University of HUDDERSFIELD

# Phantom Image Elevation Explained

Hyunkook Lee

h.lee@hud.ac.uk

Applied Psychoacoustics Lab (APL)

University of Huddersfield, UK

EPSRC
Engineering and Physical Sciences Research Council

# Contents

- Previous studies on the phantom image elevation effect.
- Experiment 1: Frequency dependency of the effect.
- Experiment 2: The role of acoustic crosstalk on the effect.

# Previous Studies

# Previous studies

- de Boer (1947): Phantom centre image is perceived to be elevated, and the elevation angle increases as the loudspeaker base angle increases. (180° → overhead region)
- Also reported by Leakey (1959).
- Confirmed by Damaske and Mellert (1969/1970).

# Previous studies

University of HUDDERSFIELD

- Previous studies reporting the elevation effect were limited in terms of sound sources or loudspeaker base angles tested.

| | **Source** | **Base angles** |
|---|---|---|
| de Boer (1947) | Not reported | 0° to 180° |
| Leakey (1959) | Speech | No formal data |
| Damaske and Mellert (1969/1970) | White noise<br>0.65 – 4.5kHz | 0° to 360° |
| Frank (2014) | Pink noise<br>Broadband | 40° |
| Lee (2015) | Pink noise<br>Broadband, octave bands | 60° |

# Previous study (Lee 2015: AES139)

University of HUDDERSFIELD

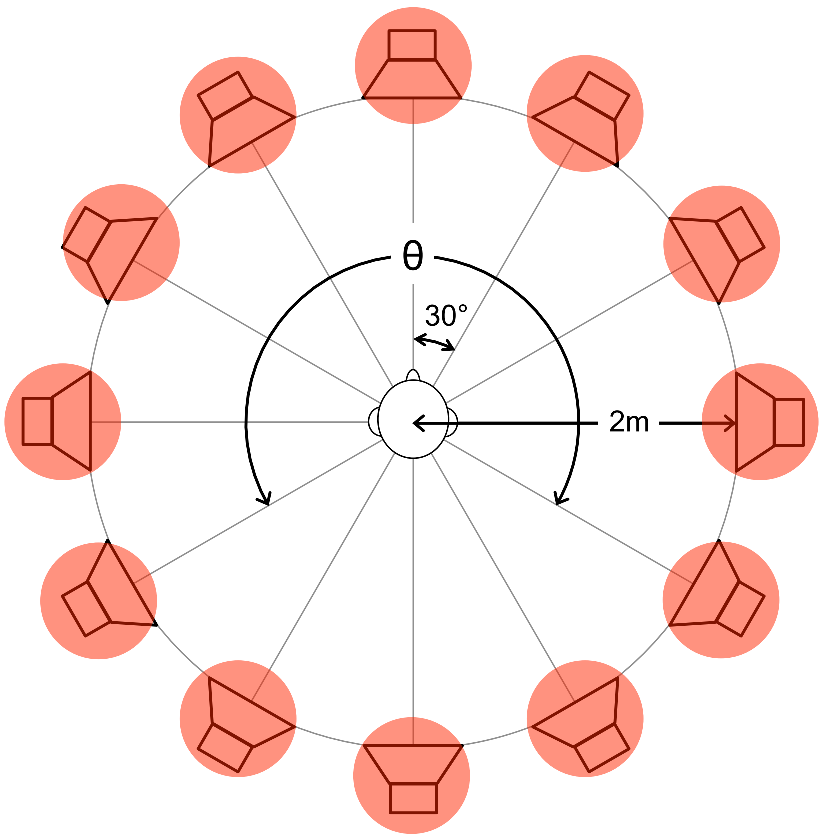
- In Lee (2015, AES139), the effect was investigated for a wide range of sound sources, with base angles covering from 0° to 360°.

- Sound sources tested
  - Speech, Helicopter, Aeroplane, Thunder, Rain, Bird, Church Bell
  - Broadband pink noises (continuous and transient)
  - Broadband white noises (continuous and transient)

# Previous study (Lee 2015: AES139)

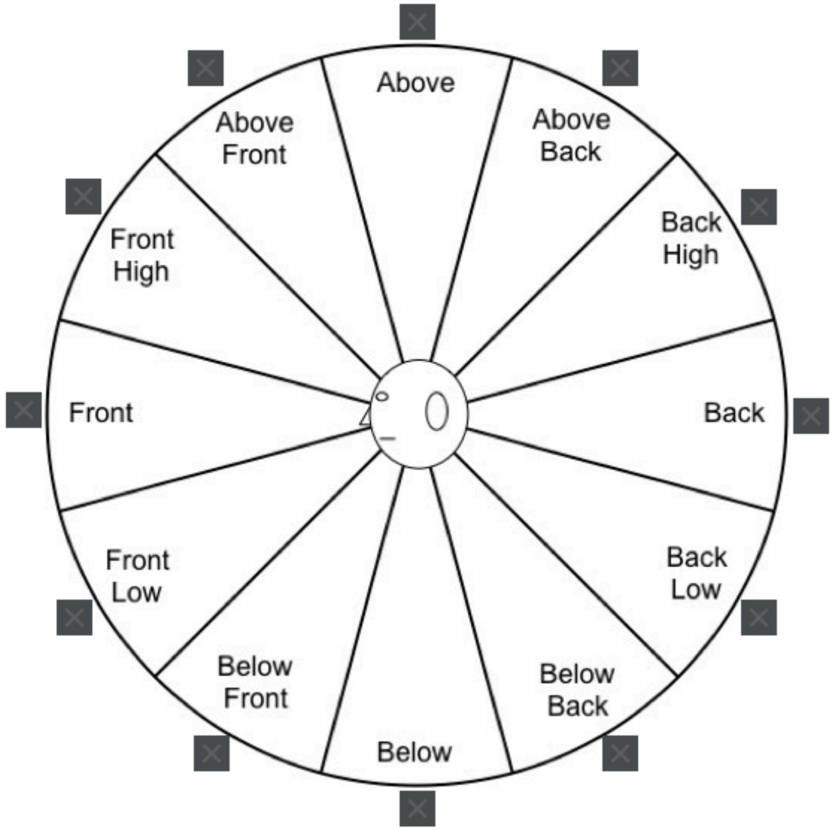
University of Huddersfield

- Loudspeaker arrangement
  - At the ear height in the horizontal plane, 0° to 360° with 30° intervals.



# Previous study (Lee 2015: AES139)

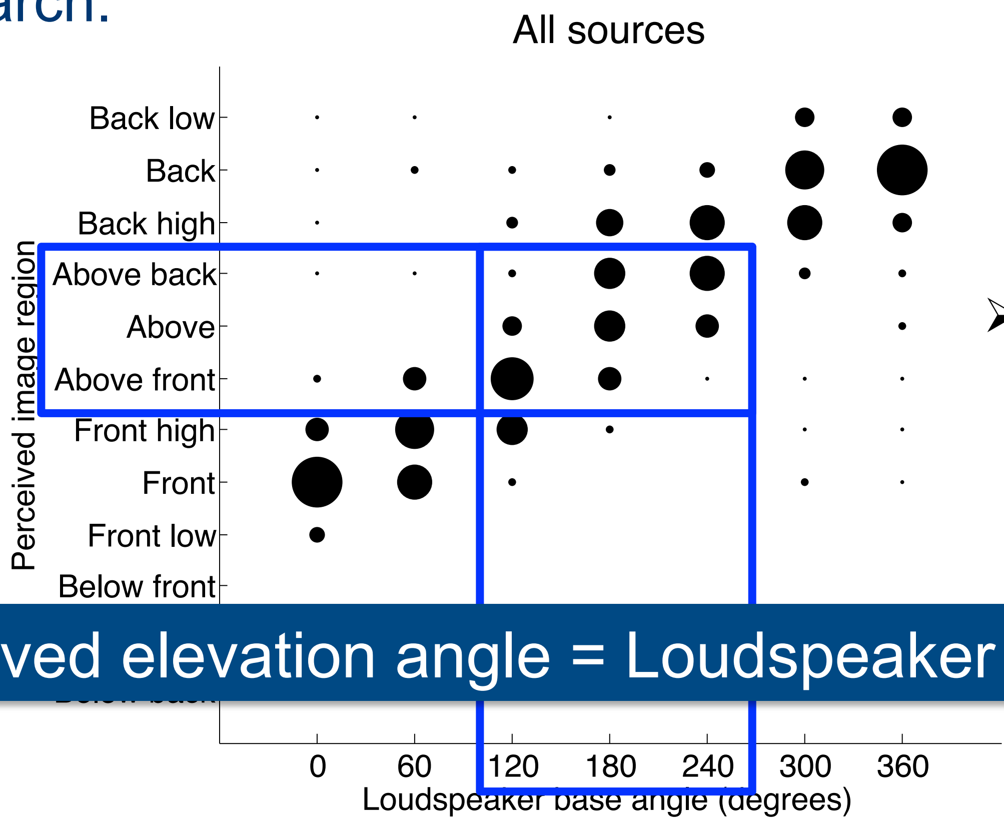
- GUI written in Max
  – Response method similar to what Blauert (1968) used in the Directional Band study (4 regions), but with a higher resolution (12 regions).



# Previous study (Lee 2015: AES139)

University of HUDDERSFIELD

- Responses for all sources
  - The general trend agrees with the suggestions from the past research.



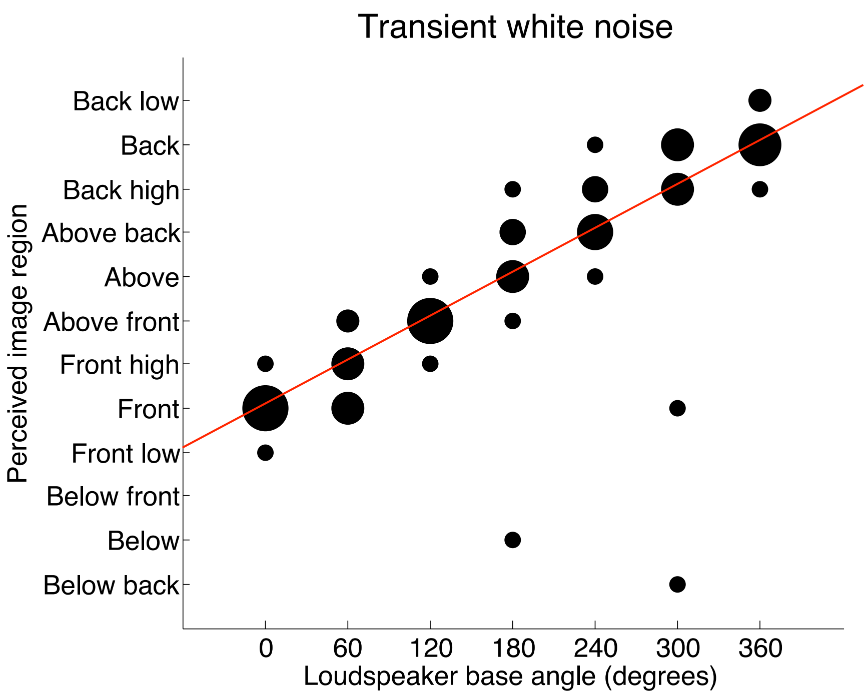

➢ Above perception for 120°, 180° and 240°

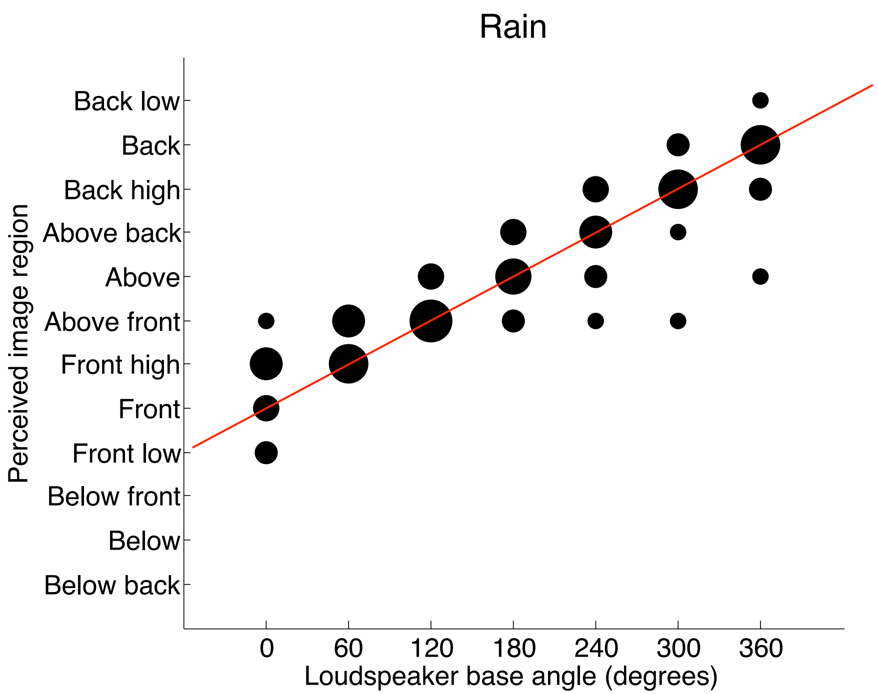
**Perceived elevation angle = Loudspeaker base angle / 2**

# Previous study (Lee 2015: AES139)

University of Huddersfield

- A significant source dependency was found.
  - Responses were most linear and consistent for sources with a broad and flat spectrum.

