A Study on the Lombard Effect in Telepresence Robotics

Ambre Davat, GIPSA-lab / LIG, CNRS, Univ. Grenoble Alpes, Grenoble INP
Gang Feng , GIPSA-lab, CNRS, Univ. Grenoble Alpes, Grenoble INP
Véronique Aubergé , LIG, CNRS, Univ. Grenoble Alpes, Grenoble INP

gipsa-lab

Grenoble INP

UNIVERSITÉ Grenoble Alpes

LIG

What is a telepresence robot ?

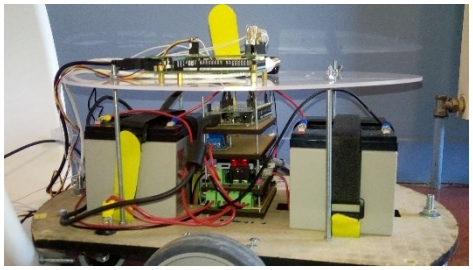
>> a **mobile** robot piloted **remotely** to **interact** with other people

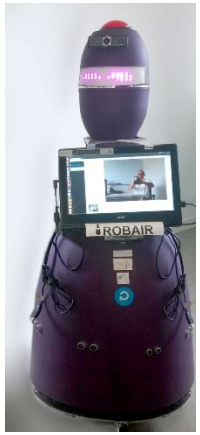
RobAIR

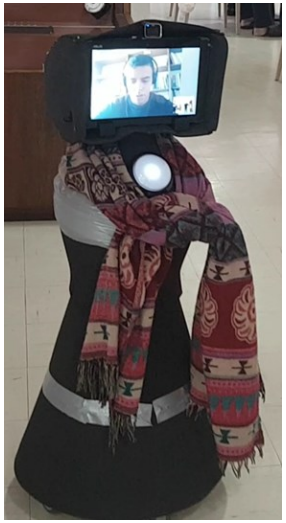
Beam+
(awabot)

RobAIR Social Touch

- Designed at the FabMSTIC
  (fablab of University Grenoble Alpes)
  & Domus (LIG's Living Lab)
- A tailor-made robot
- Experimentation
  out of the laboratory



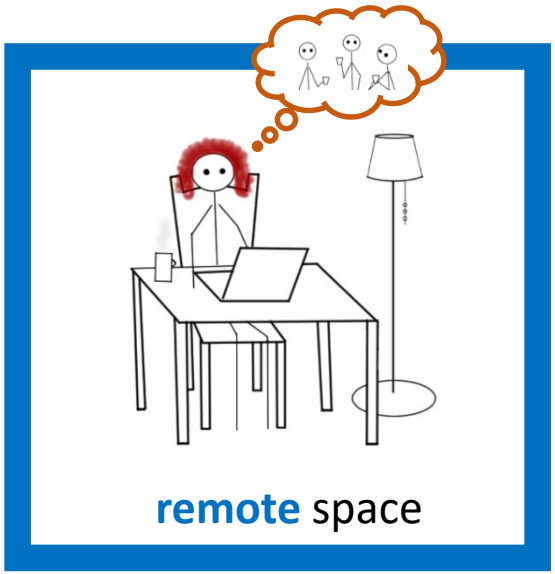




Specific issue: connection between the pilot and the robot's environment

Augmented telecommunication

>> better than traditional videoconferencing system

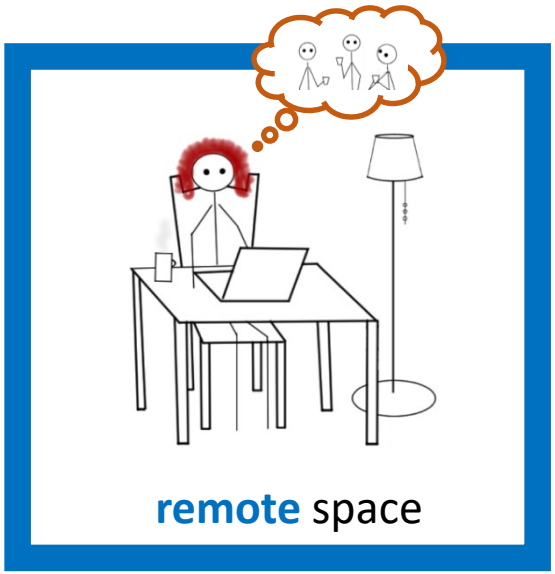


Embodied pilot
=> physical presence
in the local environment

Augmented telecommunication

>> better than traditional videoconferencing system

BUT not perfect



remote space

local space

Incomplete immersion
=> delay > 150 ms
=> no head tracking

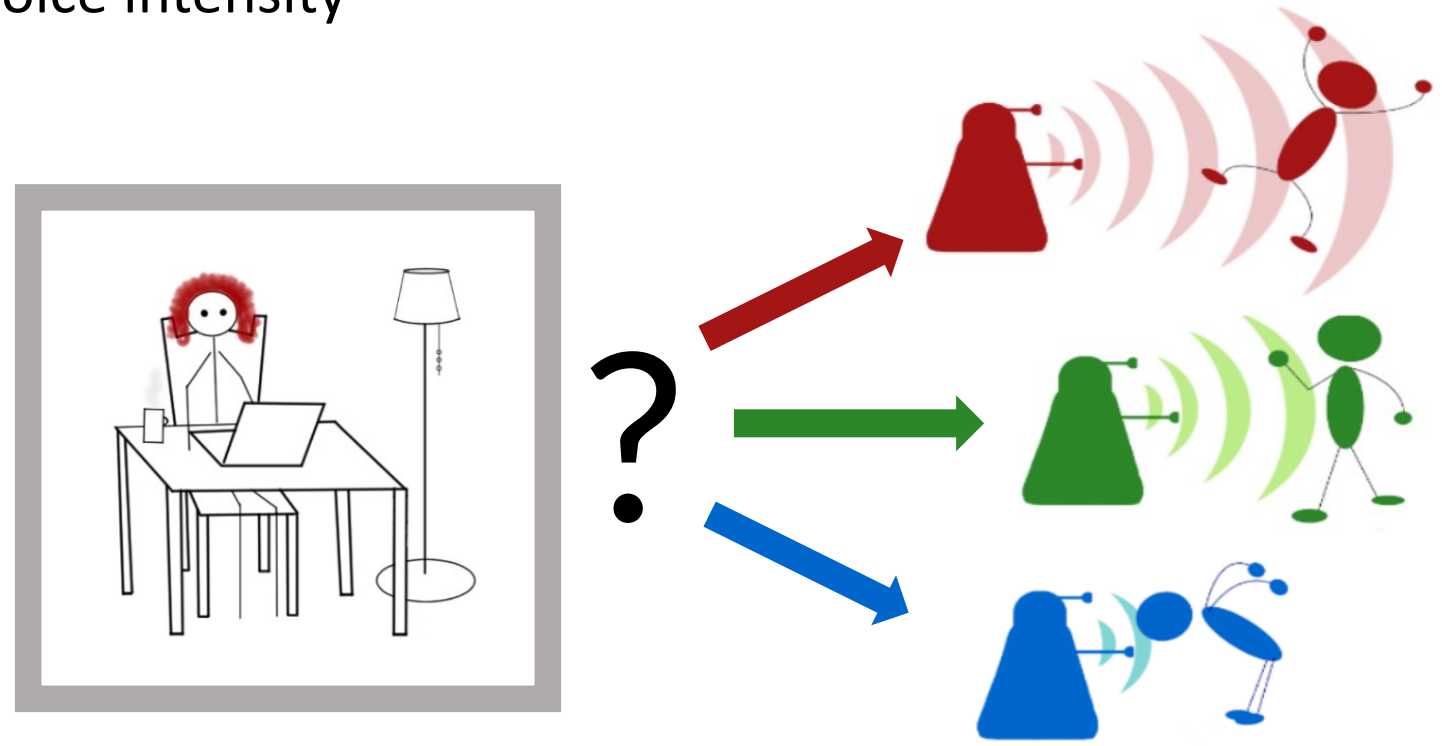
Robot ≠ clone of the pilot

voice not faithfully transmitted

(Počta & Komperda, 2016)

