



**HAL**  
open science

## Probing auditory perception in noise with reverse correlation: two recent projects

Léo Varnet

► **To cite this version:**

Léo Varnet. Probing auditory perception in noise with reverse correlation: two recent projects. Doctoral. Laboratoire des Systèmes Perceptifs, France. 2021. hal-04526762

**HAL Id: hal-04526762**

**<https://hal.science/hal-04526762>**

Submitted on 29 Mar 2024

**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.

# Probing auditory perception in noise with reverse correlation: two recent projects

**Léo Varnet**

LSP meeting, 16/12/2021

# Introduction

- **AM revcorr project:** Exploring amplitude modulation (AM) perception using auditory reverse correlation [Varnet & Lorenzi, submitted]

## Probing temporal modulation detection in white noise using intrinsic envelope fluctuations: a reverse-correlation study

Varnet, Léo<sup>1</sup> and Lorenzi, Christian<sup>1</sup>

*Laboratoire des systèmes perceptifs, Département d'études cognitives, École normale supérieure, Université Paris Sciences & Lettres (PSL), CNRS, 75005 Paris, France*

(Dated: 3 November 2021)

Part of the detrimental effect caused by a stationary noise on sound perception results from the masking of relevant amplitude modulations (AM) in the signal by random intrinsic envelope fluctuations arising from the filtering of noise by cochlear channels. This study capitalized on this phenomenon to probe AM detection strategies for human listeners using a reverse correlation analysis. Eight normal-hearing listeners were asked to detect the presence of a 4-Hz sinusoidal AM target applied to a 1-kHz tone carrier using a yes-no task with 3000 trials/participant. All stimuli were embedded in a white-noise masker. A reverse-correlation analysis applied to the data was then used to compute “psychophysical kernels” showing which aspects of the stimulus’ temporal envelope influenced the listener’s responses. These results were compared to data simulated with different implementations of a modulation

# Introduction

- **AM revcorr project**: Exploring amplitude modulation (AM) perception using auditory reverse correlation [Varnet & Lorenzi, submitted]
- **fastACI project** (2021-2023): Exploring phoneme representations using (auditory) reverse correlation. [Osses & Varnet, 2021; ...]

DAGA 2021 Wien



## Consonant-in-noise discrimination using an auditory model with different speech-based decision devices

Alejandro Osses Vecchi, Léo Varnet

Laboratoire des systèmes perceptifs, ENS, PSL University, Paris, France, Email: {alejandro.osses, leo.varnet}@ens.psl.eu

### Abstract

This study presents insights into the discrimination of two consonants presented in vowel-consonant-vowel (VCV) words embedded in speech-shaped noise (SSN) by adopting an auditory model that uses a modulation filter bank front-end followed by either of two speech back-end decision modules from the literature. These decision modules have been validated in the past for the discrimination of sentences in closed- and open-sets. Our

to which the participants had to indicate one of two possible answers (/aba/ or /ada/). This means that the task is implemented as a one-interval two-alternative forced-choice (1-I, 2-AFC) experiment. In the experiment, the level of the noises was fixed at 65 dB SPL and the signal-to-noise ratio (SNR) was adjusted on a trial-by-trial basis to track the speech level at which the participants reached a 70.7%-correct score using a weighted one-up one-down method [4] with unequal step sizes (2.41 and 1.41 dB for the up- and down steps, respectively). Hence

# Introduction

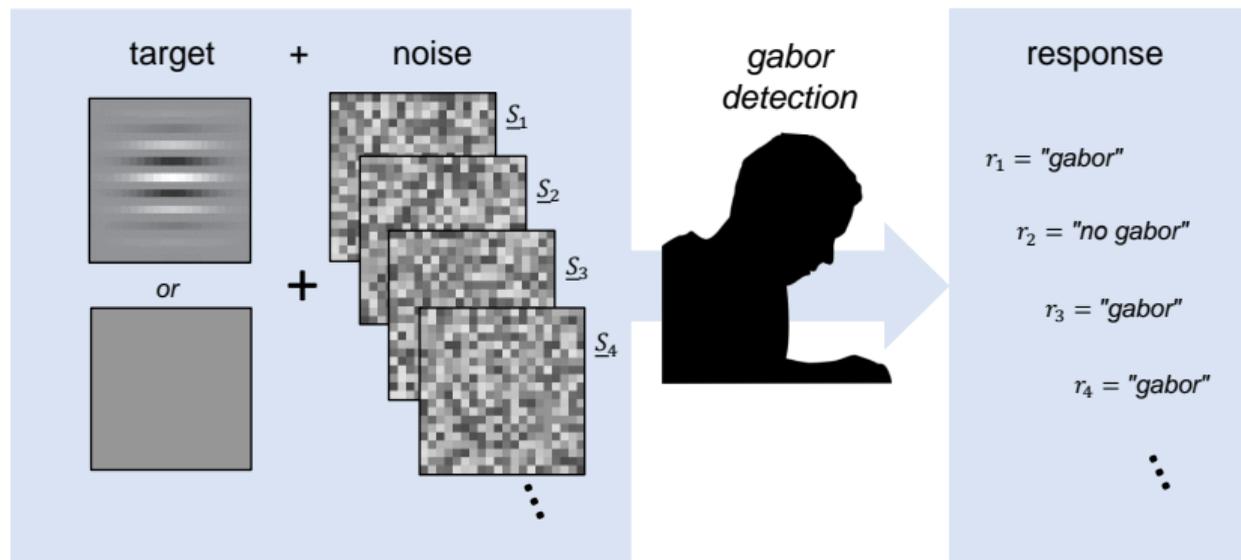
- **AM revcorr project:** Exploring amplitude modulation (AM) perception using auditory reverse correlation [Varnet & Lorenzi, submitted]
- **fastACI project** (2021-2023): Exploring phoneme representations using (auditory) reverse correlation. [Osses & Varnet, 2021; ...]
- Choose your own adventure!



## Reverse correlation ('revcorr')

**Correlational technique**  
primarily used in visual  
psychophysics experiments.

Example: visual detection of  
a Gabor target in noise  
[Solomon, 2002].

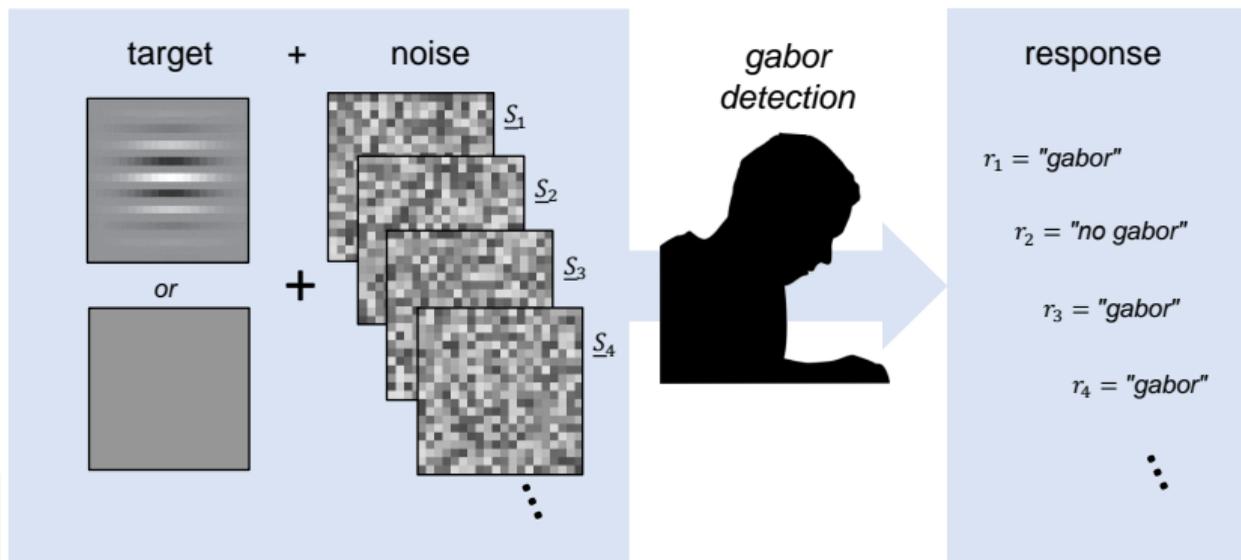


# Reverse correlation ('revcorr')

**Correlational technique** primarily used in visual psychophysics experiments.

Example: visual detection of a Gabor target in noise  
*[Solomon, 2002].*

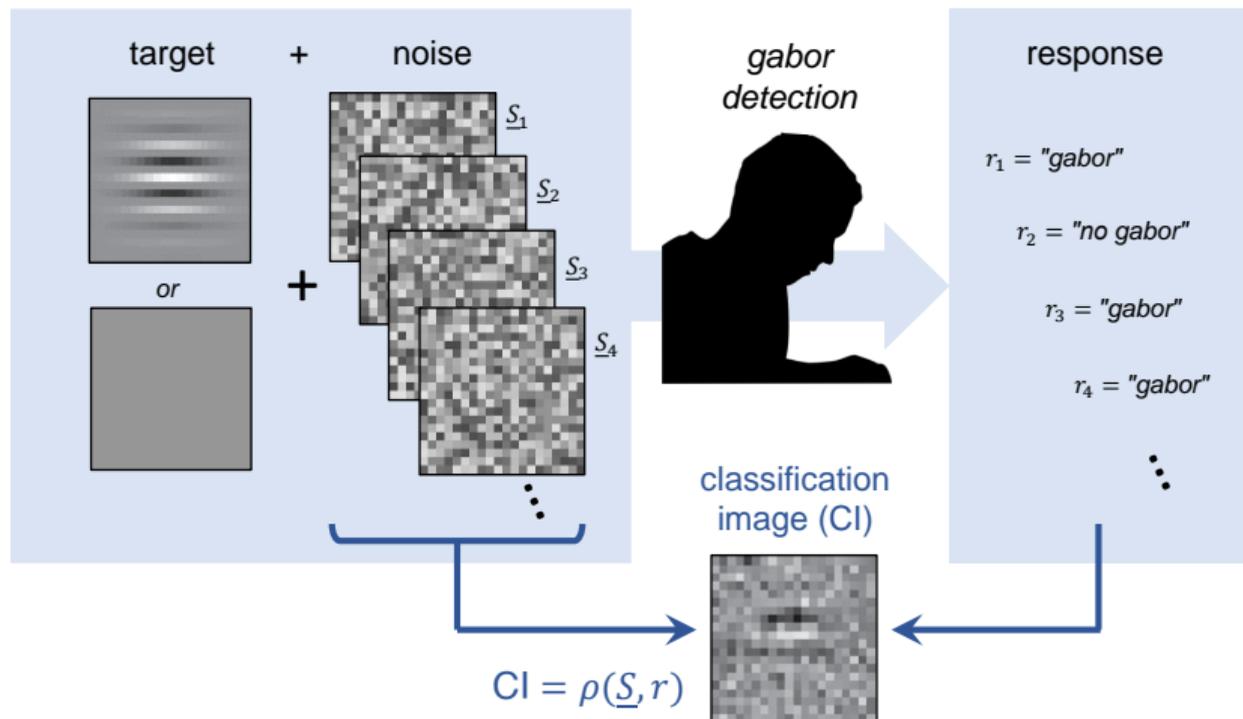
Which information is used to detect whether the target was present or not?



