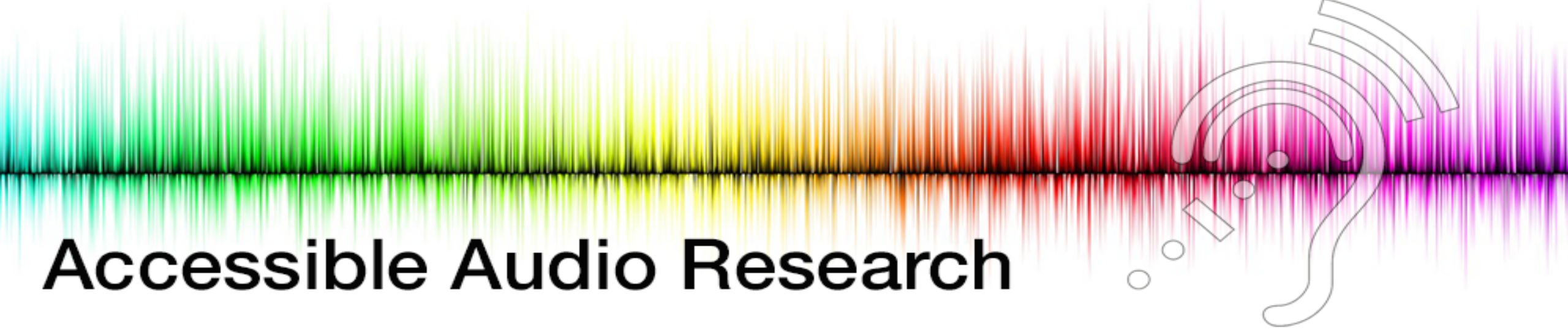


Intelligibility vs Comprehension

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]

[1] Action on Hearing Loss. (2015) Hearing Matters Report.
[2] Royal National Institute for Deaf People (RNID), "Annual survey report 2008," 2008

Current Access Needs

11 Million People

have hearing impairment in the UK in 2015 [1]

87%

struggle to understand speech on TV [2]

[1] Action on Hearing Loss. (2015) Hearing Matters Report.
[2] Royal National Institute for Deaf People (RNID), "Annual survey report 2008," 2008

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]

[1] Action on Hearing Loss. (2015) Hearing Matters Report.
[2] Royal National Institute for Deaf People (RNID), "Annual survey report 2008," 2008

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]

[3] Ofcom, *TV access services 2017: Q1 and Q2*. 2017: London, UK.
[4] Digital Production Partnership (2017), *Technical Specification for the Delivery of Television Programmes as AS-11 files*

Channel-based approaches

Clean Audio

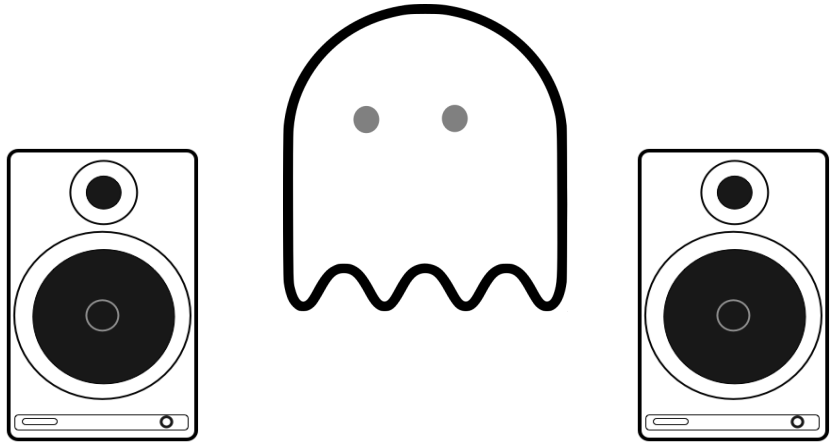
Made use of speech being (mostly) in centre channel of 5.1 broadcast [5, 6]

[5] Shirley, B.G. and P. Kendrick, *The Clean Audio Project: Digital TV as assistive technology*. Journal of Technology & Disability, 2006. **18**(1): p. 31-41.
[6] ETSI, *ETSI TS101154 v1.9.1 Digital Video Broadcasting (DVB); Specification for the use of Video and Audio Coding in Broadcasting Applications based on the MPEG-2 Transport Stream*, in *Annexe E.4 Coding for Clean Audio SA services*. 2009, ETSI: FRANCE.

Channel-based approaches

Clean Audio

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



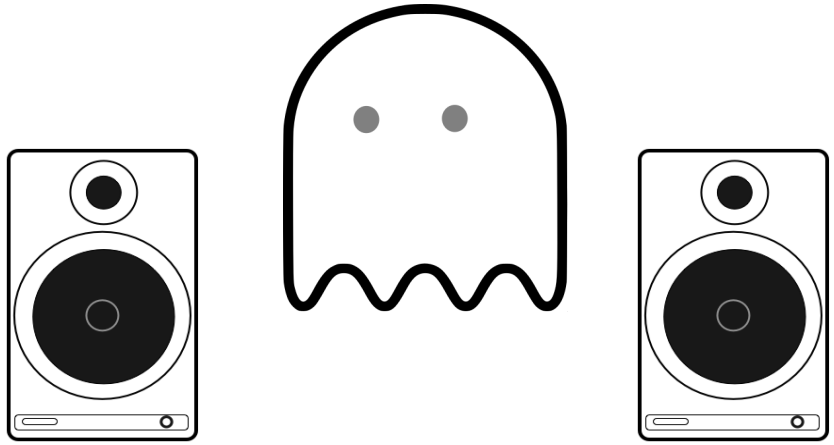
Centre speaker improves intelligibility compared with phantom centre [7]

[7] Shirley, B., Kendrick, P., & Churchill, C. (2007). The effect of stereo crosstalk on intelligibility: comparison of a phantom stereo image and a central loudspeaker source. *Journal of the Audio Engineering Society*, *55*(10), 852-863.