Unfaithful transmission of the voice

Ex: voice intensity



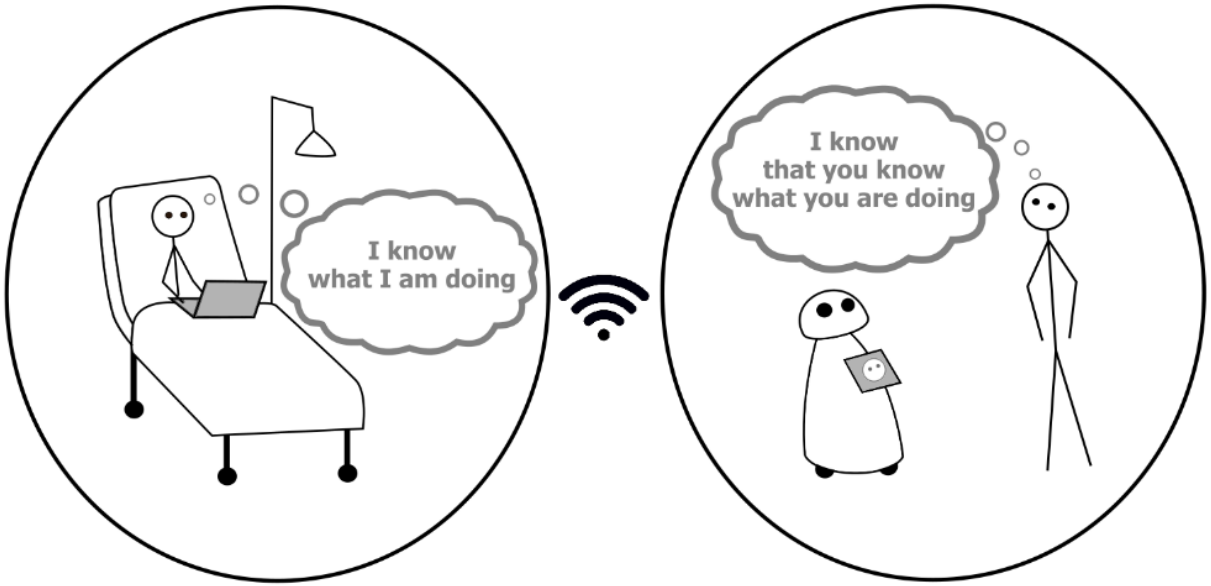
Risk of misunderstanding and break of social rules

Solution

>> **binaural acoustic immersion**

>> **calibration of the audio channel**

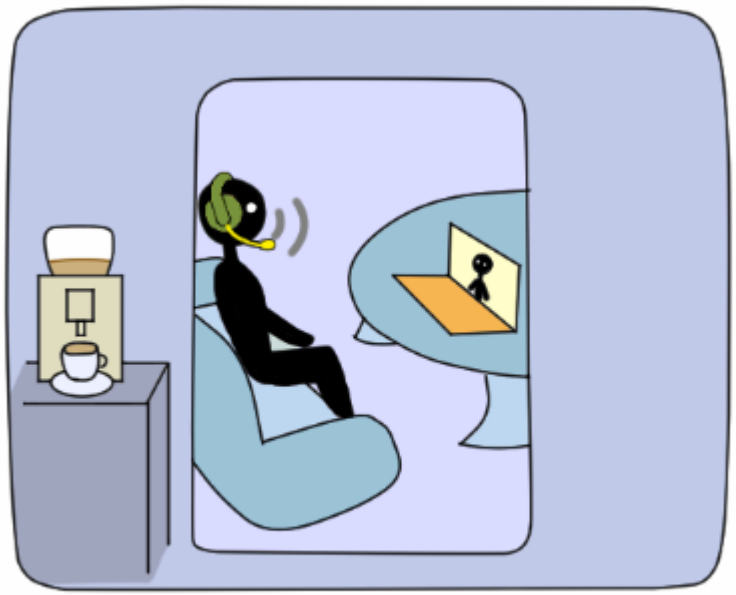
- sound intensity heard by the pilot = sound intensity in the local space
- voice intensity of the robot = voice intensity of the pilot



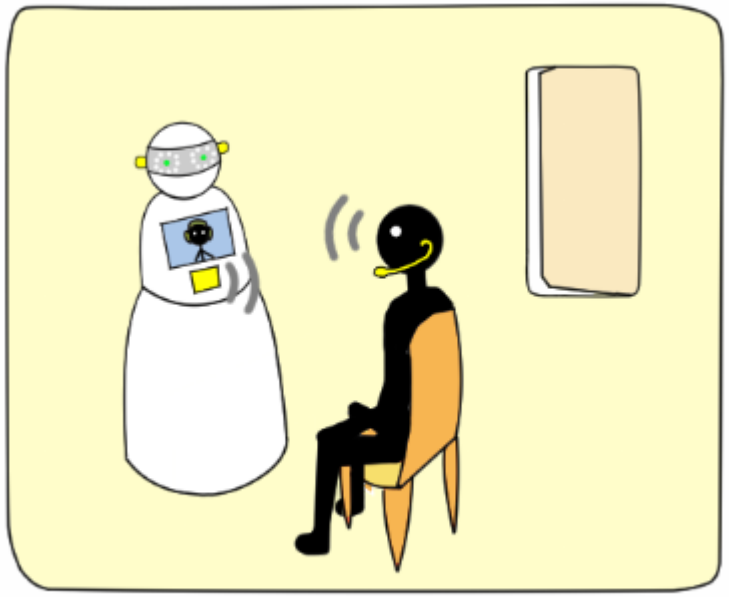
proprioceptive & interproprioceptive control of voice intensity

Problematic

Does **noise** in telepresence have different effects than in face-to-face interaction?



remote space



local space

The Lombard Effect

= automatic regulation to noise
(Lombard, 1911)

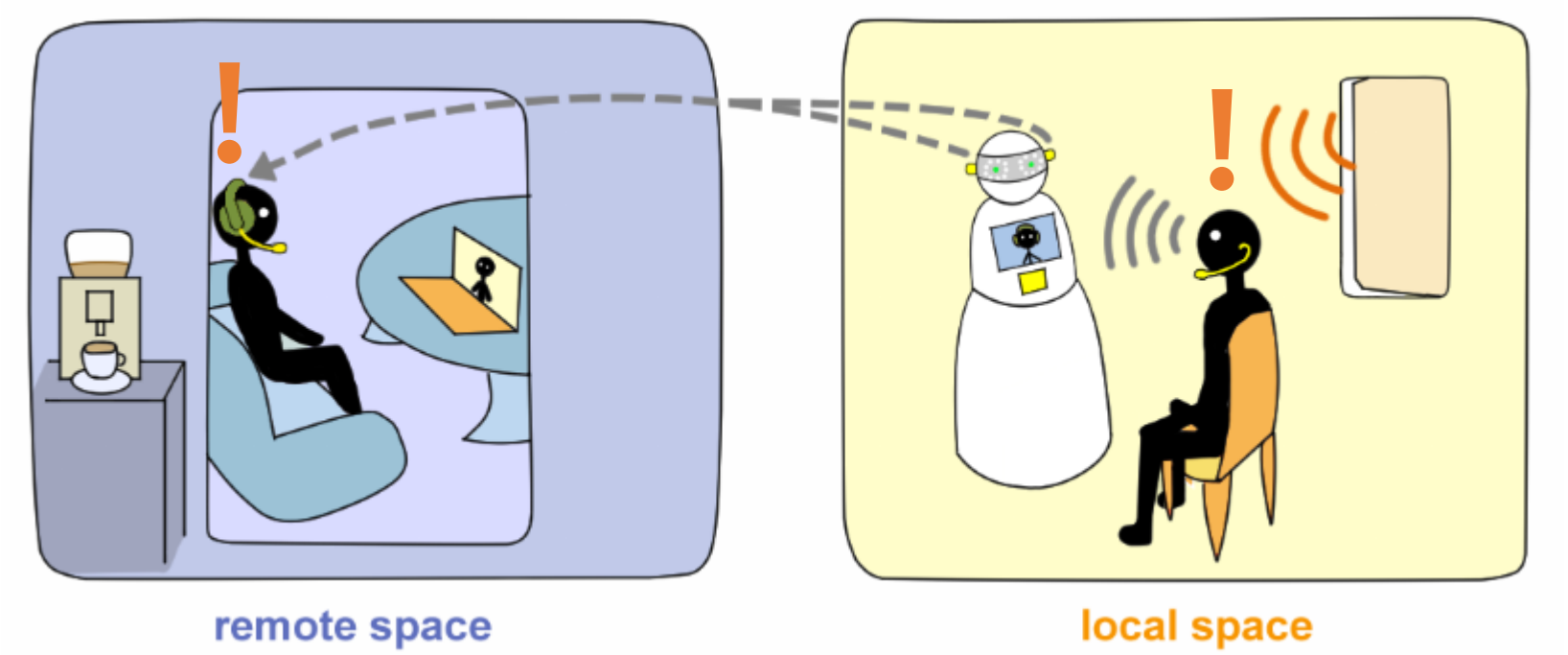
- Affects **intensity**, pitch, spectral energy, speech rate, articulatory movements… (Cooke et al., 2014)
- Common to other animals (Brumm & Zollinger, 2011)
- Stronger during interactive tasks (Garnier et al. 2010)
- Hard to consciously inhibit (Pick et al. 1989)

Noise sources in Telepresence Robotics

Two cases:

=> Noise in the **local space** : both subjects can hear it.

