



Auditory Classification Images: A Psychophysical Paradigm to Explore Listening Strategies in Phoneme Perception

Léo Varnet, Alejandro Osses Vecchi

ARO 2022

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



ଏ ଲାଠିକି ଲାଠିକି ଲାଠିକି
ଲାଠିକି ଲାଠିକି ଲାଠିକି
ଲାଠିକି ଲାଠିକି ଲାଠିକି
ଲାଠିକି ଲାଠିକି ଲାଠିକି
ଲାଠିକି ଲାଠିକି ଲାଠିକି
ଲାଠିକି ଲାଠିକି ଲାଠିକି
ଲାଠିକି ଲାଠିକି ଲାଠିକି

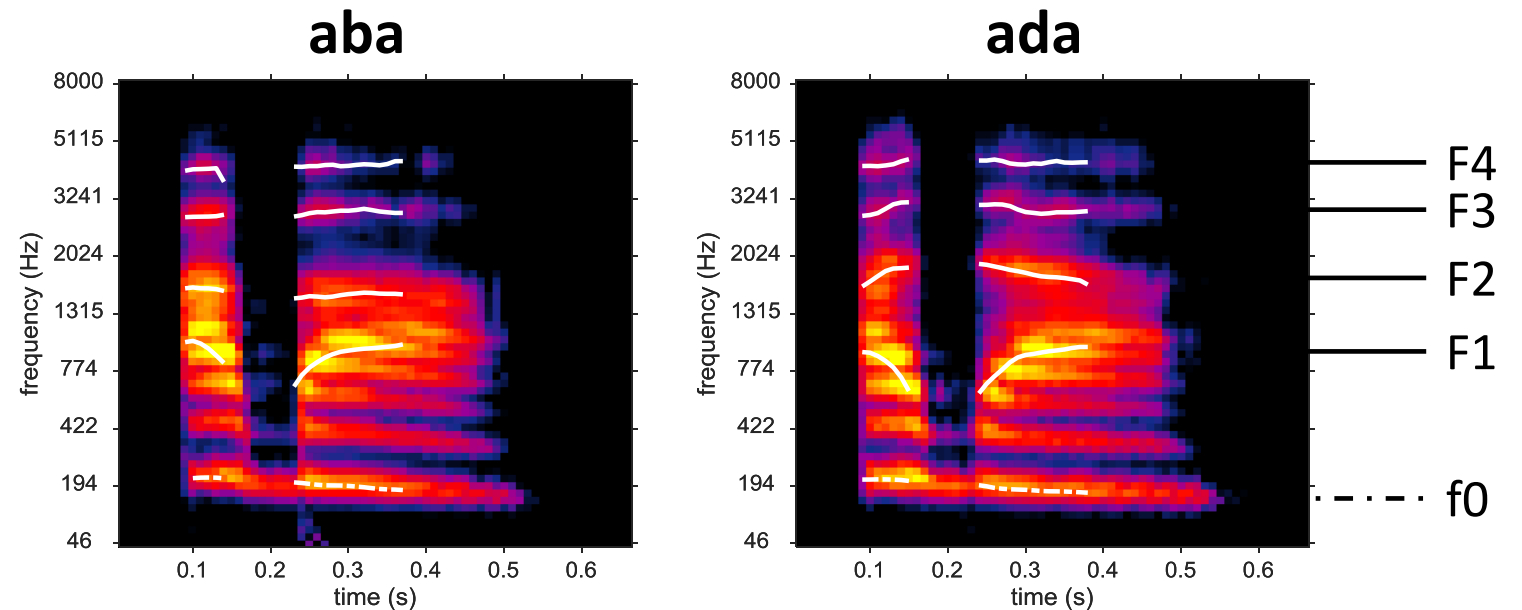


Cracking the speech code

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

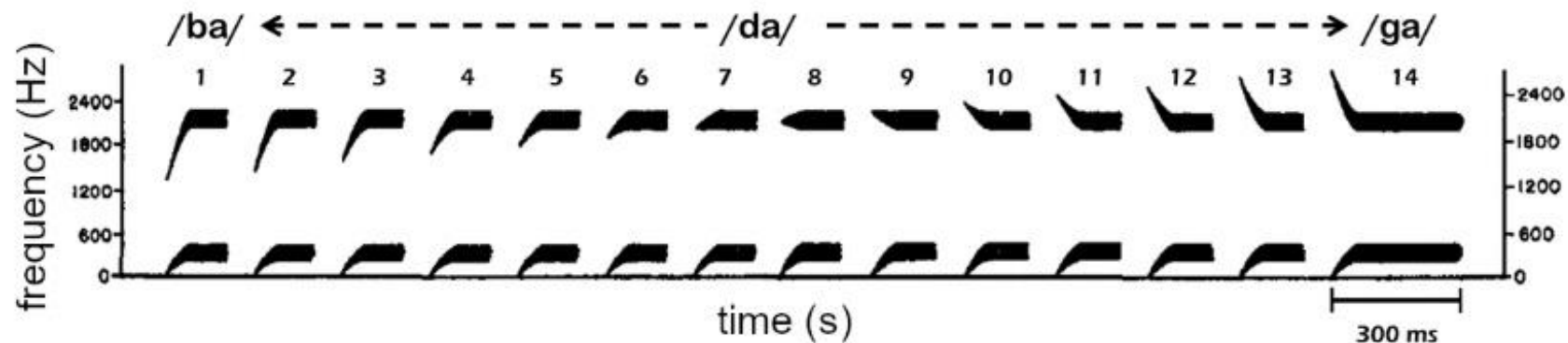