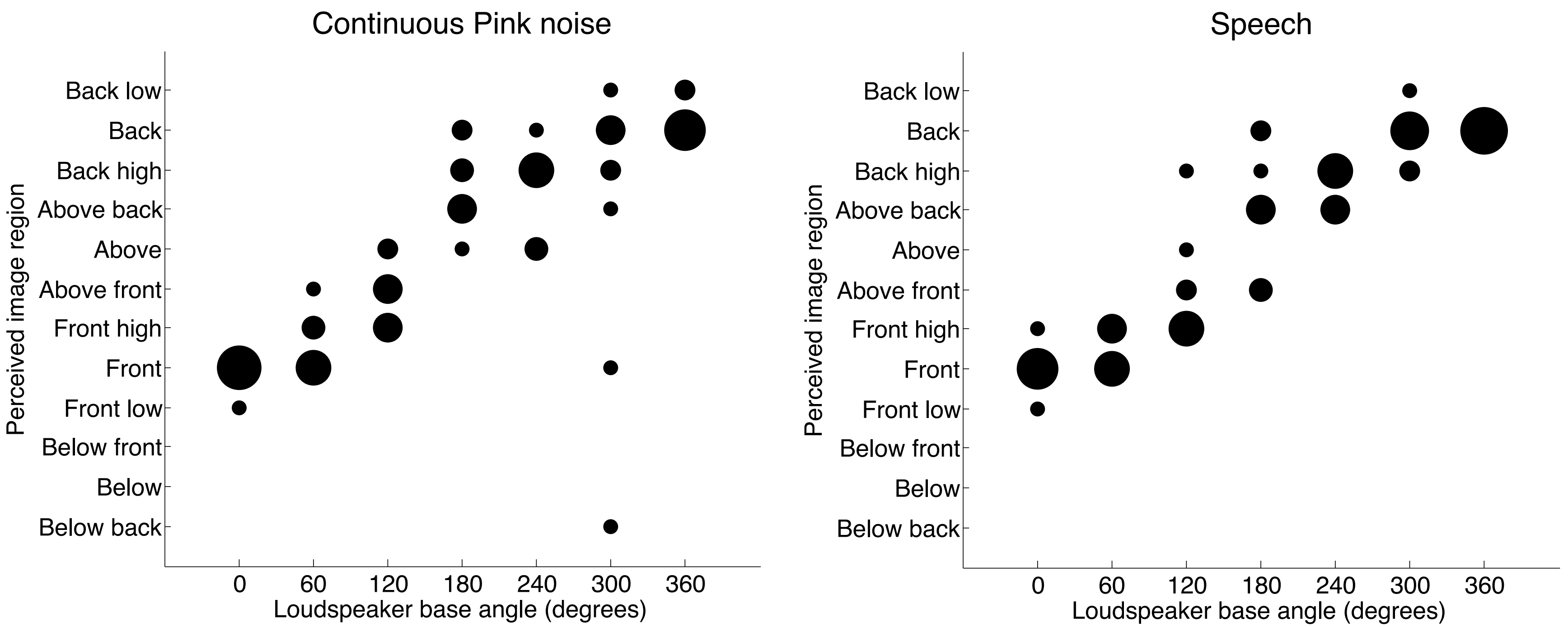


# Previous study (Lee 2015: AES139)

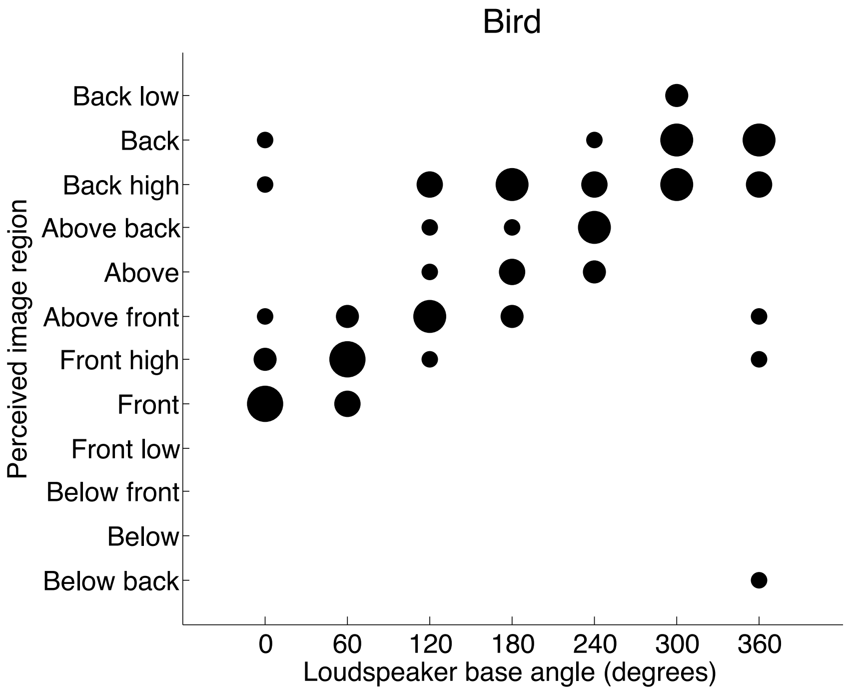
University of HUDDERSFIELD

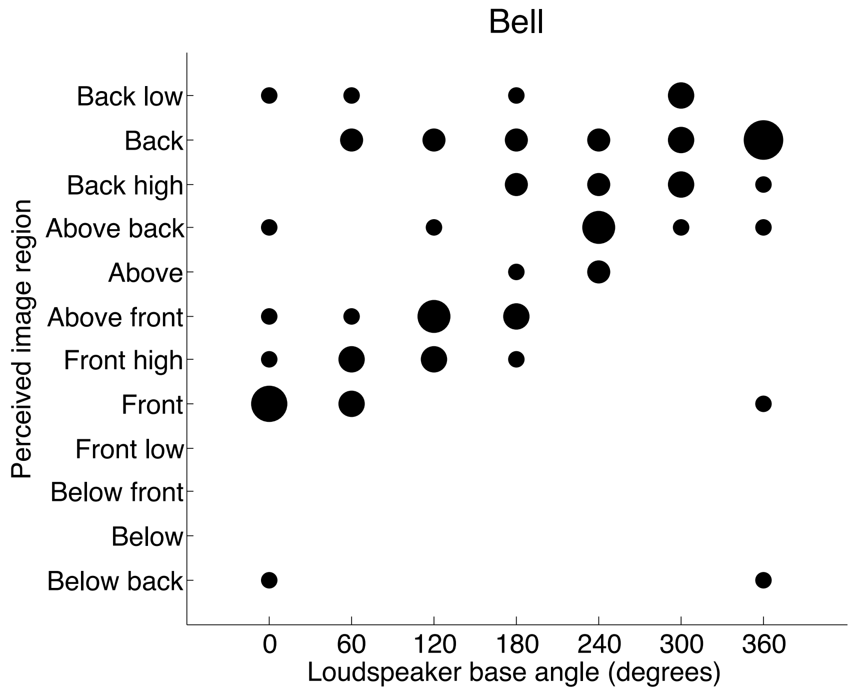
- A significant source dependency was found.
  - The elevation effect was weaker for sources with more low frequency dominance. (no strong directly above perception)



# Previous study (Lee 2015: AES139)

- A significant source dependency was found.
  - Responses were most inconsistent for sources with narrow spectrum or steady-state nature.

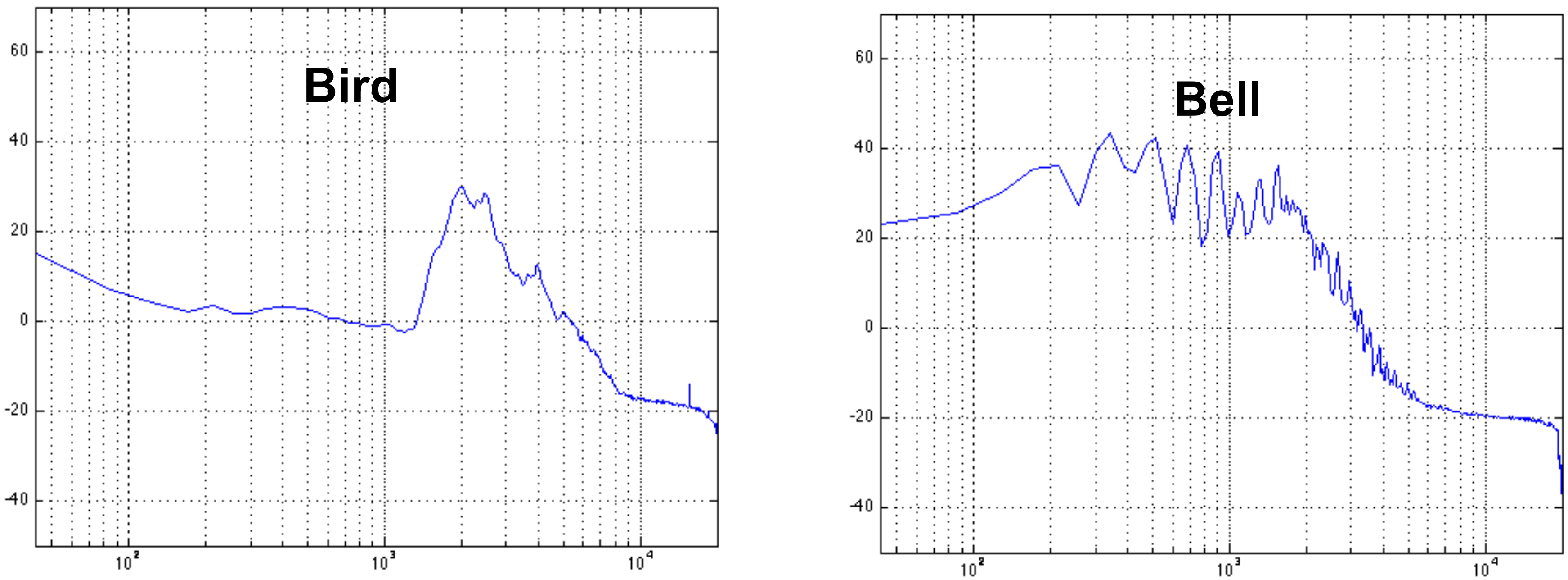


# Previous study (Lee 2015: AES139)

University of Huddersfield

- A significant source dependency was found.
  - Responses were most inconsistent for sources with narrow spectrum or steady-state nature.



# Experiment 1:
# Frequency Dependency of
# the Phantom Image Elevation Effect

# Experiment 1

- To investigate which octave frequency bands are most responsible for the phantom elevation effect.
  - Especially for the directly “above” perception with the 180° loudspeaker base angle.

- Octave band and broadband pink noise bursts stimuli
  - 63Hz, 125Hz, …, 16kHz.
  - $16^{th}$ order linear phase Butterworth filter.
  - 5ms onset/offset; 200ms ongoing; 500ms intervals.

