

Effects of incongruent auditory and visual room-related cues on sound externalization

Carvajal, Juan Camilo Gil; Santurette, Sébastien; Cubick, Jens; Dau, Torsten

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):

Carvajal, J. C. G., Santurette, S., Cubick, J., & Dau, T. (2015). Effects of incongruent auditory and visual room-related cues on sound externalization. Poster session presented at Tenth anniversary symposium of the international laboratory for Brain, Music, and Sound Research, Montréal, Canada.

DTU Library

Technical Information Center of Denmark

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Introduction

Sounds presented via headphones are typically perceived inside the head. However, the illusion of a sound source located out in space away from the listener's head can be generated with binaural headphone-based auralization systems by convolving anechoic sound signals with a binaural room impulse response (BRIR) measured with miniature microphones placed in the listener's ear canals. Sound externalization of such virtual sounds can be very convincing and robust but there have been reports that the illusion might break down when the listening environment differs from the room in which the BRIRs were recorded [1,2,3]. This may be due to incongruent auditory cues between the recording and playback room during sound reproduction [2]. Alternatively, an expectation effect caused by the visual impression of the room may affect the position of the perceived auditory image [3]. Here, we systematically investigated whether incongruent auditory and visual room-related cues affected sound externalization in terms of perceived distance, azimuthal localization, and compactness.

Methods

- Eighteen naïve listeners were divided into two groups:
 - A:** Blindfolded during testing but with *auditory awareness* of the room provided by an in-room noise source
 - V:** Shielded from room-related acoustic input but with *visual awareness* of the room
- All listeners also tested with all cues (**VA**) available
- Three playback rooms:
 - IEC_M:** Medium-sized IEC standard ($V \approx 100 \text{ m}^3$, $T_{60} = 0.4 \text{ s}$) in which all BRIRs were recorded
 - Rev_S:** Small reverberant ($V = 43.2 \text{ m}^3$, $T_{60} = 2.8 \text{ s}$)
 - Dry_L:** Large anechoic ($V = 330.4 \text{ m}^3$, $T_{60} < 0.01 \text{ s}$)