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]

Standardised in:

DVB Digital Video Broadcasting

ETSI World Class Standards

IPTV

EBU

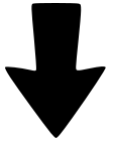
NorDig

[5] Shirley, B., Kendrick, P., & Churchill, C. (2007). The effect of stereo crosstalk on intelligibility: comparison of a phantom stereo image and a central loudspeaker source. *Journal of the Audio Engineering Society*, *55*(10), 852-863.

Channel-based approaches

Clean Audio

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



Clean audio algorithm using IRT ‘centre cut’ approach

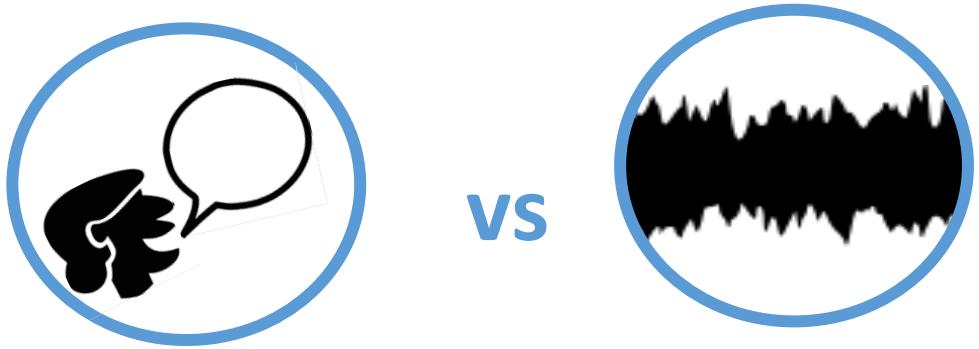
[8] *D4.4 – Pilot-B Evaluations and recommendations*, in *HBB4ALL, Connected TV Accessibility*. 2016.

Channel-based approaches

Speech Enhancement

Early Work: based on blind source separation

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



Speech
Useful

Non-speech
Masker

[9] Armstrong, M (2011). Audio Processing and Speech Intelligibility: a literature review, *BBC R&D White Paper WHP190*

[10] Torcoli, Matteo, and Christian Uhle. "On the Effect of Artificial Distortions on Objective Performance Measures for Dialog Enhancement." *Audio Engineering Society Convention 141*. Audio Engineering Society, 2016.

Channel-based approaches

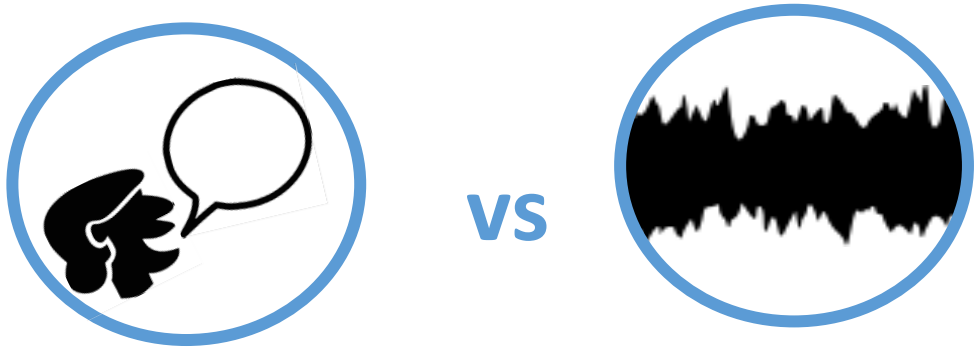
Speech Enhancement

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]



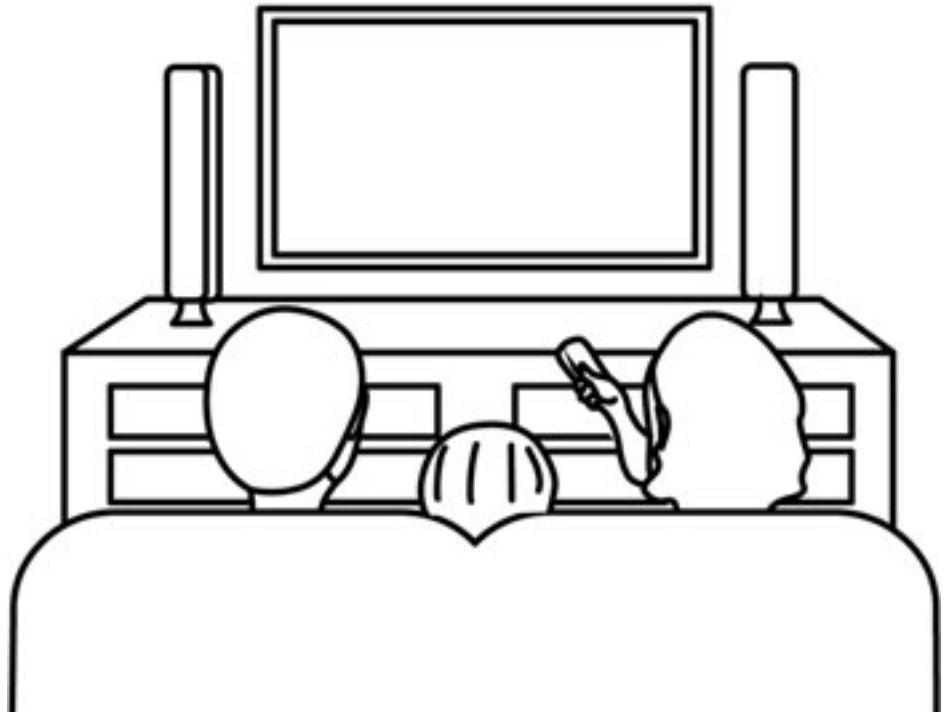
Speech
Useful

Non-speech
Masker

[9] Armstrong, M (2011). Audio Processing and Speech Intelligibility: a literature review, *BBC R&D White Paper WHP190*

[10] Torcoli, Matteo, and Christian Uhle. "On the Effect of Artificial Distortions on Objective Performance Measures for Dialog Enhancement." *Audio Engineering Society Convention 141*. Audio Engineering Society, 2016.