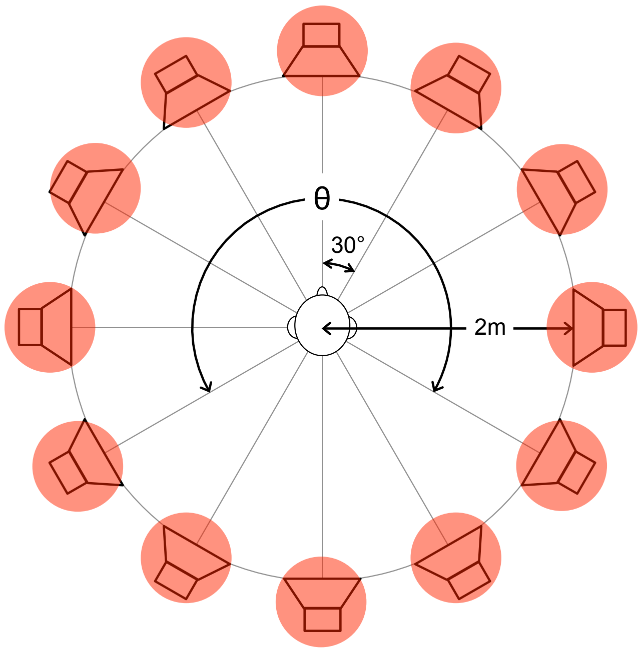
- 20 subjects
  - Music technology students and spatial audio researchers.
  - Experienced in spatial audio evaluation, but not trained for this task.

# Experiment 1

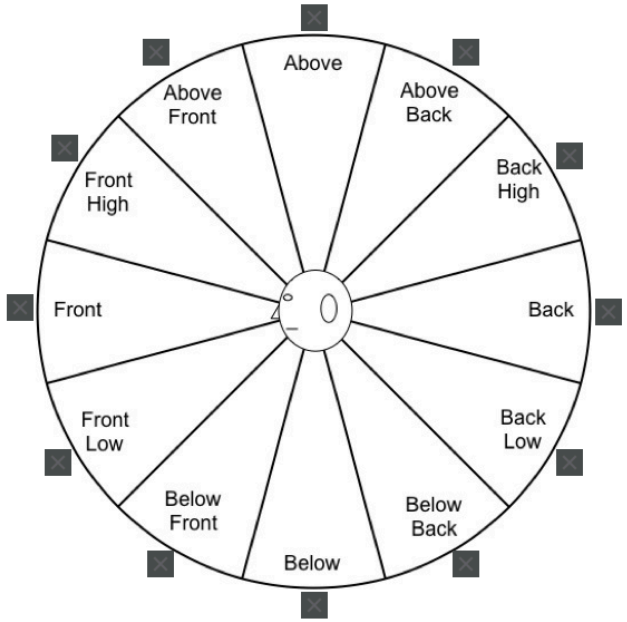
University of HUDDERSFIELD

- Phantom centre images with 7 loudspeaker base angles.
- 12 Genelec 8040As horizontally arranged in a circle.
- Head rotation was strictly now allowed.