# Reverse correlation ('revcorr')

**Correlation** between the specific noise field in each trial and the response of the observer.

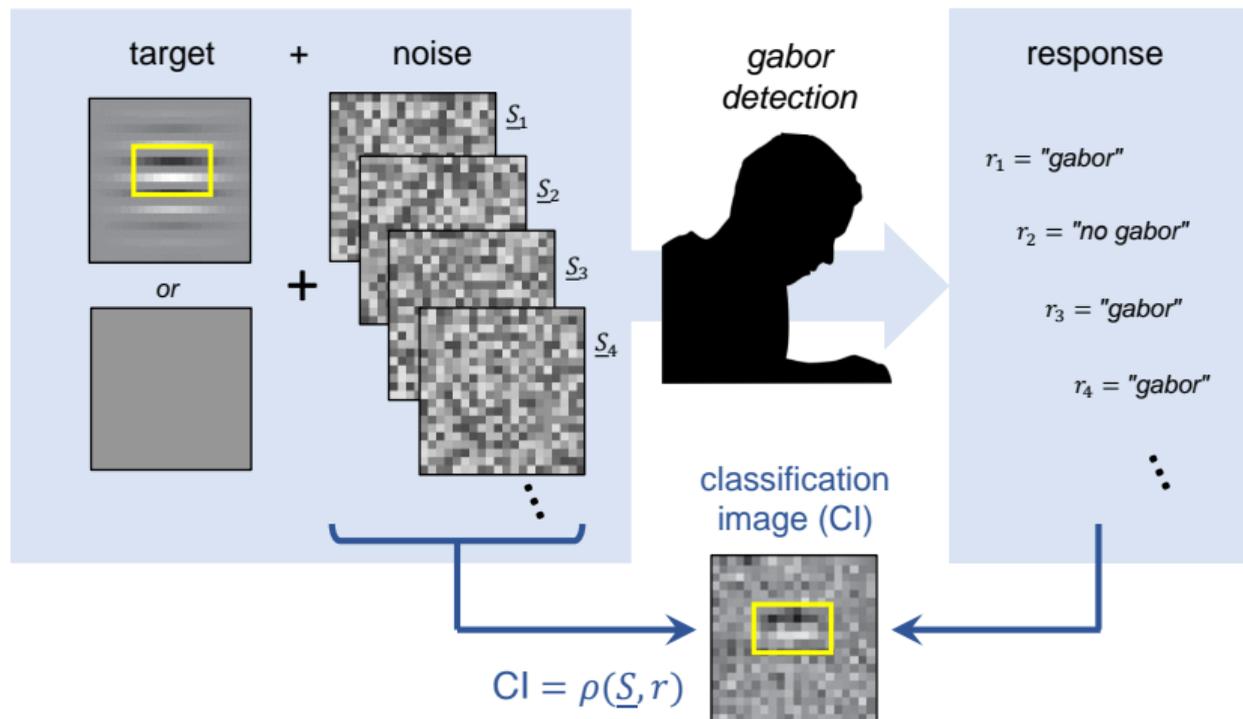
The resulting correlation matrix shows how the presence of noise on each pixel interferes with the decision.



# Reverse correlation ('revcorr')

**Correlation** between the specific noise field in each trial and the response of the observer.

The resulting correlation matrix shows how the presence of noise on each pixel interferes with the decision.

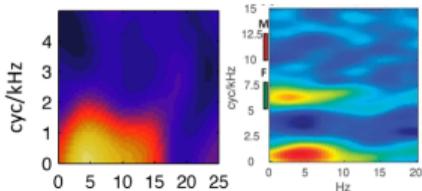


# Auditory revcorr studies (full diagram on <https://dbao.leo-varnet.fr/>)

high level

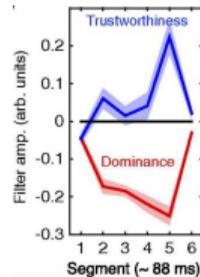
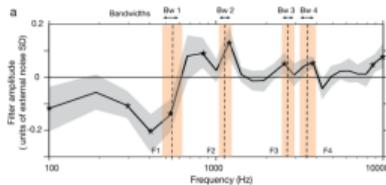
## sentence recognition

[Venezia et al., 2016, 2019]



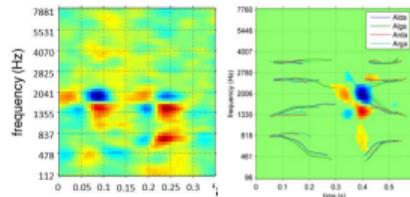
## paralinguistics

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



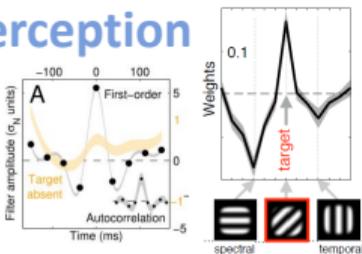
## phoneme categorization

[Varnet et al., 2013, 2015]



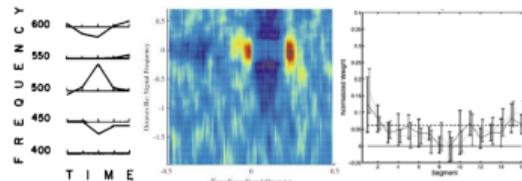
## modulation perception

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



## pure-tone detection & loudness perception

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



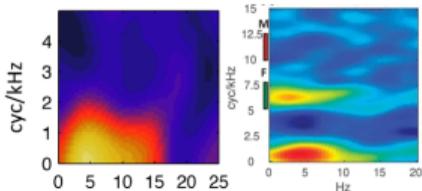
low level

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

high level

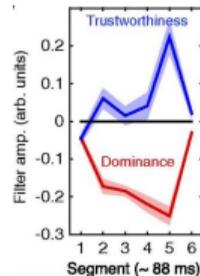
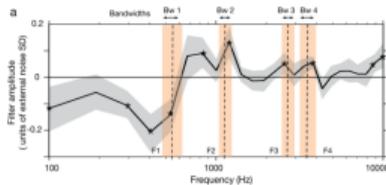
## sentence recognition

[Venezia et al., 2016, 2019]



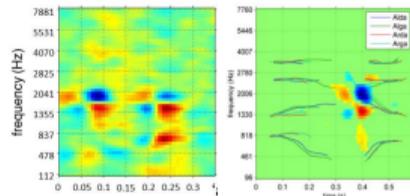
## paralinguistics

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



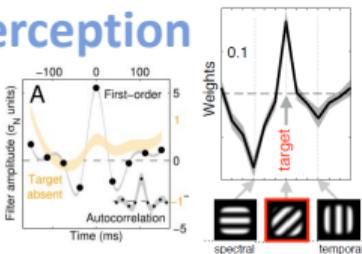
## phoneme categorization

[Varnet et al., 2013, 2015]



## modulation perception

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

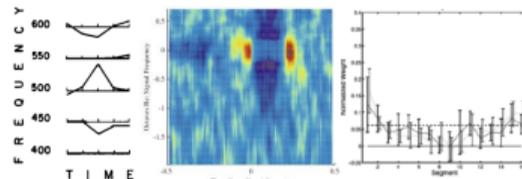


**YOU ARE  
HERE**



## pure-tone detection & loudness perception

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

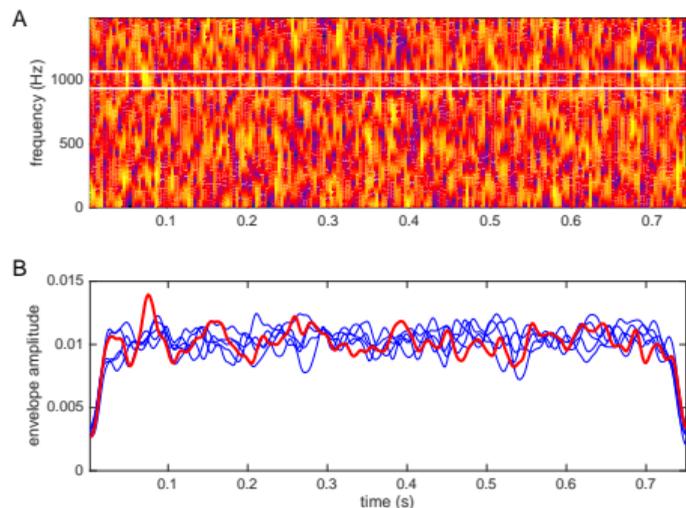


low level

# Intrinsic fluctuations and revcorr

- Part of the deleterious effect of a steady noise on AM perception is due to the confusion of useful modulations in the signal with **intrinsic envelope fluctuations** arising from the filtering of noise into cochlear critical bands [Dau et al., 1997, 1999]
- Revcorr techniques are particularly suitable for exploring the effect of noise on perception

Using the revcorr method to **probe internal representations of AM** by relating random envelope fluctuations of masking noise with the response of listeners on a trial-by-trial basis

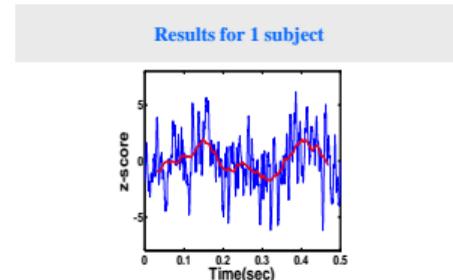
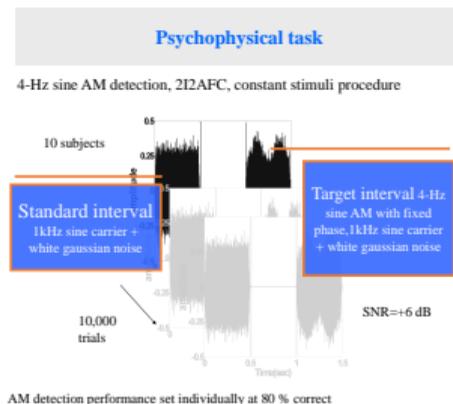


Envelopes for 6 realizations of a white noise, bandpass-filtered between 935 and 1068 Hz

# AM perception and revcorr

Three recent attempts:

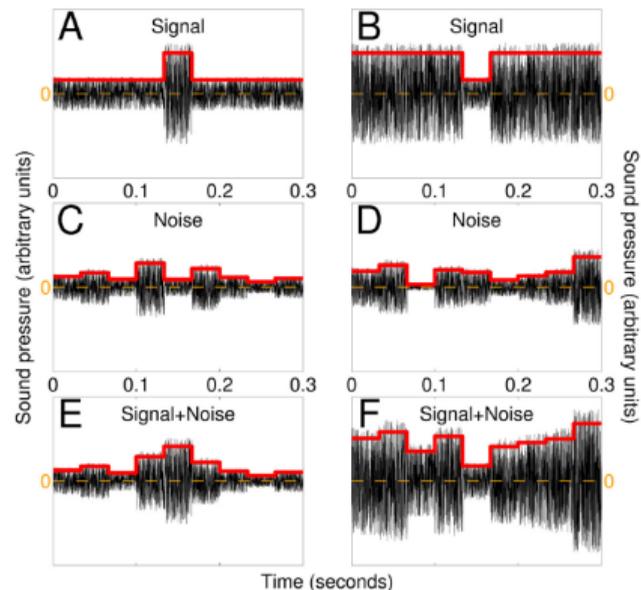
- **Ardoit, Mamassian, Lorenzi (2007)**. “Internal representation of amplitude modulation revealed by reverse correlation,” 30th ARO midwinter meeting. **(white noise)**



# AM perception and revcorr

Three recent attempts:

- **Ardoit, Mamassian, Lorenzi (2007)**. “Internal representation of amplitude modulation revealed by reverse correlation,” 30th ARO midwinter meeting. (**white noise**)
- **Joosten, Shamma, Lorenzi, Neri (2016)**. “Dynamic Reweighting of Auditory Modulation Filters,” PLoS Comp Biol 12(7). (**dimensional noise**)

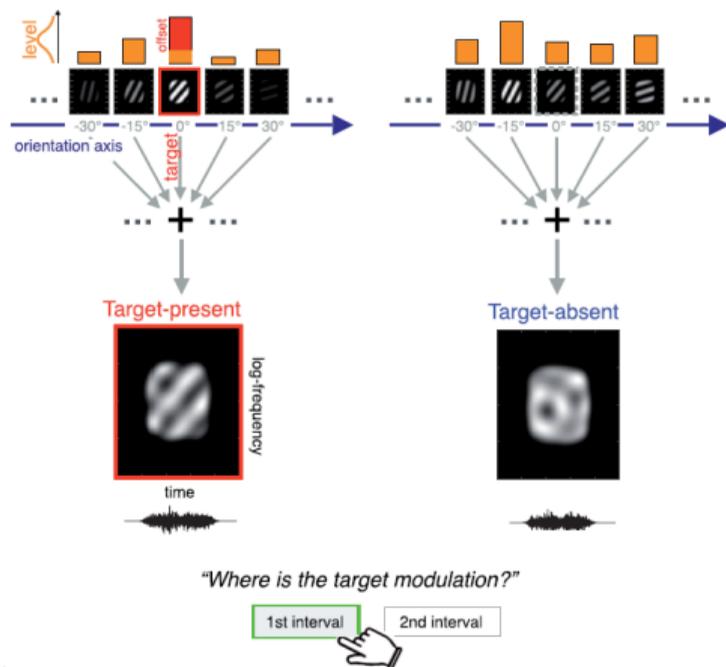


**Fig 1. AM modulations as targets and interferers.** The signal to be detected consisted of a centered 33-ms square-pulse change in AM envelope that was either larger (A) or smaller (B) than baseline modulation depending on whether listeners were engaged in the increment or decrement task respectively. (C,D) Noise was applied in the form of Gaussian modulations around baseline value every 33-ms to generate signal +noise stimuli (E-F). Black traces show sound waves, red traces show amplitude modulations. See [Materials and Methods](#) for details.