(VCVs from the Oldenburg Logatome Corpus [[Wesker et al., 2005](#)], gammatone-based TF representation)

- Many acoustical differences between « aba » and « ada » (e.g. formant trajectories)
- Which ones are actually used by the auditory system?



Cracking the speech code

One solution: using **reduced speech** [Fletcher, 1922; Li & Allen, 2012; ...], or **synthetic speech** [Liberman et al., 1952; ...]



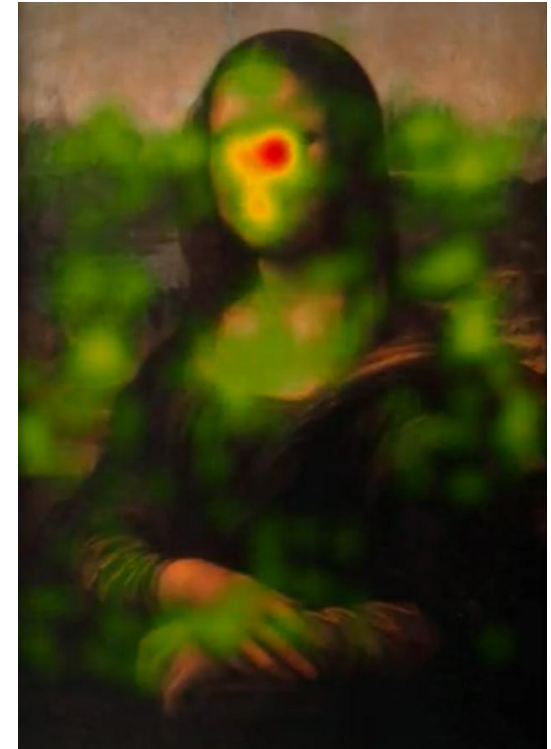
Proof that the F2 onset is a cue for categorizing speech sounds into /b/ or /d/?

Problem: Results may not be representative of **natural** speech perception...