Loudspeaker setup



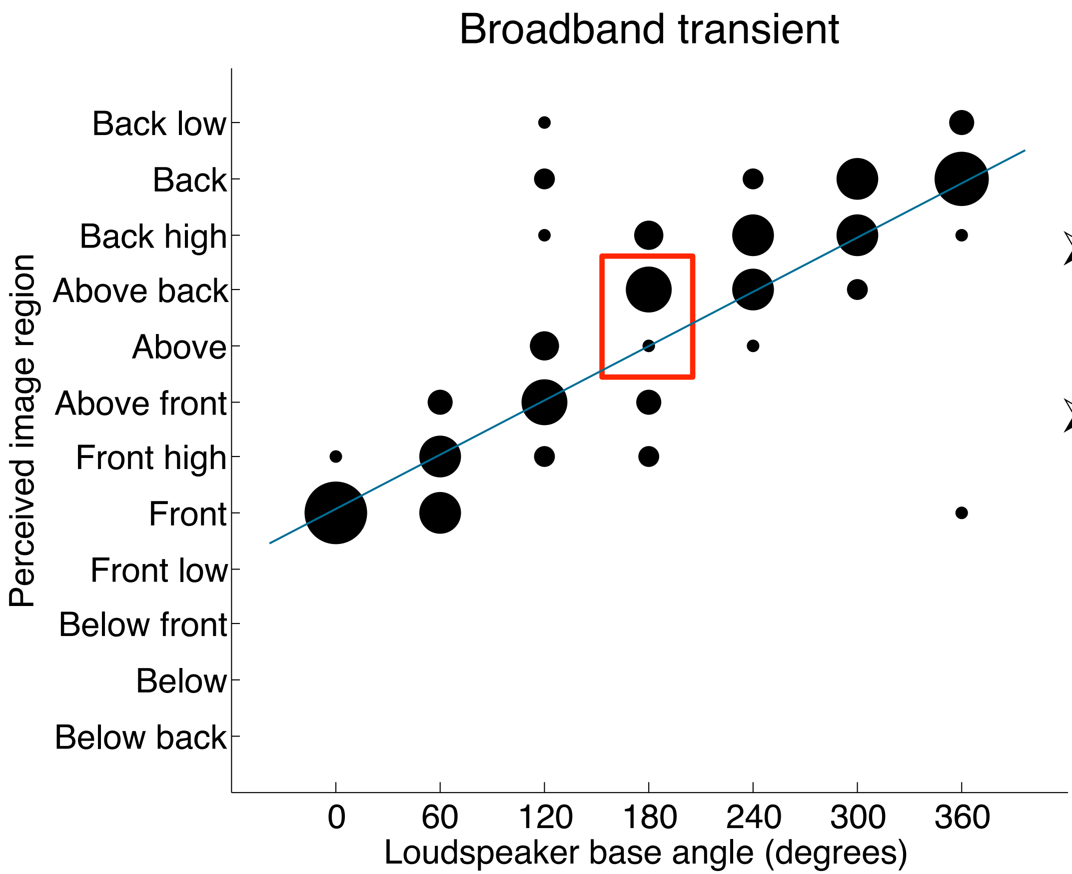

Response method



# Results

University of HUDDERSFIELD

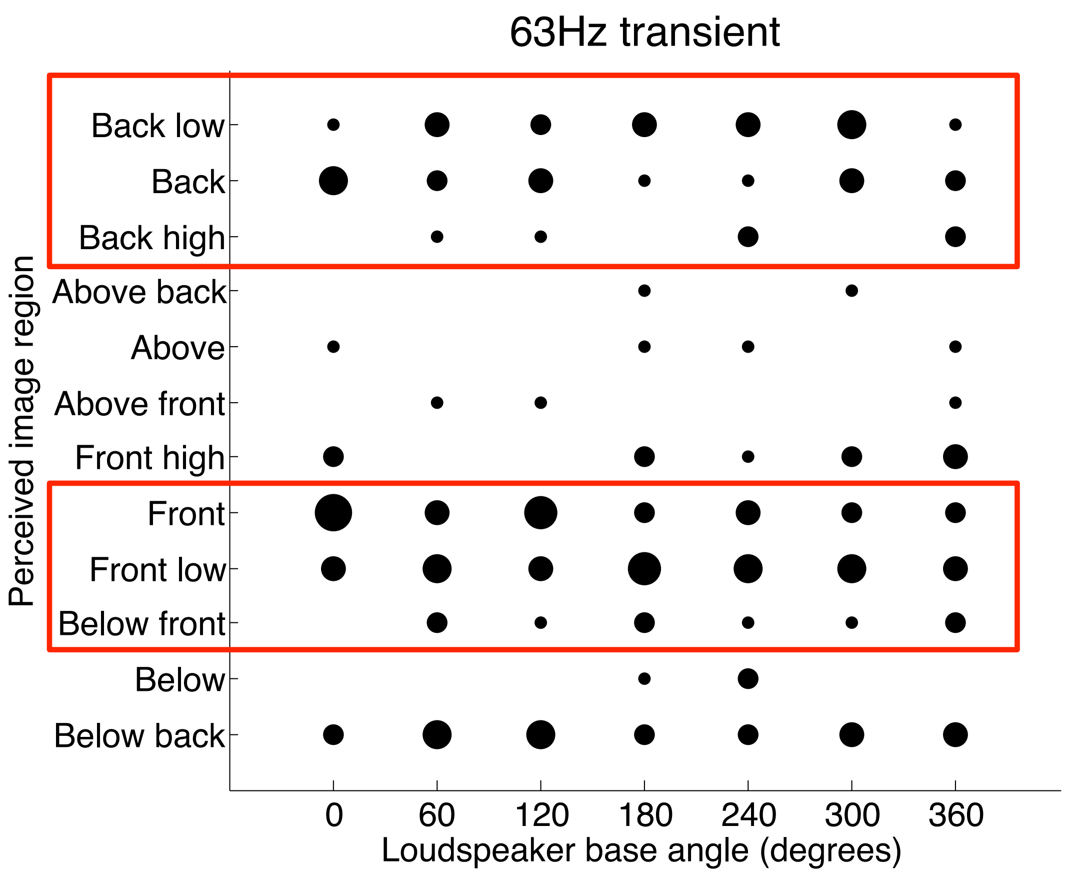
- ## Broadband




- A linear pattern in general.
- “above back” most dominant for 180°

University of HUDDERSFIELD

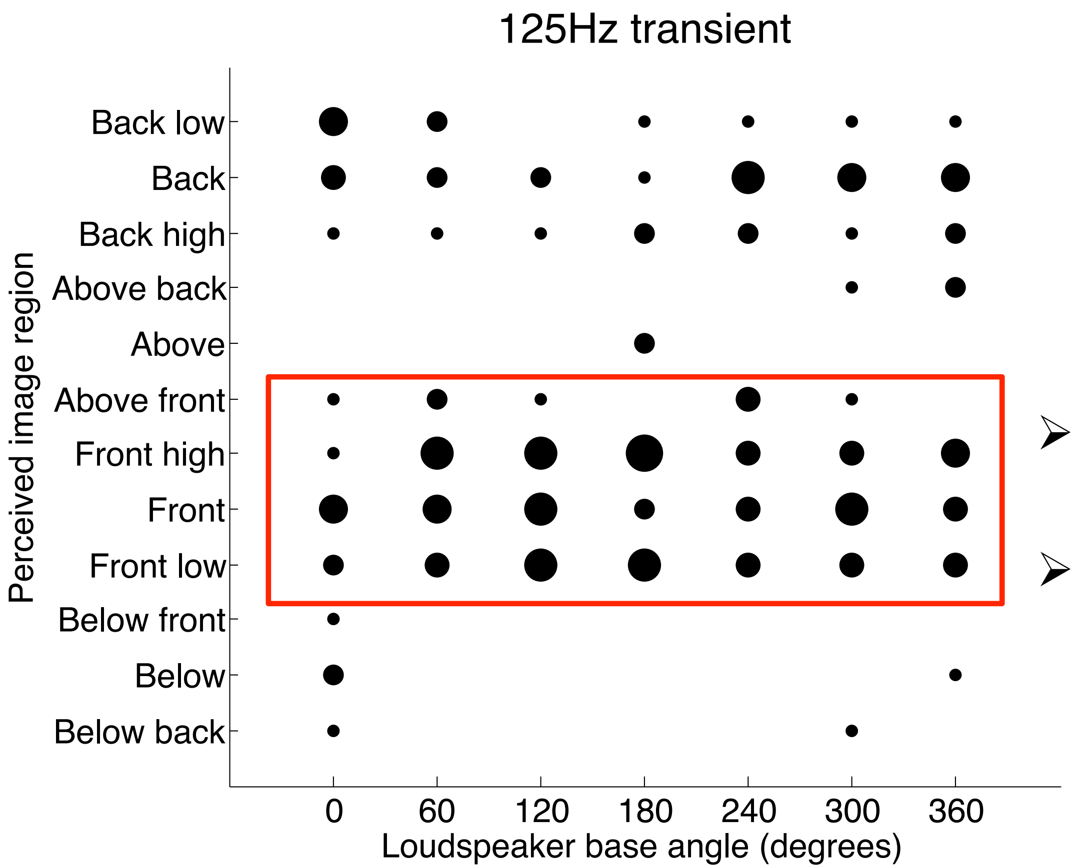
- 63Hz band




- Front-Back confusion
- Pitch-Height effect

# Results

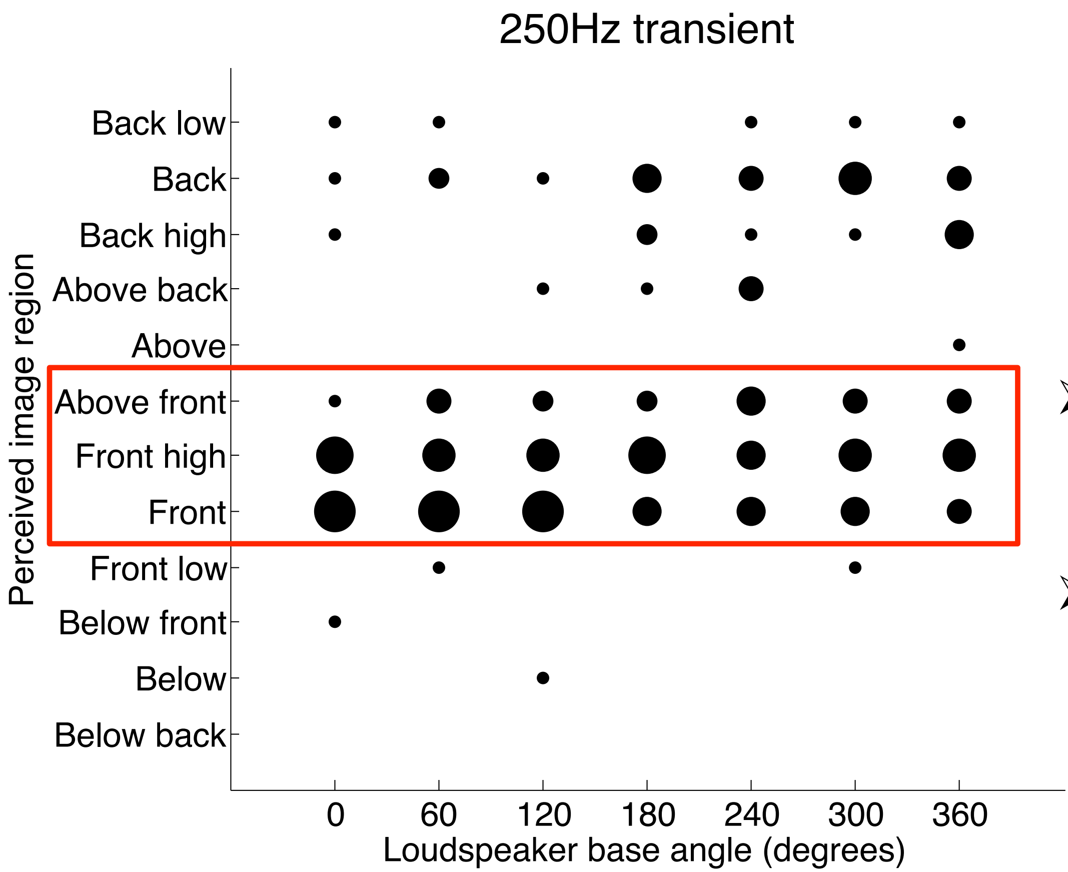
- 125Hz band




- Some pitch-height effect
- Some elevation effect

# Results

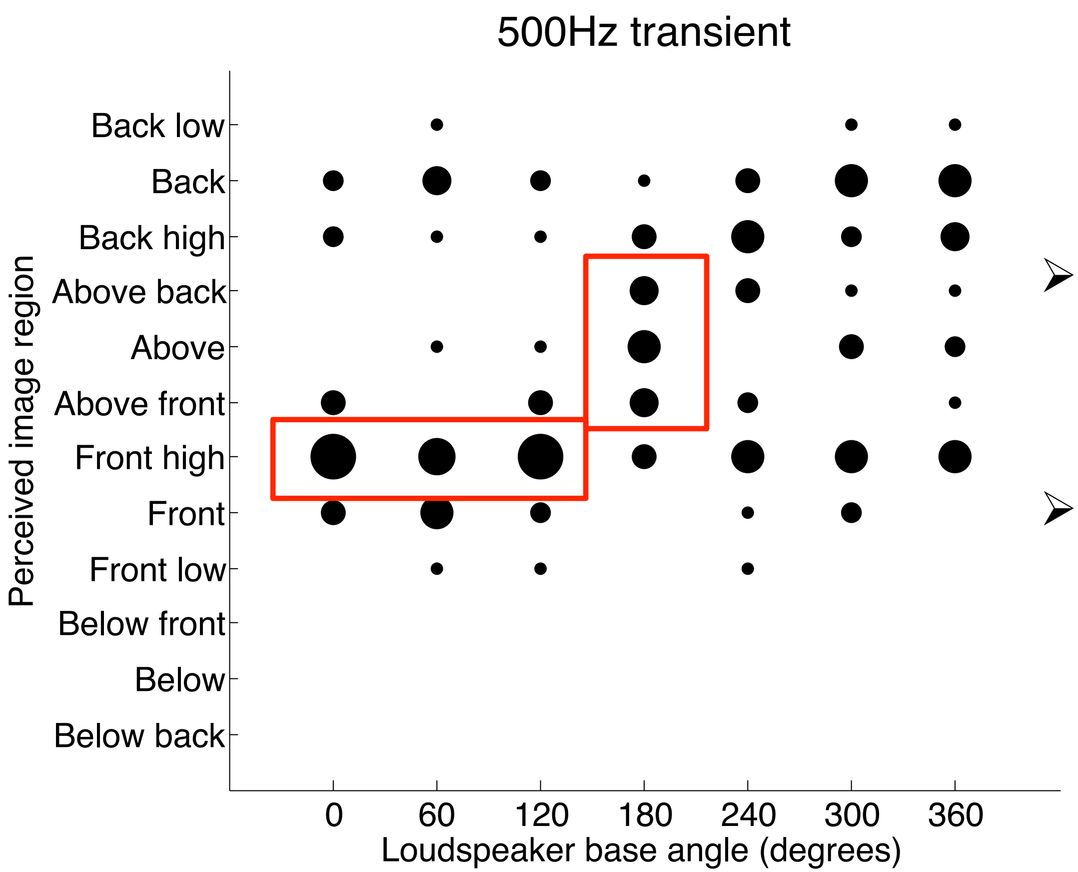
- 250Hz band




- ➢ Stronger elevation effect than 125Hz.
- ➢ No direct "above" for 180°.

# Results

- 500Hz band

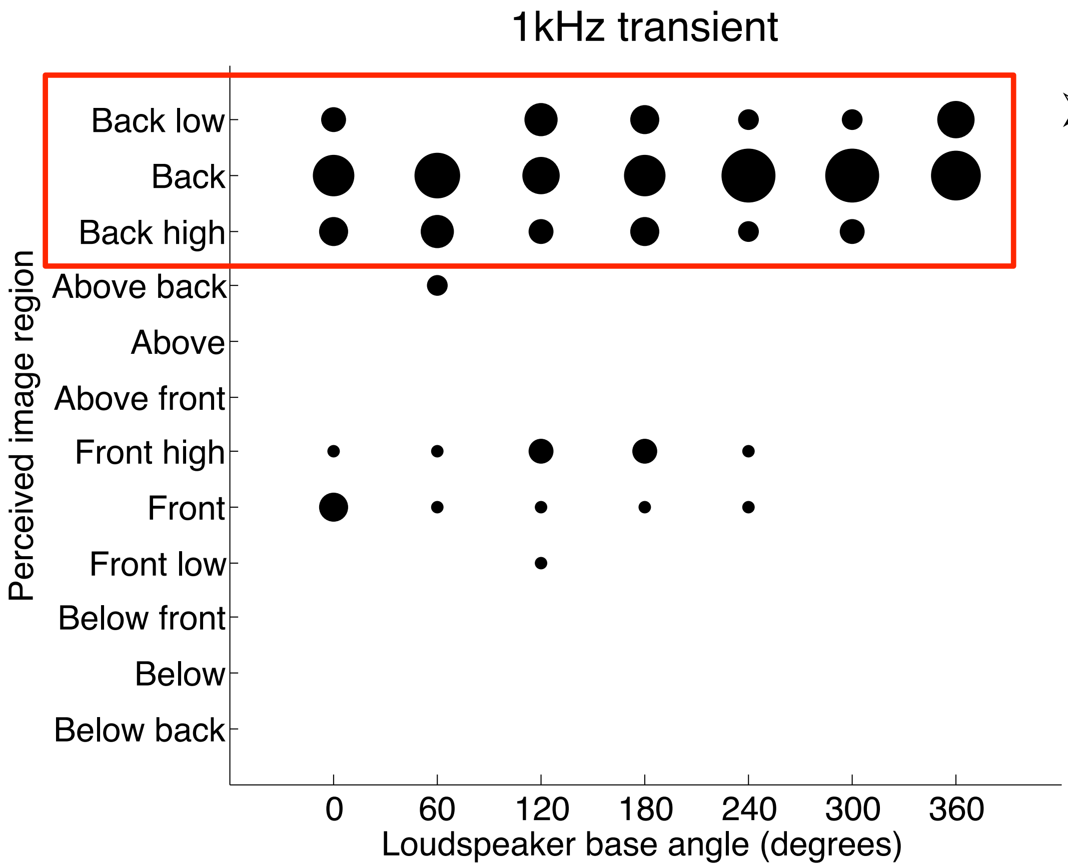



- Stronger “front high” perception for 0°, 60° and 120°.
- Strong ”above” perception for 180°.

# Results

University of HUDDERSFIELD

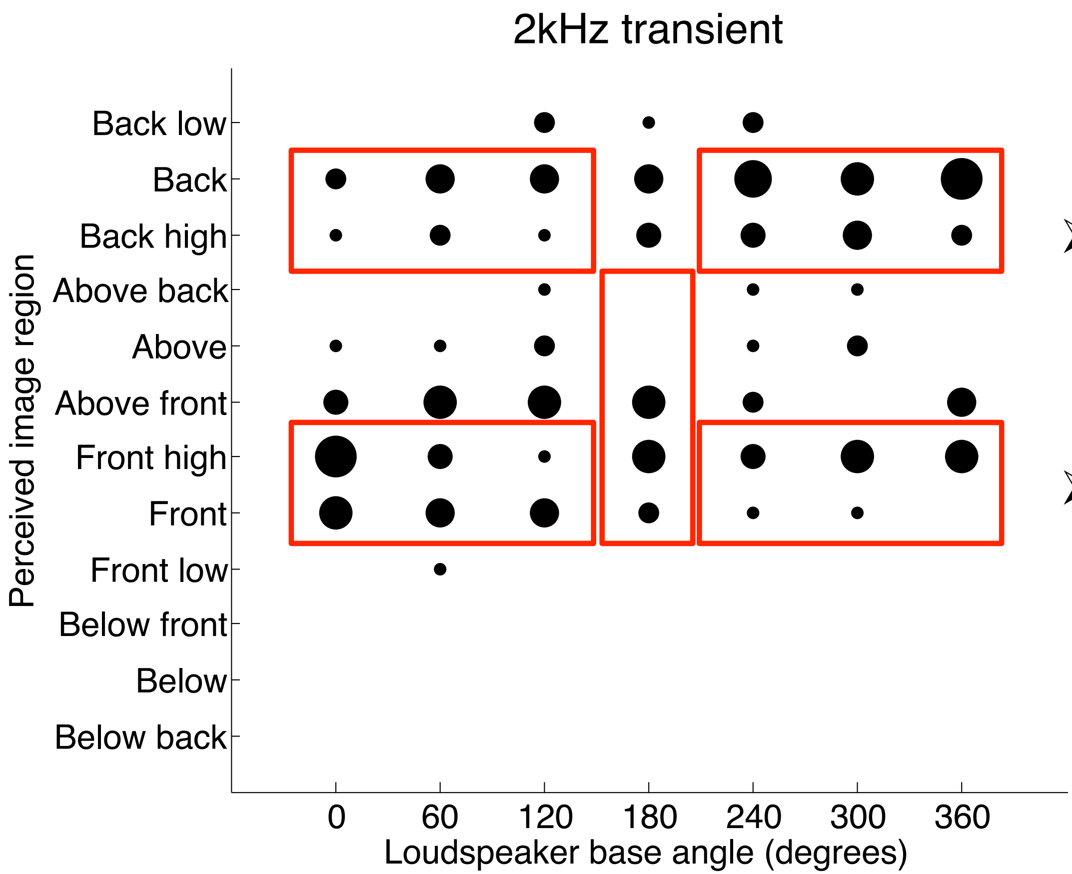
- 1kHz band




- Strong “Back” perception → Supports Blauert’s Directional Band theory.