# AM perception and revcorr

Three recent attempts:

- **Ardoit, Mamassian, Lorenzi (2007).** “Internal representation of amplitude modulation revealed by reverse correlation,” 30th ARO midwinter meeting. **(white noise)**
- **Joosten, Shamma, Lorenzi, Neri (2016).** “Dynamic Reweighting of Auditory Modulation Filters,” PLoS Comp Biol 12(7). **(dimensional noise)**
- **Ponsot, Varnet, Wallaert, Daoud, Shamma, Lorenzi, Neri (2021).** “Mechanisms of Spectrotemporal Modulation Detection for Normal- and Hearing-Impaired Listeners,” Trends Hear, 25. **(dimensional noise)**

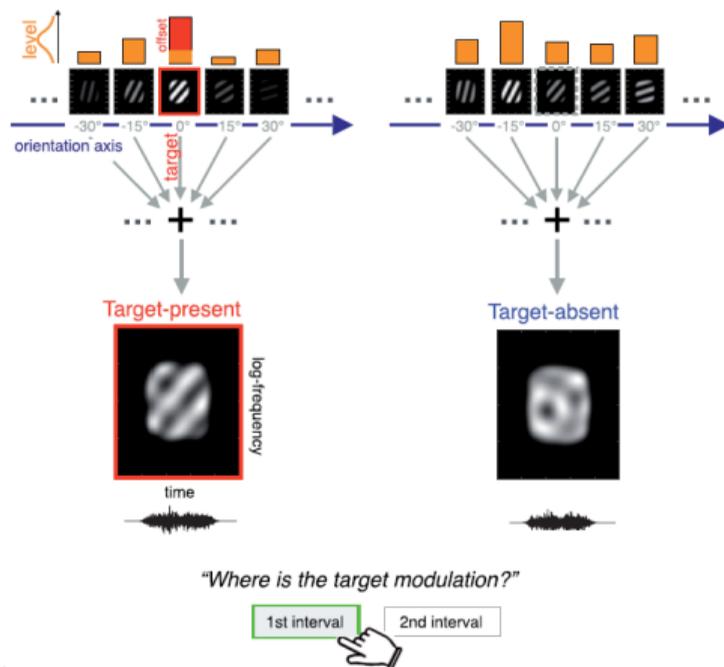


# AM perception and revcorr

Three recent attempts:

- **Ardoit, Mamassian, Lorenzi (2007)**. “Internal representation of amplitude modulation revealed by reverse correlation,” 30th ARO midwinter meeting. **(white noise)**
- **Joosten, Shamma, Lorenzi, Neri (2016)**. “Dynamic Reweighting of Auditory Modulation Filters,” PLoS Comp Biol 12(7). **(dimensional noise)**
- **Ponsot, Varnet, Wallaert, Daoud, Shamma, Lorenzi, Neri (2021)**. “Mechanisms of Spectrotemporal Modulation Detection for Normal- and Hearing-Impaired Listeners,” Trends Hear, 25. **(dimensional noise)**

When applying reverse correlation to study AM processing, “adding acoustic white-noise is inappropriate, [...] because the envelope fluctuations it induces are difficult to control and exercise adequately.” [\[Joosten et al., 2016\]](#)



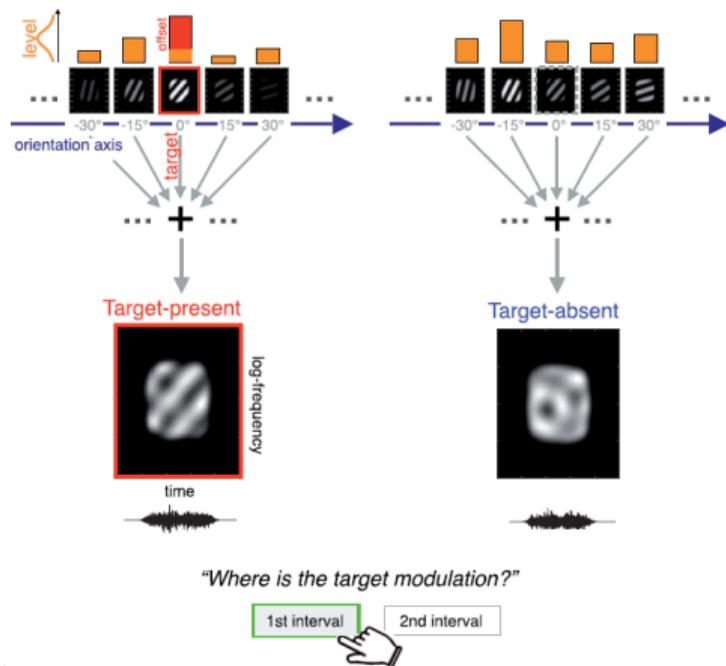
3

# AM perception and revcorr

Three recent attempts:

- **Ardoit, Mamassian, Lorenzi (2007)**. “Internal representation of amplitude modulation revealed by reverse correlation,” 30th ARO midwinter meeting. **(white noise)**
- **Joosten, Shamma, Lorenzi, Neri (2016)**. “Dynamic Reweighting of Auditory Modulation Filters,” PLoS Comp Biol 12(7). **(dimensional noise)**
- **Ponsot, Varnet, Wallaert, Daoud, Shamma, Lorenzi, Neri (2021)**. “Mechanisms of Spectrotemporal Modulation Detection for Normal- and Hearing-Impaired Listeners,” Trends Hear, 25. **(dimensional noise)**

When applying reverse correlation to study AM processing, “adding acoustic white-noise is inappropriate, [...] because the envelope fluctuations it induces are difficult to control and exercise adequately.” [Joosten et al., 2016] ... or is it?



3

# Materials & methods

## Signals:

Unmod. tone vs. 4-Hz AM

Duration = 0.75 s

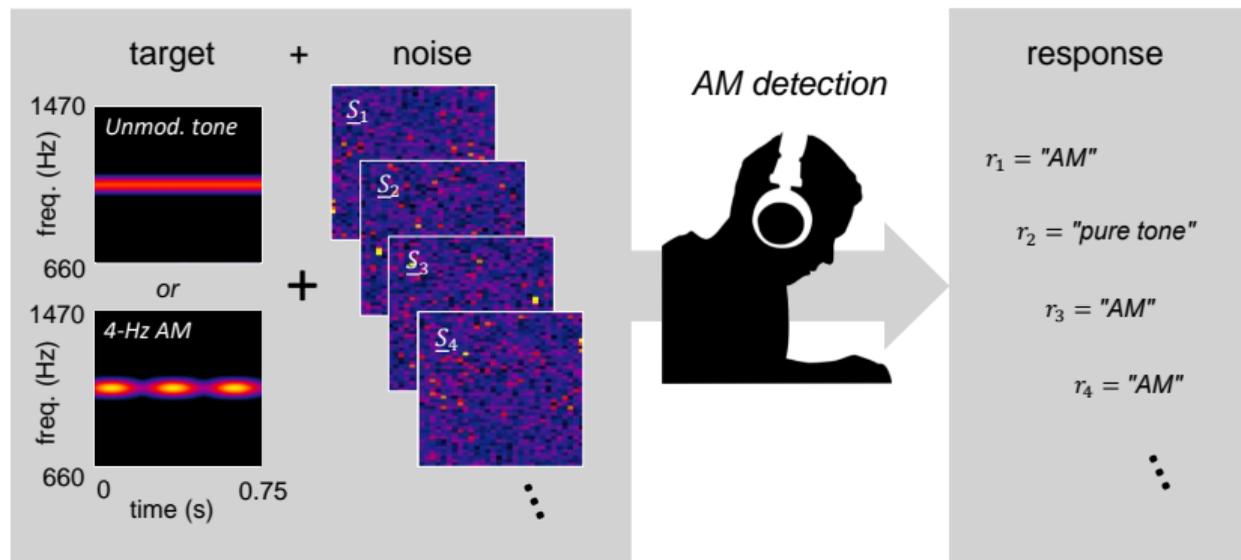
Carrier freq. = 1 kHz

Fixed AM phase

**Task:** yes/no detection task  
in white noise (-10 dB SNR)

Modulation depth (m)  
adapted from trial to trial to  
target 70.7% correct  
responses

**Participants:** 9 NH listeners



# Materials & methods

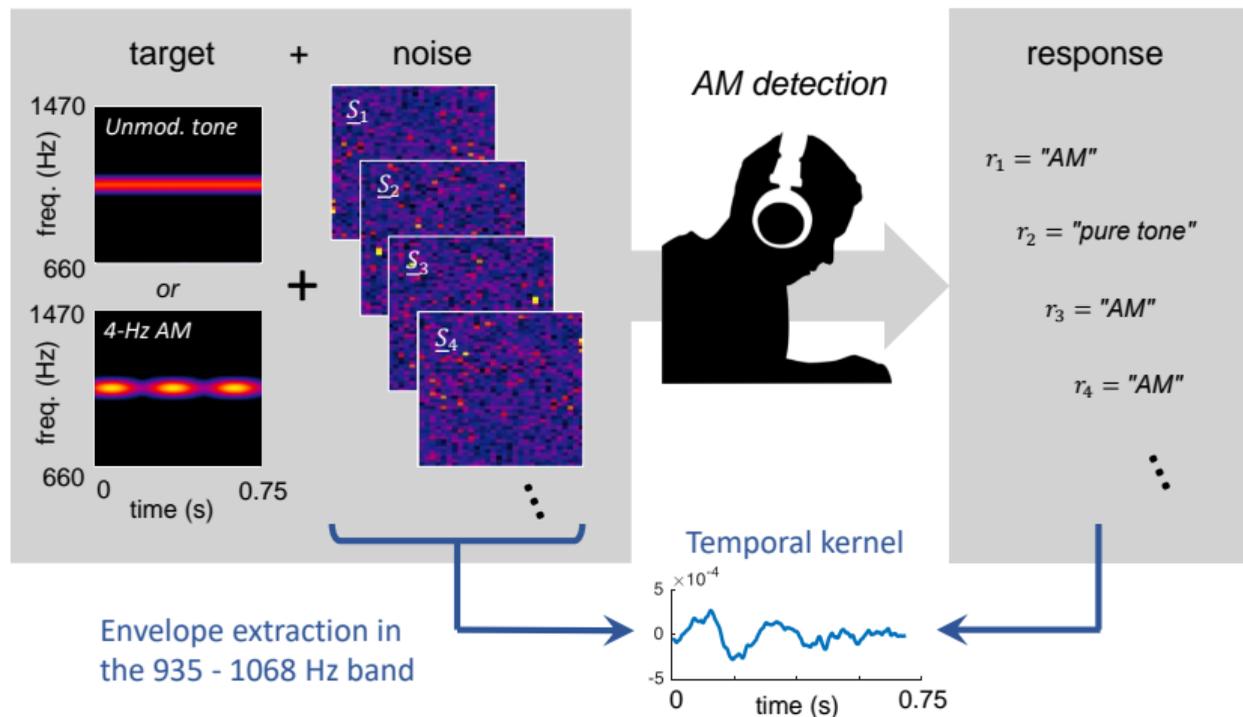
## Signals:

Unmod. tone vs. 4-Hz AM  
 Duration = 0.75 s  
 Carrier freq. = 1 kHz  
 Fixed AM phase

**Task:** yes/no detection task  
 in white noise (-10 dB SNR)

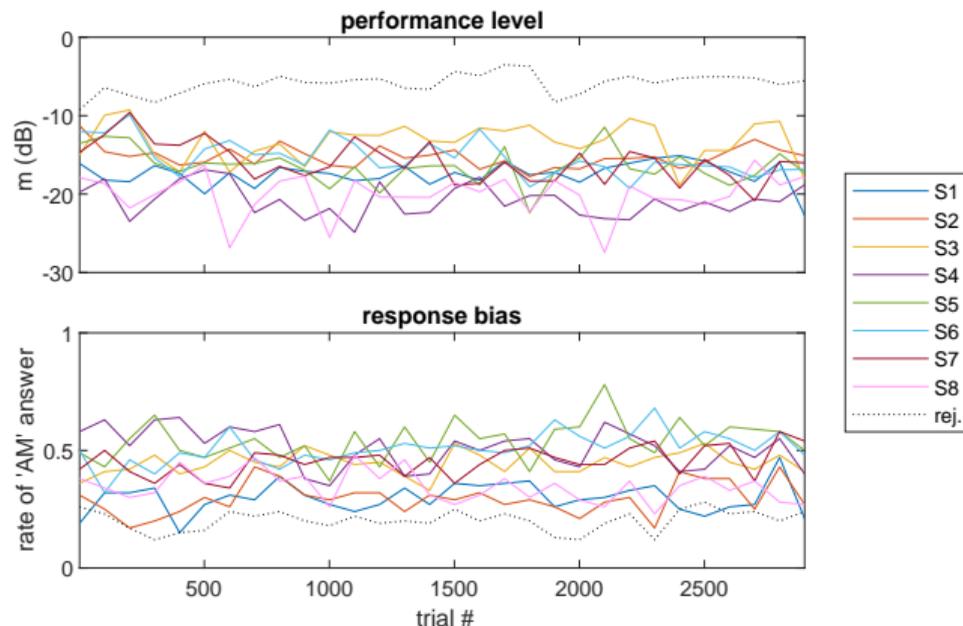
Modulation depth (m)  
 adapted from trial to trial to  
 target 70.7% correct  
 responses