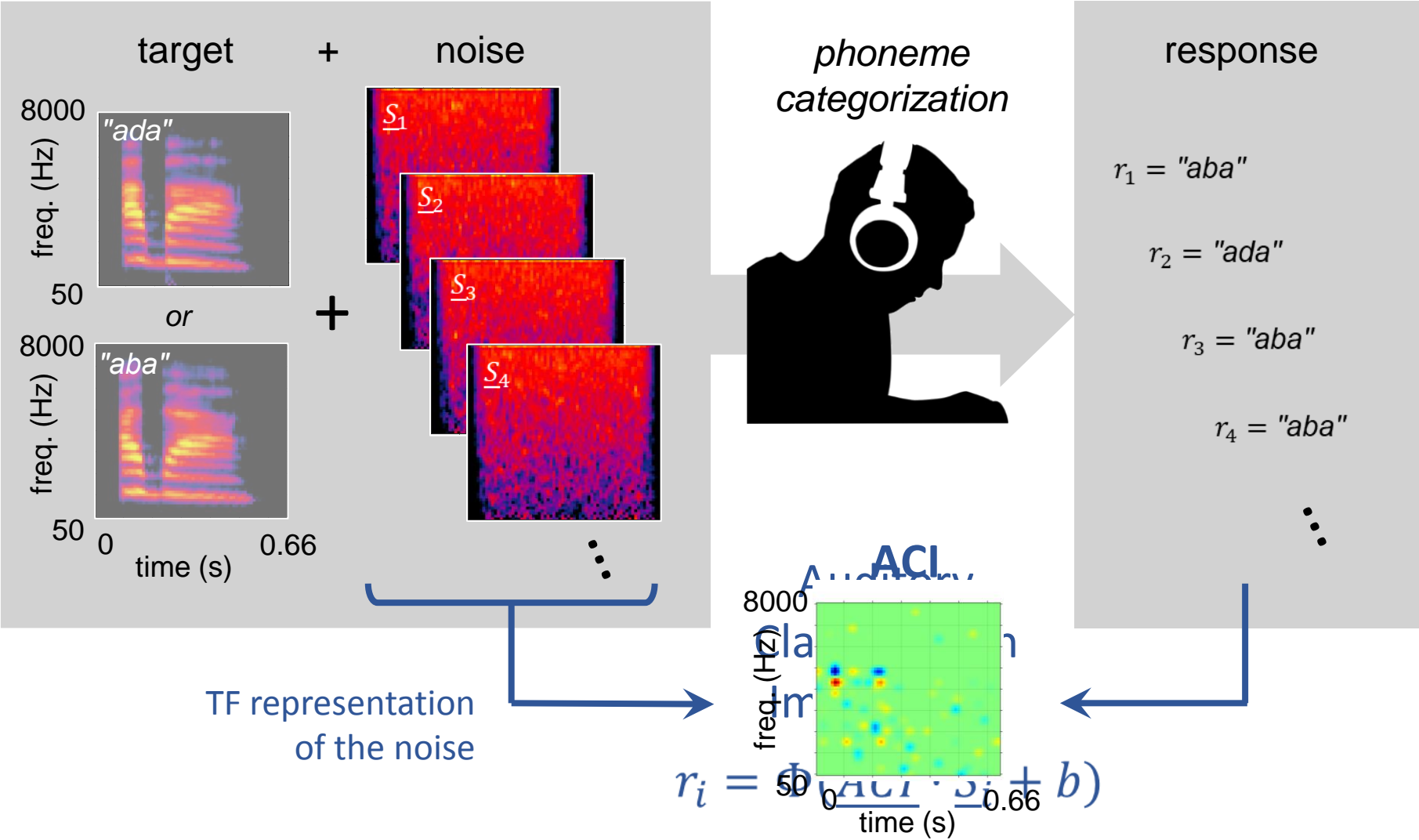
Auditory Classification Images (ACI)

- The need for an “**ear-tracker**” *[Varnet et al., 2013...]*
- Based on the **reverse correlation approach** developed in visual psychophysics:

Introducing **random fluctuations** in the stimuli and measure **how they affect the behavioural responses** of the participant on a **trial-by-trial basis**.



Using the reverse correlation paradigm to identify the acoustic cues underlying phoneme perception in noise [Varnet et al. 2013, 2015, 2016, 2019...].