# Results

- 2kHz band

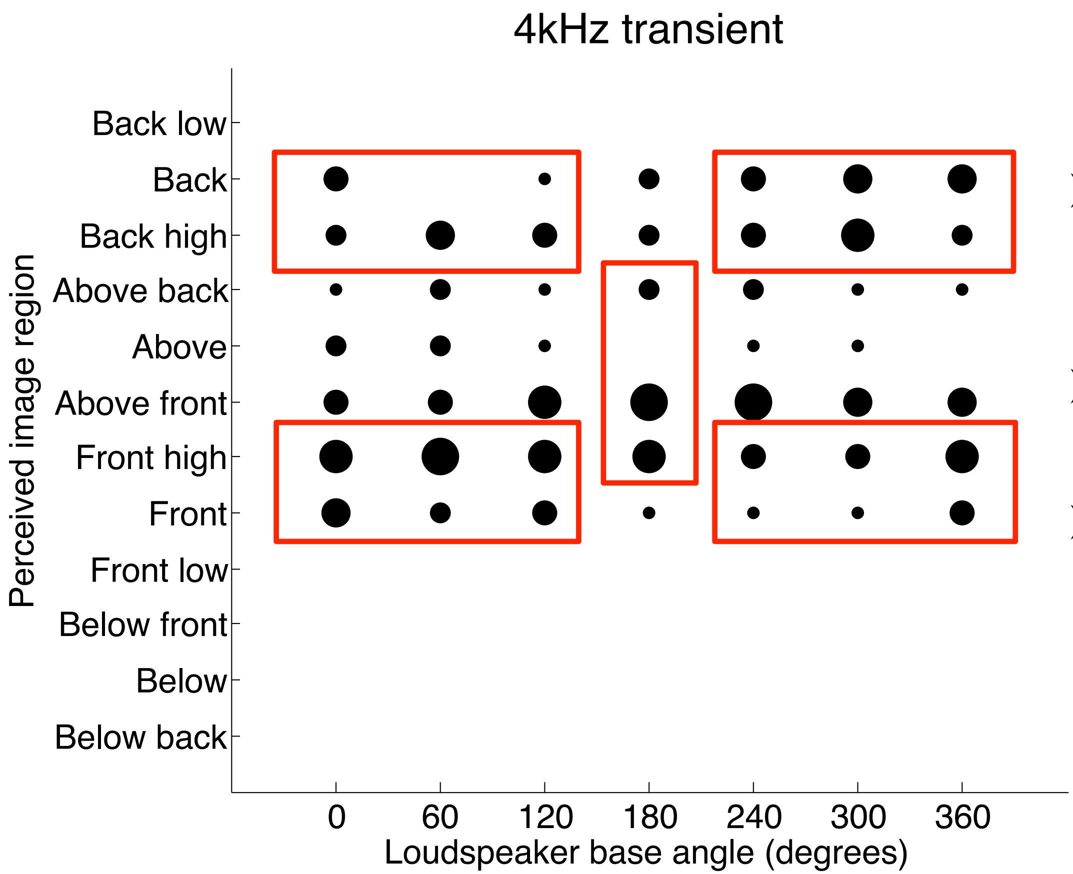



➢ Mostly “above front” and “front high” but no directly “above” for 180°

➢ Some Front-Back confusion

# Results

- 4kHz band

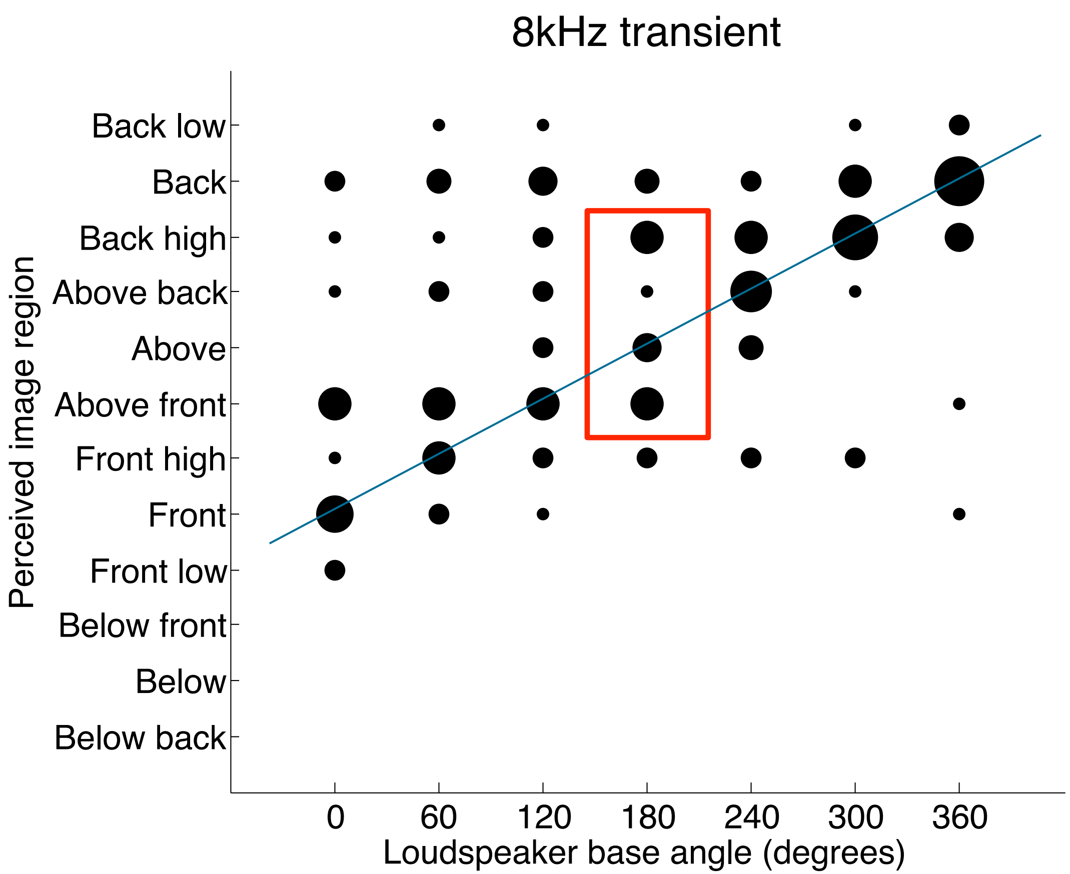



- Strong “above front” and “front high” for 180°
- No direct “above” for 180°.
- Some Front-Back confusion

# Results

University of HUDDERSFIELD

- 8kHz band

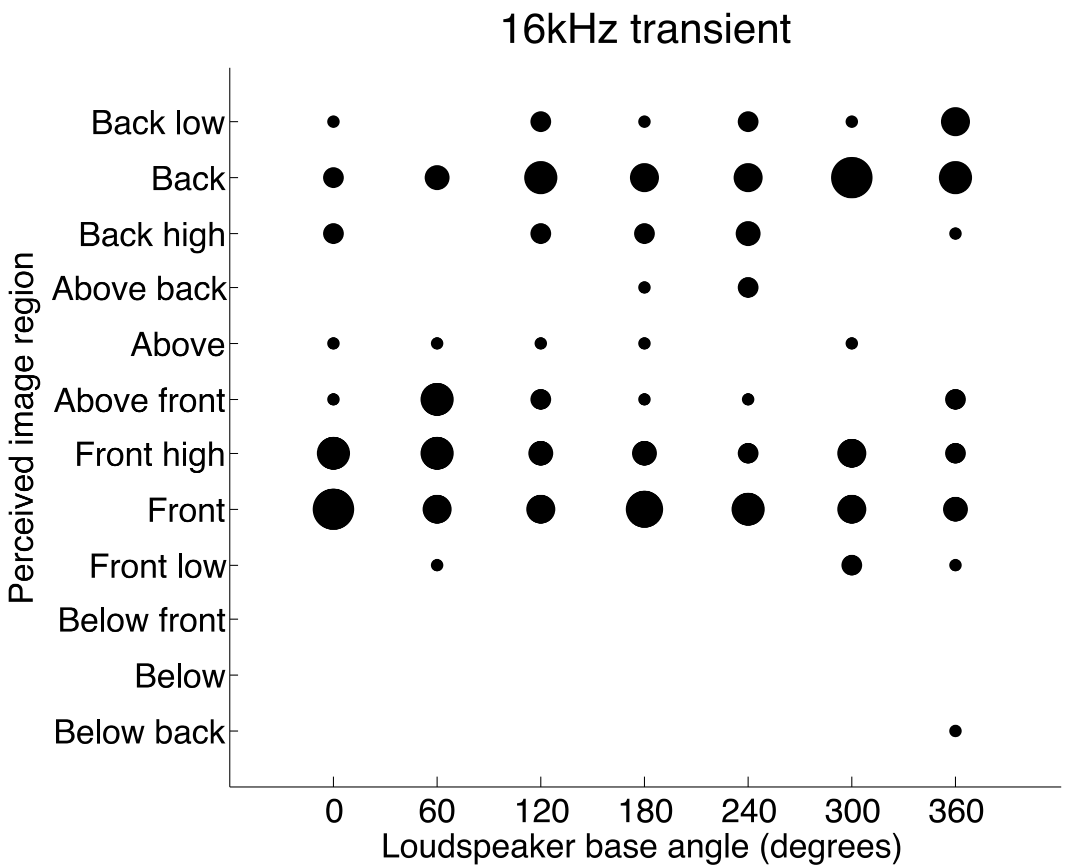



- Generally had the most linear pattern among all bands, but still with large spreads.
- “Above front” & “above”, strong “back high” for 180°.

# Results



- 16kHz band




- No strong elevation in general.
- Front-Back confusion.

University of HUDDERSFIELD

# Experiment 2:
# The Role of Acoustic Crosstalk for the Phantom Image Elevation Effect

# Previous explanations

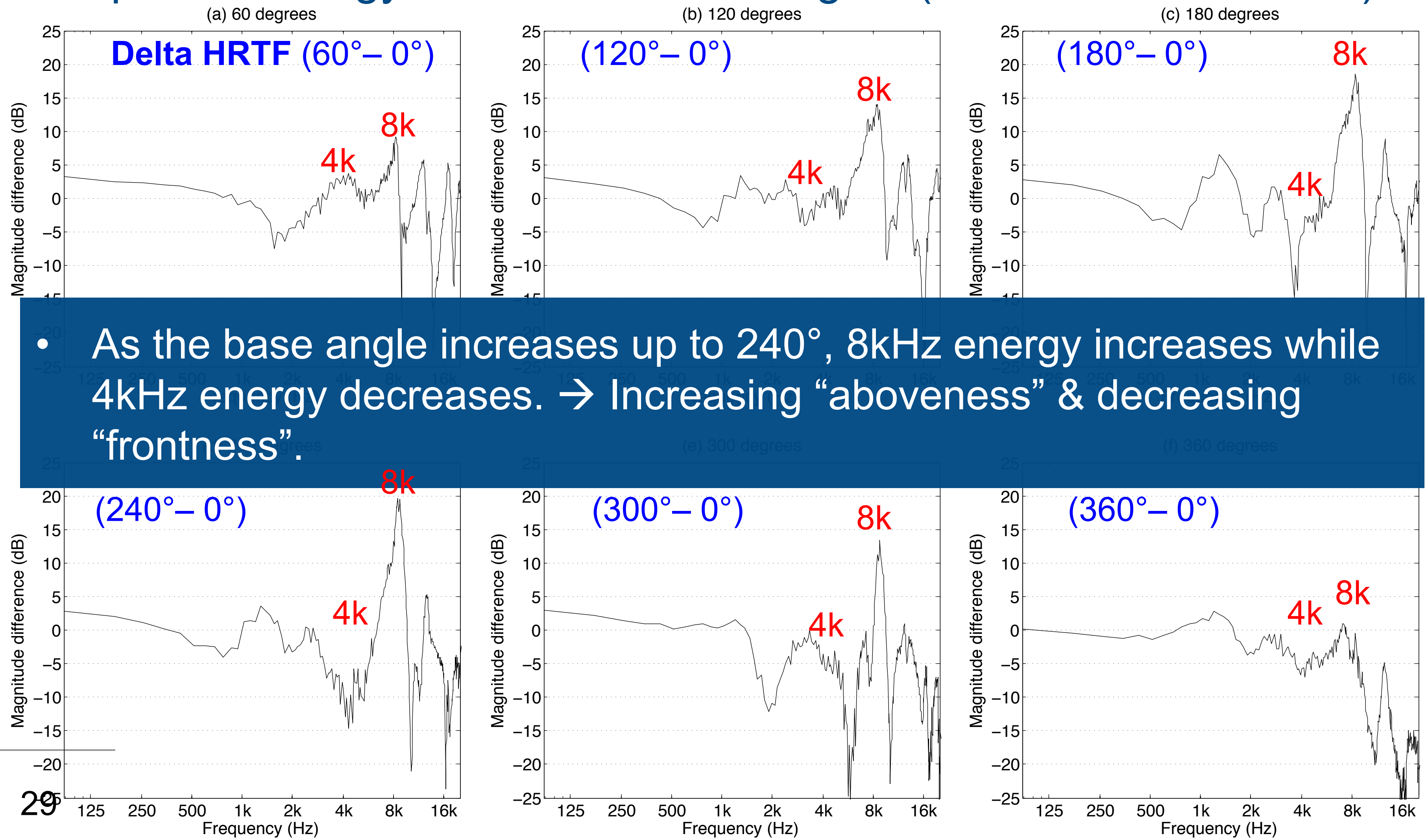
- de Boer (1947), Leakey (1959)
  - ITD matching between a real elevated source and a horizontal phantom source when rotating the head.
  - But the effect is perceived “without” head rotation.

