

## **Paris II – Simulating the perception of soundscapes, speech-, AM-, and FM-sounds**

**Alejandro Osses V.<sup>1</sup>, Léo Varnet, Christian Lorenzi**

<sup>1</sup>E-mail: [ale.a.osses@gmail.com](mailto:ale.a.osses@gmail.com)

Laboratoire des Systèmes Perceptifs, École Normale Supérieure (ENS),  
PSL University, Paris, France

3-4 February 2022

14<sup>th</sup> General Meeting AABBA, Vienna, Austria

### **Who are we?**

- At ENS, you may already know: Christian Lorenzi, Daniel Pressnitzer, Alain de Cheveigné, and Shihab Shamma
- The Modulation Group: Christian and Léo
- The two most important projects:
  - Hearing biodiversity: the HEARBIODIV project
  - Individualised speech perception: the fastACI project

# Hearing biodiversity: the HEARBIODIV project



Christian Lorenzi

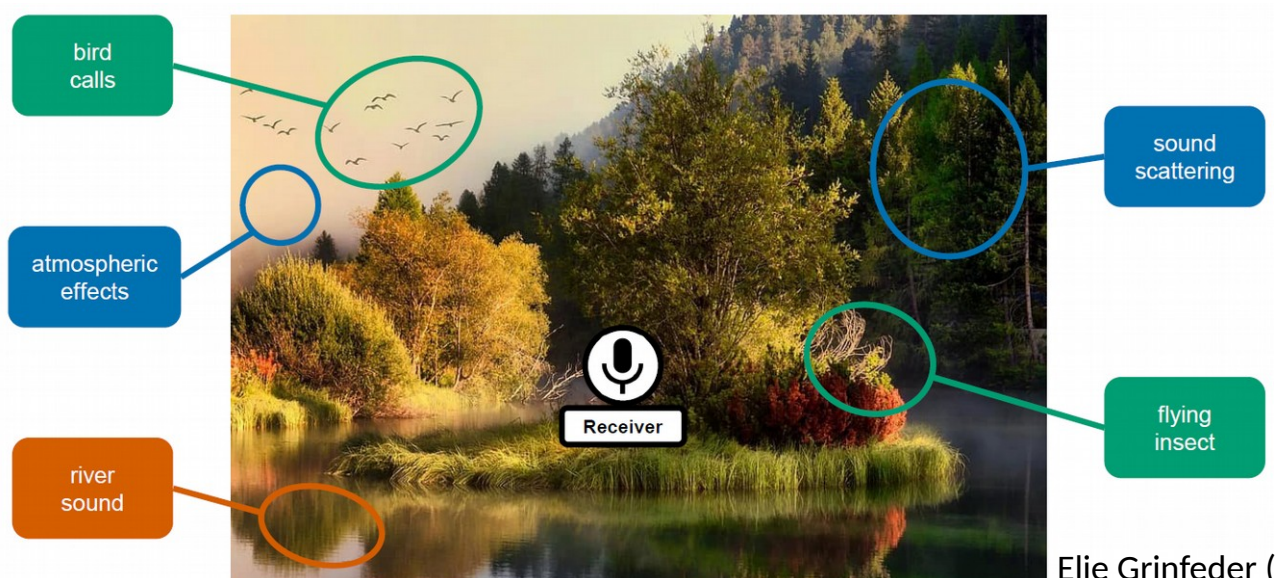
Frédéric Apoux, Elie Grinfeder, Richard McWalter & Nicole Miller