=> Noise in the **remote space** : only the pilot can hear it.

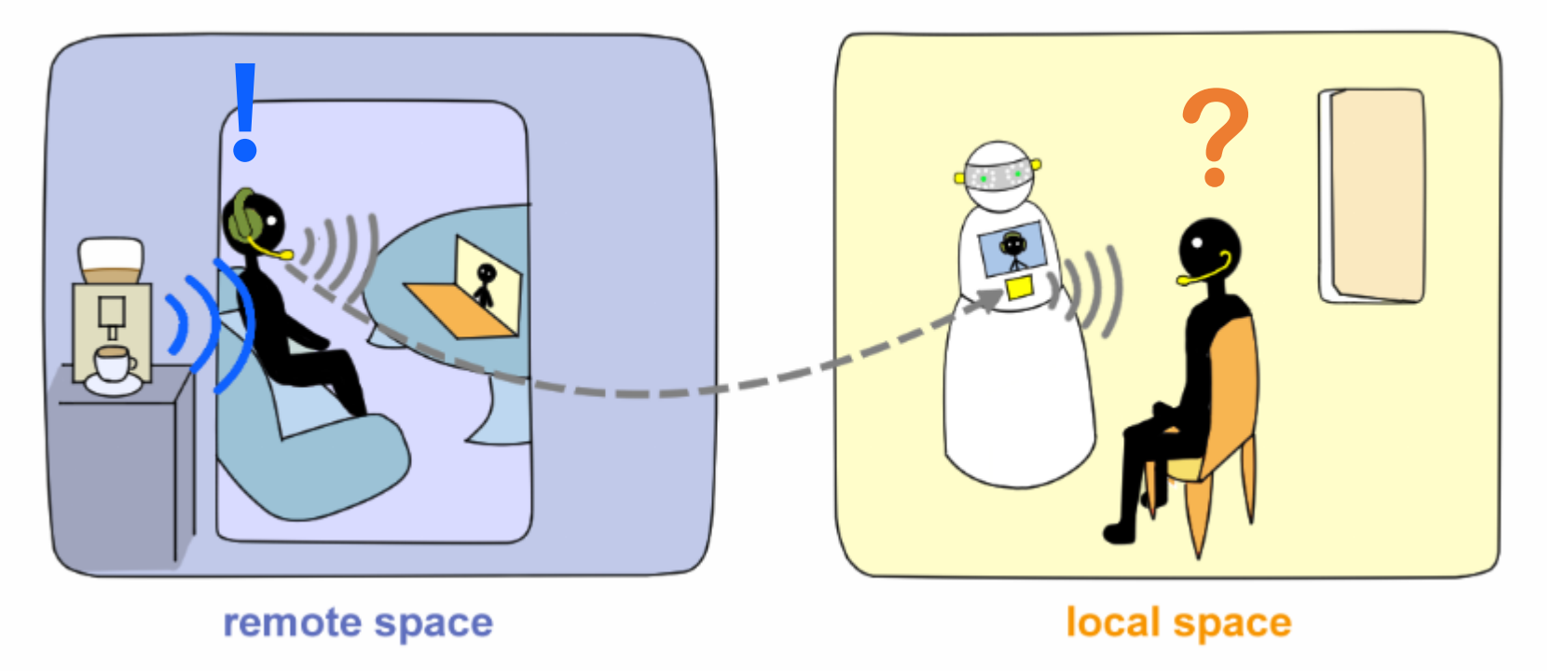


Noise sources in Telepresence Robotics

Two cases:

=> Noise in the **local space** : both subjects can hear it.

=> Noise in the **remote space** : only the pilot can hear it.



Our Approach

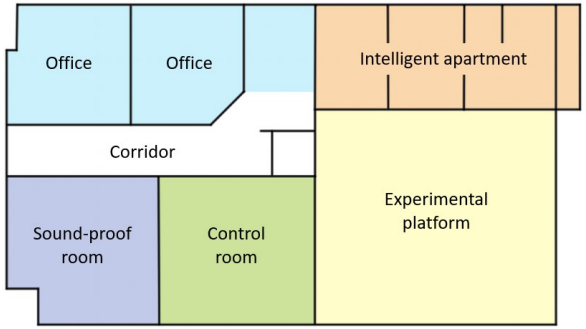
An experiment to study how two interlocutors using a **telepresence robot** adapt to different noise sources

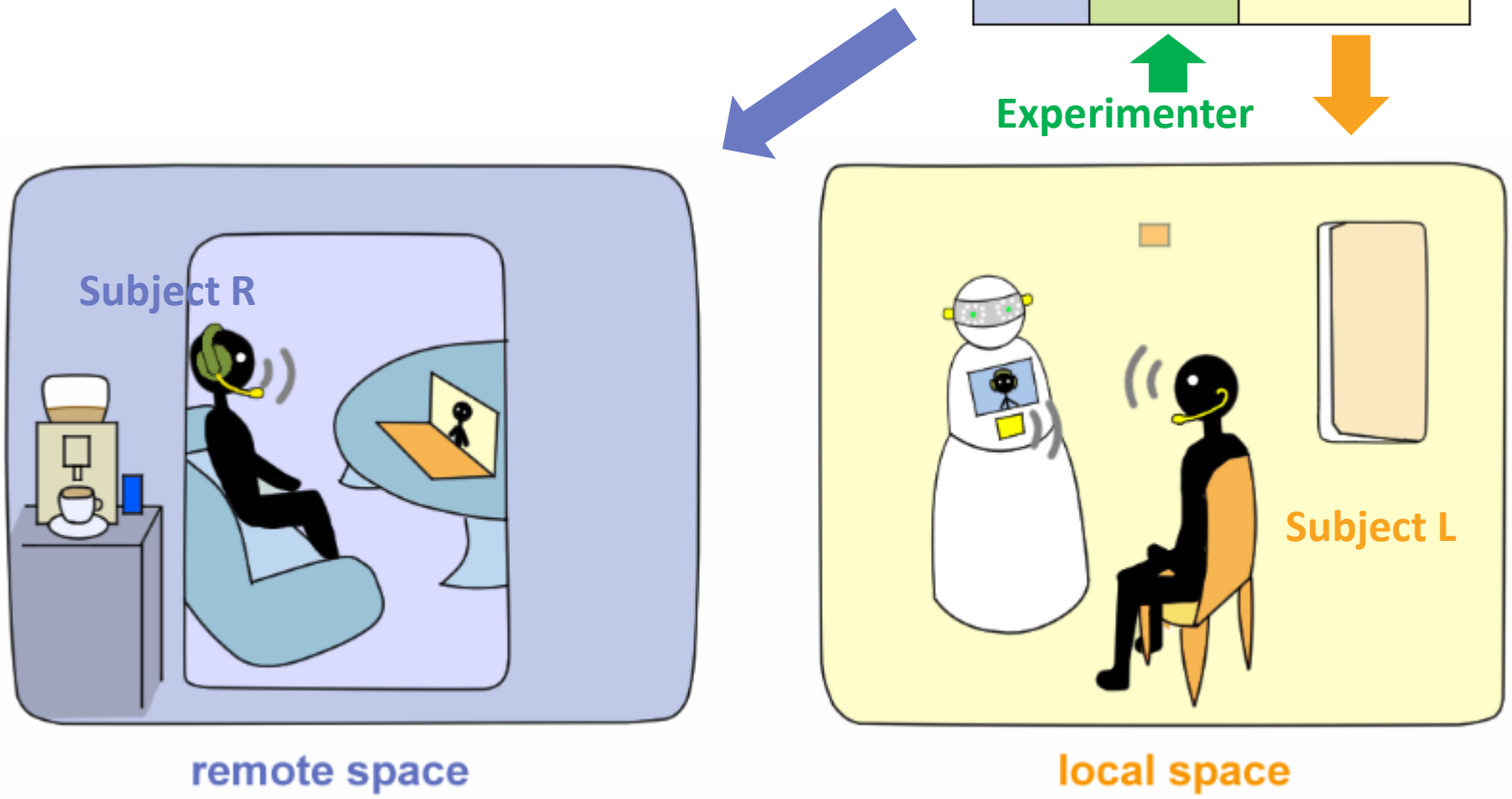
- measuring the Lombard Effect of **both** subjects

Standard case : both can hear the noise
Problematic case : only one of them can hear the noise

- individual or system-wide regulation to noise ?