- Blauert (1997)
  - Spectral energy distribution of ear input signals
  - Directional bands (8kHz for ‘aboveness’ and 4kHz for ‘frontness’)

# Previous explanations

University of HUDDERSFIELD

- Spectral energy distribution of ear signal (Phantom minus Real)




- As the base angle increases up to 240°, 8kHz energy increases while 4kHz energy decreases. → Increasing "aboveness" & decreasing "frontness".

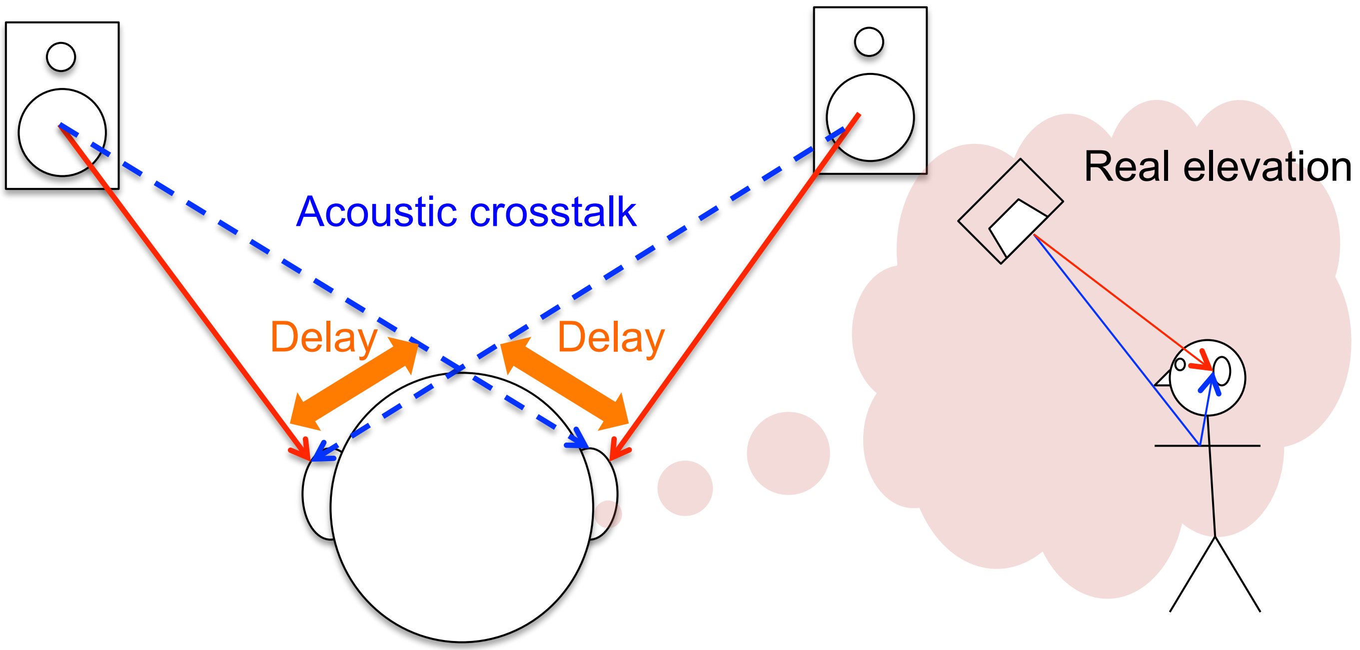
# Previous explanations



- However, spectral energy distribution does not explain the phantom image elevation for **low frequencies**.
  - E.g. 250Hz and 500Hz bands.

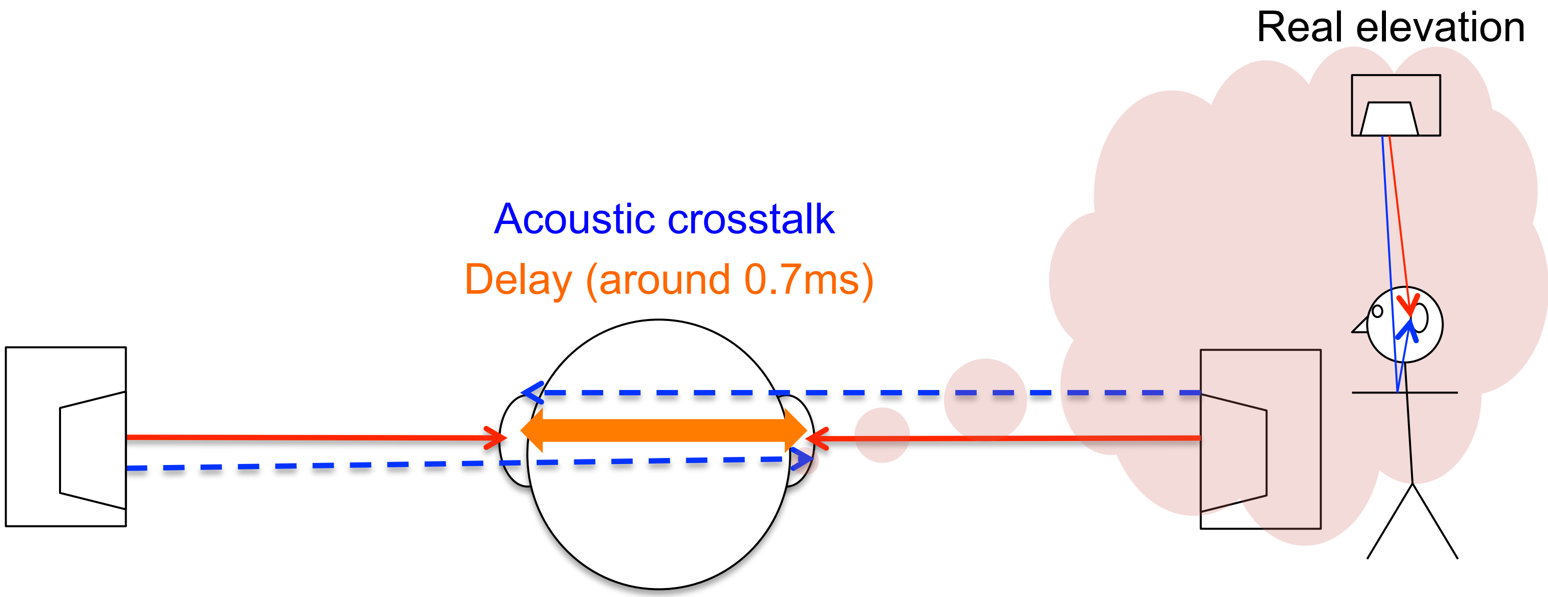
# New hypothesis

- A new explanation from a **cognitive** viewpoint (Lee 2015)
  - The brain interprets the head shadowed **acoustic crosstalk** as a torso reflection for a real elevated source.
  - **Below 3kHz, torso reflection delay** contributes to HRTF (Algazi et al. 2001)



# New hypothesis

- A new explanation from a **cognitive** viewpoint (Lee 2015)
  - As the **loudspeaker base angle** increases, **acoustic crosstalk delay** increases (max. around 0.7ms for 180°)
  - As the **real source elevation angle** increases, **torso reflection delay** increases (max. around 0.7ms for a source right above).



# Experiment 2

University of HUDDERSFIELD

- The role of acoustic crosstalk was investigated in terms of
  - Frequency range
  - Delay time
- Binaural simulation with individual BRIRs.
  - 5 subjects from Experiment 1.
  - Individual BRIRs were captured in the ITU-R BS1116 room.
  - Each test was repeated 10 times for each subject.
- 5 different sound sources
  - Rain & Thunder.
  - White noise burst, 500Hz octave band & 8kHz octave band.

# Experiment 2

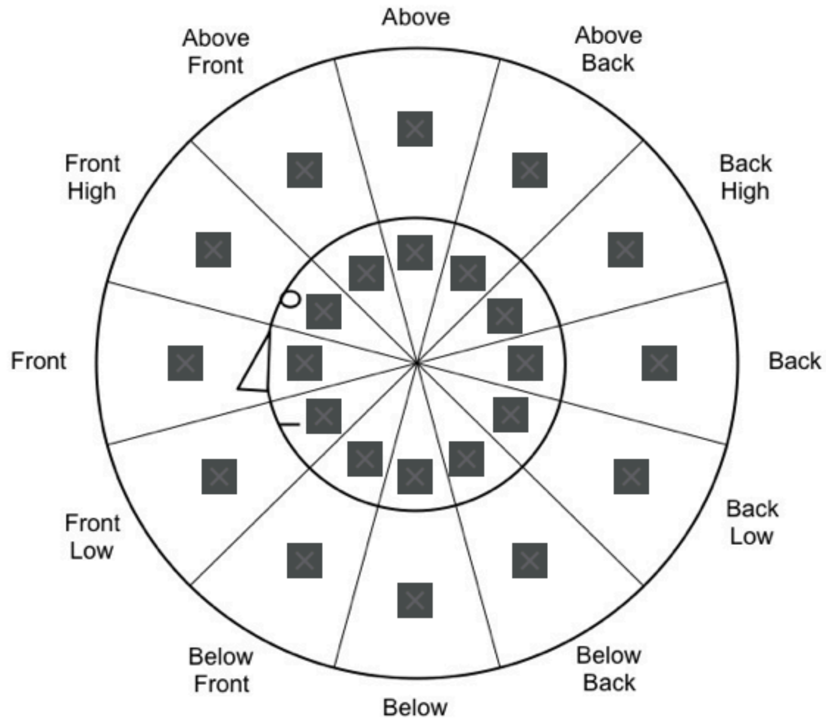
University of HUDDERSFIELD

- Comparing 5 different conditions.

  - XT on = Full binaural rendering of the 180° base angle condition.
  - XT off = With the interaural crosstalk completely removed.
  - XT 3k LPF = With the crosstalk low-pass filtered at 3kHz.
  - XT 3k HPF = With the crosstalk high-pass filtered at 3kHz.
  - XT 0ms = The crosstalk delay made as 0ms.