Bernie Krause, Jérôme Sueur & Régis Ferrière

<https://anr.fr/Project-ANR-20-CE28-0011>

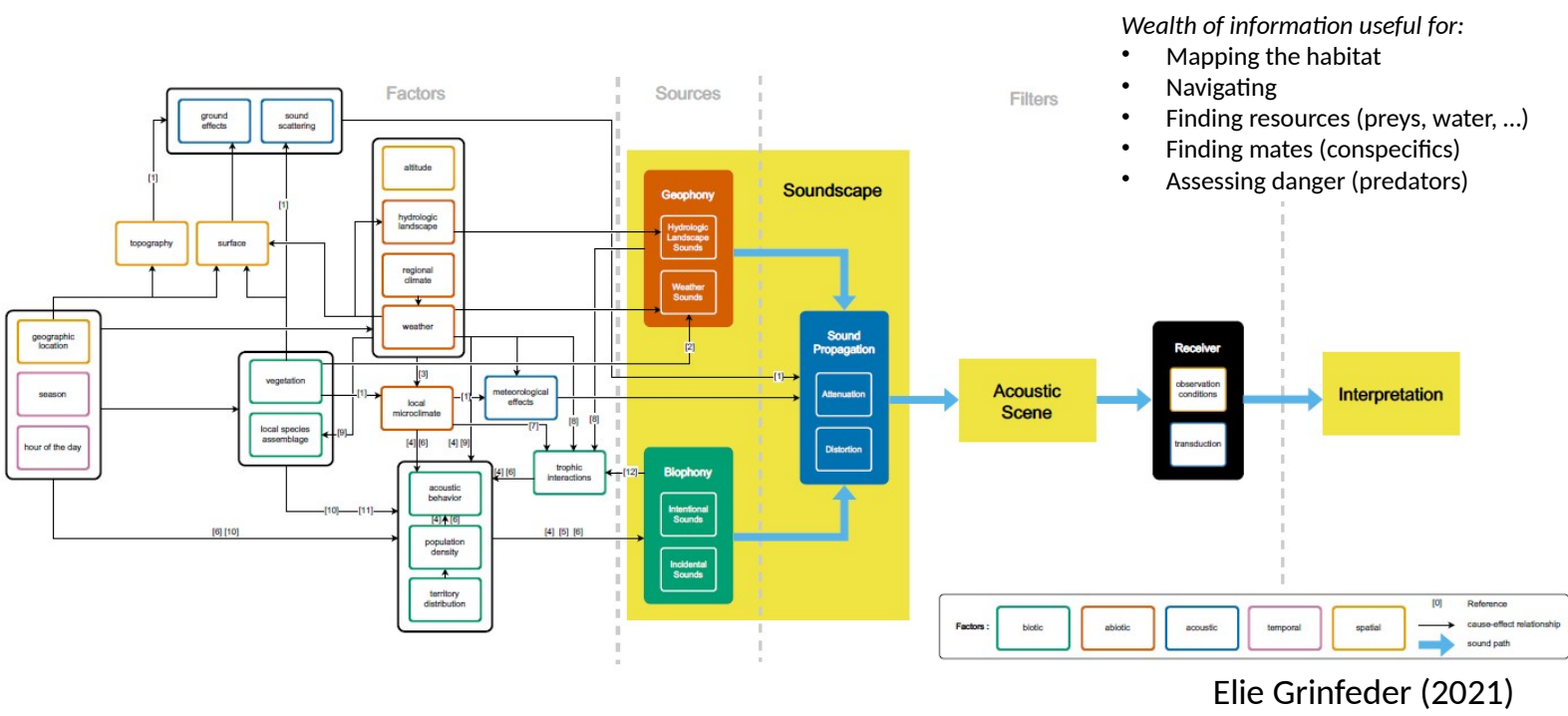
4

- **Natural soundscapes** correspond to complex *patterns* of **biological** and **geophysical** sounds *shaped* by the **sound-propagation characteristics** in the habitats, *with respect* to a point of observation ...



Elie Grinfeder (2021)

- **Natural soundscapes** are **structured** by ecological and geophysical processes and the rules of sound propagation (Farina, 2017)



## Lack of knowledge about how we perceive natural soundscapes

Main research questions adressed via psychophysical and modelling studies :

- 1) Are we able to **discriminate differences across soundscapes** corresponding to changes in **habitat** (where), **season** and **moment of the day** (when)?
  - If so, how do we achieve this (what are the **cues** & sensory **mechanisms**)?
- 2) What auditory cues and processes distinguish **biological** from **geophysical sounds** ?
- 3) Are we able to discriminate levels of **biodiversity** when listening to soundscapes?
  - Again, if so, how do we achieve this (what are the **cues** & sensory **mechanisms**)?
- 4) How does **sensorineural hearing loss** affect these capacities?

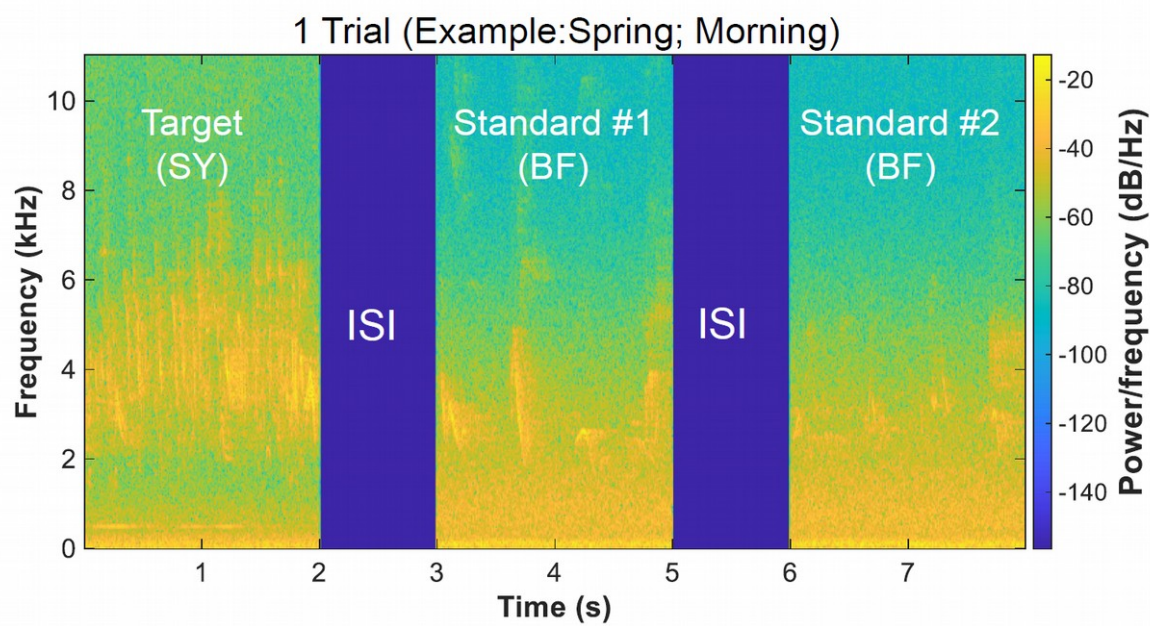


# The Hearbiodiv Database: 13 terrestrial biomes



Nicole  
Miller  
(2021)