Object-based Audio

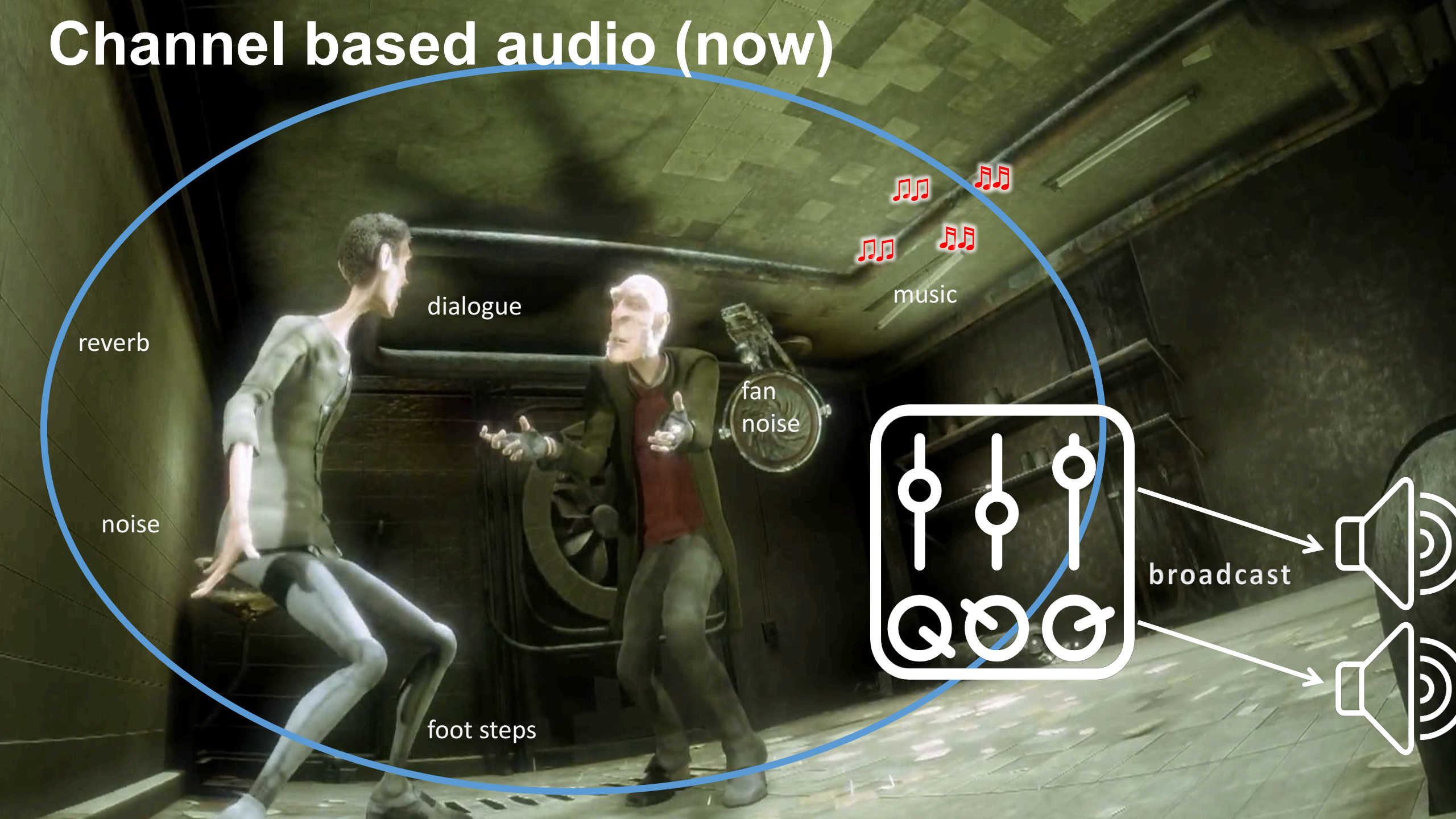


New opportunity for accessible TV

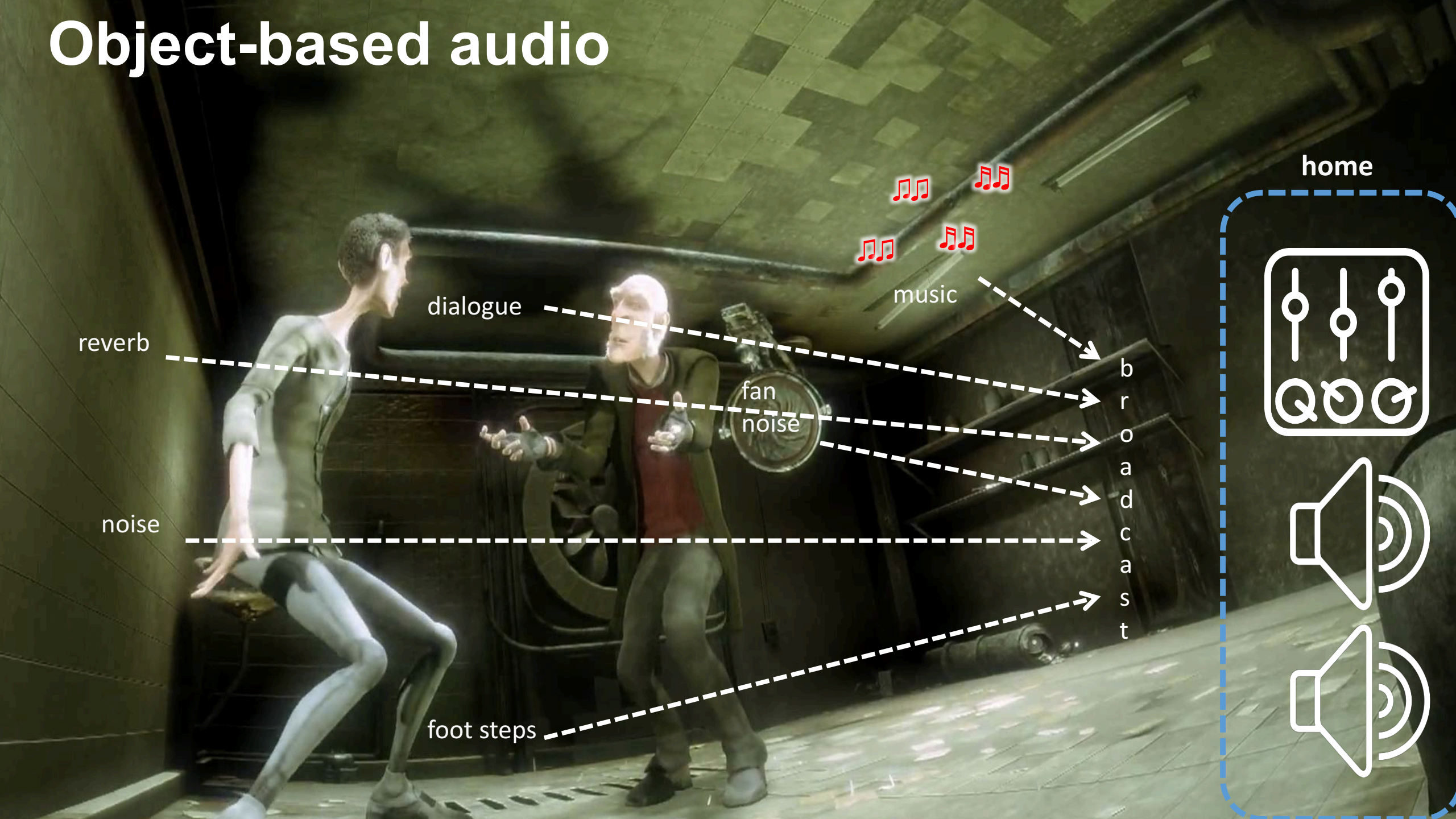
- ‘Next-generation’ Object-based audio
- Personalisation for accessibility

What is object-based audio?

Channel based audio (now)



Object-based audio



Object-based audio: An Analogy

Traditional Broadcasting

Object-based audio: An Analogy

Traditional Broadcasting

Object-based audio: An Analogy

Traditional Broadcasting