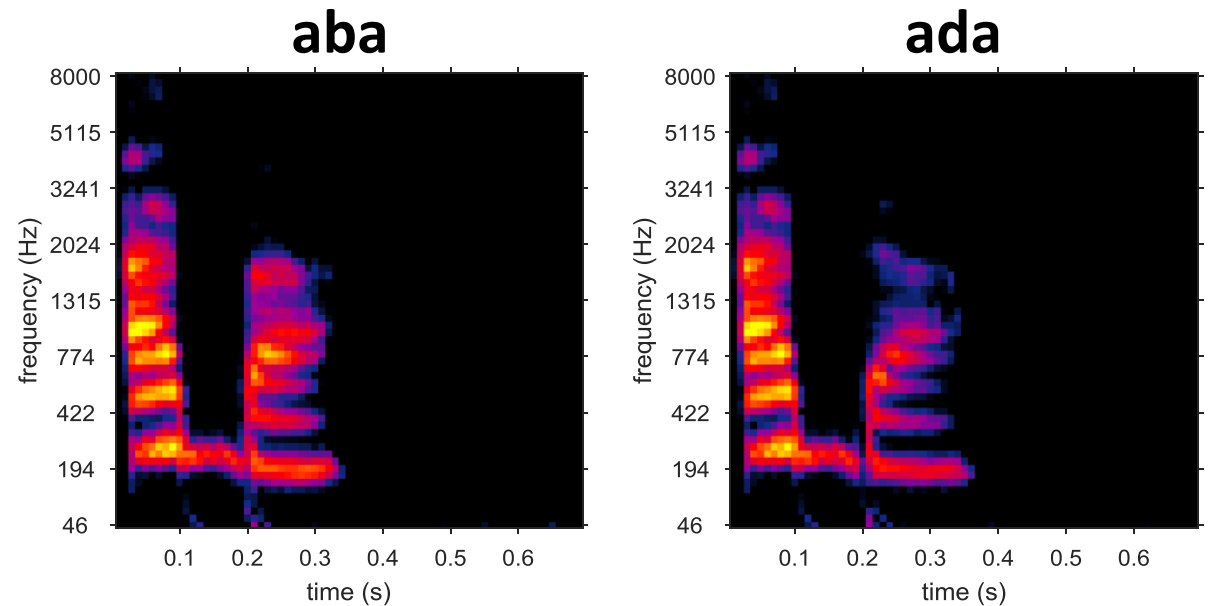
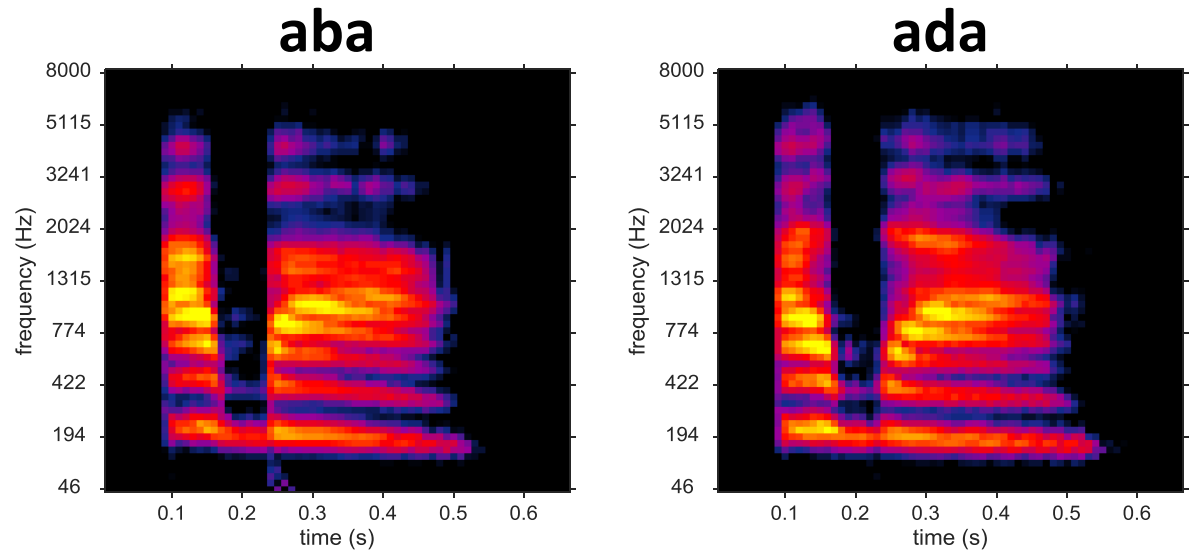
Methods

Stimuli: Target sounds (65 dB SPL) in an additive Gaussian noise.