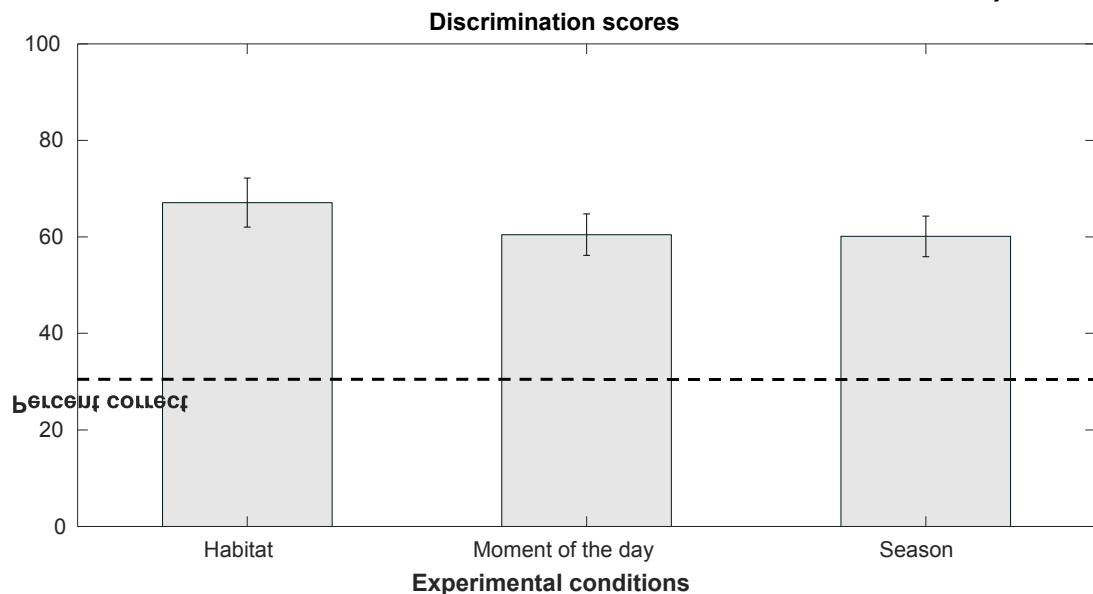
## Psychophysical experiments



# Psychophysical experiments : Preliminary results

## Experiment 1: Discrimination of Habitat, Moment of the day, Season

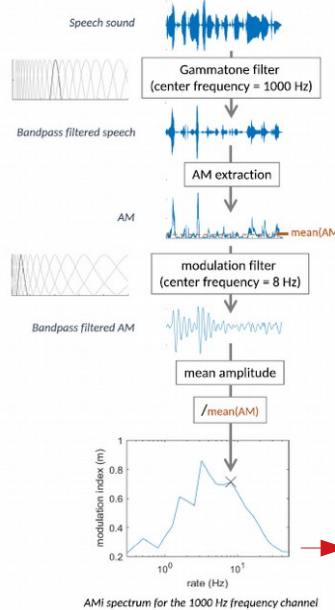
N=11 NHy participants



## Modelling studies

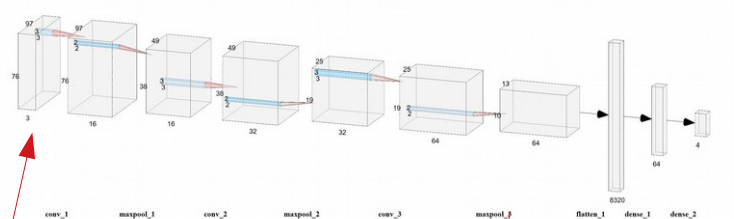
Input  
Waveforms

### Modulation filter bank model



Fred Apoux et al. (2022)

### (CNN) Convolutional Neural Network



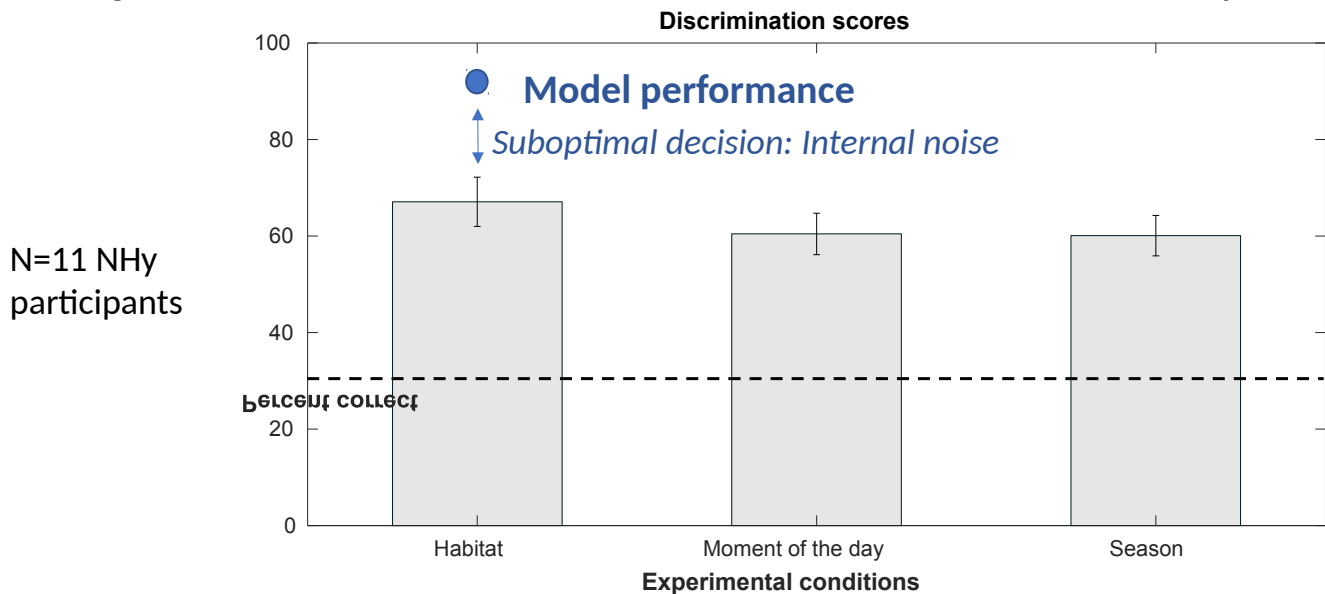
Classifier: Probabilities of belonging to each category

AM metric from the Internal representation ("slow AM cues")

Varnet et al. (2017); King et al. (2019); Thoret et al. (2020)