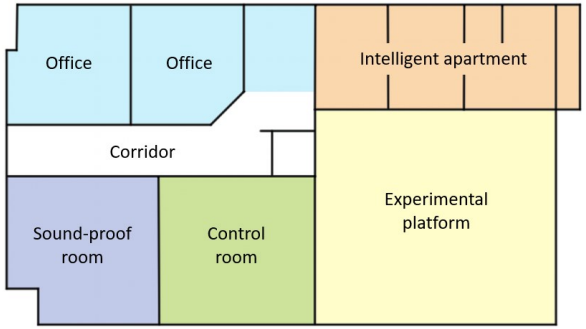
Test conditions

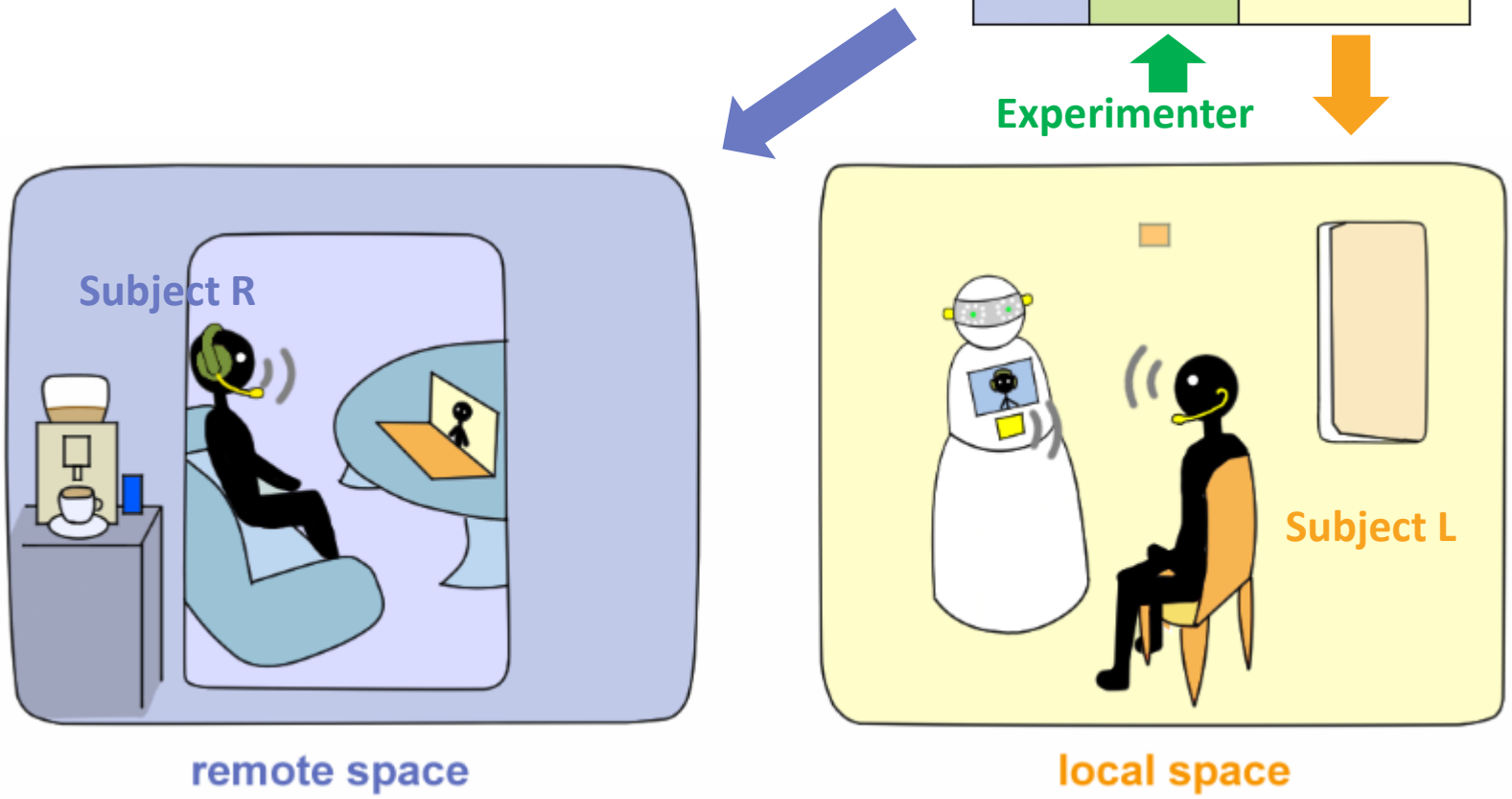




Test conditions

A - quiet in both rooms

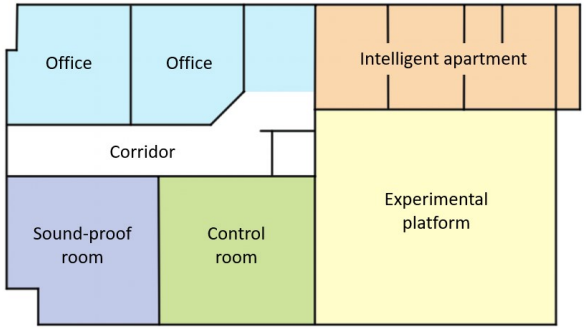


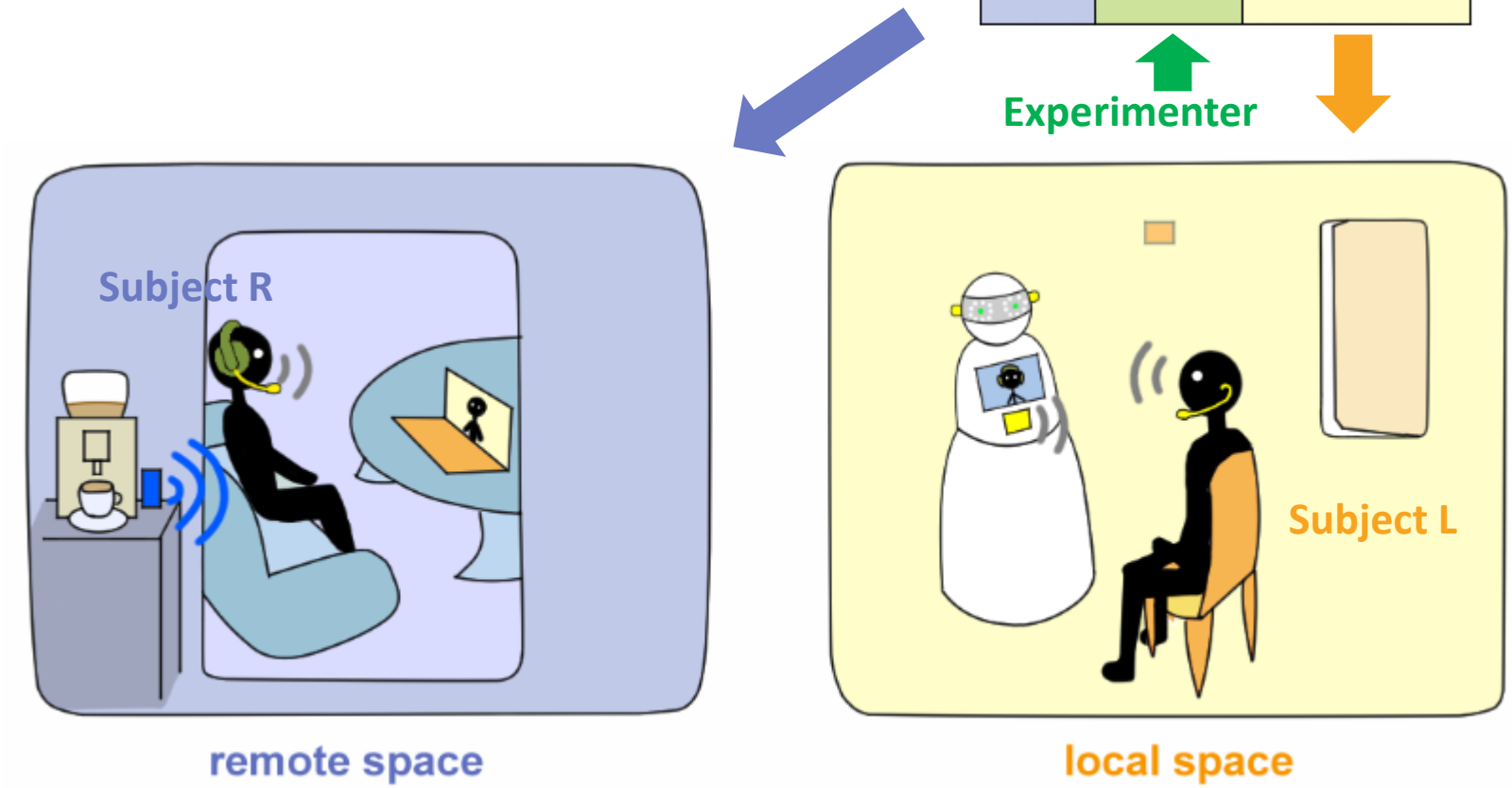


Test conditions

A - quiet in both rooms

B - noise in the remote space



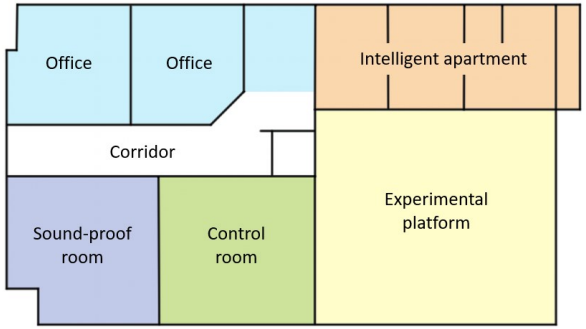


Test conditions

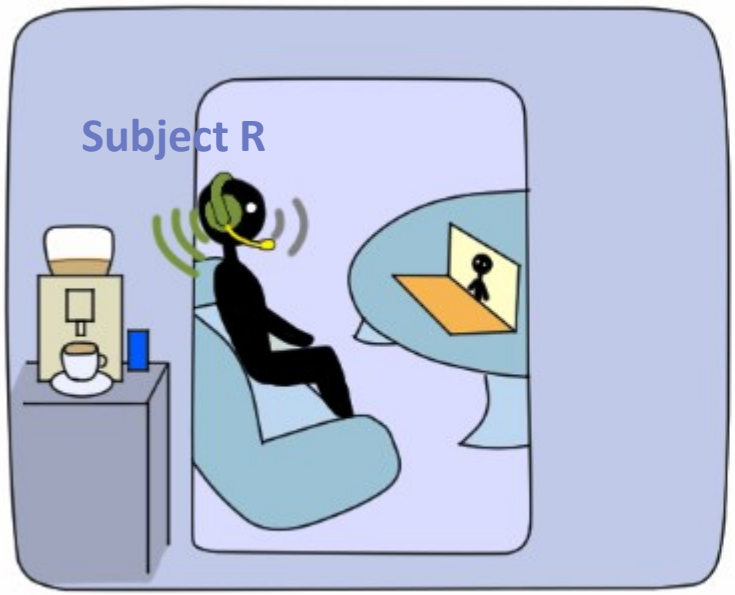
A - quiet in both rooms