# Experiment 2

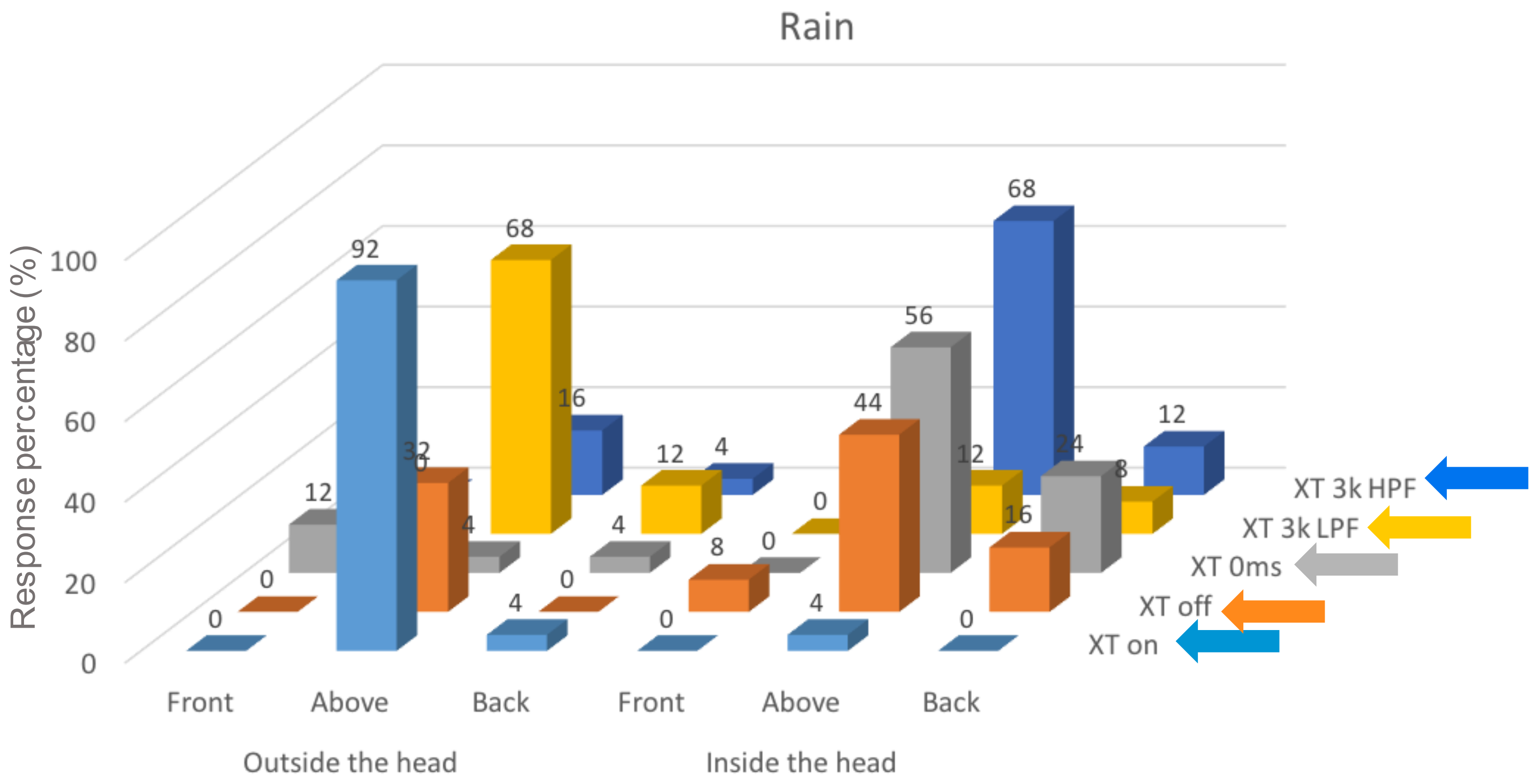
- Response method
  - Elevated position in the median plane.
  - Outside-the-head vs. Inside-the-head.



# Results



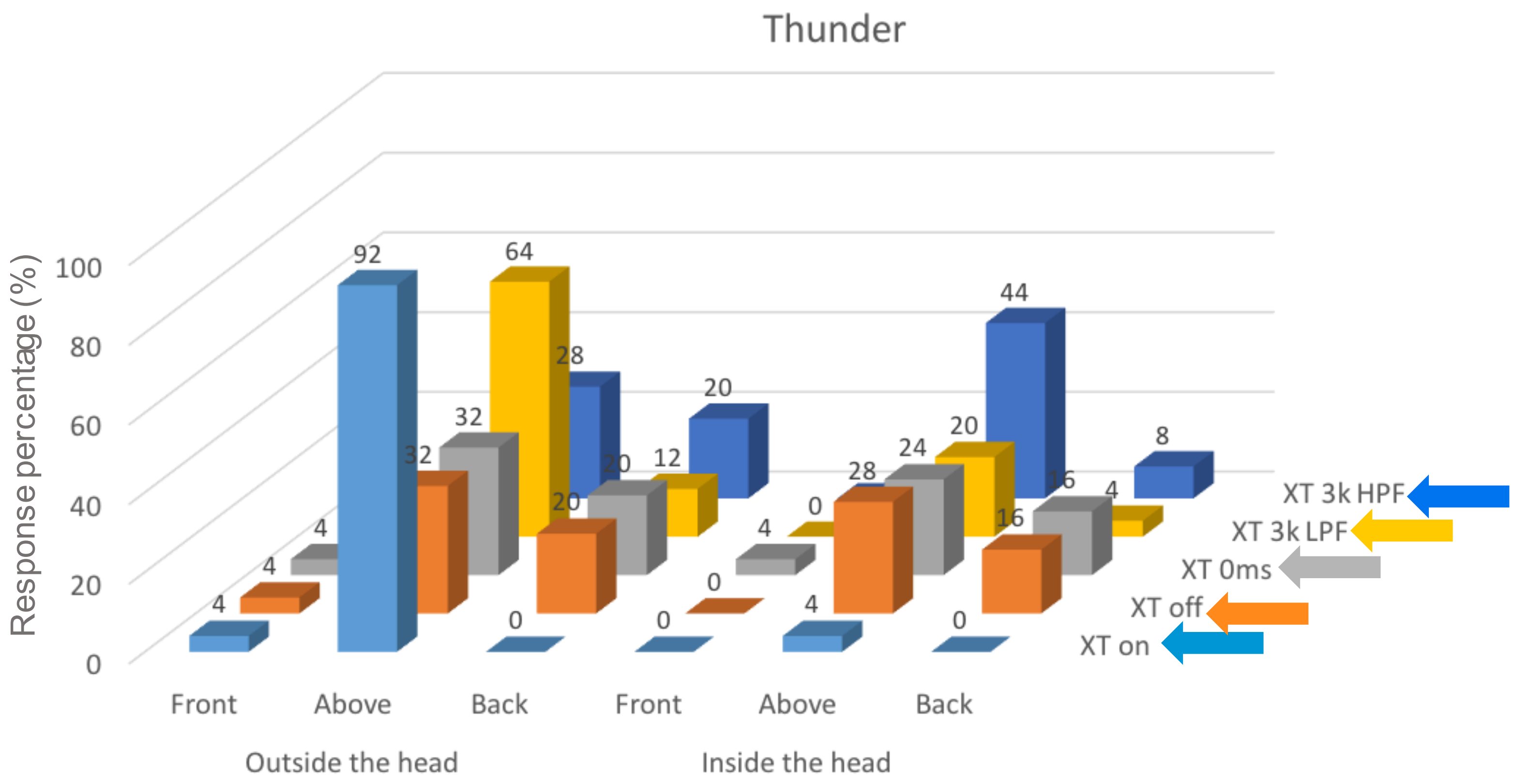
- “Above Outside the head” with “XT on” and “XT 3k LPF”.



# Results

University of HUDDERSFIELD

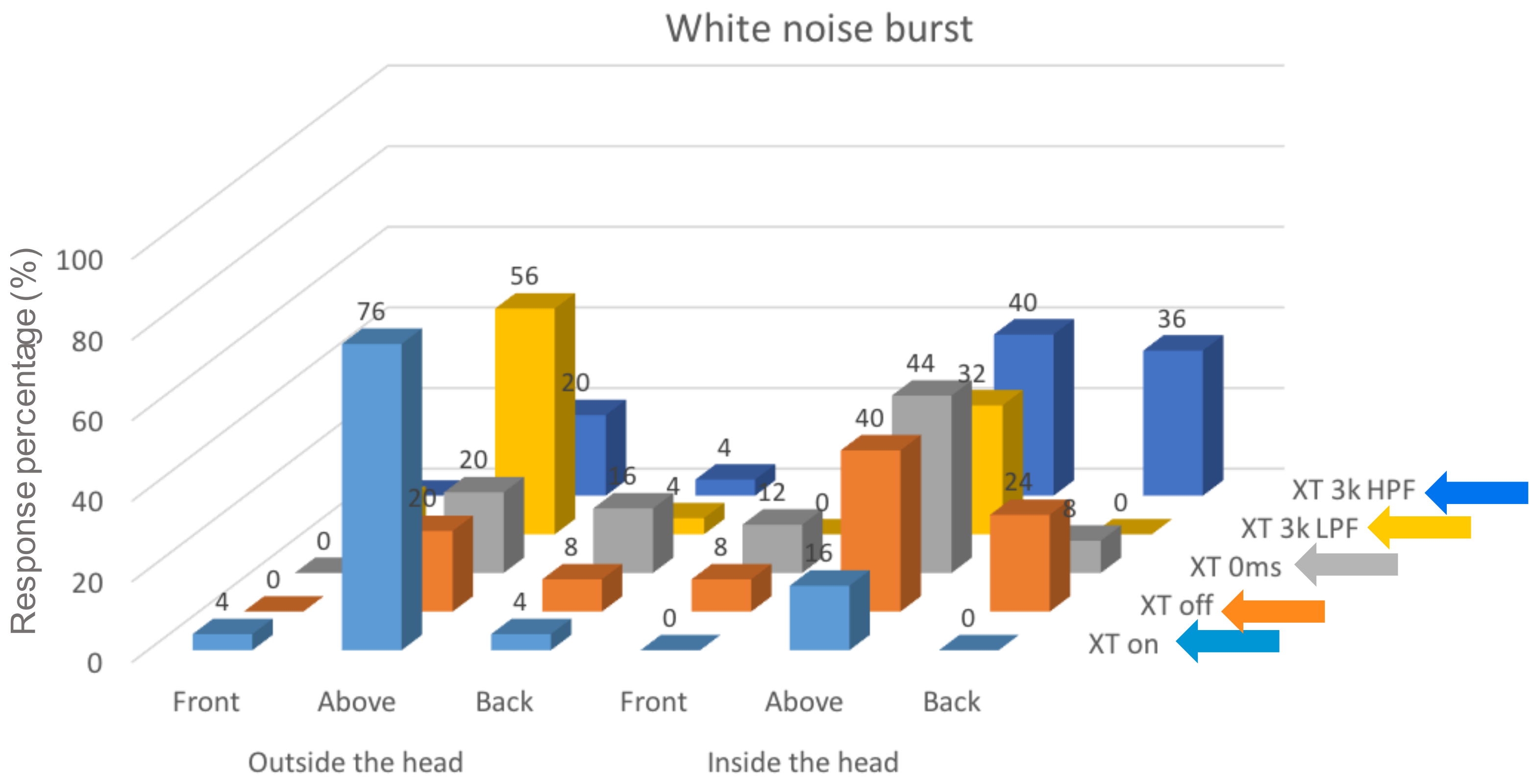
- "Above Outside the head" with "XT on" and "XT 3k LPF".



# Results

University of HUDDERSFIELD

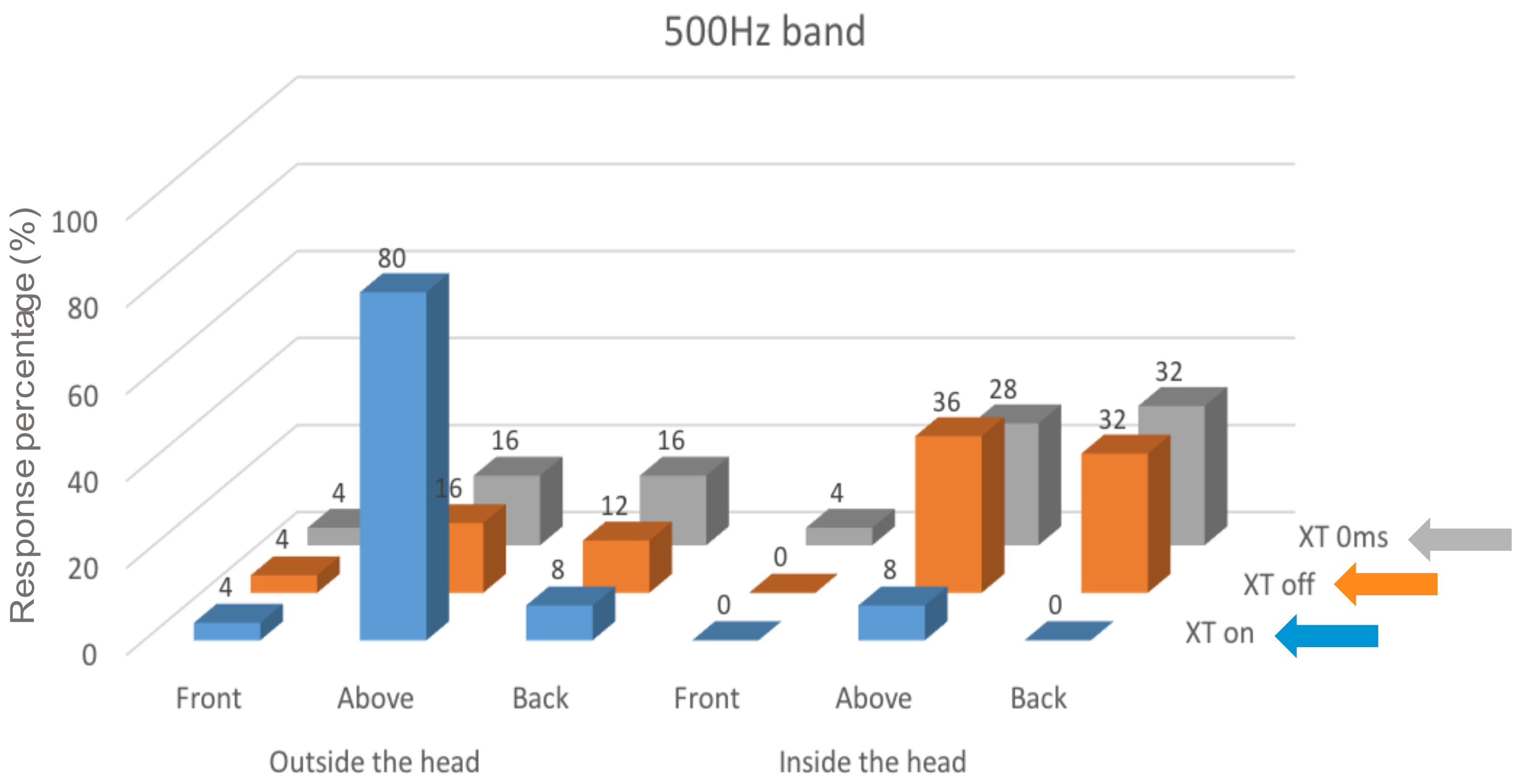
- “Above Outside the head” with “XT on” and “XT 3k LPF”.