Task: Categorizing stimuli as /aba/ or /ada/

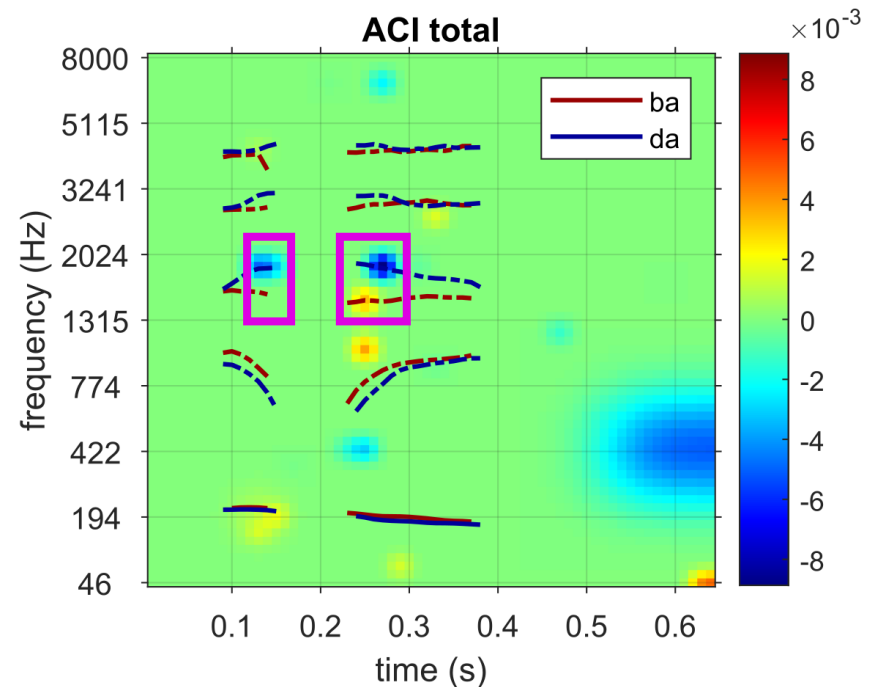
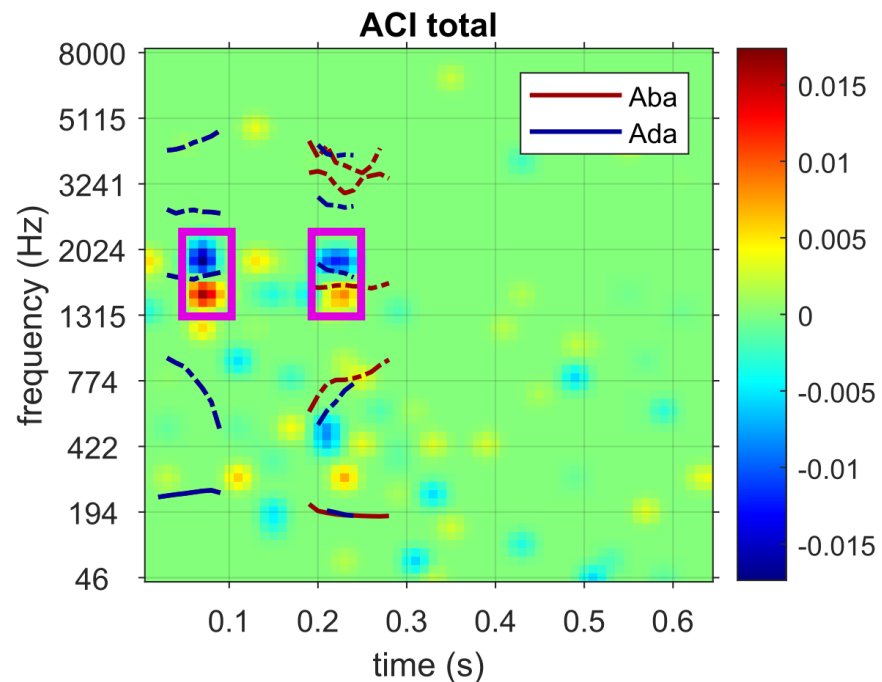
SNR: Adapted continuously to ensure a correct response rate of **75 %**

Roving: **no roving**



ACIs (one participant, two different /aba/-/ada/ pairs)