B - noise in the remote space

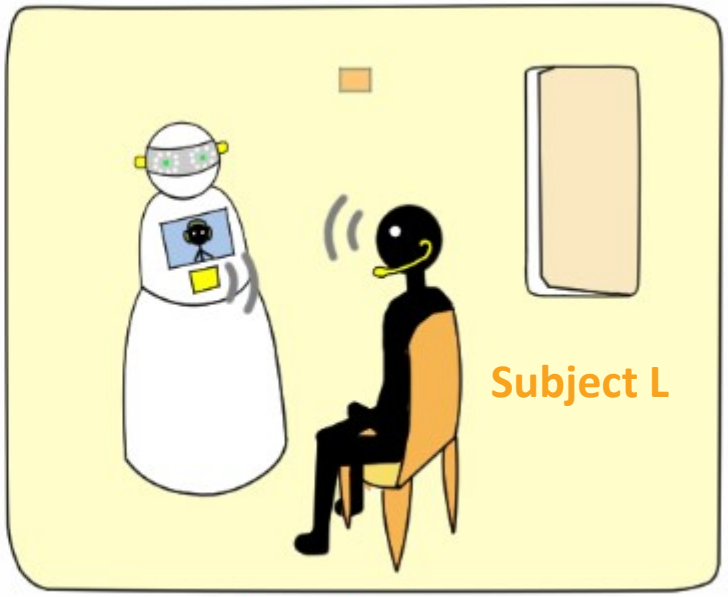
C - « virtual noise » in the local space



Experimenter



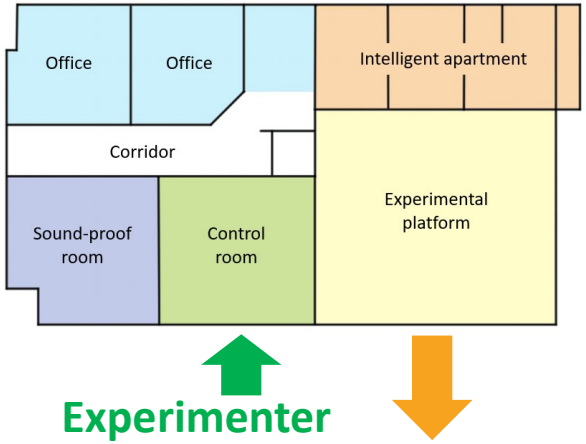
remote space

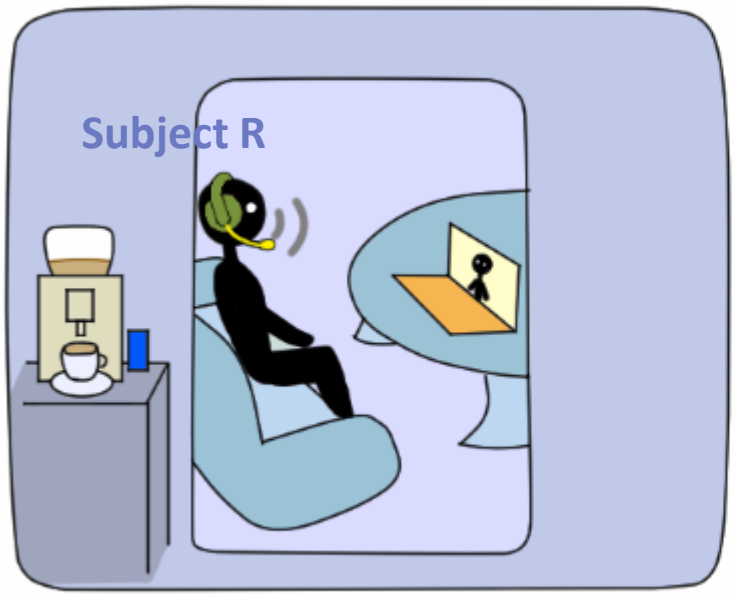


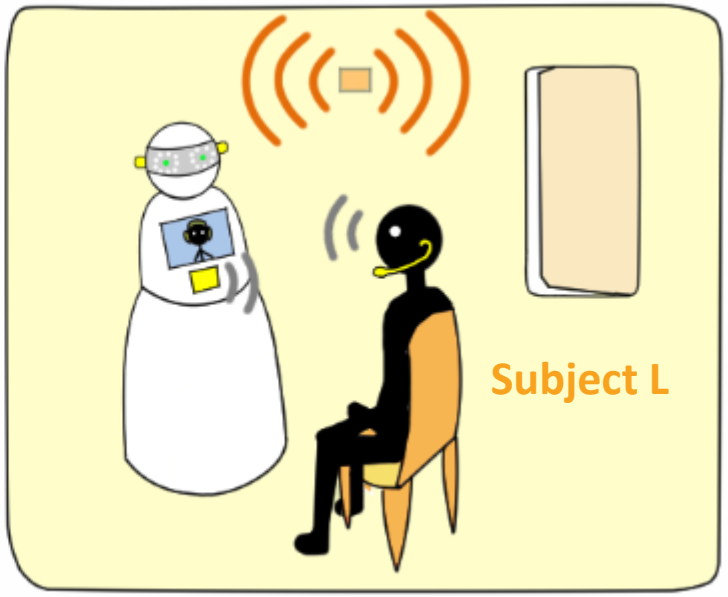
local space

Test conditions

A - quiet in both rooms
B - noise in the remote space
C - « virtual noise » in the local space
D - noise in the local space