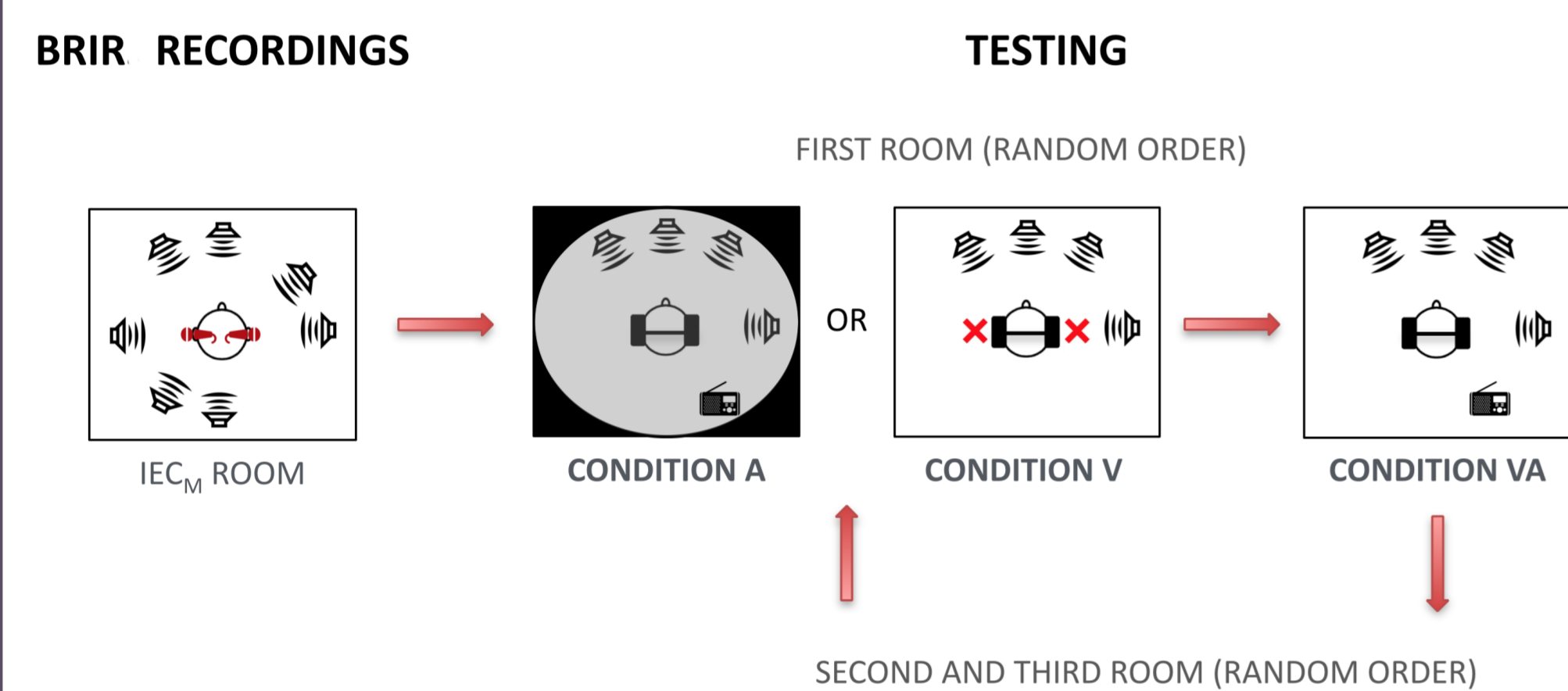


Fig.1 Overview of the experimental procedure followed by test subjects.

- Seven azimuthal positions (2, 3, 6, 7, 9, 11, and 12 o'clock) were reproduced
- Loudspeakers visible at positions 1, 3, 11, and 12
- Subjective rating scales for distance, azimuth, and compactness perception