**Participants:** 9 NH listeners



# Performance over the course of the experiment

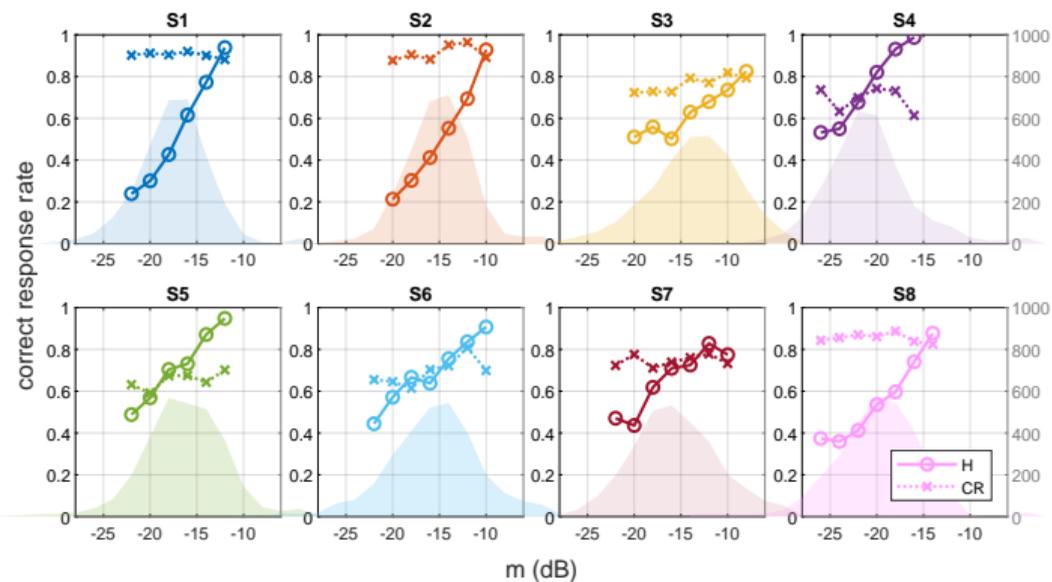
- 3.000 trials per participant (≈ 3h including breaks)
- Stable behavior in terms of performance and bias
- High performance level (mean  $m \approx -16$  dB, at -10 dB SNR)
- Most participants are unbiased!



## Performance as a function of modulation depth

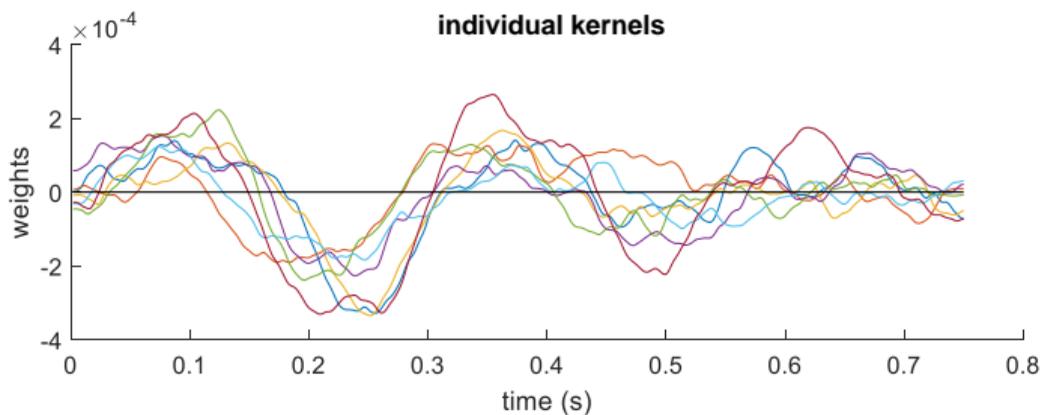
Performance in target-present trials (percentage of Hits) strongly depends on modulation depth.

Performance in target-absent trials (percentage of Correct Rejections) does not.



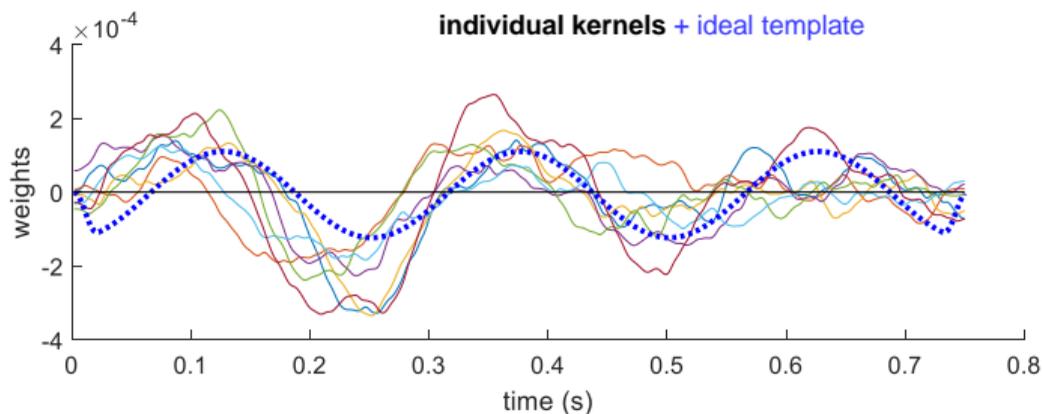
# Temporal kernel (= pattern of intrinsic fluctuations of the noise consistently associated with an 'AM' response)

- Similar pattern for all 8 participants
- **Strong 4-Hz component**



# Temporal kernel (= pattern of intrinsic fluctuations of the noise consistently associated with an 'AM' response)

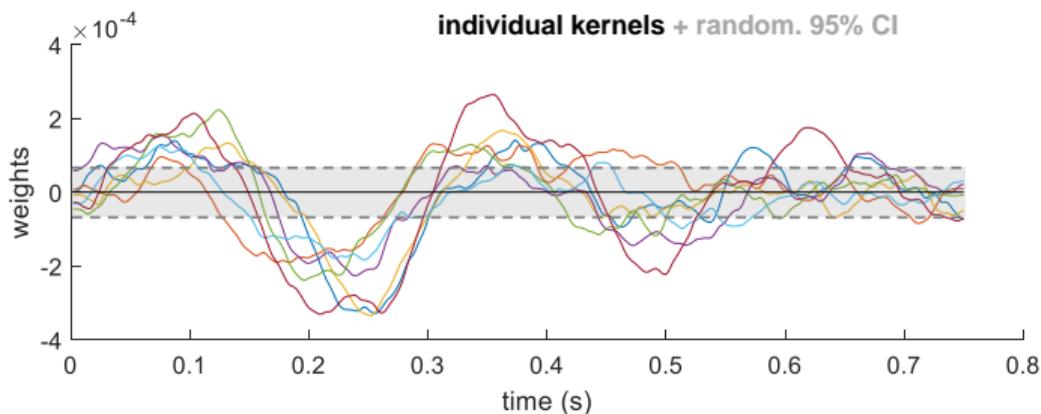
- Similar pattern for all 8 participants
- **Strong 4-Hz component**, very close to the **AM target**



⇒ When the noise intrinsic envelope shows a strong 4-Hz component, it is more likely to be confused with the AM target to be detected

# Temporal kernel (= pattern of intrinsic fluctuations of the noise consistently associated with an 'AM' response)

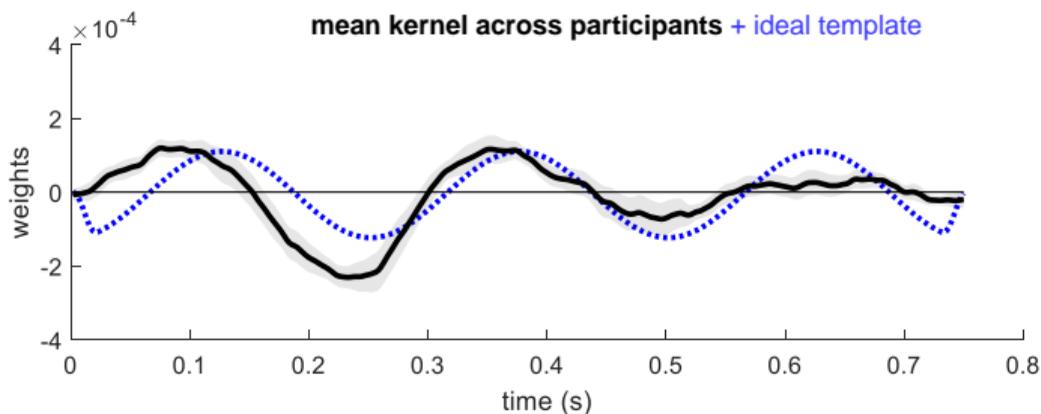
- Similar pattern for all 8 participants
- **Strong 4-Hz component**, very close to the **AM target**



⇒ When the noise intrinsic envelope shows a strong 4-Hz component, it is more likely to be confused with the AM target to be detected

# Temporal kernel (= pattern of intrinsic fluctuations of the noise consistently associated with an 'AM' response)

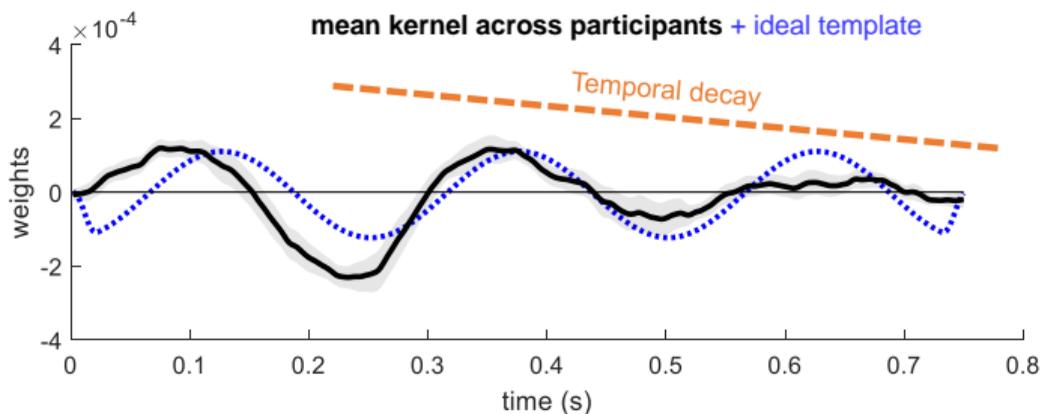
- Similar pattern for all 8 participants
- **Strong 4-Hz component**, very close to the **AM target**



⇒ When the noise intrinsic envelope shows a strong 4-Hz component, it is more likely to be confused with the AM target to be detected

# Temporal kernel (= pattern of intrinsic fluctuations of the noise consistently associated with an 'AM' response)

- Similar pattern for all 8 participants
- **Strong 4-Hz component**, very close to the **AM target**

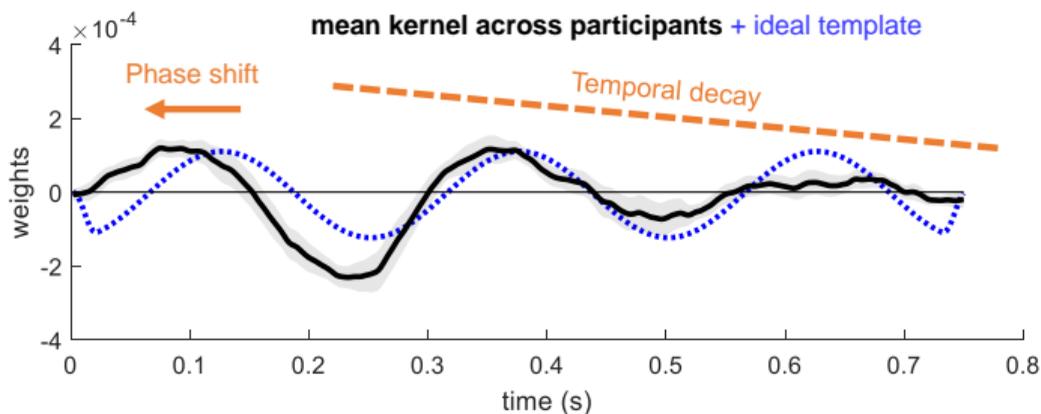


⇒ When the noise intrinsic envelope shows a strong 4-Hz component, it is more likely to be confused with the AM target to be detected