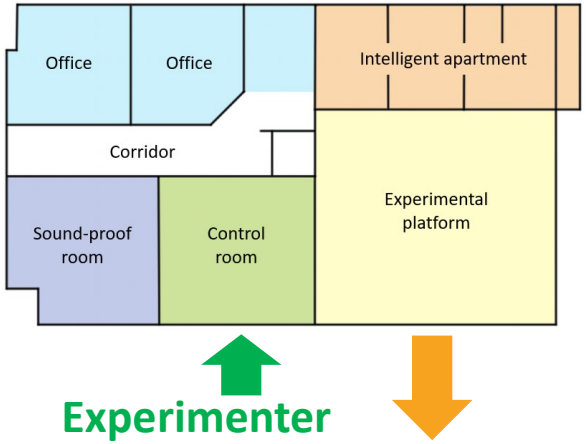


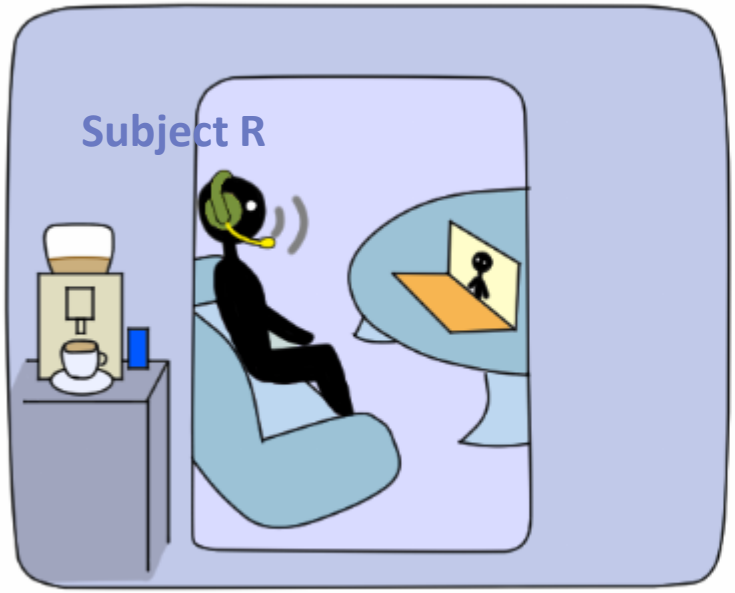


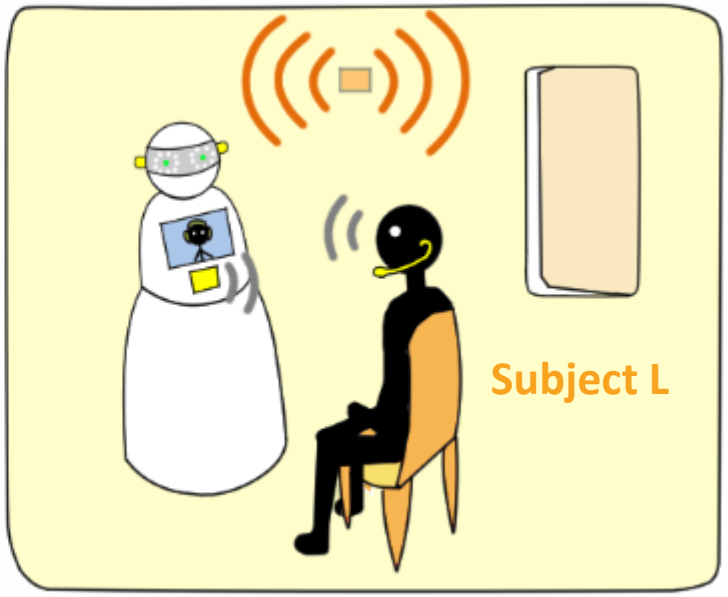
Test conditions

≈ 55 dB(A)

- **A** - quiet in both rooms
- **B** - noise in the remote space
- **C** - « virtual noise » in the local space
- **D** - noise in the local space







Expectations

Increase of voice intensity:

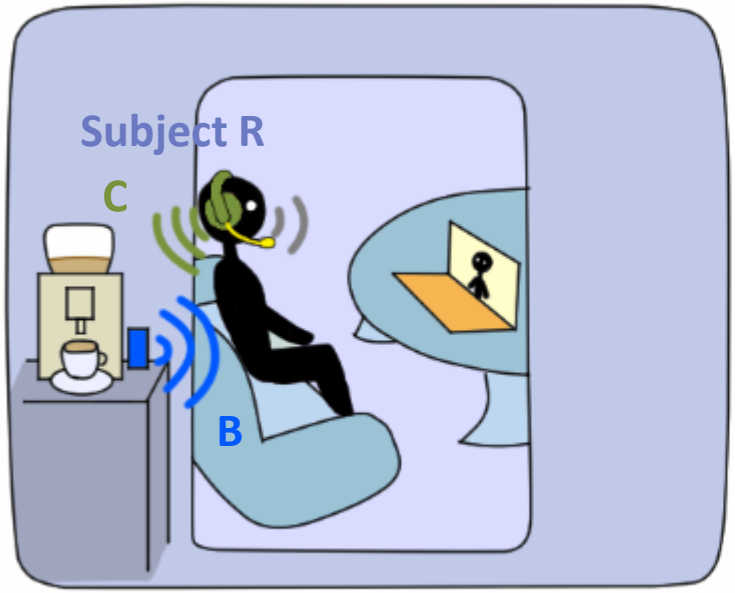
- if the subject can hear the noise
- when the other subject speaks louder

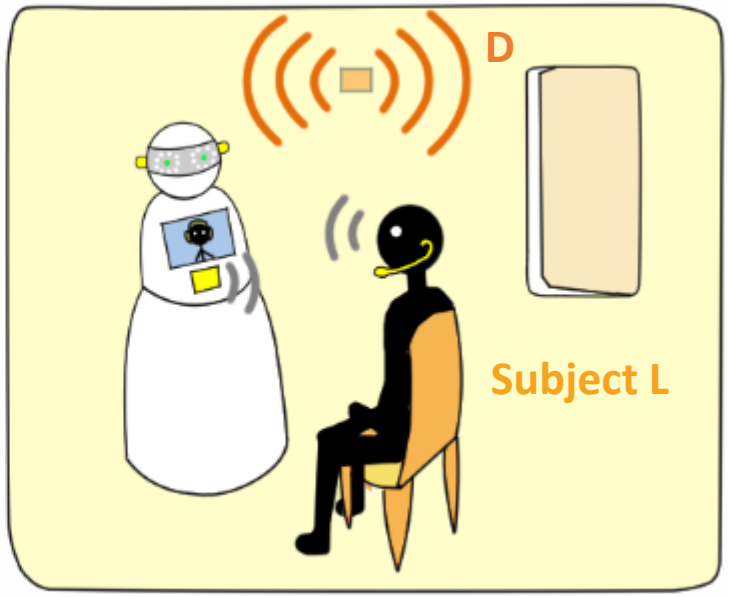
  = entrainment effect (Székely et al., 2015)

A priori variations of voice intensity

Subject

Noise	R	L
B	↗	→
C	↗	→
D	↗	↗





Pretext task: test of the robot's interface

Subjects L

Asks a list of questions & instructions

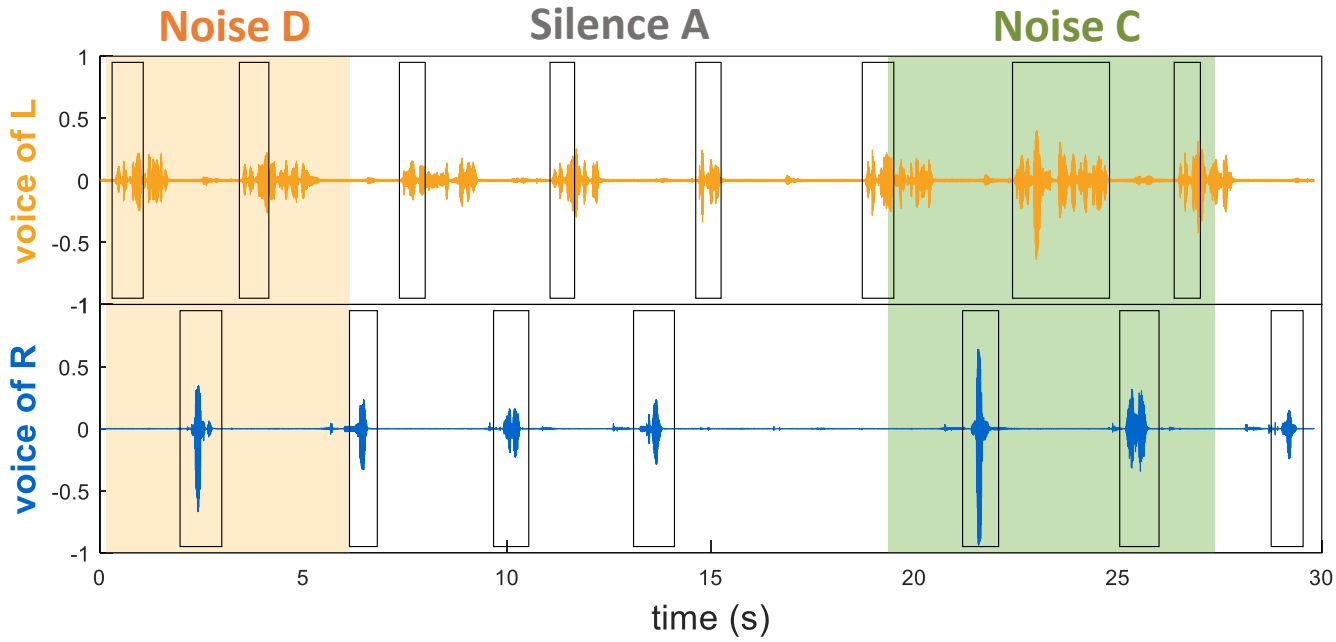
Ex: « **What is the color** of the sky? »

« Change the camera. »

Subjects R

Answers

The sky is **blue.** »



Noise apparently **incidental** (B: coffee machine & C/D: drill)