Objects

Metadata

Object-based audio: An Analogy

Traditional Broadcasting

Objects

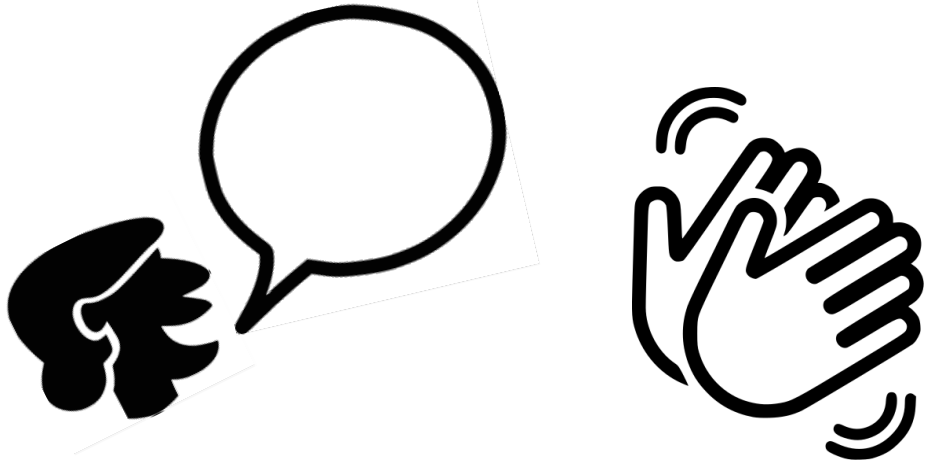
Renderer

Metadata

Early Object-based approaches

Dialogue Enhancement

based on Spatial Audio Coding [11]



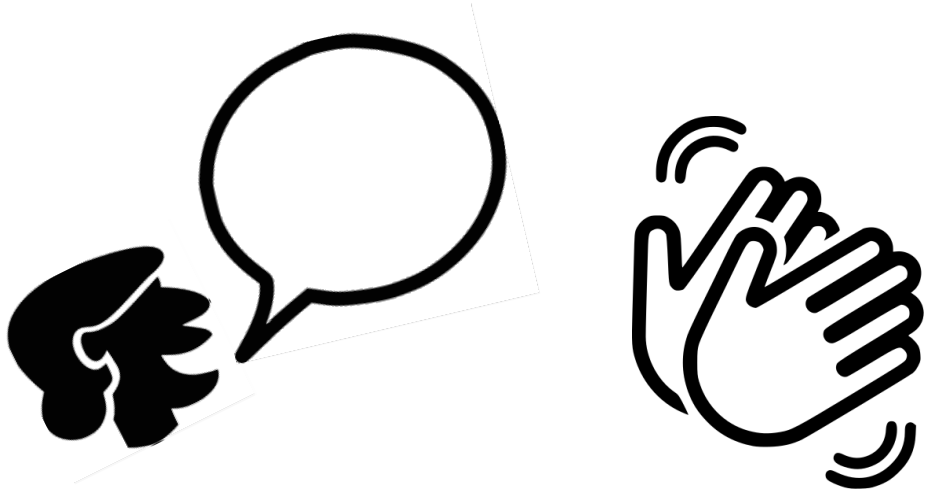
34%

[11] Paulus, Jouni, et al. "MPEG-D spatial audio object coding for dialogue enhancement (SAOC-DE)." Audio Engineering Society Convention 138. Audio Engineering Society, 2015.