**Differences with the ideal template** inform us about the underlying auditory processes

# Temporal kernel (= pattern of intrinsic fluctuations of the noise consistently associated with an 'AM' response)

- Similar pattern for all 8 participants
- **Strong 4-Hz component**, very close to the **AM target**

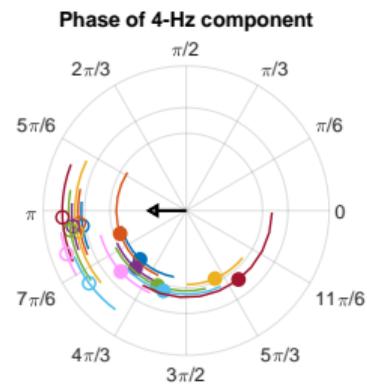
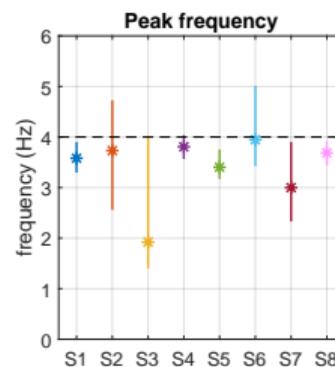
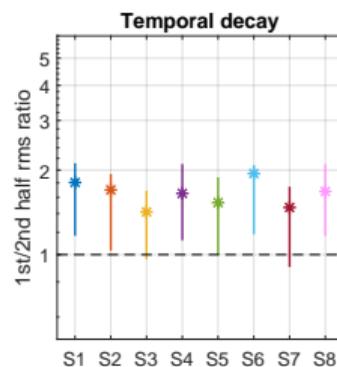


⇒ When the noise intrinsic envelope shows a strong 4-Hz component, it is more likely to be confused with the AM target to be detected

**Differences with the ideal template** inform us about the underlying auditory processes

# Kernel metrics

Deviations to the ideal template were quantified and assessed at the individual level using 3 metrics derived from the kernels (+ 95% bootstrap intervals).



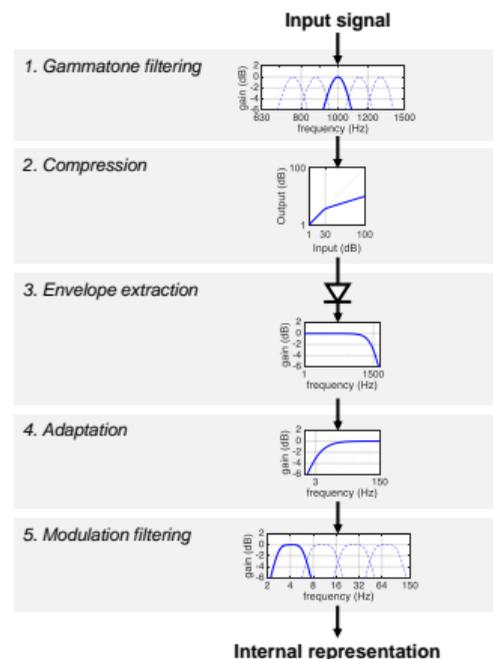
Identifying the source of these suboptimalities will require a model of the human auditory system...

# Modulation Filterbank Model

The **Modulation Filterbank** model: a simple, widely-accepted front-end model of the auditory system. [e.g., *Dau et al., 1997; Ewert & Dau, 2000; King et al., 2019*]



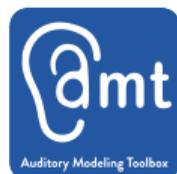
Auditory Modeling Toolbox  
 (<https://www.amtoolbox.org/>)  
 [Majdak et al., 2021; Osses et al., 2021]



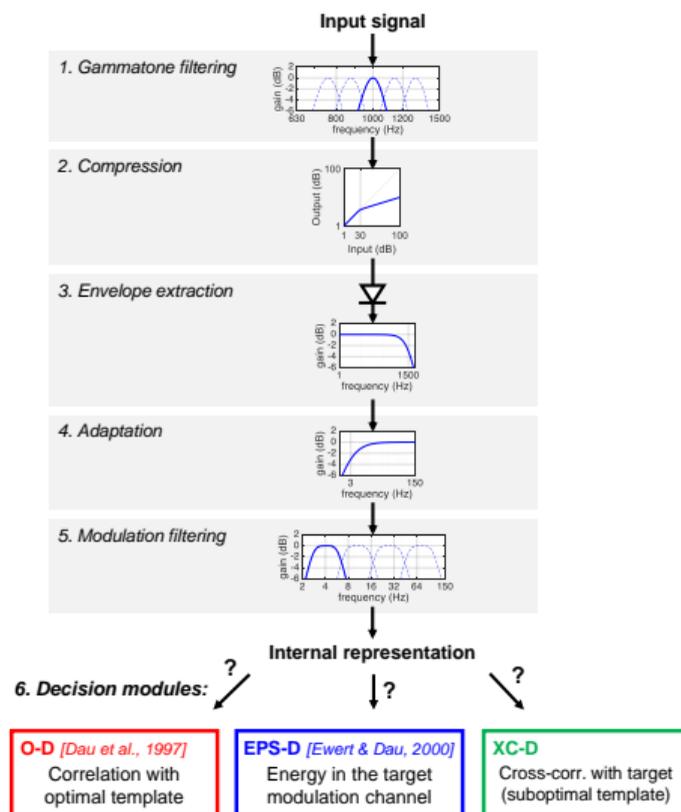
# Modulation Filterbank Model

The **Modulation Filterbank** model: a simple, widely-accepted front-end model of the auditory system. [e.g., *Dau et al., 1997*; *Ewert & Dau, 2000*; *King et al., 2019*]

However, no general agreement on how the auditory system converts this internal representation into a decision statistics.

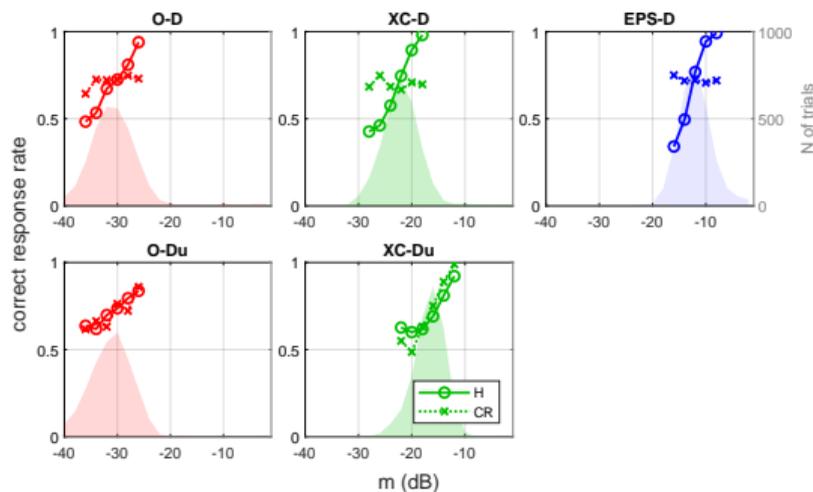


Auditory Modeling Toolbox  
 (<https://www.amtoolbox.org/>)  
 [Majdak et al., 2021; Osses et al., 2021]



# Performance as a function of modulation depth

O-D and XC-D base their decision on the comparison with a template of the target. Is this template updated throughout the experiment or not?

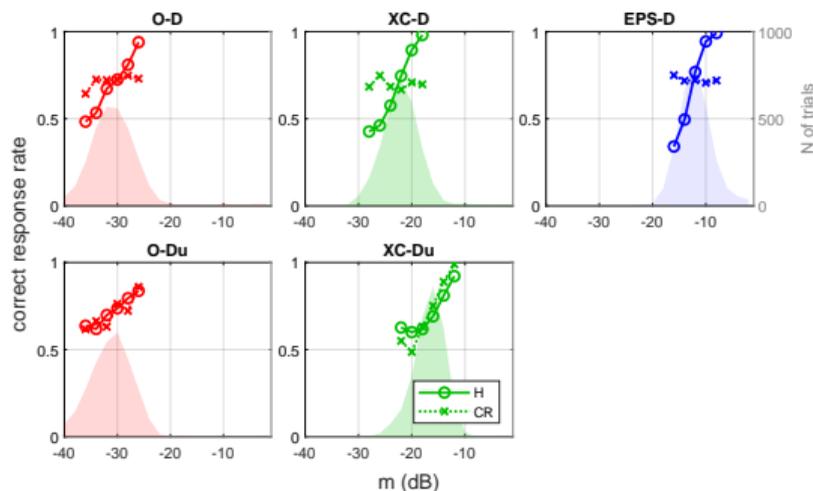


# Performance as a function of modulation depth

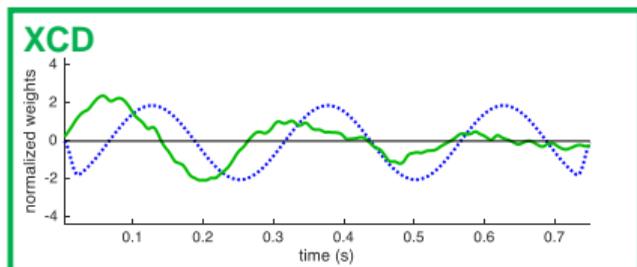
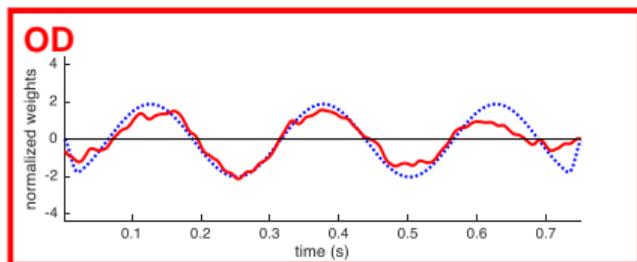
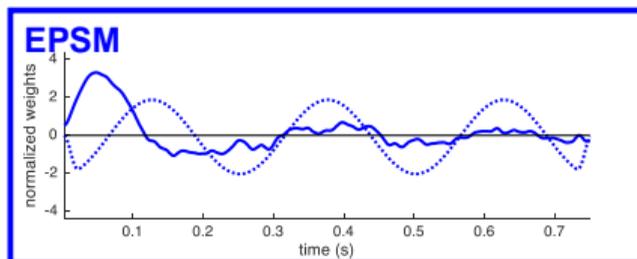
O-D and XC-D base their decision on the comparison with a template of the target. Is this template updated throughout the experiment or not?

Only non-updating models can account for the fact that the performance in target-absent trials does not depend on the performance in target-present trials.

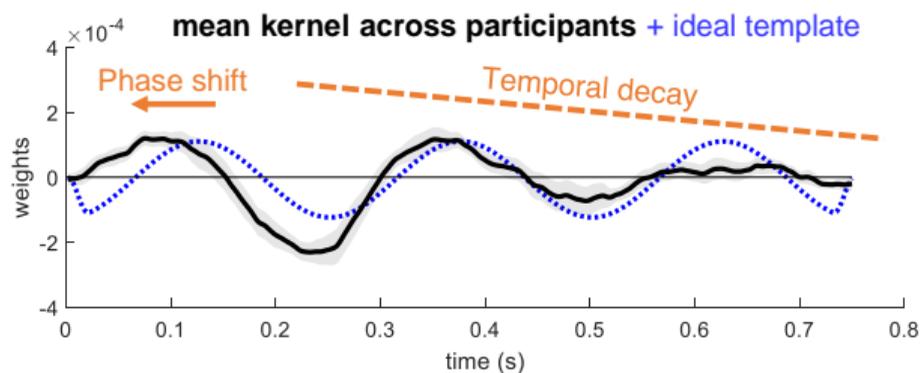
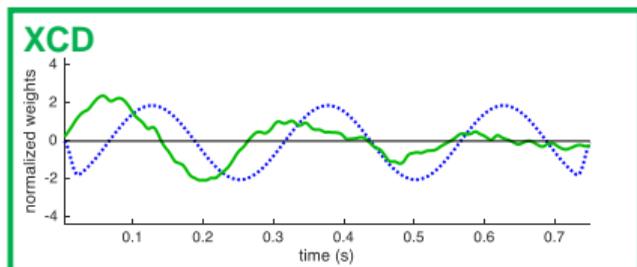
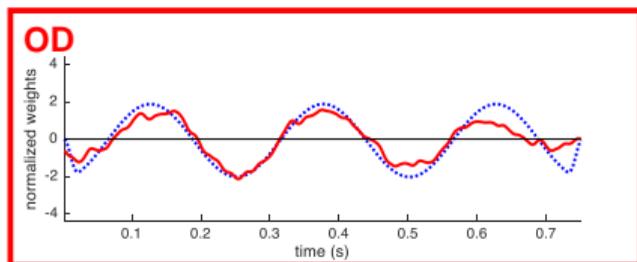
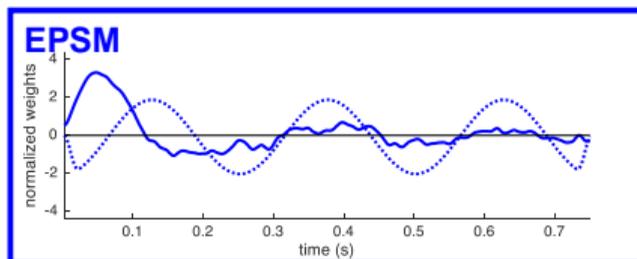
⇒ supports the idea that human participants used a fixed decision rule over the duration of the experiment.