Data

Total number of extracts in each condition

	R	L
A	1521	1322
B	182	127
C	161	139
D	163	152

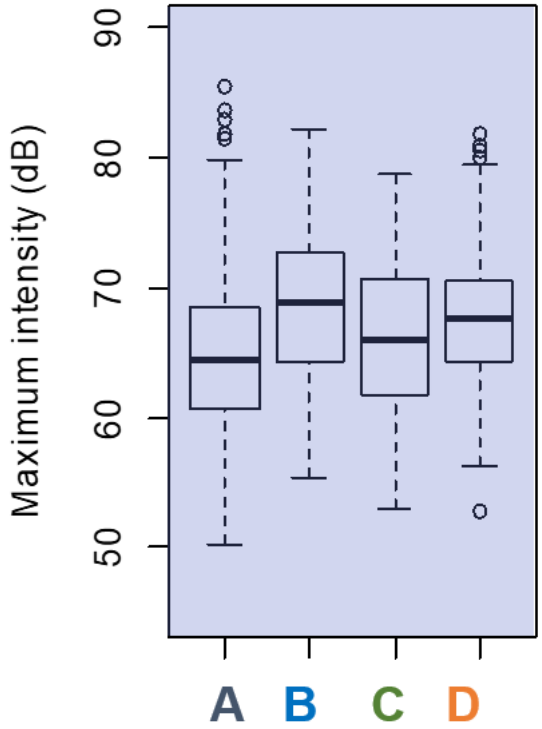
≈ 130 questions
+ 13 instructions for each session (≈ 20 min)

14 pairs of subject
(25/28 were native French speakers)

4 recordings:

- Voice of both subjects (headworn microphones) => Measures
- Microphone of the laptop of subject R => Noise B
- Monitoring signal of subject R => Noise C
- Pair of microphones of the robot => Noise D

Results



Global results for subjects R (remote pilots)

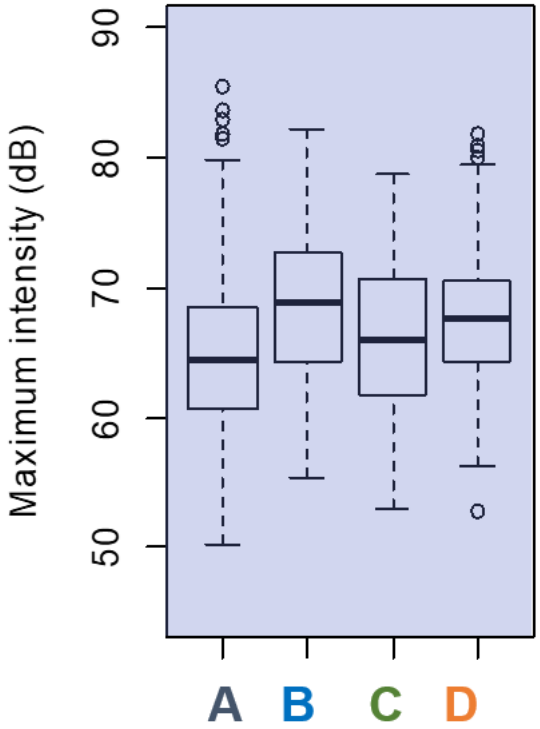
$$max\ intensity \sim noise\ condition + (1|keyword) + (1|subject)$$

	Subjects R			
Condition	**A**	**B**	**C**	**D**
Maximum intensity (dB)	64.61	+ 2.98	+ 1.19	+ 2.38
Standard error (dB)	1.11	0.36	0.37	0.36
Number of extracts	1263	160	146	150
Statistical significance	$\chi^2(1) = 98.47\ ;\ p < 2.2\ e^{-16}$			

Linear mixted effects model

(Winter Bodo 2013)

Results

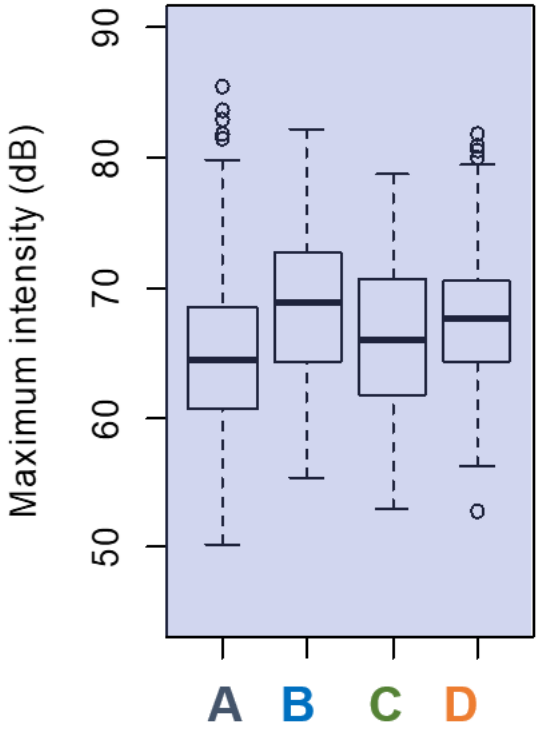


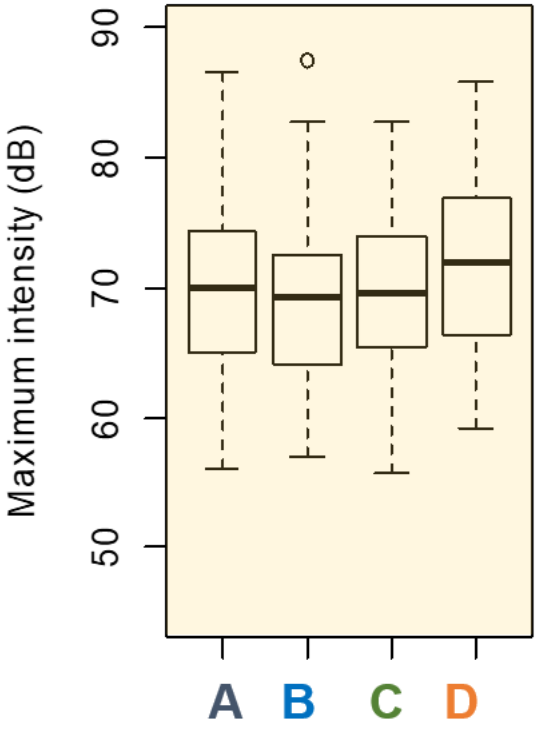
Subjects R (remote pilots)

- Augmentation of voice intensity in each noise condition
- B (Direct noise of a coffee machine) > { C, D } (Drill noise heard with headphones)
- D (R & L hear the noise) > C (Only R hears the noise)

=> entrainment effect ?

Results



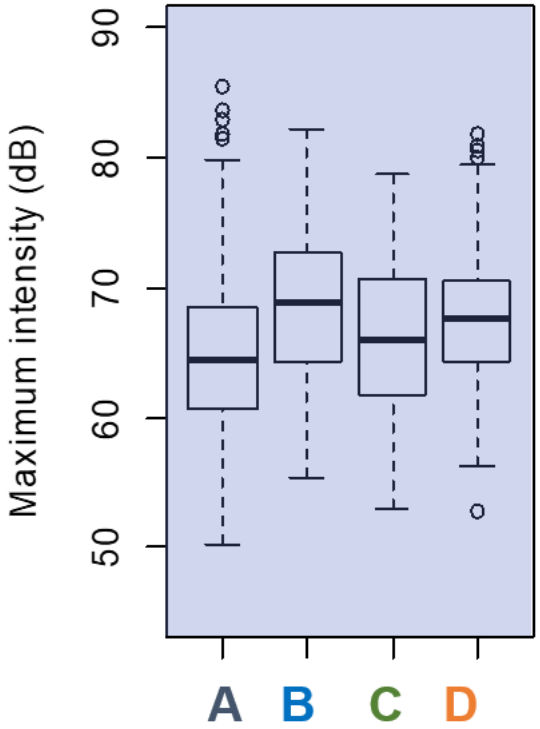


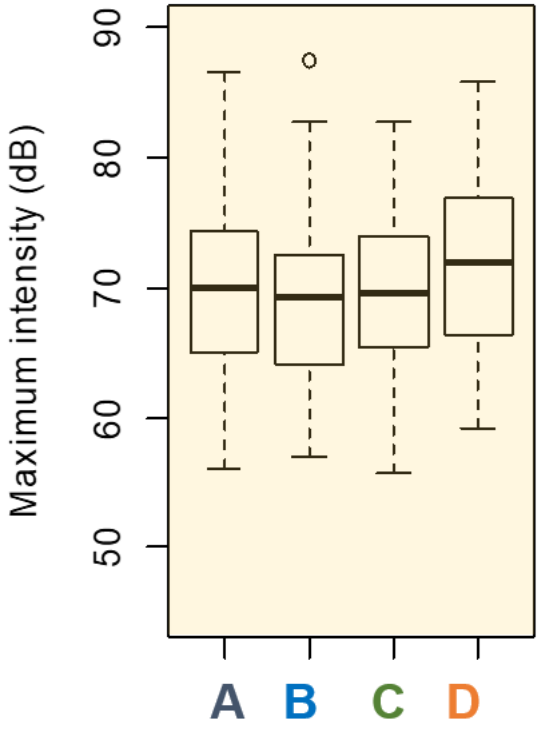
Subjects R (remote pilots)