# Results

University of HUDDERSFIELD

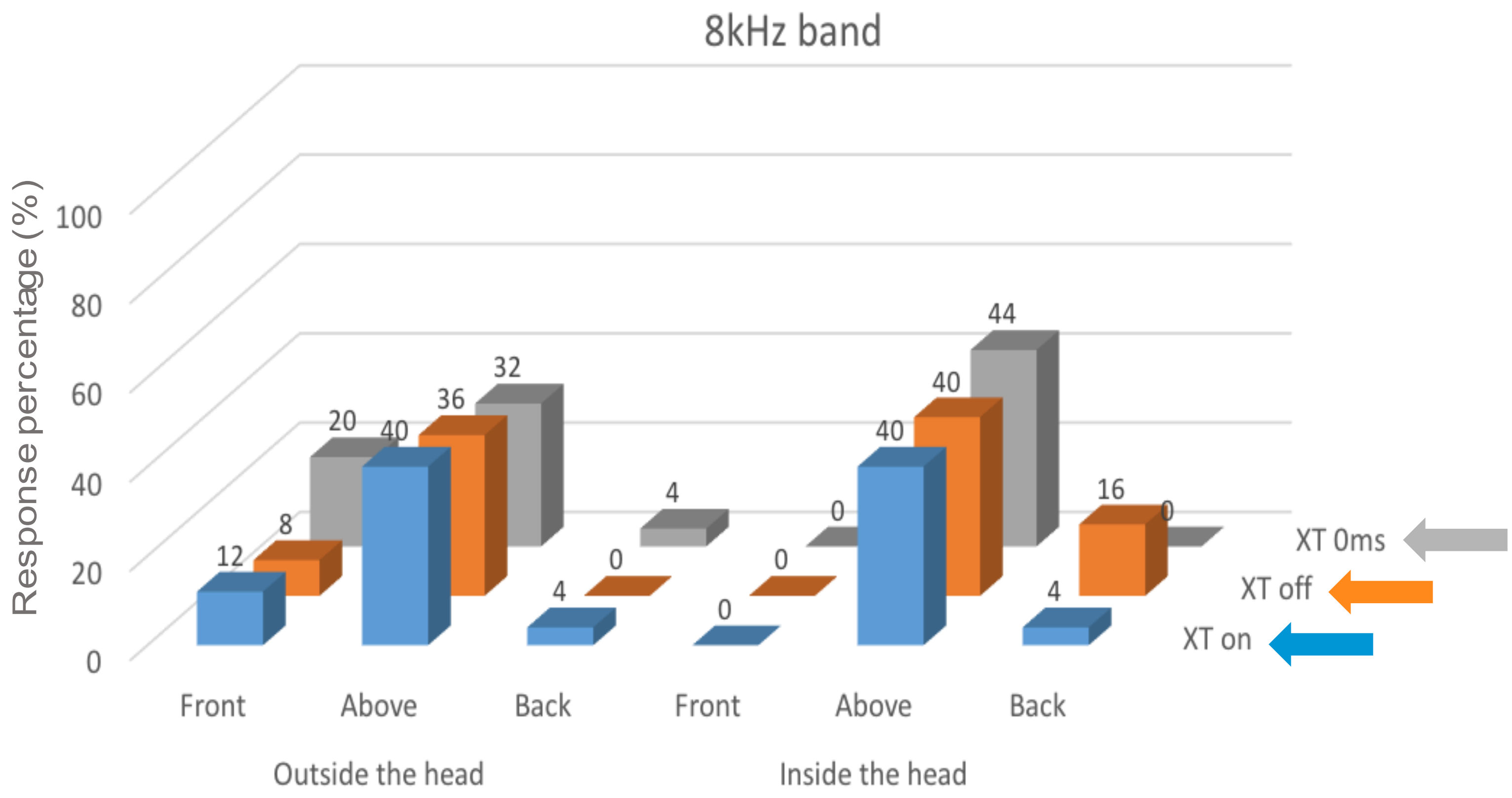
- “Above Outside the head” with “XT on” for the 500Hz band.



# Results

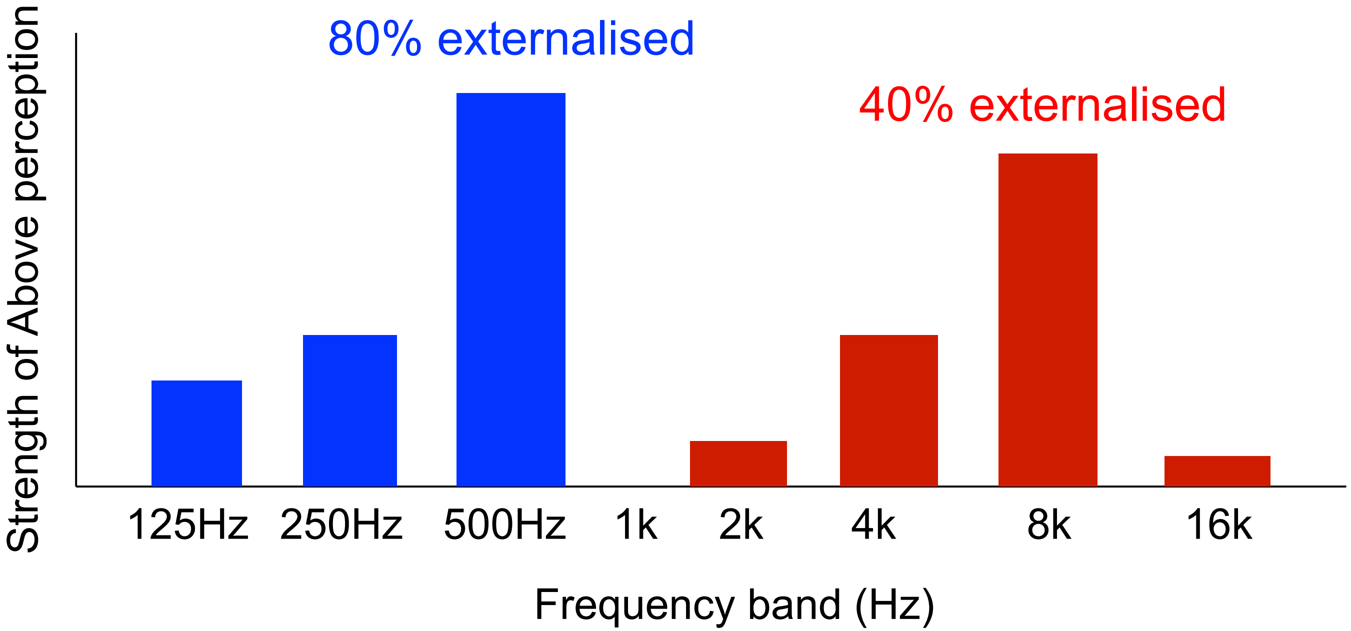
University of HUDDERSFIELD

- “Above Inside the head” with “XT on” for the 8kHz band.



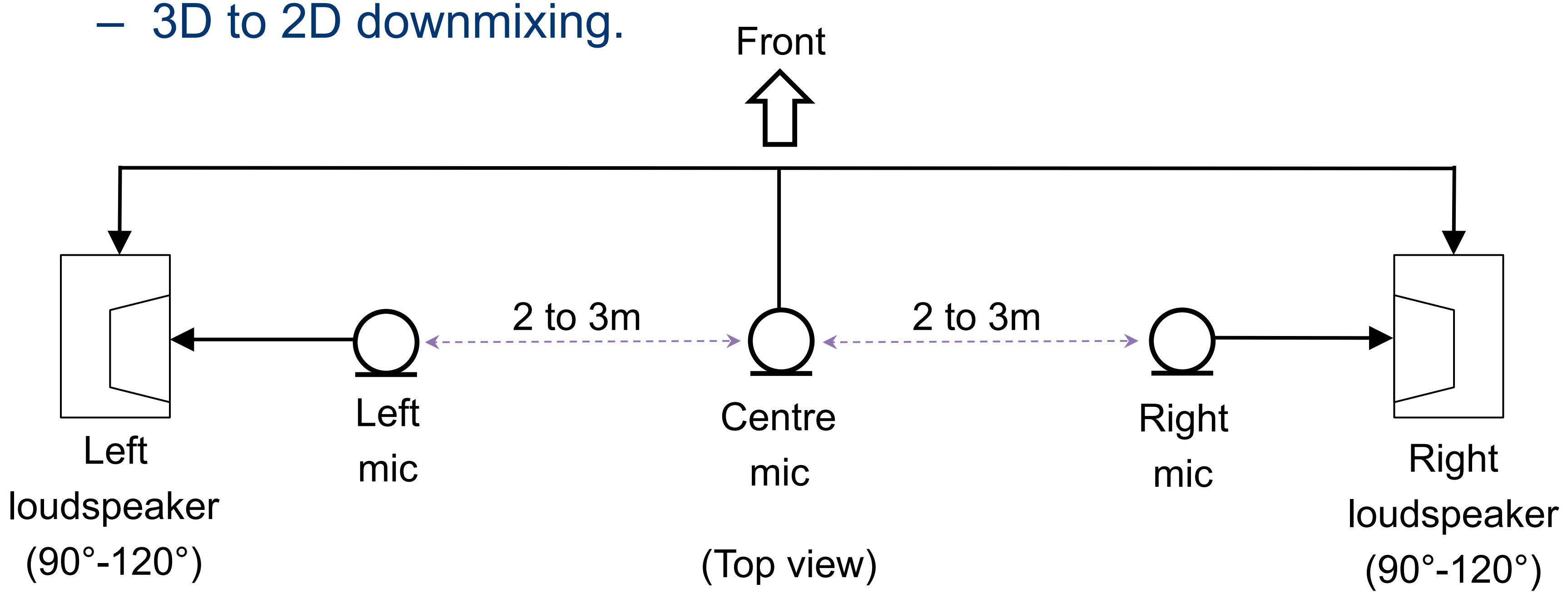
# New hypothesis

- Duplex theory of phantom image elevation for the 180° base angle.
  - Low frequencies < 3 kHz: Cognitive effect (Crosstalk delay)
  - High frequencies > 3 kHz: Hard-wired effect (i.e., Directional bands)



# Practical implications

- Exploiting the results for
  - 3D image rendering and mic technique without height channels.
  - 2D to 3D upmixing.
  - 3D to 2D downmixing.



# Conclusions

University of Huddersfield

- The phantom image elevation effect is most dominant with the 500Hz and 8kHz octave bands.
- At frequencies below 3kHz, the delay time of acoustic crosstalk plays the main role for the effect (cognitive effect).
- The 500Hz band is perceived above, mostly outside the head.
- The 8kHz band is perceived above, mostly inside the head.

University of Huddersfield

# Thank you for listening

Hyunkook Lee
h.lee@hud.ac.uk