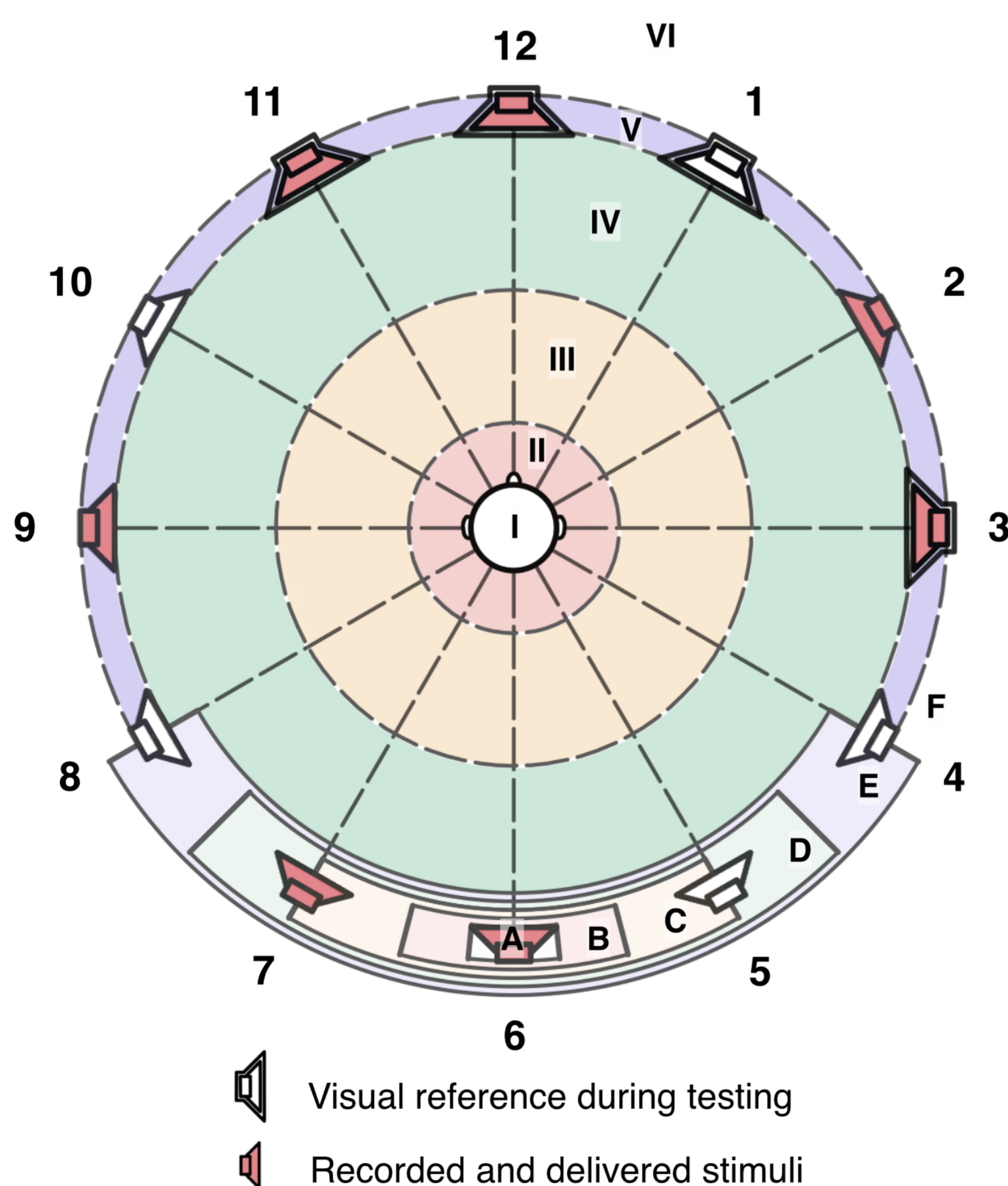


Fig.2 Loudspeaker setup and subjective rating scales used for the experiments. For distance judgments, listeners could provide ratings of 0 (inside the head), 1, 2, 3, 4 (at the loudspeaker), or 5, corresponding to areas I, II, III, IV, V, and VI, respectively. For compactness judgments, listeners could provide ratings of 0 (most compact), 1, 2, 3, 4, or 5 (least compact), corresponding to areas A, B, C, D, E, and F, respectively.