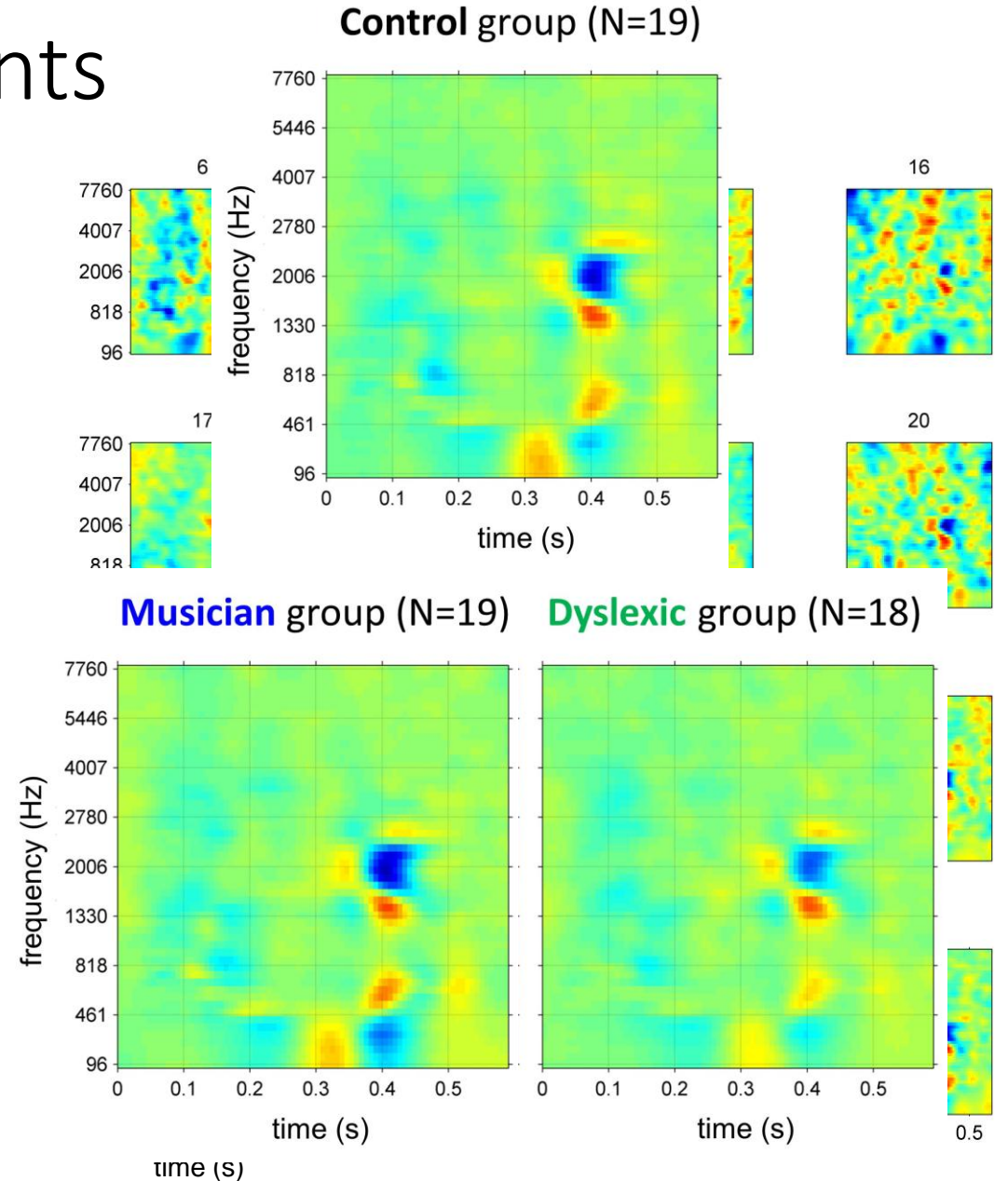
Clusters of **positive** and **negative** weights corresponding to the acoustic cues.



- Confirms that **F2 onset** is a cue for classifying phonemes into /b/ or /d/ [*Lieberman et al., 1954*]
- Other (unexpected) cues are involved in this task.