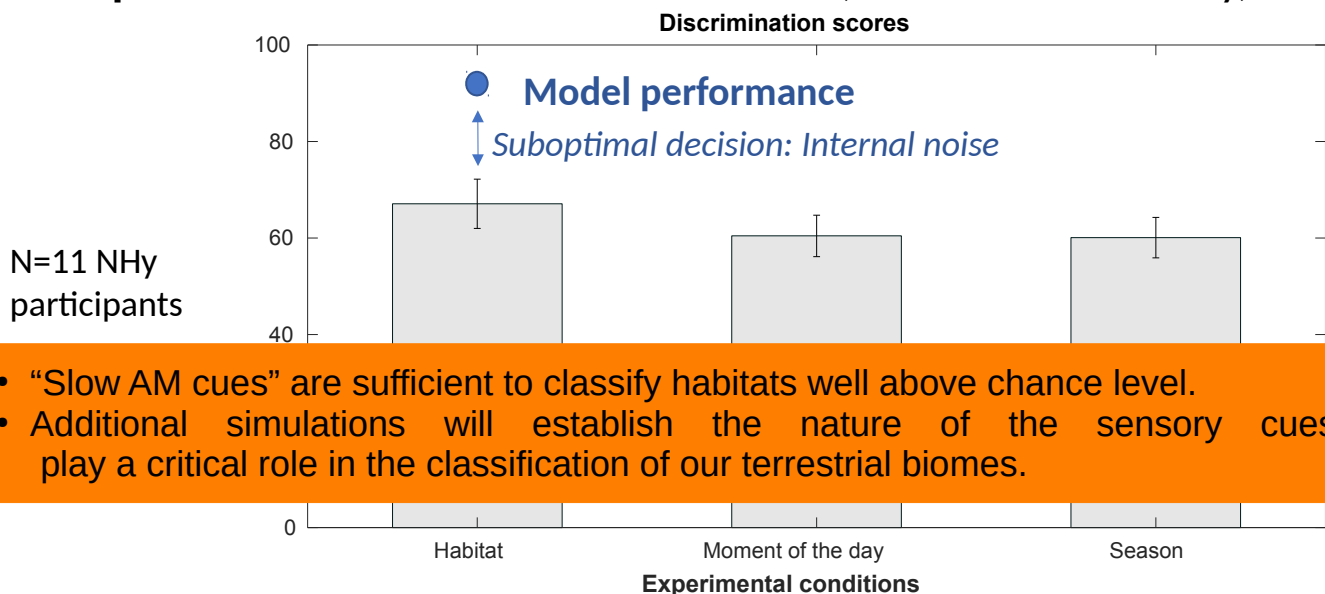
## Modelling studies

### Experiment 1: Discrimination of Habitat, Moment of the day, Season



## Modelling studies

### Experiment 1: Discrimination of Habitat, Moment of the day, Season



- “Slow AM cues” are sufficient to classify habitats well above chance level.
- Additional simulations will establish the nature of the sensory cues that play a critical role in the classification of our terrestrial biomes.

More about this in the ARO poster by Apoux et al. (ID: 721)

**C Lorenzi**



# The Hearbiodiv project

**J Sueur**



**R Ferriere**



**L Varnet**



**FM Le Tourneau**



Modelling  
Soundscapes

Perception of  
Biophony

Perception of  
Biodiversity

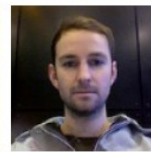
**E Grinfeder**



**N Miller**



**R McWalter**



**F Apoux**



**E Thoret**



Luciano Bertoni

Coll. A Axel; B Krause (USA)



# Exploring phoneme representations and their adaptability using fast Auditory Classification Images – fastACI

Léo Varnet (PI), Alejandro Osses Vecchi

Christian Lorenzi, Laurianne Cabrera, Willy Serniclaes

<https://anr.fr/Project-ANR-20-CE28-0004>

25

## Cracking the speech code

Speech is a **complex code** (acoustics → phonetics)

Which **acoustic cues** allow the listener to differentiate one phoneme from another?

No easy answer, due to the spectrotemporal complexity of natural speech



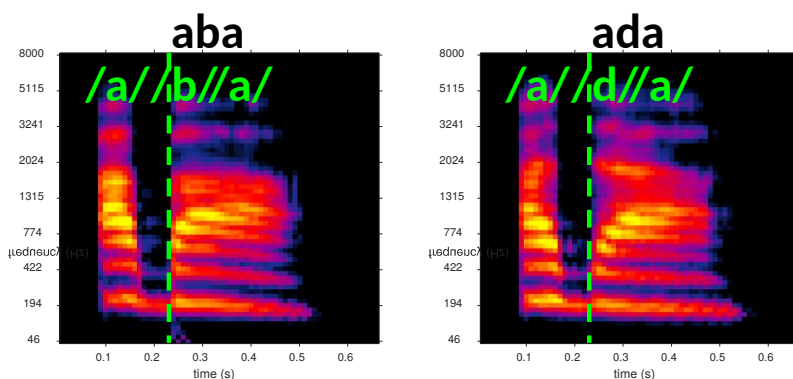


# Cracking the speech code

How do we distinguish « aba » from « ada »?

(VCVs from the Oldenburg Logatome Corpus [Wesker et al., 2005], gammatone-based TF representation)

- Many acoustical differences between « aba » and « ada »

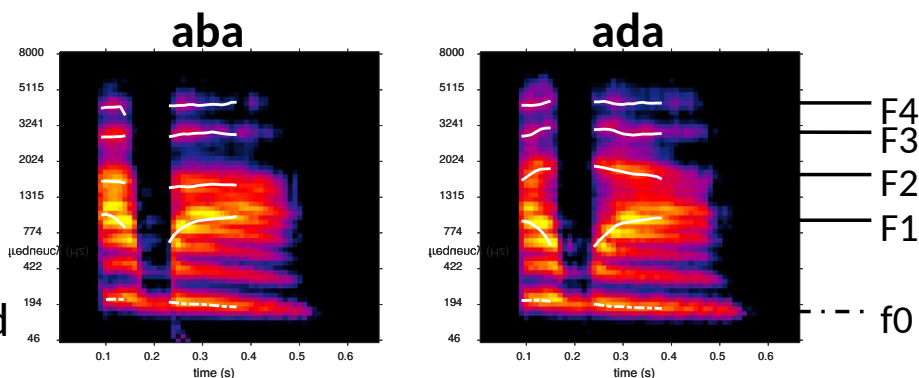


# Cracking the speech code

How do we distinguish « aba » from « ada »?

