Assessment of the impact of spatial audiovisual coherence on source unmasking

Julian Palacino, Mathieu Paquier, Vincent Koehl, Frédéric Changenet, Etienne Corteel

▶ **To cite this version:**

Julian Palacino, Mathieu Paquier, Vincent Koehl, Frédéric Changenet, Etienne Corteel. Assessment of the impact of spatial audiovisual coherence on source unmasking. AES (Audio Engineering Society). 140th AES Convention, Jun 2016, Paris, France. 2016. hal-01326842

HAL Id: hal-01326842
https://hal.univ-brest.fr/hal-01326842

Submitted on 6 Jun 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Assessment of the impact of spatial audiovisual coherence on source unmasking

Julian Palacino, julian.palacino@univ-brest.fr, Mathieu Paquier, mathieu.paquier@univ-brest.fr, Vincent Koehl, vincent.koehl@univ-brest.fr,
Frédéric Changenet, frederic.changenet@radiofrance.fr, and Etienne Corteel, etienne.corteel@sonicemotion.com

140th Convention
june 2016

ABSTRACT

The present study aims at evaluating the contribution of spatial audiovisual coherence for sound source unmasking for live music mixing. Sound engineers working with WFS technologies for live sound mixing have reported that their mixing methods have radically changed. Using conventional mixing methods, the audio spectrum is balanced in order to get each instrument intelligible inside the stereo mix. In contrast, when using WFS technologies, the source intelligibility can be achieved thanks to spatial audiovisual coherence and/or sound spatialization (and without using spectral modifications). The respective effects of spatial audiovisual coherence and sound spatialization should be perceptually evaluated.
As a first step, the ability of naive and expert subjects to identify a spatialized mix was evaluated by a discrimination task. For this purpose, live performances (rock, jazz and classic) were played back to subjects with and without stereoscopic video display and VBAP or WFS audio rendering. Two sound engineers realized the audio mixing for three pieces of music and for both audio technologies in the same room where the test have been carried out.

TEST

Hypothesis :
- The audiovisual coherence allows source unmasking
- The source spatialization allows source unmasking

UBO 3D room (controlled environment)
3D stereoscopic projection (HD)

- 2 conditions: Audio / Audiovisual
- 3 excerpts: Classic / Jazz / Rock
- 2 sound engineers: Sound ing. 1, Sound ing. 2
- 4 versions: WFS Coherent / WFS Non Coherent / VBAP Coherent / VBAP Non Coherent

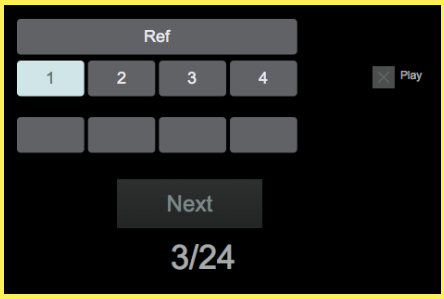
Preliminary listening : Very similar !

Subjects are able to identify the differences to evaluate it?

DISCRIMINATION TEST

Protocol

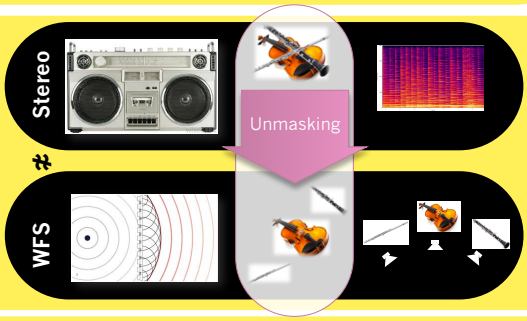
- 2 sessions (Audio, Audiovisual)
- 3 Excerpts x 2 sound ing. x 4 versions



X 13 Experts (ISB master's students)

X 13 Naives

radio france → sonic emotion ABSOLUTE 3D

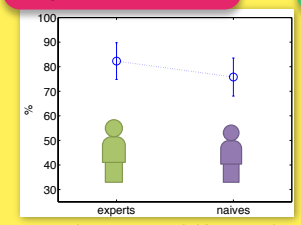


EDISON 3D

RESULTS

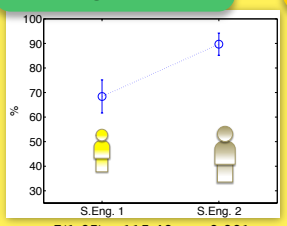
good answers rate > 60 %

Expertise :



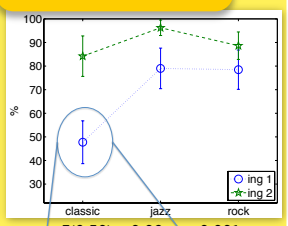
F(1,25) = 1.54 p = 0.23 and > 70%

Sound Ing. :



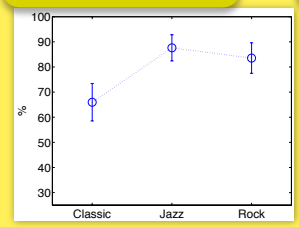
F(1,25) = 115.40; p < 0.001

Sound Ing. and excerpt:



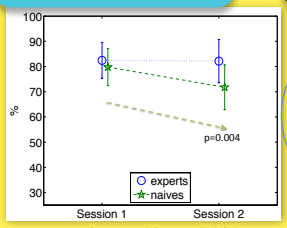
F(2,50) = 9.92; p < 0.001

Excerpt :



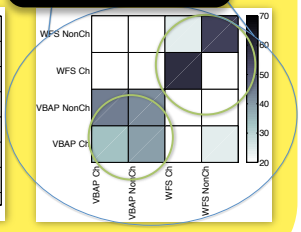
F(1,25) = 34.09; p < 0.001
Bonferroni post hoc (p<0.001)

Session and expertise :



F(1,25) = 4.90; p = 0.036

Confusion Matrix



CONCLUSIONS

- The expertise does not influence the discrimination.
- Subjects are able to detect small differences between versions,
- Mixing choices affect discrimination.
- Classic : Conservative mixing conform to the style aesthetics.
- Experts have equivalent scores in audio an audiovisual conditions.
- Assessment by naives is degraded in audiovisual condition.
- Results are significantly different between sound ing. (especially for the classic excerpt)
- WFS and VBAP are differentiated

Subjects are able to perform discrimination task

EVALUATION TEST