Early Object-based approaches

Dialogue Enhancement

based on Spatial Audio Coding [11]



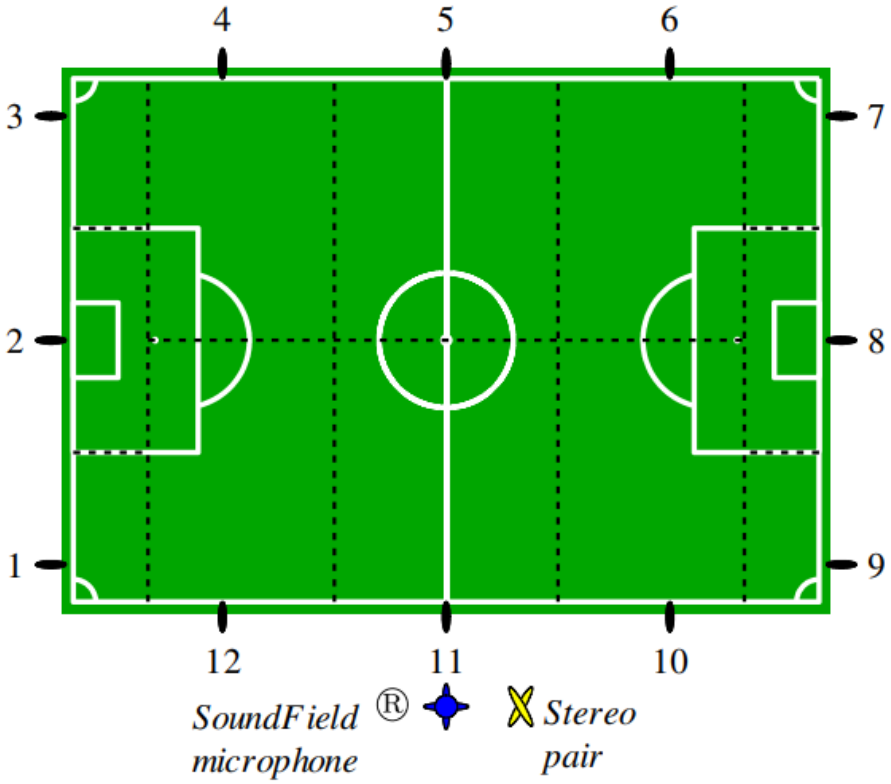
34% → **81%**

Improvement speech recognition in applause noise

[11] Paulus, Jouni, et al. "MPEG-D spatial audio object coding for dialogue enhancement (SAOC-DE)." Audio Engineering Society Convention 138. Audio Engineering Society, 2015.

Early Object-based approaches

FascinatE [12]



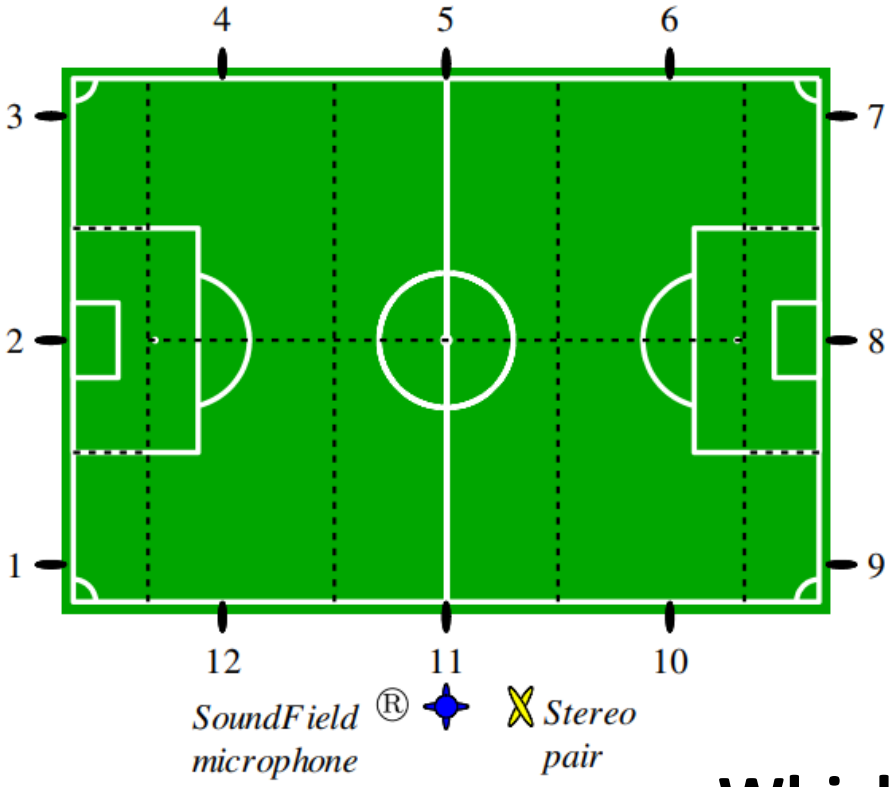
Individual user control of:

- on pitch sounds
- crowd level
- commentary

[12] Oldfield, Robert, Ben Shirley, and Jens Spille. "Object-based audio for interactive football broadcast." *Multimedia Tools and Applications* 74.8 (2015): 2717-2741.

Early Object-based approaches

FascinatE [12]



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?

[12] Oldfield, Robert, Ben Shirley, and Jens Spille. "Object-based audio for interactive football broadcast." *Multimedia Tools and Applications* 74.8 (2015): 2717-2741.