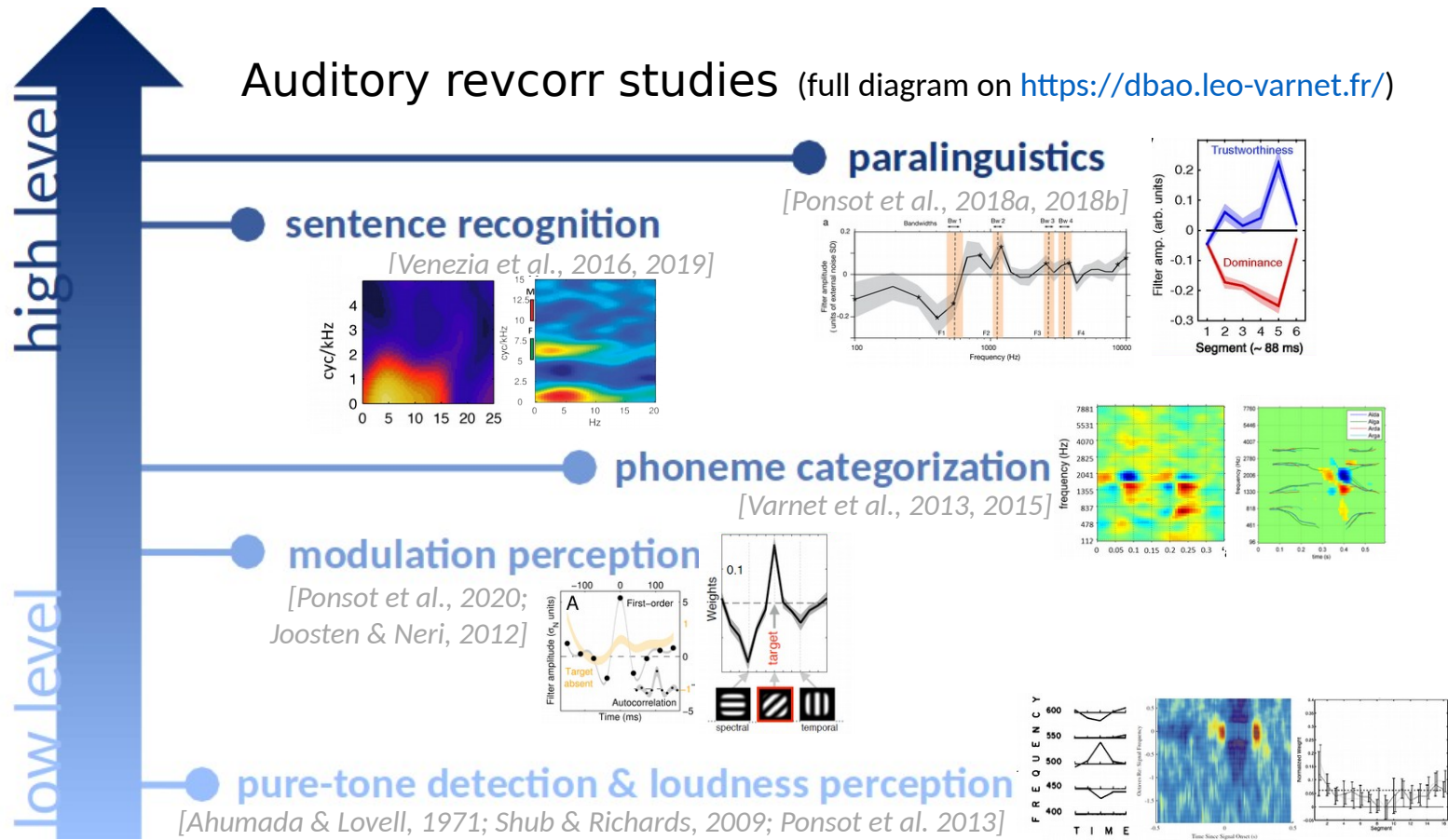
(VCVs from the Oldenburg Logatome Corpus [Wesker et al., 2005], gammatone-based TF representation)

- Many acoustical differences between « aba » and « ada » (e.g. formant trajectories)



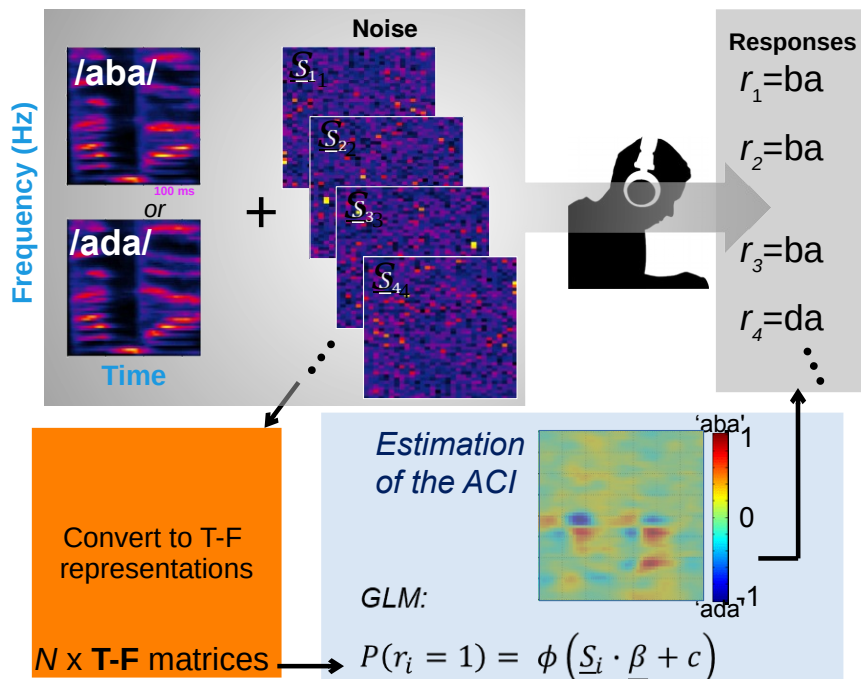
- Which ones are actually used by the auditory system?