Results: Distance perception

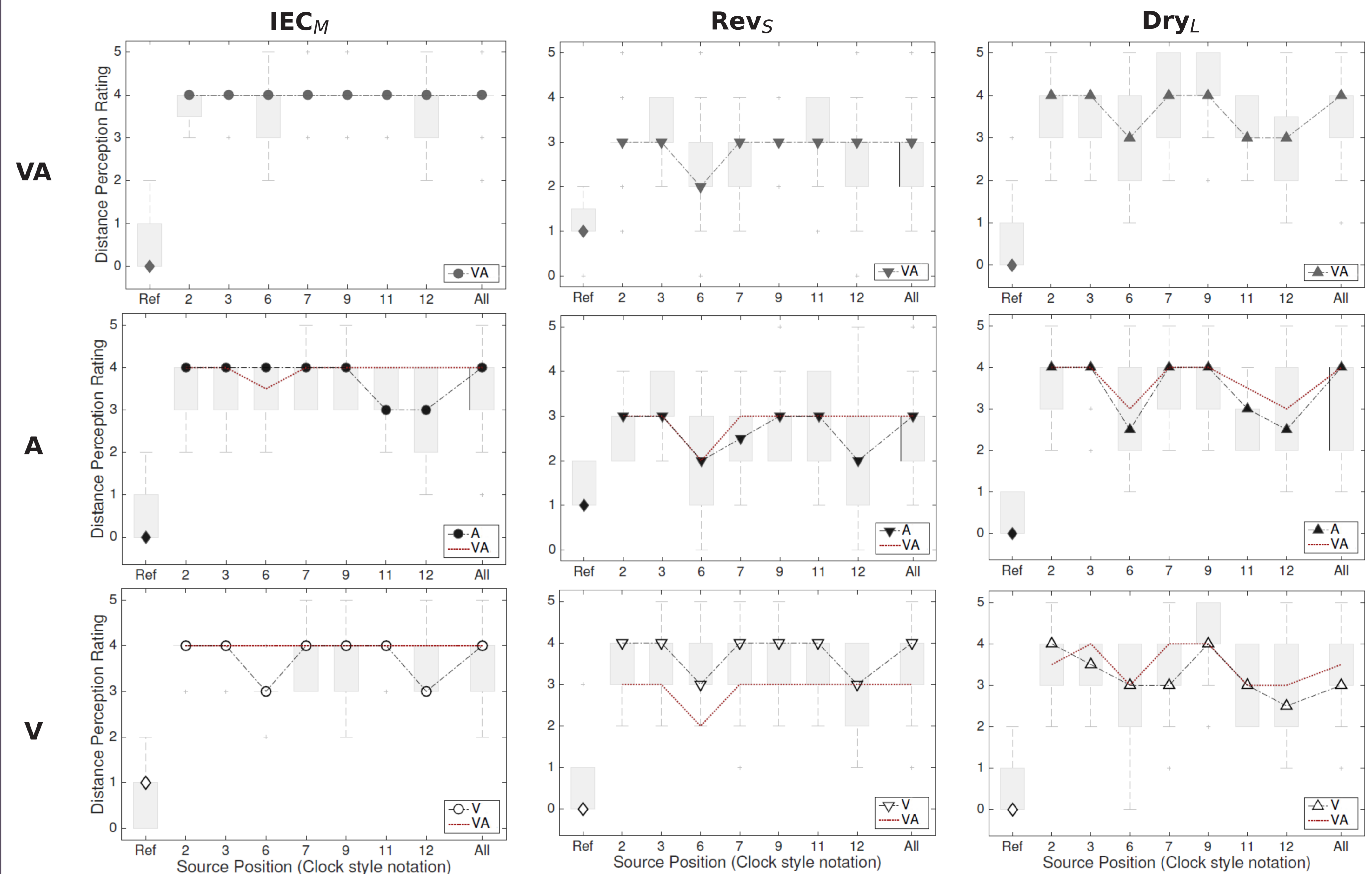


Fig.3 Distance perception ratings obtained in the three playback rooms under conditions VA, A, and V for each tested source position. The boxplots indicate median values across subjects as well as first and third quartiles. 'Ref' stands for a dry diotic reference signal, 'All' for the median rating across all source positions.

- Significant effect of listening environment: lower VA ratings in Rev_S for all positions, in Dry_L for positions 3, 6, 11, and 12
- In Rev_S, no difference between A and VA conditions, but significantly higher ratings for V than VA for most positions
- In Dry_L, no significant differences between A, V, and VA conditions for most positions