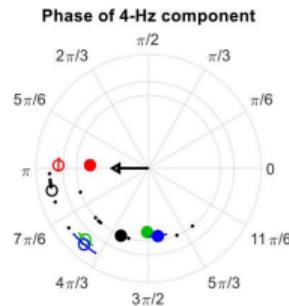
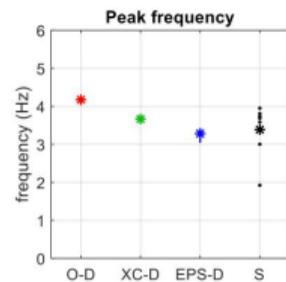
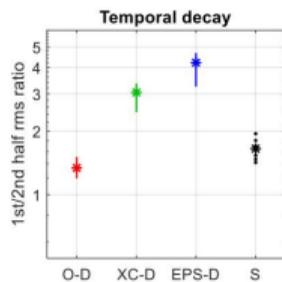
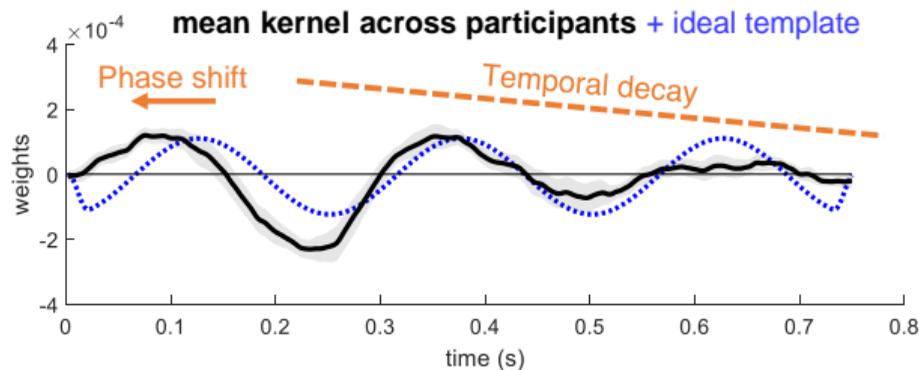
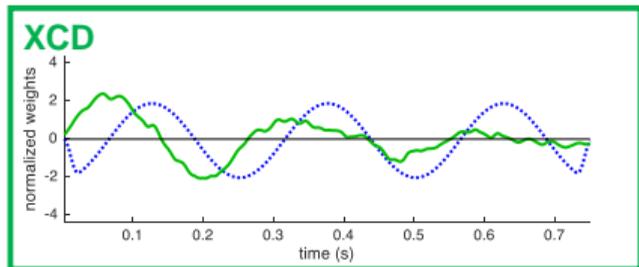
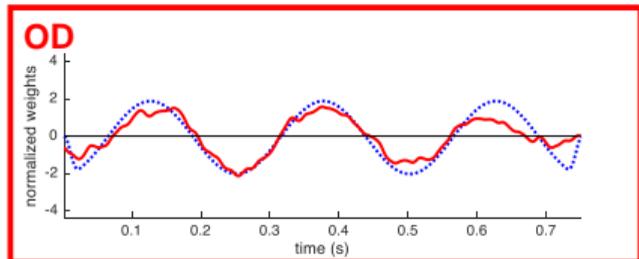
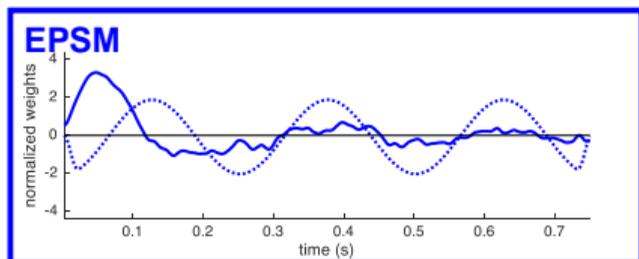
## Simulated kernels



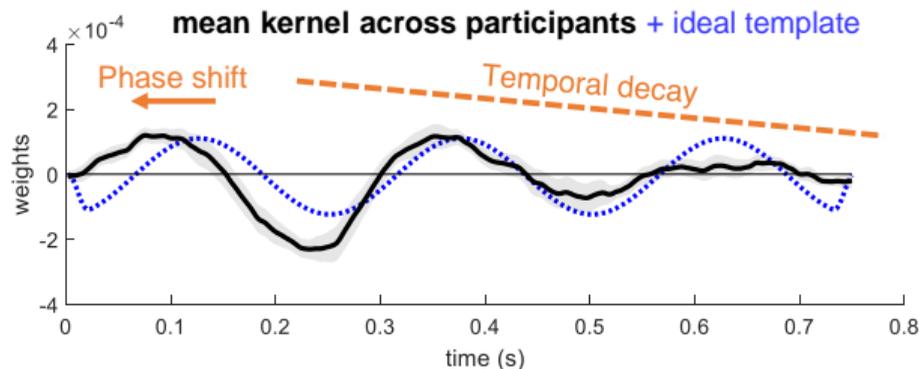
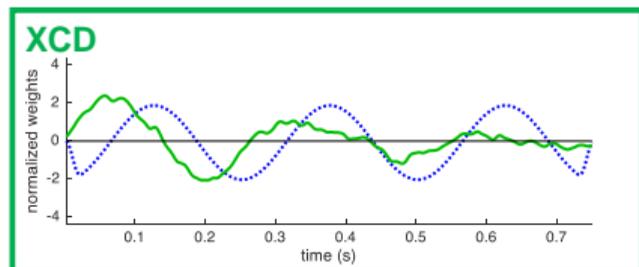
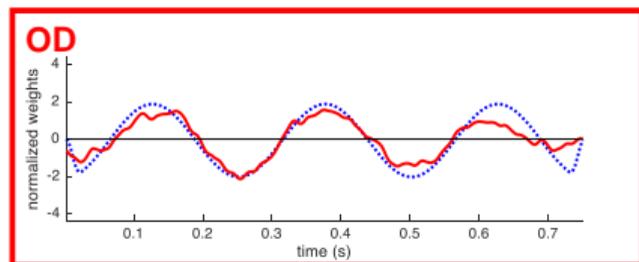
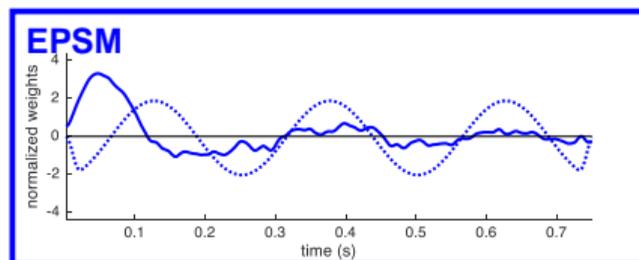
# Simulated kernels



# Simulated kernels



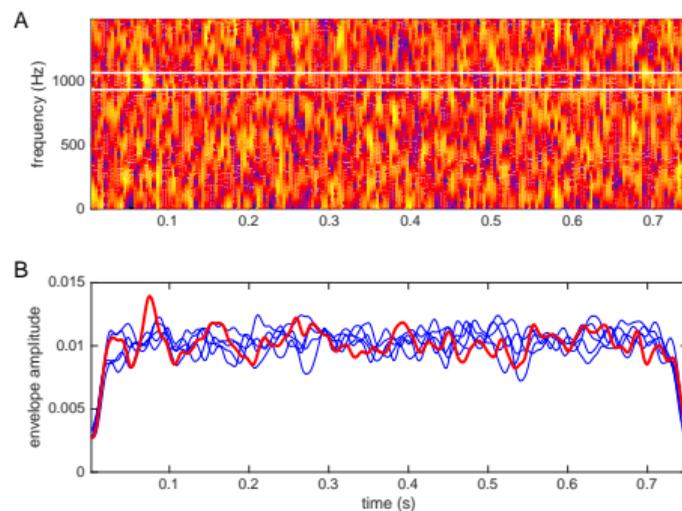
# Simulated kernels



Only **XCD** can reproduce the characteristics of the human kernel, suggesting that the auditory system uses a suboptimal template-matching strategy for this task

# Conclusions

- Direct illustration of the **effect of intrinsic fluctuations on perception**
  - The **auditory revcorr** technique can be used to **probe internal representations and decision strategies** in real and simulated listeners
- Listeners are able to encode a 4-Hz AM target with its phase, but they use a **suboptimal detection strategy**

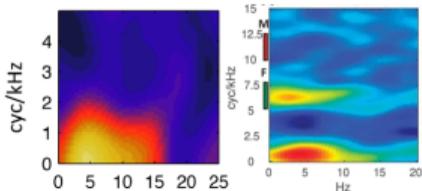


# Auditory revcorr studies (full diagram on <https://dbao.leo-varnet.fr/>)

high level

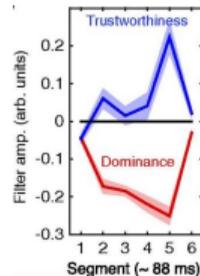
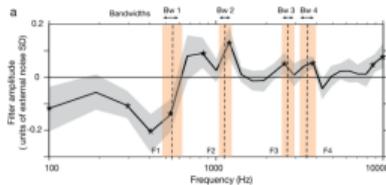
## sentence recognition

[Venezia et al., 2016, 2019]



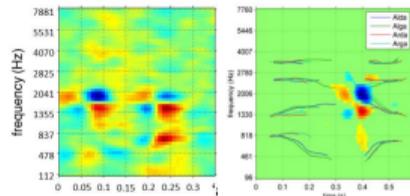
## paralinguistics

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



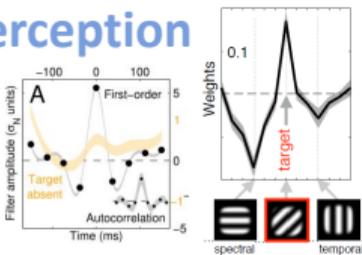
## phoneme categorization

[Varnet et al., 2013, 2015]



## modulation perception

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

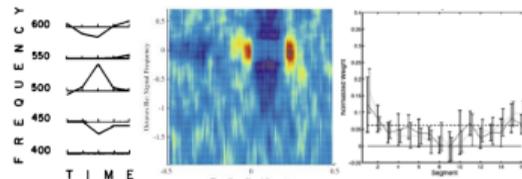


**YOU ARE  
HERE**



## pure-tone detection & loudness perception

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



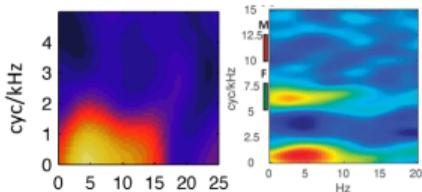
low level

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

high level

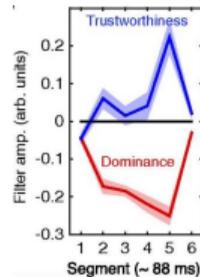
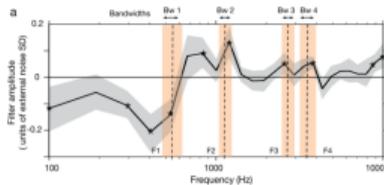
## sentence recognition

[Venezia et al., 2016, 2019]



## paralinguistics

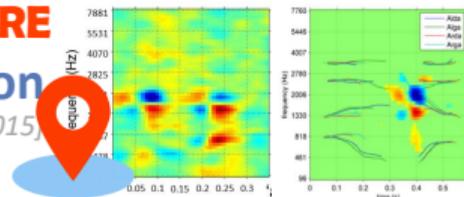
[Ponsot et al., 2018a, 2018b]