ACIs for a group of participants

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



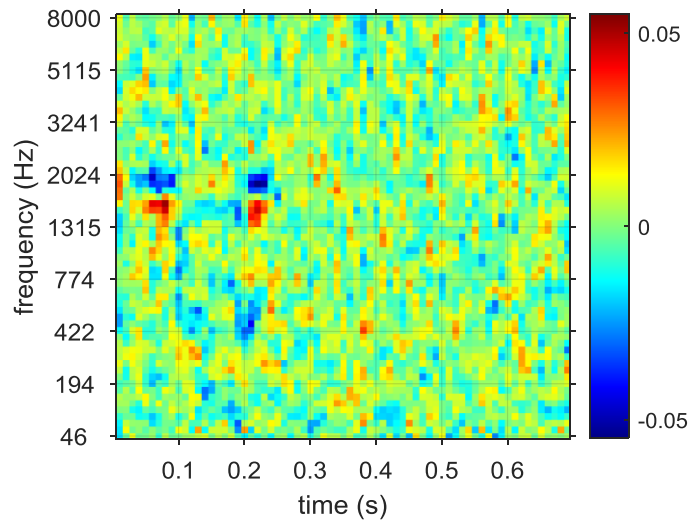
The evolution of ACIs

Analysis of the same data using...

Classic revcorr analysis

[Varnet et al., 2013]

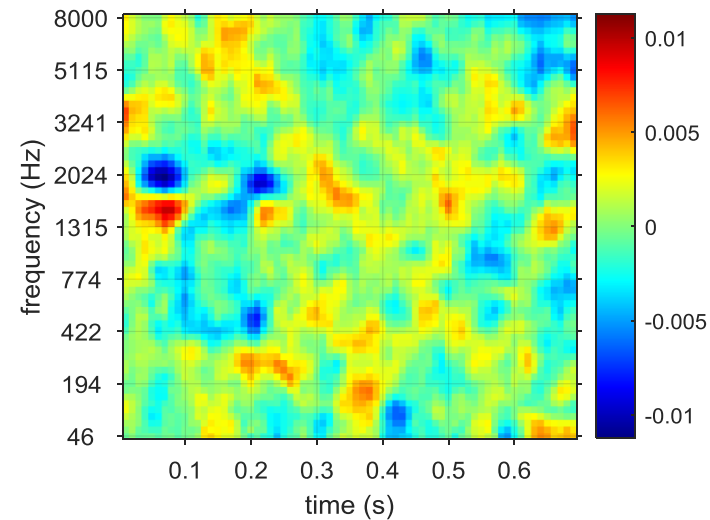
N = **10 000 trials** (\approx 8h)



Ridge regression

[Varnet et al., 2015, 2016...]

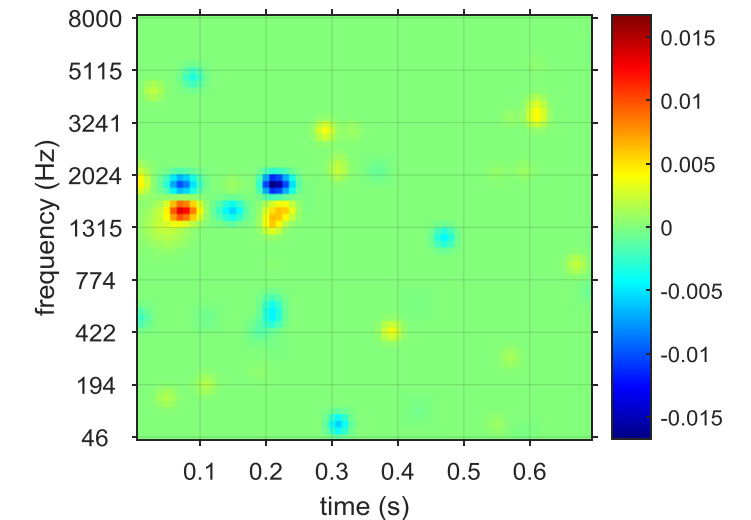
N = **5000 trials** (\approx 4h)



Lasso on Gaussian pyramid

[Osses & Varnet, 2021]

N = **3000 trials** (\approx 2,5h)



Auditory Classification Images

- A purely behavioral approach can reveal listening strategies in auditory categorization tasks in noise (in particular for speech categorization tasks)



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

- An interesting approach is to compare the ACIs of real listeners and automatic speech recognition models (cf. Alejandro's presentation)

Thank you for your attention!