Results: Azimuthal localization

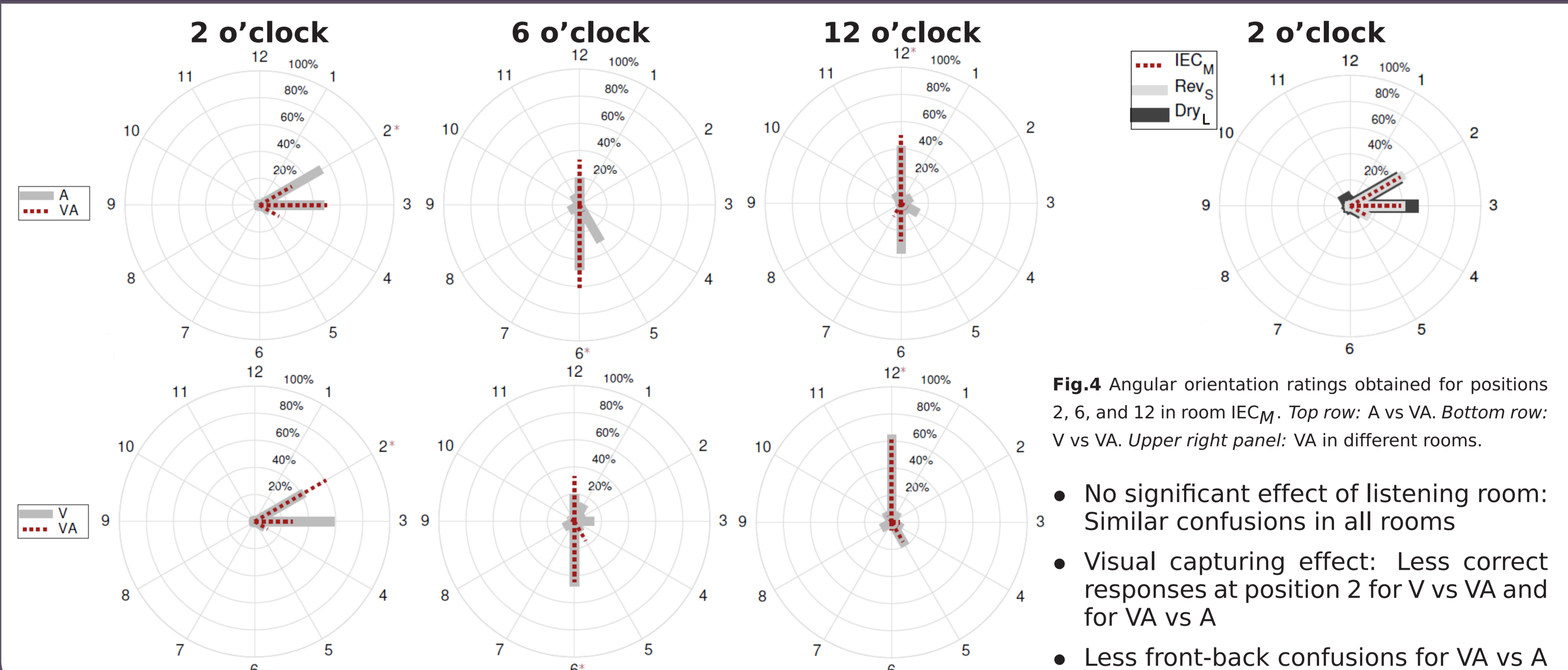


Fig.4 Angular orientation ratings obtained for positions 2, 6, and 12 in room IEC_M. Top row: A vs VA. Bottom row: V vs VA. Upper right panel: VA in different rooms.

- No significant effect of listening room: Similar confusions in all rooms
- Visual capturing effect: Less correct responses at position 2 for V vs VA and for VA vs A
- Less front-back confusions for VA vs A