**YOU ARE HERE**

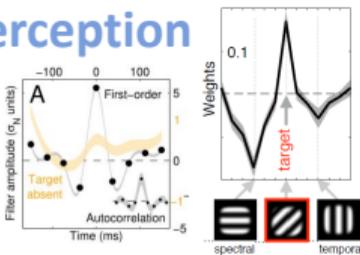
## phoneme categorization

[Varnet et al., 2013, 2015]



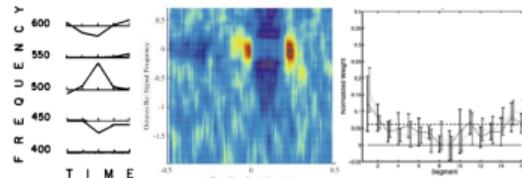
## modulation perception

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



## pure-tone detection & loudness perception

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

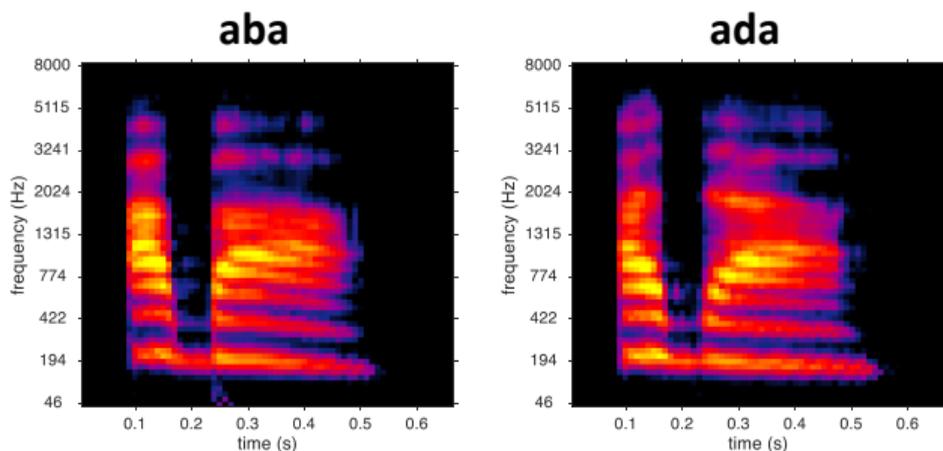


low level

## fastACI project

How do we distinguish “aba” from “ada”?

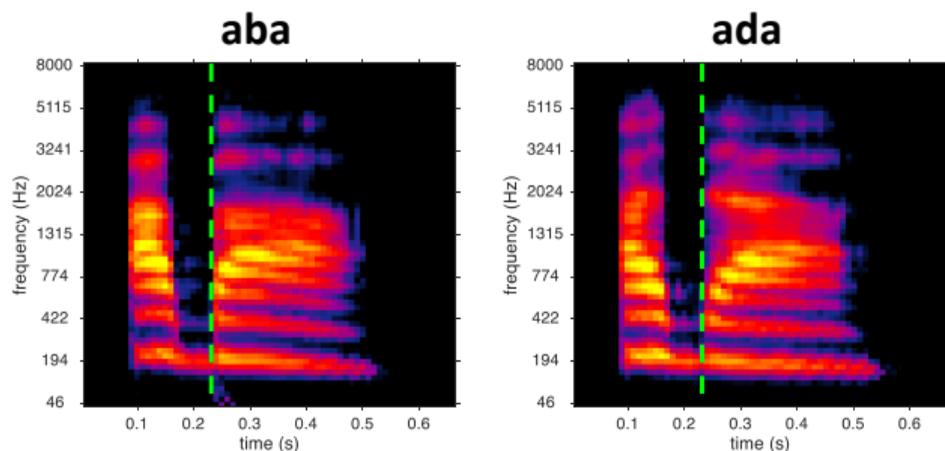
(VCVs from the Oldenburg Logatome Corpus [Wesker et al., 2005], gammatone-based TF representation)



## fastACI project

How do we distinguish “aba” from “ada”?

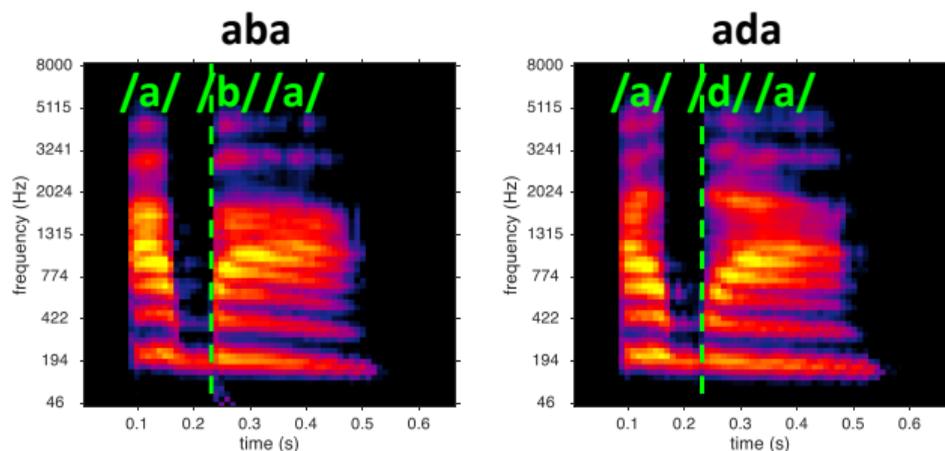
(VCVs from the Oldenburg Logatome Corpus [Wesker et al., 2005], gammatone-based TF representation)



## fastACI project

How do we distinguish “aba” from “ada”?

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

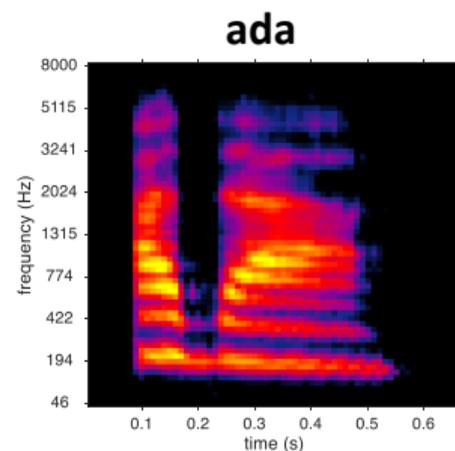
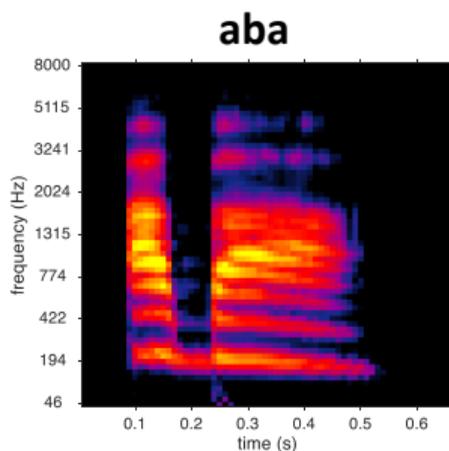


# fastACI project

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”

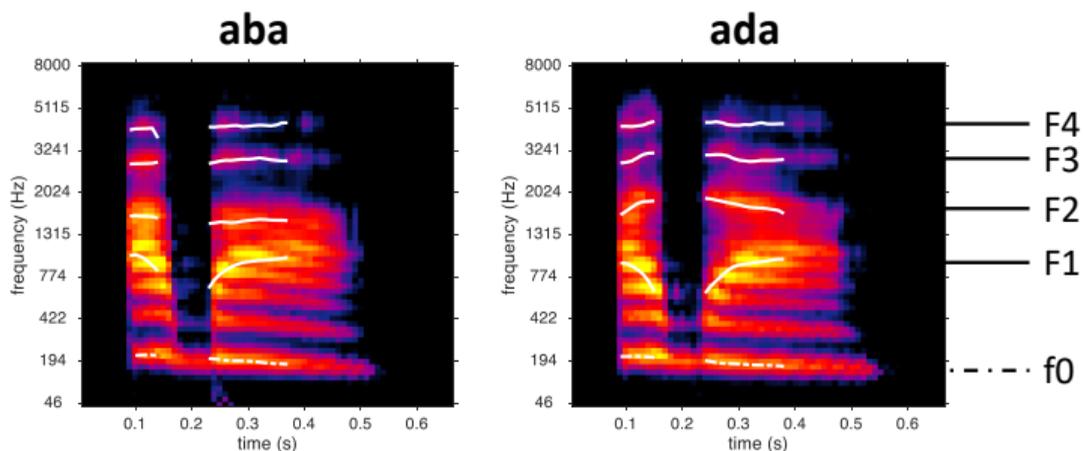


## fastACI project

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)

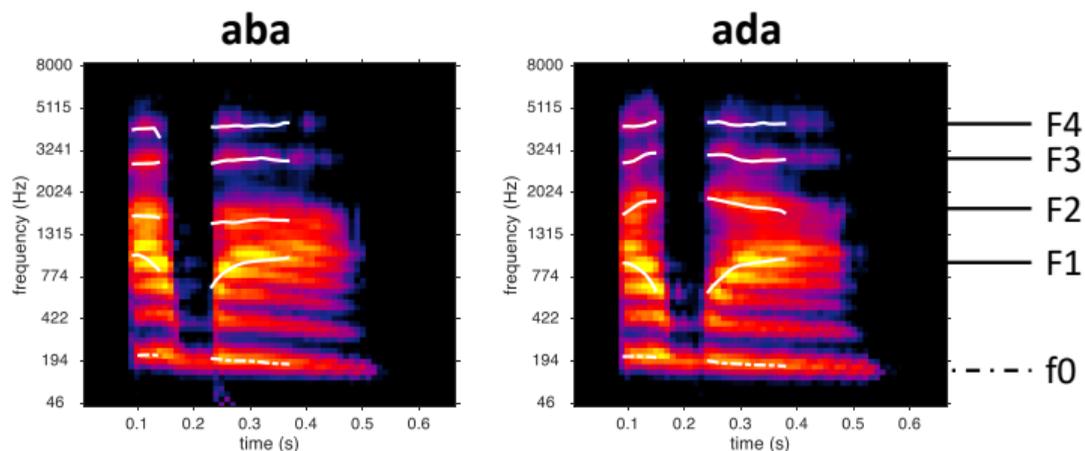


## fastACI project

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?

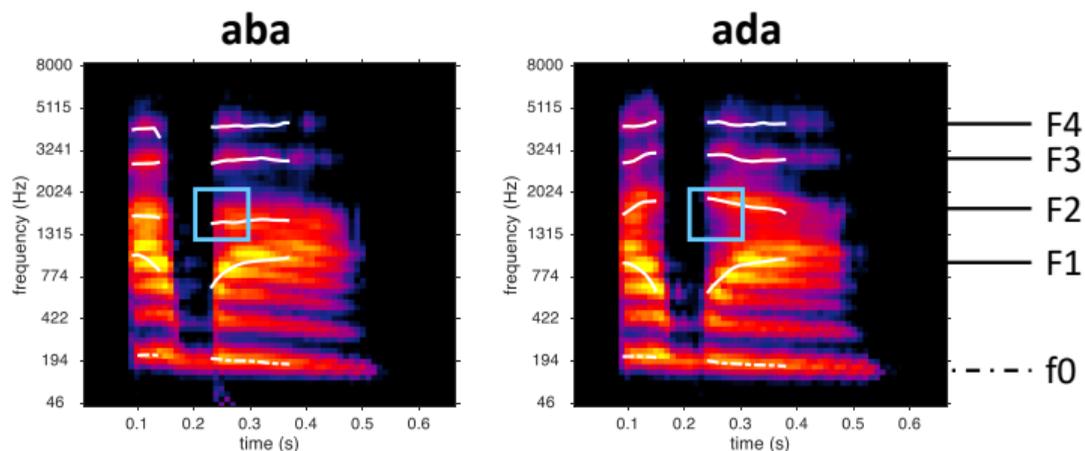


## fastACI project

How do we distinguish “aba” from “ada”?