Intelligibility vs. Comprehension

Intelligibility: *Proportion of words correctly heard*

Speech
Useful

Non-speech
Masker

Intelligibility vs. Comprehension

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

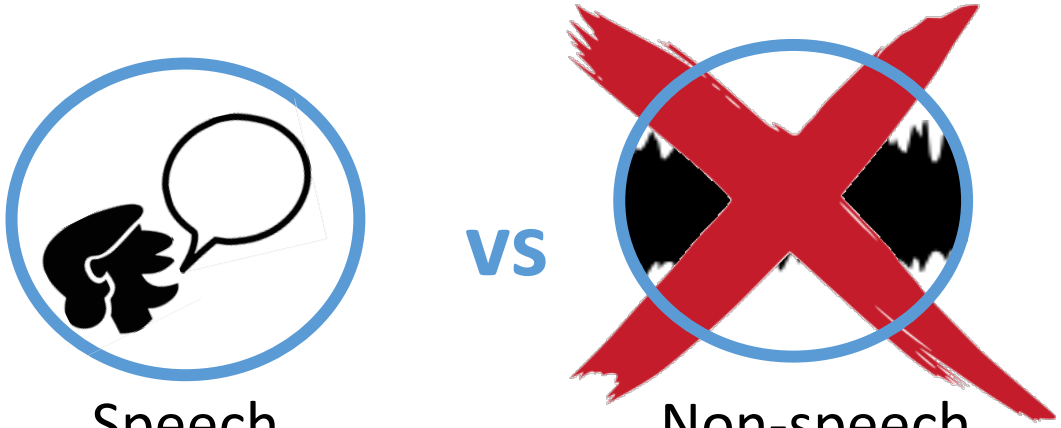
Speech
Useful

Non-speech
Masker

Comprehension: ***Proportion of content understood***

Intelligibility vs. Comprehension

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



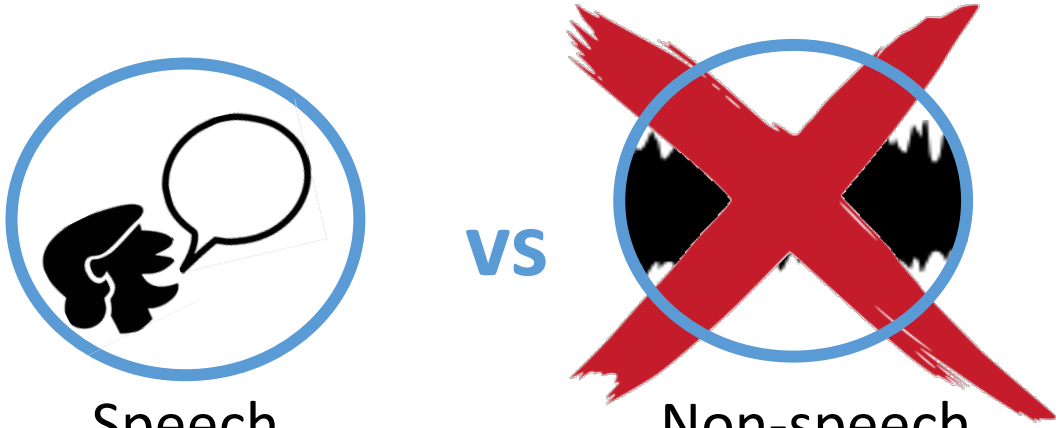
Speech
Useful

Non-speech
Masker

Comprehension: ***Proportion of content understood***

Intelligibility vs. Comprehension

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



Speech
Useful

vs

Non-speech
Masker

Signalling

Continuity

Narratively Important

Comprehension: ***Proportion of content understood***

Effect of non-speech sounds

Normal Hearing[13]

36%

Without

Sound Effects

[13] Ward, Lauren, et al. "The effect of situation-specific non-speech acoustic cues on the intelligibility of speech in noise." *Proc. Interspeech 2017* (2017): 2958-2962.

Effect of non-speech sounds

Normal Hearing[13]

36%

Without

Sound Effects

62%

With

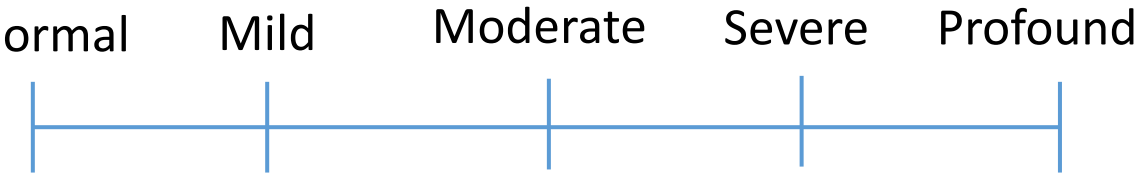
Sound Effects

[13] Ward, Lauren, et al. "The effect of situation-specific non-speech acoustic cues on the intelligibility of speech in noise." *Proc. Interspeech 2017* (2017): 2958-2962.

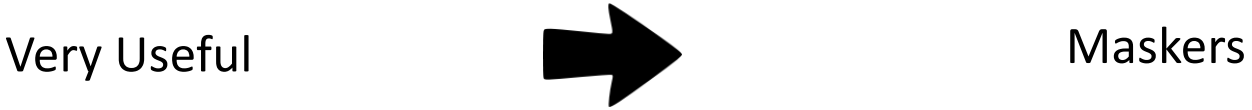
Effect of non-speech sounds

Hard of Hearing[14]