Results: Compactness

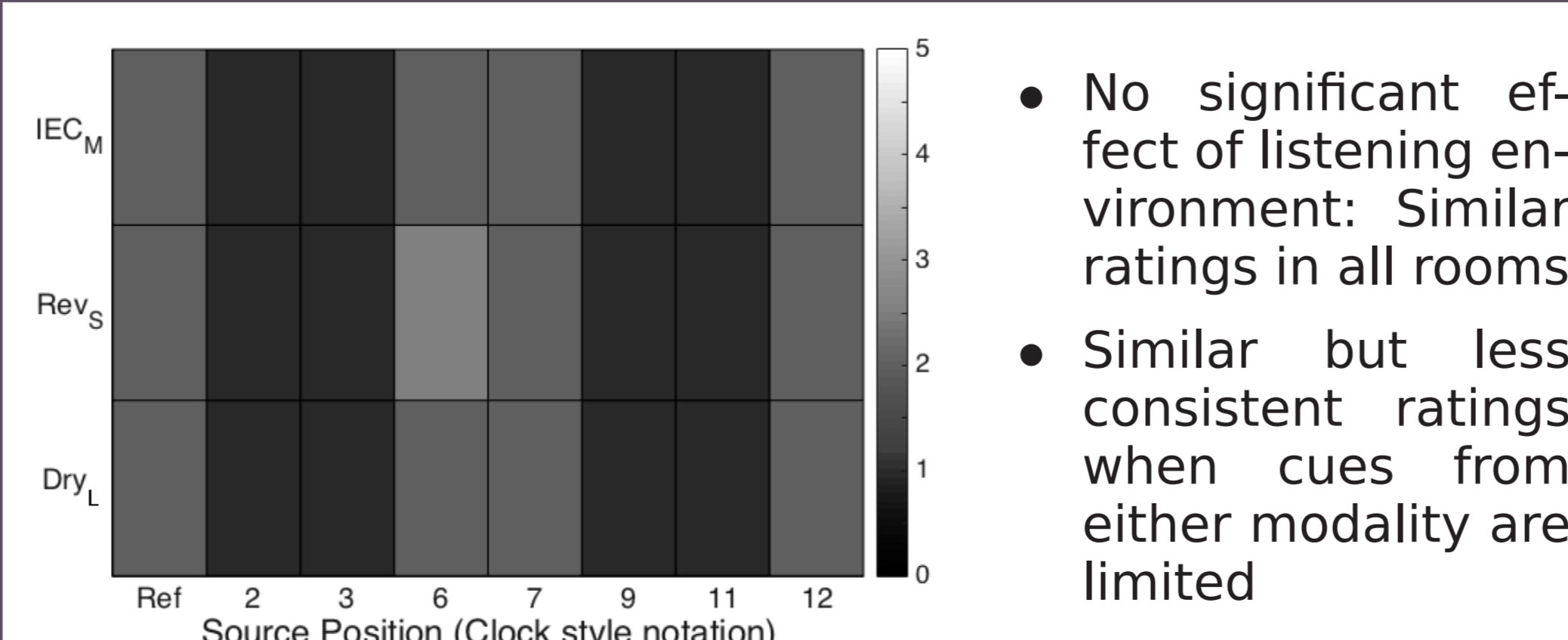


Fig.5 Compactness ratings obtained in condition VA in the three playback rooms.

References:

- Begault DR (1999). "Auditory and non-auditory factors that potentially influence virtual acoustic imagery." Proc. Audio Eng. Soc. 16:13-26.
- Werner S, Siegel A (2011). "Effects of binaural auralization via headphones on the perception of acoustic scenes." Proc. ISAAR 3:215-222.
- Udesen J, Piechowiak T, Gran F (2015). "The effect of vision on psychoacoustic testing with headphone-based virtual sound." J. Audio Eng. Soc. 63:552-561.

Conclusions

- A mismatch between recording and playback room is detrimental to virtual sound externalization.
- The auditory modality governs externalization in terms of perceived distance when cues from the recording and playback room are incongruent.
- In incongruent listening situations, the more reverberant the playback room is, the more critical the auditory impression of the room becomes.
- The visual impression of the room does not affect perceived distance, but source-related visual cues help resolve localization ambiguities and improve compactness perception.

Presented at: Tenth Anniversary Symposium of the International Laboratory for Brain, Music, and Sound Research — October 22, 2015, Montréal, Canada
 Contact: Juan Camilo Gil Carvajal — s130655@student.dtu.dk
 Sébastien Santurette — ses@elektro.dtu.dk
 Jens Cubick — jecu@elektro.dtu.dk
 Torsten Dau — tdau@elektro.dtu.dk