# Auditory revcorr studies (full diagram on <https://dbao.leo-varnet.fr/>)



## Phoneme discrimination using the ACI method

- The reverse correlation method allows to experimentally assess the cues that are relevant in a discrimination task: ACIs



### Experimental design:

- Noise fixed, variable speech level
- 1-up, 1-down weighted procedure (Kaernbach, 1991) to track 70.7% correct responses: down step = 1 dB, up step = 2.4 dB
- Little level roving of  $\pm 2.5$  dB

Before: transformed up-down (Levitt, 1971)  
Before: No roving

### Participant's response

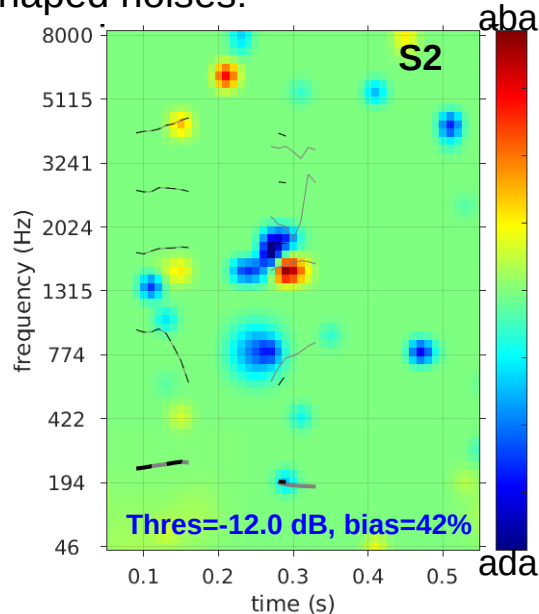
	'ba'	'da'
Presented 'ba' ("target")	Hit	Miss
Presented 'da' ("no target")	False Alarm	Correct Rejection

**ACI is built from (H+FA) - (M+CR)**

One weighting function per participant, <sup>49</sup> the Auditory Classification Image (ACI)

## Phoneme discrimination using the ACI method

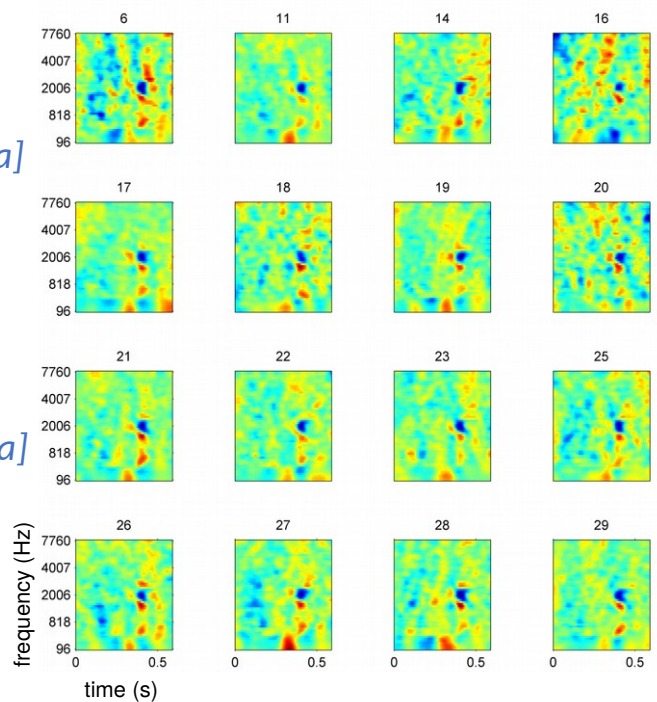
- An ACI from a previous study (Osses & Varnet 2021, DAGA).
  - Aba / ada discrimination, female speaker
  - Speech-shaped noises:



50

## ACIs for a group of participants

- ACI for a /da/-/ga/ categorisation task measured on 16 individuals [Varnet et al., 2015a]
- Comparison between groups of listeners:
  - **Musicians** [Varnet et al., 2015b]
  - **Listeners with dyslexia** [Varnet et al. 2016a]
  - Hearing-impaired listeners [Varnet et al., 2019].



# Methods

The method is nice but it requires too many trials

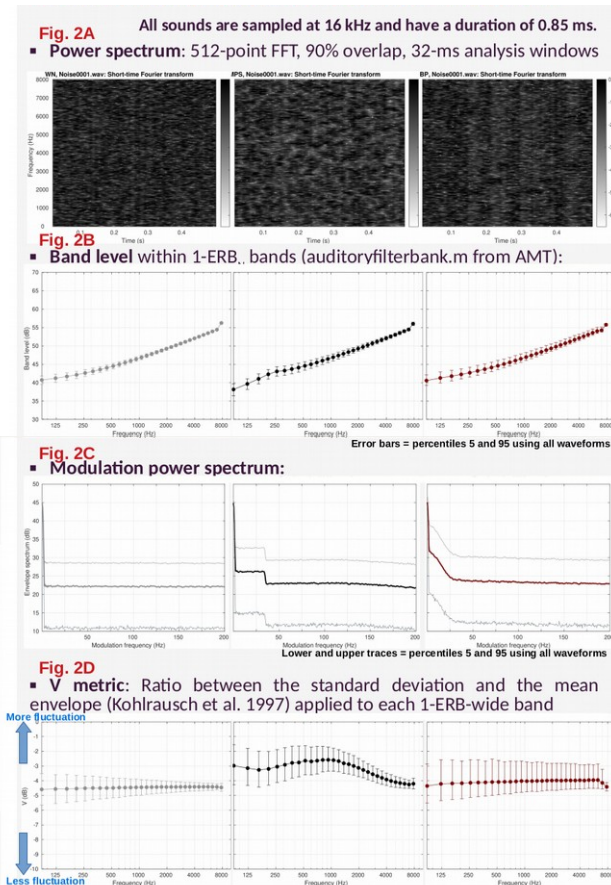
## During data collection:

**Stimuli:** Target sounds (65 dB SPL) in background noise

**Task:** Is it /aba/ or /ada/?

**SNR:** Adaptive procedure to reach a score of 70.7%

**Roving:**  $\pm 2.5$  dB



# Methods

The method is nice but it requires too many trials

## During data collection:

**Stimuli:** Target sounds (65 dB SPL) in background noise

**Task:** Is it /aba/ or /ada/?

**SNR:** Adaptive procedure to reach a score of 70.7%

**Roving:**  $\pm 2.5$  dB

## During the ACI derivation

**T-F conversion:** FFT, ERB spaced

**Task:** Fitting responses to /aba/ or /ada/ (binary decision)

**GLM fitting procedure:** Normalisation, fitting algorithm - priors

We plan to collect data  $N=10$  listeners

But, while waiting to test the real listeners...