Audiometric Thresholds



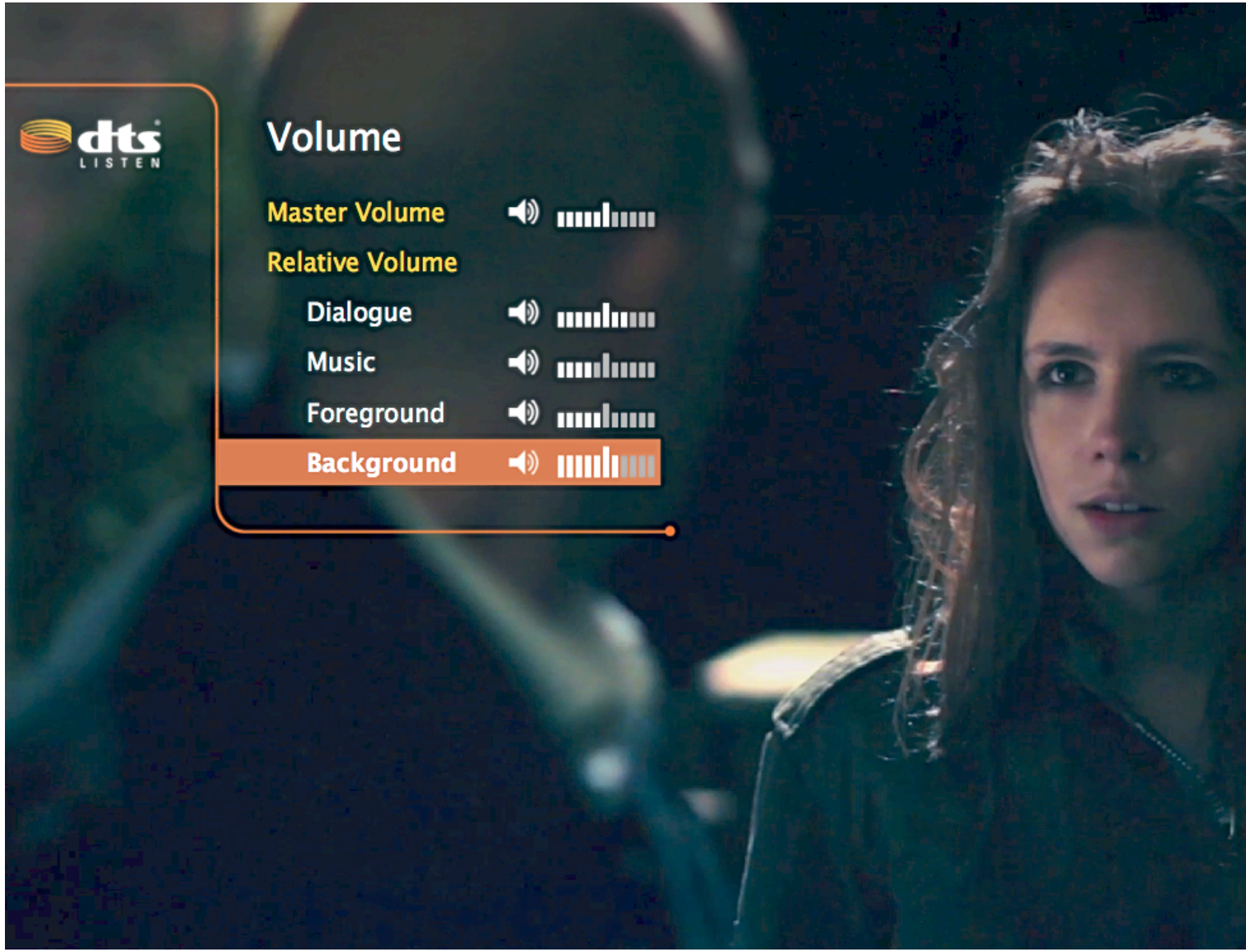
Usefulness of Sound Effects



[14] Ward, L, and Shirley, B.G. "Television Dialogue; Balancing Audibility, Attention and Accessibility." *Conf. on Accessibility in Film, Television and Interactive Media, York, UK*. 2017.

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



[15] Shirley, Ben Guy, et al. "Personalized object-based audio for hearing impaired TV viewers." *Journal of the Audio Engineering Society* 65.4 (2017): 293-303.

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

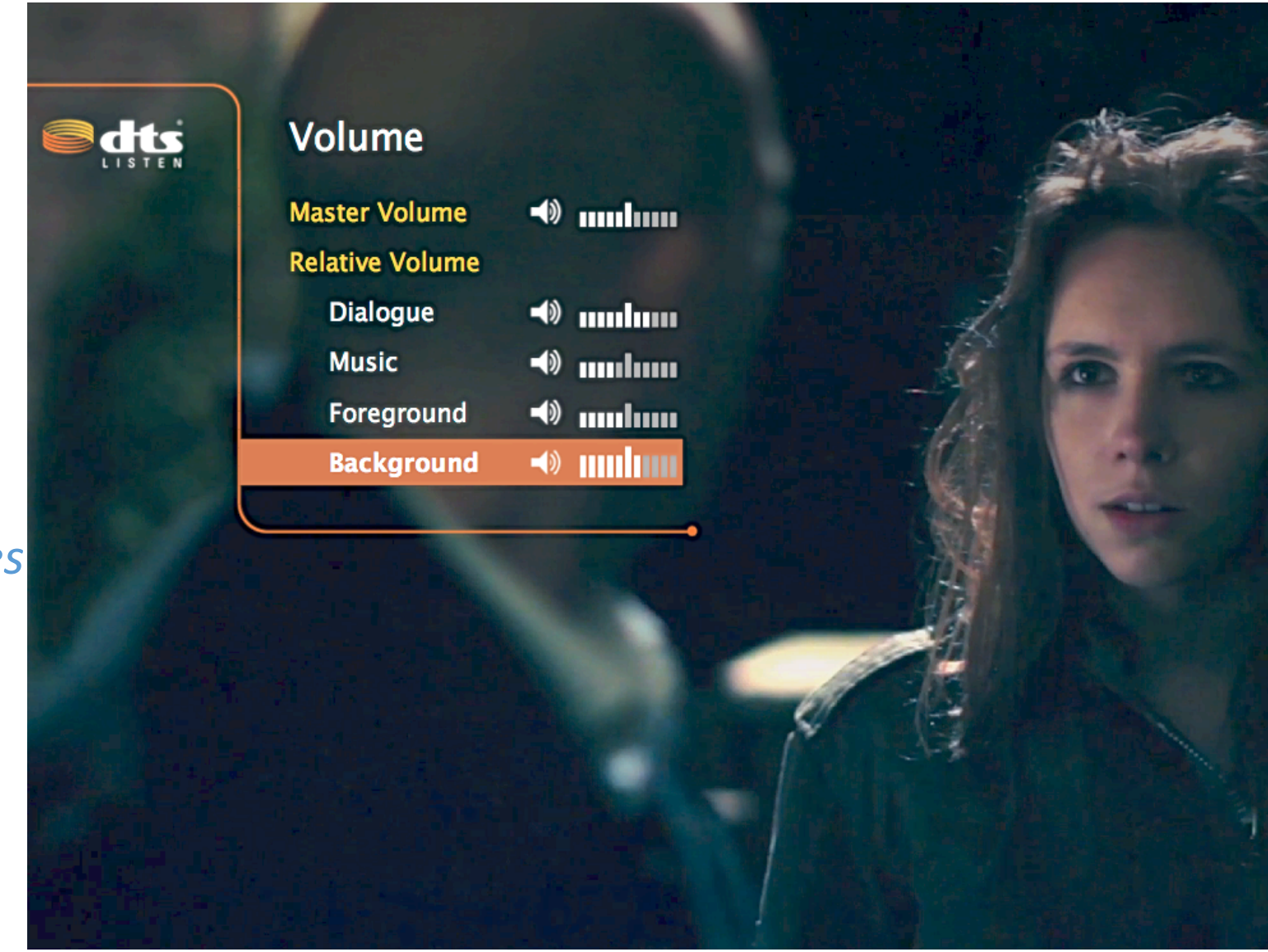
Exploration of user-preferences for audio-object categories volumes