- Augmentation of voice intensity in each noise condition
- B (Direct noise of a coffee machine) > { C, D } (Drill noise heard with headphones)
- D (R & L hear the noise) > C (Only R hears the noise)

Subjects L			
A	**B**	**C**	**D**
69.19	- 0.06	- 0.30	+ 2.24
1.86	0.21	0.21	0.20
1125	118	129	143
$\chi^2(1) = 132.59\,;\, p < 2.2\, e^{-16}$			

Results





Subjects R (remote pilots)

- Augmentation of voice intensity in each noise condition
- B (Direct noise of a coffee machine) > { C, D } (Drill noise heard with headphones)
- D (R & L hear the noise) > C (Only R hears the noise)

Subjects L (local interlocutors)

- Augmentation of voice intensity only in condition D

=> entrainment effect

Conclusion

Lombard effect:

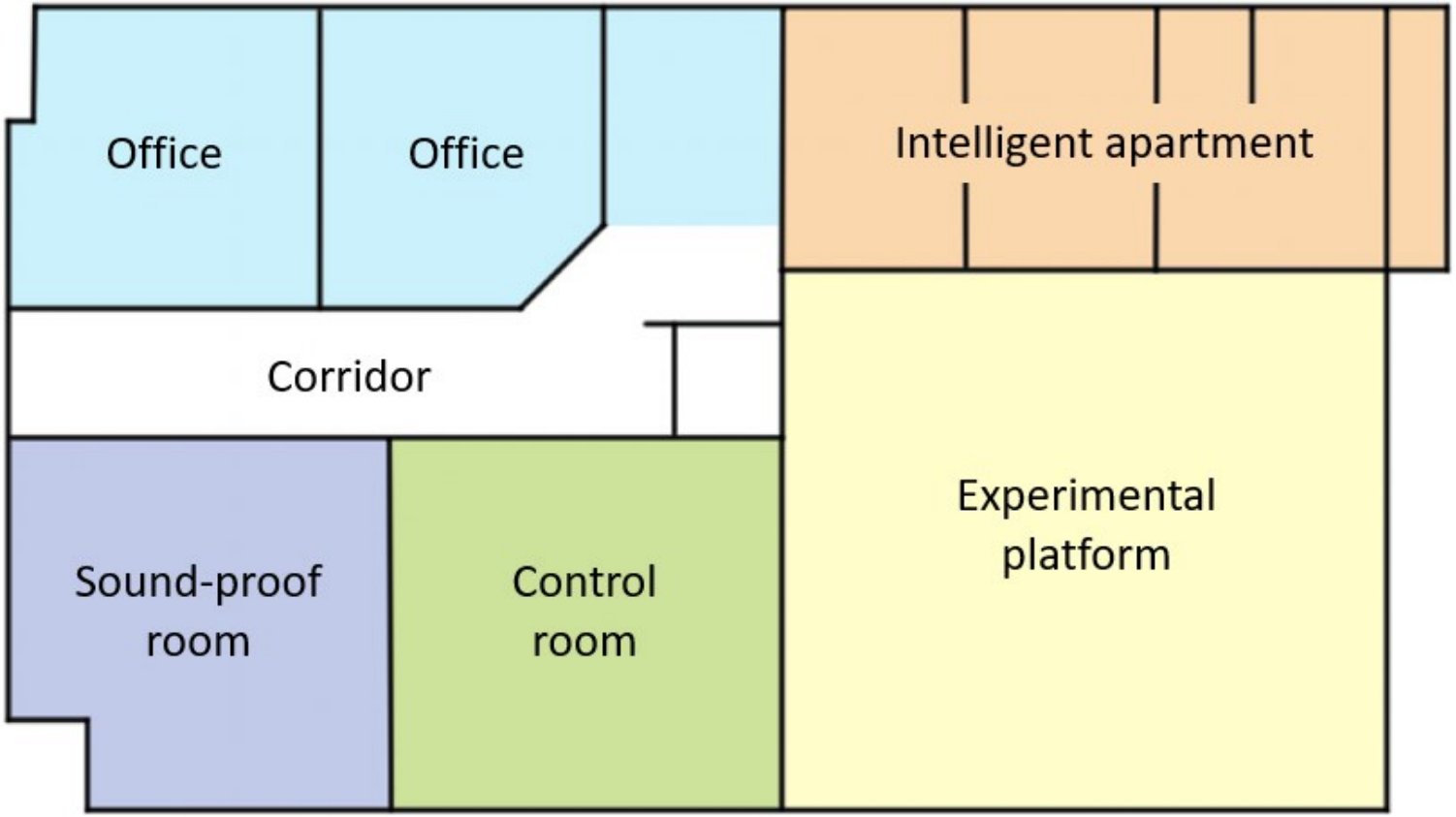
- in every noisy condition
- smaller than in (Winkworth & Davis 1997)
  (+1 to 3 dB instead of +10 dB), but significative

Perspective

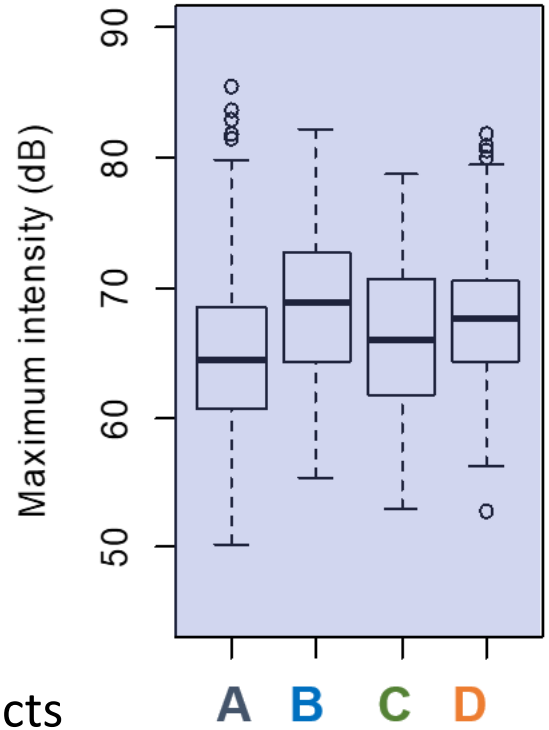
- Using different noises (duration, type, intensity)
- Check if people are able to perceive these small intensity variations
  => If so, how do they understand them?

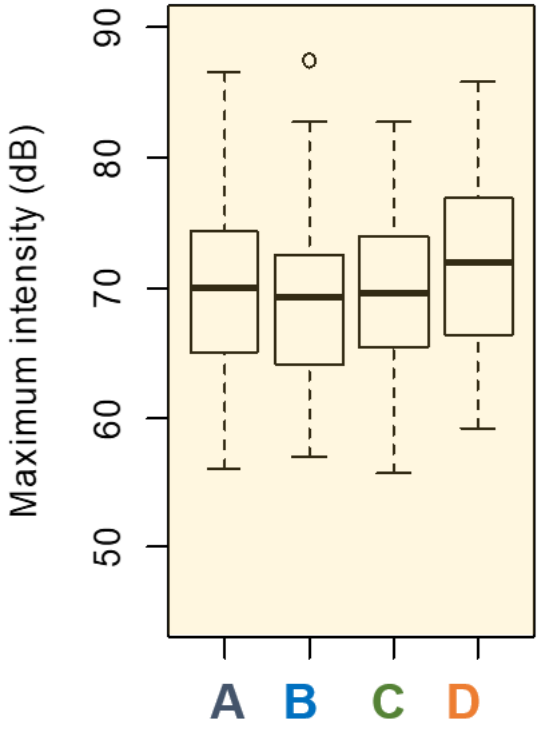
Thank you for your attention!

Living Lab Domus



Results





14 pairs of subjects

R model : $max\ intensity \sim noise\ condition + (1|keyword) + (1|subject)$

	Subjects R				Subjects L			
Condition	**A**	**B**	**C**	**D**	**A**	**B**	**C**	**D**
Maximum intensity (dB)	64.61	+ 2.98	+ 1.19	+ 2.38	69.19	- 0.06	- 0.30	+ 2.24
Standard error (dB)	1.11	0.36	0.37	0.36	1.86	0.21	0.21	0.20
Number of extracts	1263	160	146	150	1125	118	129	143
Statistical significance	$\chi^2(1) = 98.47\ ;\ p < 2.2\ e^{-16}$				$\chi^2(1) = 132.59\ ;\ p < 2.2\ e^{-16}$			