And thanks to our collaborators:



**Christian
Lorenzi**



**Emmanuel
Ponsot**



**Laurianne
Cabrera**



**Diane
Lazard**



**Michel
Hoen**



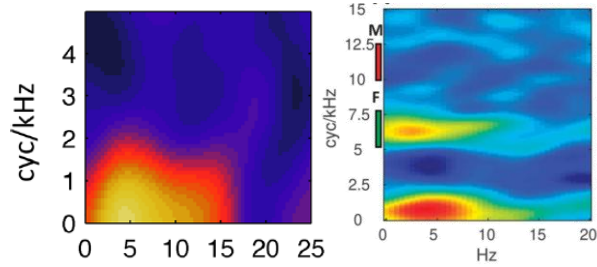
**Fanny
Meunier**

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

high level

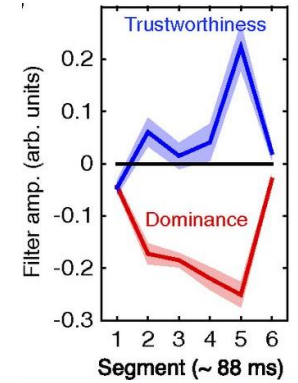
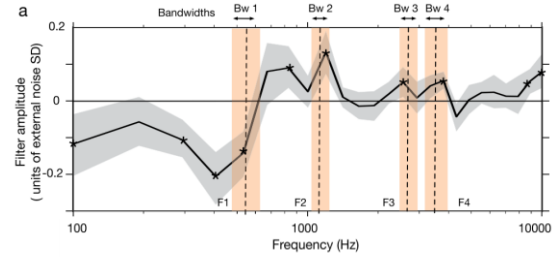
sentence recognition

[Venezia et al., 2016, 2019]



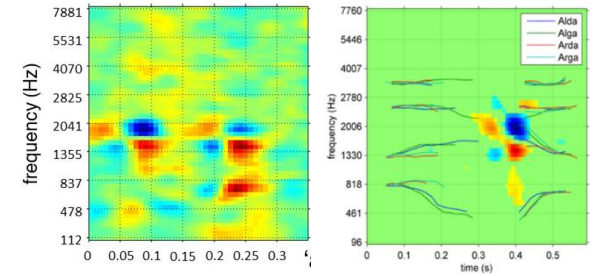
paralinguistics

[Ponsot et al., 2018a, 2018b]



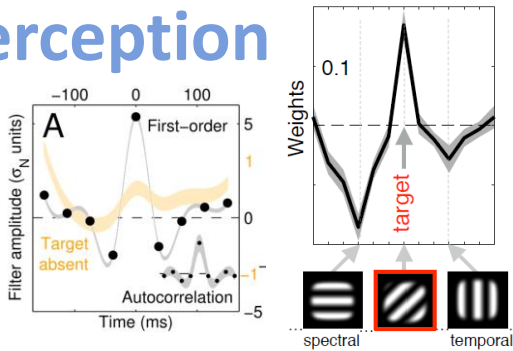
phoneme categorization

[Varnet et al., 2013, 2015]



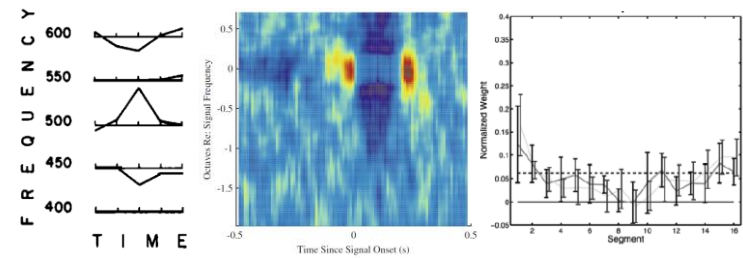
modulation perception

[Ponsot et al., 2020;
Joosten & Neri, 2012;
Varnet & Lorenzi, 2022]



pure-tone detection & loudness perception

[Ahumada & Lovell, 1971; Shub & Richards, 2009; Ponsot et al. 2013]



low level

ACIs for different types of noise

Comparing the effect of 3 different types of noise:

- white noise
- bump noise
- noise low-pass filtered in the MPS domain

[Varnet et Osses, in prep.]