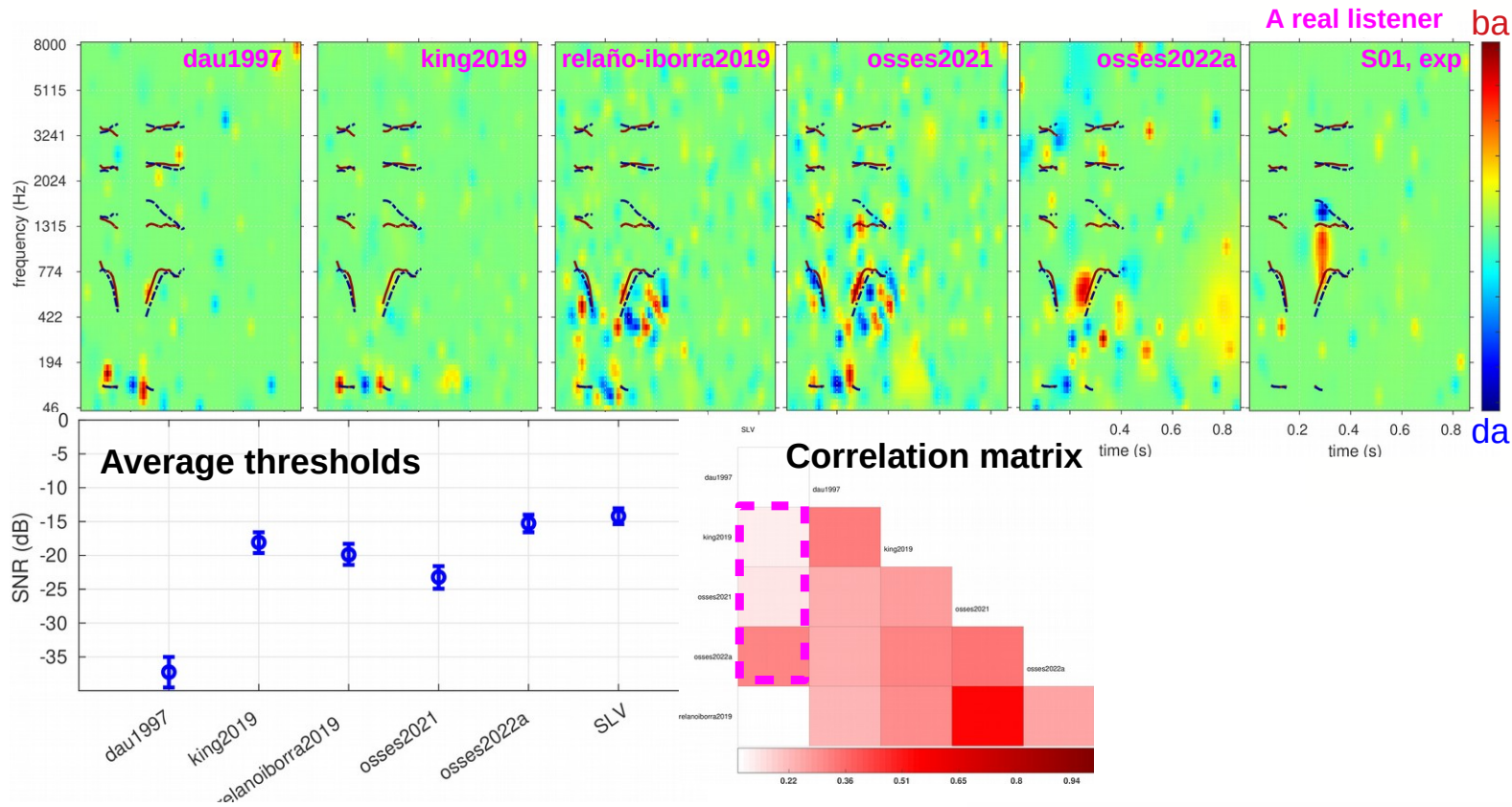


**fastACI toolbox v1.0:** a MATLAB toolbox for investigating auditory perception using reverse correlation (<https://github.com/aosses-tue/fastACI>, cf. **ARO poster 800**)



# Artificial listeners

/aba/ - /ada/ in white noise

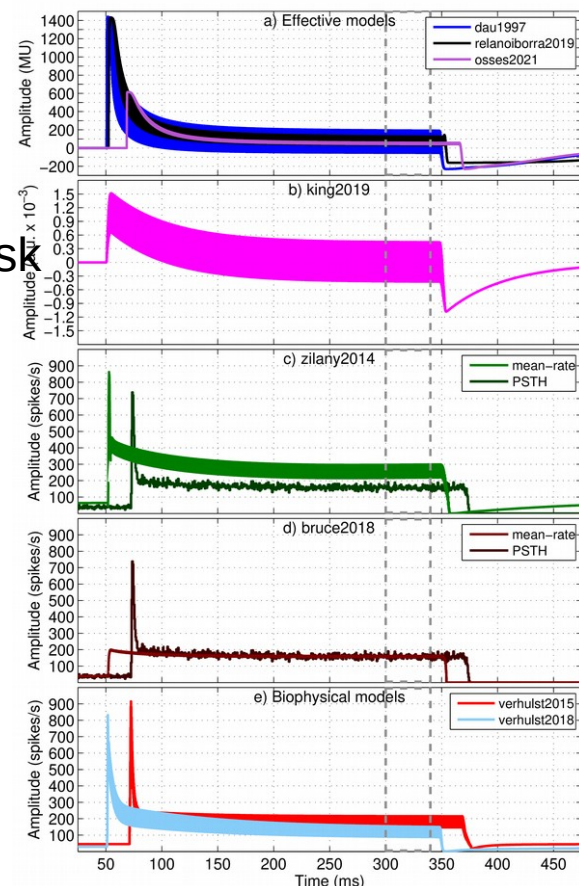


## Summary

- Reverse correlation method:  
Trial by trial analysis for a supra-threshold task
- Between-participant variability:  
Many analyses are possible
- Auditory models: Can be strategically used,  
but a warning:  
**They can provide very different results**