Feedback from hard of hearing participants

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

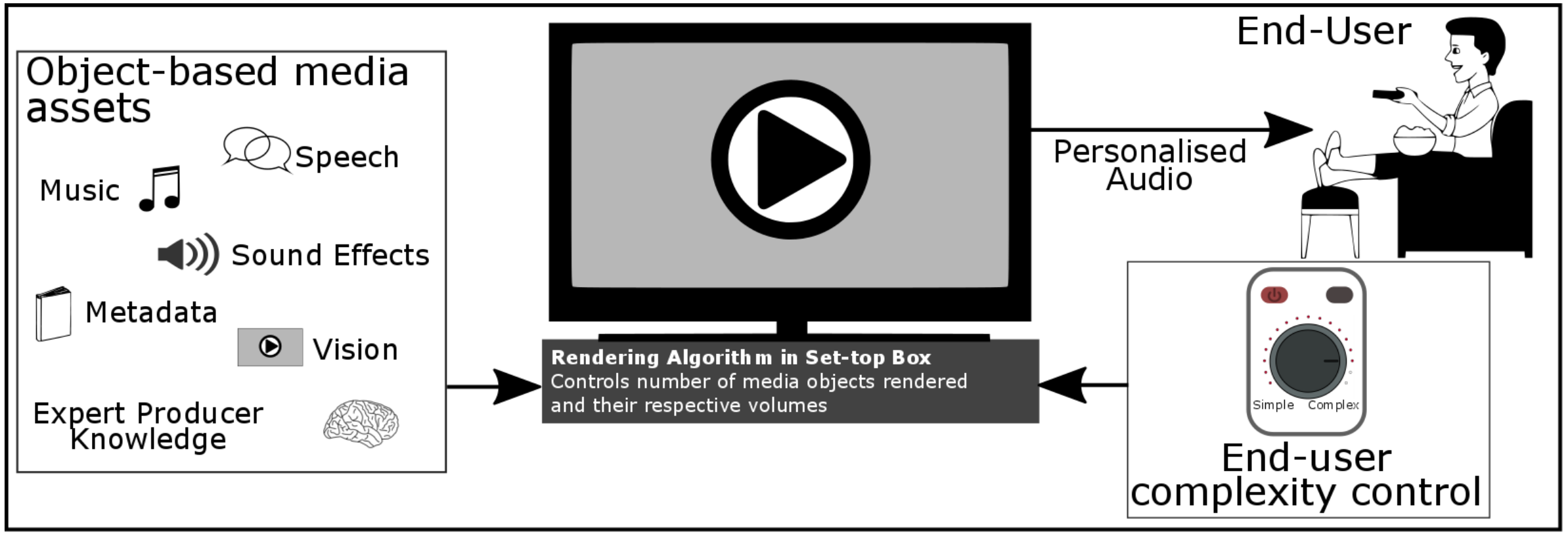
"useful and interesting to be able to adjust different aspects of sound"

"very straightforward, very good, when can I have one?"



[15] Shirley, Ben Guy, et al. "Personalized object-based audio for hearing impaired TV viewers." *Journal of the Audio Engineering Society* 65.4 (2017): 293-303.

Accessible Broadcast Audio Customisation



Assessing quality for new access services

It's complicated…

Assessing quality for new access services

It's complicated…

Object-based personalisation facilitates useful solutions but

…evaluation becomes a complex problem

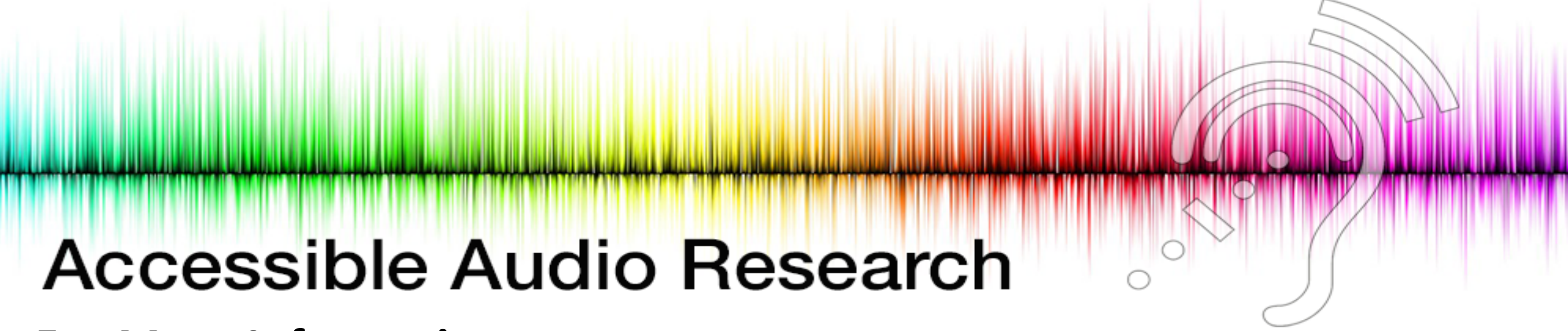
Assessing quality for new access services

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

University of Salford MANCHESTER