehension

June 2018
Object-based audio: An Analogy

Traditional Broadcasting

Objects

Intelligibility vs. Comprehension

June 2018
Object-based audio: An Analogy

Traditional Broadcasting

Objects

Metadata

Shirley & Ward – Intelligibility vs. Comprehension

June 2018
Object-based audio: An Analogy

Traditional Broadcasting

Objects

Metadata

Renderer

Shirley & Ward – Intelligibility vs. Comprehension

June 2018
Early Object-based approaches

Dialogue Enhancement

based on Spatial Audio Coding [11]

34%

Early Object-based approaches

Dialogue Enhancement

based on Spatial Audio Coding [11]

34% → 81%

Improvement speech recognition in applause noise

Early Object-based approaches

FascinatE [12]

Individual user control of:
- on pitch sounds
- crowd level
- commentary

Early Object-based approaches

Individual user control of:
- on pitch sounds
- crowd level
- commentary

Which raised the question:

Is speech really the only important thing for understanding the narrative of media?

Intelligibility vs. Comprehension

**Intelligibility**: Proportion of words correctly heard

- Speech
  - Useful
- Non-speech
  - Masker
Intelligibility vs. Comprehension

**Intelligibility:** Proportion of words correctly heard

Speech

*Useful*

**Comprehension:** Proportion of content understood

Non-speech

*Masker*
Intelligibility vs. Comprehension

**Intelligibility:** *Proportion of words correctly heard*

**Comprehension:** *Proportion of content understood*
Intelligibility vs. Comprehension

**Intelligibility:** Proportion of words correctly heard

**Comprehension:** Proportion of content understood

Shirley & Ward – *Intelligibility vs. Comprehension*  
June 2018
Effect of non-speech sounds

Normal Hearing[13]

36%

Without Sound Effects

Effect of non-speech sounds

Normal Hearing\textsuperscript{[13]}

36\%  
Without Sound Effects

$\rightarrow$

62\%  
With Sound Effects


Shirley & Ward – \textit{Intelligibility vs. Comprehension}  
June 2018
Effect of non-speech sounds
Hard of Hearing

Audiometric Thresholds

Usefulness of Sound Effects

Implementation of an object based clean audio solution for hearing impaired viewers using DTS:X and MDA

Exploration of user-preferences for audio-object categories volumes

Implementation of an object based clean audio solution for hearing impaired viewers using DTS:X and MDA [15]

Exploration of user-preferences for audio-object categories volumes

“first time I have been able to understand dialogue without subtitles in a very long time”

“useful and interesting to be able to adjust different aspects of sound”

“very straightforward, very good, when can I have one?”

Accessible Broadcast Audio Customisation

Object-based media assets
- Speech
- Sound Effects
- Music
- Vision
- Metadata
- Expert Produced Knowledge

Rendering Algorithm in Set-top Box
Controls number of media objects rendered and their respective volumes

End-user complexity control

Personalised Audio

End-User

Shirley & Ward – Intelligibility vs. Comprehension
June 2018
Assessing quality for new access services

It’s complicated…

Shirley & Ward – *Intelligibility vs. Comprehension*
Assessing quality for new access services

It’s complicated…

Object-based personalisation facilitates useful solutions but

…evaluation becomes a complex problem
Assessing quality for new access services

It’s complicated…

Object-based personalisation facilitates useful solutions

but

…evaluation becomes a complex problem

No ‘one size fits all’ solution for accessibility means no ‘one size
fits all’ solution for evaluating quality
Accessible Audio Research

For More Information
hub.salford.ac.uk/accessibleaudio

Lauren Ward
Postgraduate Researcher
General Sir John Monash Scholar
L.Ward7@edu.salford.ac.uk
@thepengineer

Dr. Ben Shirley
Senior Lecturer in Audio Technology
B.G.Shirley@salford.ac.uk
@bengshirley