Fig. 10 from Osses, Varnet, Carney,  
Dau, Bruce, Verhulst, & Majdak (2022, ArXiv)

Fully reproducible figures with AMT 1.1



# Thank you for your attention!

And thanks to our  
collaborators:



Christian  
Lorenzi



Emmanuel  
Ponsot



Laurianne  
Cabrera



Diane  
Lazard



Michel  
Hoen



Fanny  
Meunier

9:30 AM

## Symposium 7: Individual Listening Strategies: From Physiological to Perceptual Responses in Experiments and Simulations

Sun 9:30 AM - 11:30 AM

### Individual Listening Strategies: From Physiological to Perceptual Responses in Experiments and Simulations

Sun 9:30 AM - 11:30 AM • Young Investigator Symposia

[hide presenters](#)

[abstract/session info](#)

[add to calendar](#)

**Alejandro Osses Vecchi, Ph.D.** *Chair*  
École Normale Supérieure, PSL University, CNRS

**Amarins Heeringa, Ph.D.** *Presenter*  
Carl Von Ossietzky University  
Auditory Nerve Fiber Discrimination and Representation of Naturally-Spoken Vowels in Noise  
9:30 AM - 9:45 AM

**Dyan Ramekers, Ph.D.** *Presenter*  
University Medical Center Utrecht  
Variability in Auditory Nerve Responses in Animals: How to Extrapolate to Individual Human CI Users?  
9:45 AM - 10 AM

**Margriet van Gendt, M.Sc.** *Presenter*  
Leiden University Medical Center, the Netherlands  
Power Law Adaptation in the Electrically Stimulated Auditory Nerve  
10 AM - 10:15 AM

**Gerard Encina-Llamas, Ph.D.** *Presenter*  
Technical University of Denmark  
Re-Interpreting Experimental Results With Computer Model Simulations: The Case of a Peripheral Compression Study.  
10:15 AM - 10:30 AM

**Laurianne Cabrera, Ph.D.** *Presenter*  
Université de Paris - CNRS  
The Impact of Internal Noise on Amplitude Modulation Detection for Children and Young Adults: Empirical and Simulation Data  
10:30 AM - 11 AM

**Léo Varnet, Ph.D.** *Presenter*  
Laboratoire des Systèmes Perceptifs (ENS Paris/CNRS)  
Auditory Classification Images: A Psychophysical Paradigm to Explore Listening Strategies in Phoneme Perception  
11 AM - 11:15 AM

**Alejandro Osses Vecchi, Ph.D.** *Presenter*  
École Normale Supérieure, PSL University, CNRS  
Trial-by-Trial Analysis of Phoneme-in-Noise Perception using a Model of Monaural Auditory Processing  
11:15 AM - 11:30 AM

## Almost all references...

- Kaernbach, C. (1991). Simple adaptive testing with the weighted up-down method. *Percept. Psychophys.*, 49(3), 227–229. <https://doi.org/10.3758/BF03214307>
- King, A., Varnet, L., & Lorenzi, C. (2019). Accounting for masking of frequency modulation by amplitude modulation with the modulation filter-bank concept. *J. Acoust. Soc. Am.*, 145, 2277–2293. <https://doi.org/10.1121/1.5094344>
- Osses, A., & Varnet, L. (2021). Consonant-in-noise discrimination using an auditory model with different speech-based decision devices. *DAGA*, 298–301.
- Osses, A., Varnet, L., Carney, L., Dau, T., Bruce, I., Verhulst, S., & Majdak, P. (2022). A comparative study of eight human auditory models of monaural processing. *ArXiv*, 1–24. <https://arxiv.org/abs/2107.01753>
- Osses, A., & Varnet, L. (2021). *fastACI toolbox: the MATLAB toolbox for investigating auditory perception using reverse correlation (v1.0)*. <https://doi.org/10.5281/zenodo.5500139>
- Osses, A., & Kohlrausch, A. (2021). Perceptual similarity between piano notes: Simulations with a template-based perception model. *J. Acoust. Soc. Am.*, 149, 3534–3552. <https://doi.org/10.1121/10.0004818>
- Varnet, L., Knoblauch, K., Serniclaes, W., Meunier, F., & Hoen, M. (2015). A psychophysical imaging method evidencing auditory cue extraction during speech perception: A group analysis of auditory classification images. *PLoS One*, 10(3), 1–23. <https://doi.org/10.1371/journal.pone.0118009>
- Varnet, L., Langlet, C., Lorenzi, C., Lazard, D., & Micheyl, C. (2019). High-frequency sensorineural hearing loss alters cue-weighting strategies for discriminating stop consonants in noise. *Trends in Hearing*, 23, 1–18. <https://doi.org/10.1177/2331216519886707>
- Varnet, L., Meunier, F., Trollé, G., & Hoen, M. (2016). Direct viewing of dyslexics' compensatory strategies in speech in noise using Auditory Classification Images. *PLoS One*, 11(4), 1–17. <https://doi.org/10.1371/journal.pone.0153781>
- Varnet, L., Ortiz-Barajas, M., Guevara, R., Gervain, J., & Lorenzi, C. (2017). A cross-linguistic study of speech modulation spectra. *J. Acoust. Soc. Am.*, 142, 1976–1989. <https://doi.org/10.1121/1.5006179>
- Varnet, L., Wang, T., Peter, C., Meunier, F., & Hoen, M. (2015). How musical expertise shapes speech perception: Evidence from auditory classification images. *Sci. Rep.*, 5, 14489. <https://doi.org/10.1038/srep14489>