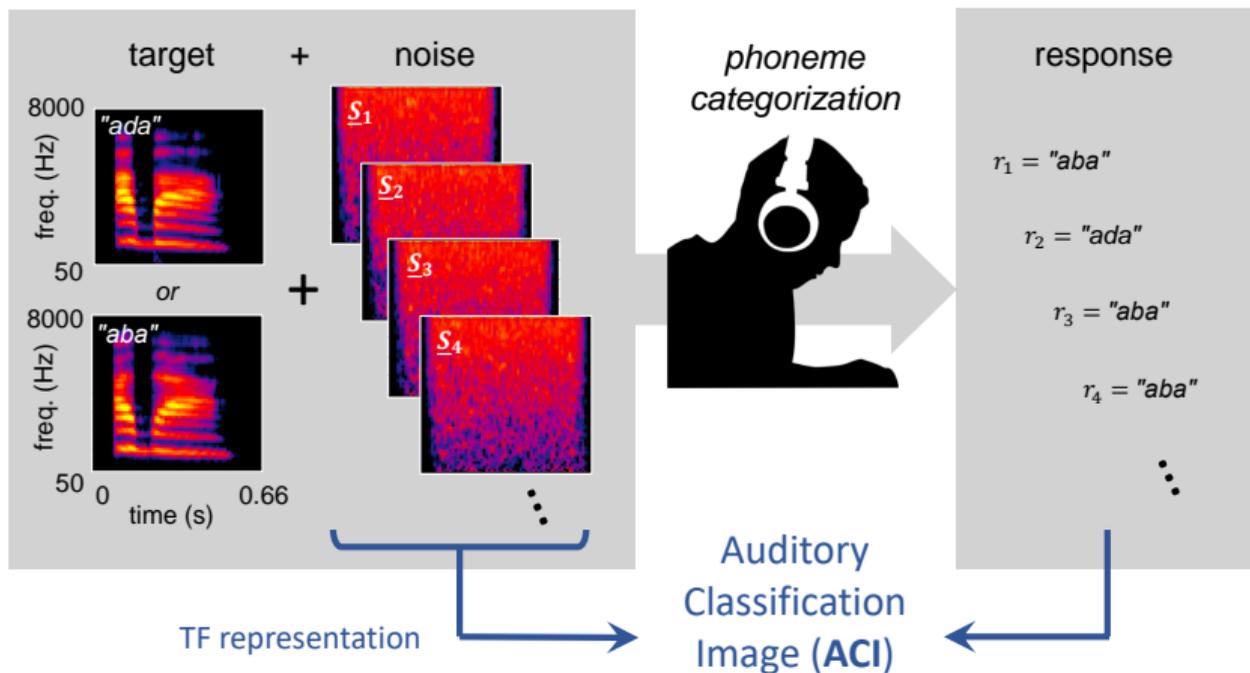
(VCVs from the Oldenburg Logatome Corpus [Wesker et al., 2005], gammatone-based TF representation)

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

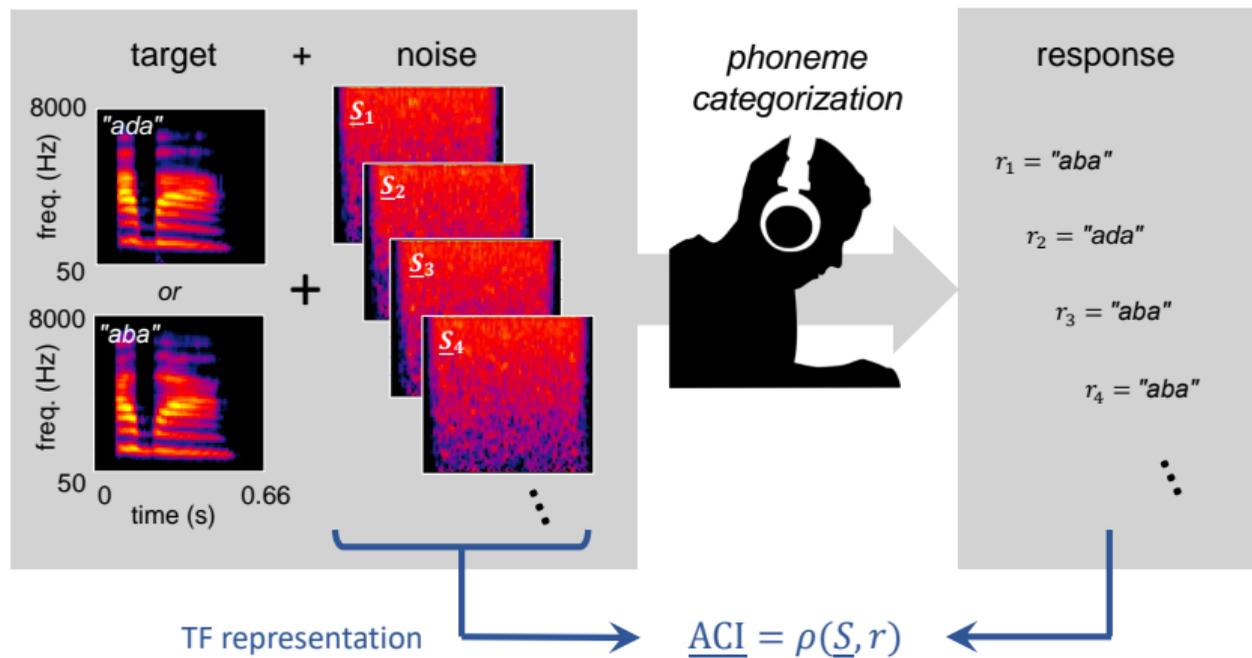


The F2 onset is a cue for categorizing speech sounds into /b/ or /d/ [Liberman et al., 1952; ...]

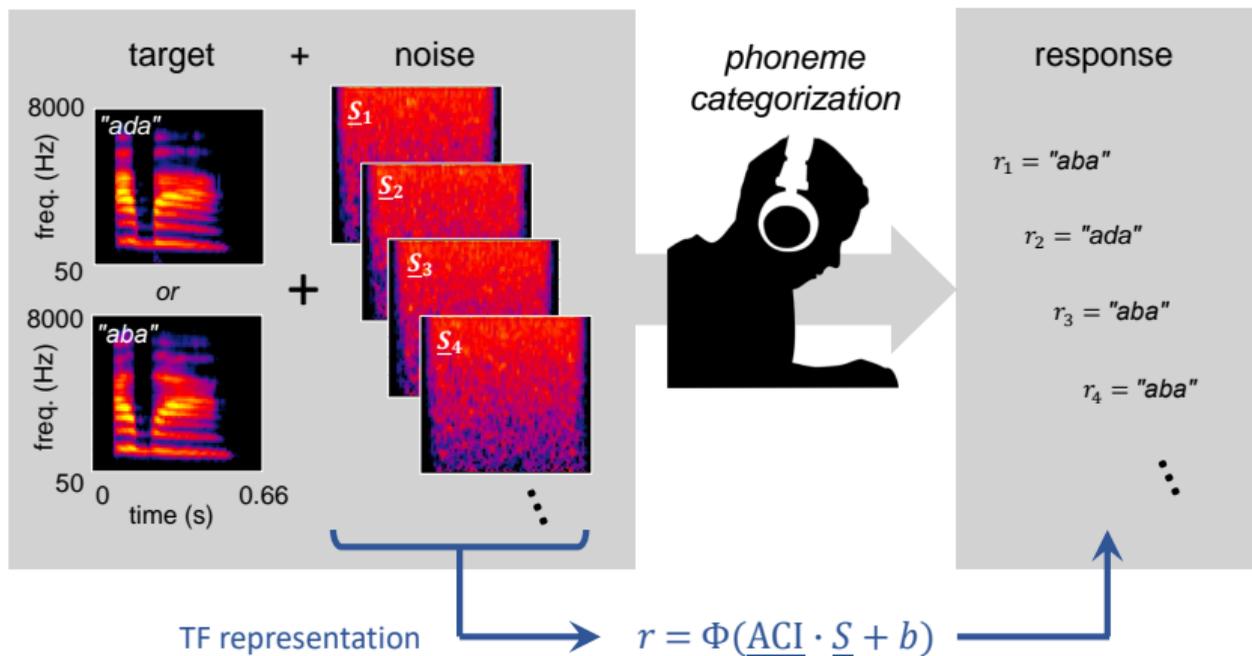
Using the revcorr paradigm to identify the acoustic cues underlying phoneme perception (in noise)  
 [Varnet et al. 2013, 2015, 2016, 2019...].



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



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



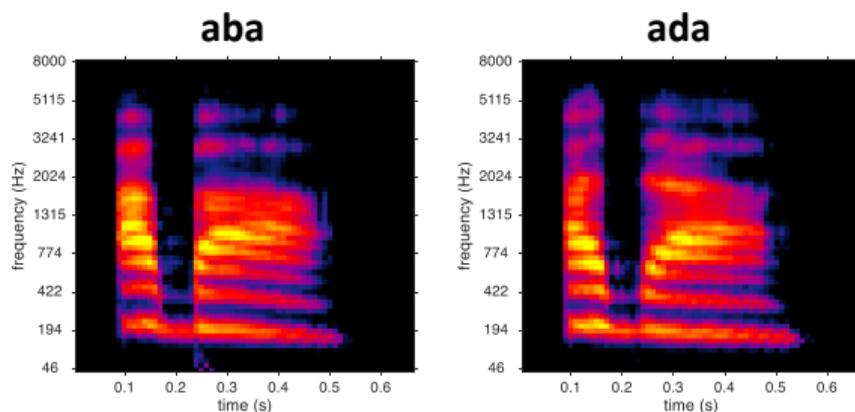
# Methods

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

**Task:** Indicate whether the target was /aba/ or /ada/

**SNR:** adapted continuously to ensure a correct response rate of 70.7% (weighted up-down rule).

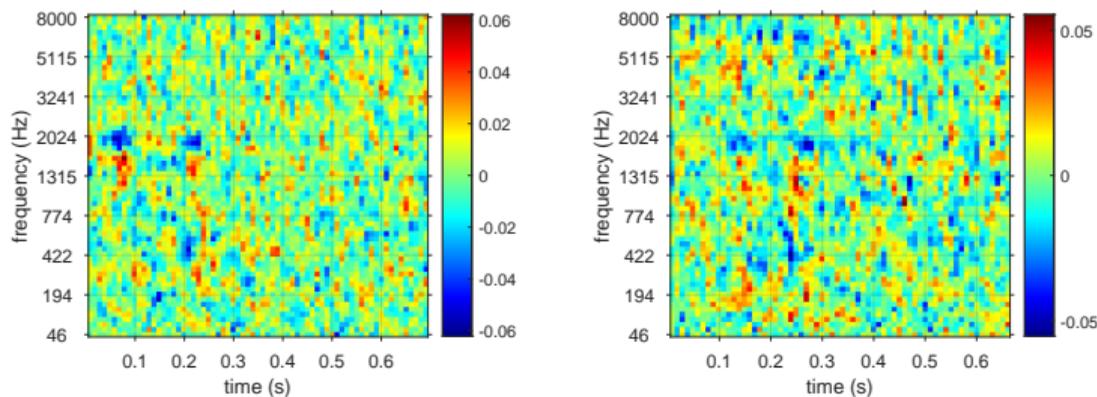
**Roving:**  $\pm 2.5$  dB.



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

## ACIs

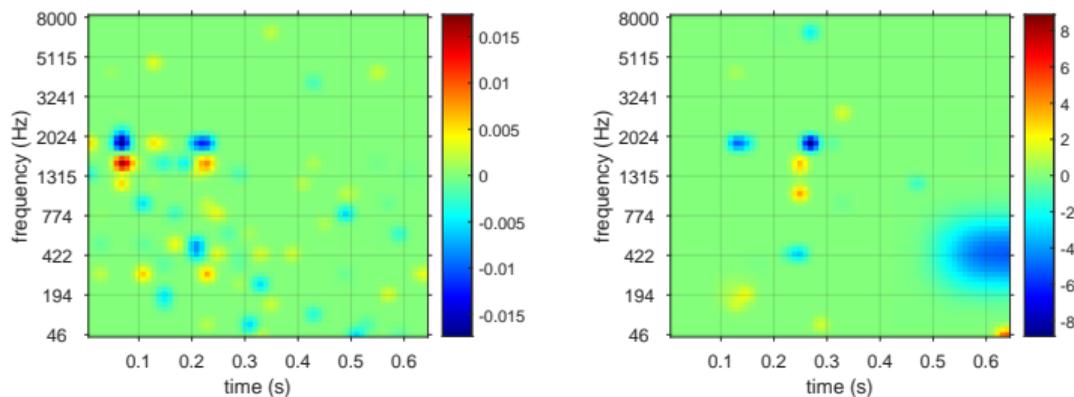
ACIs obtained from the classification data of participant LV, for two different experiments (left: [Varnet et al. 2013]; right: 2021 pilot)



Clusters of **positive** and **negative** weights corresponding to the acoustic cues.  
**Standard correlation matrix approach**

## ACIs

ACIs obtained from the classification data of participant LV, for two different experiments (left: [Varnet et al. 2013]; right: 2021 pilot)

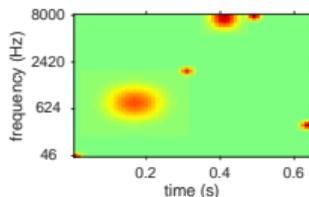


Clusters of **positive** and **negative** weights corresponding to the acoustic cues.

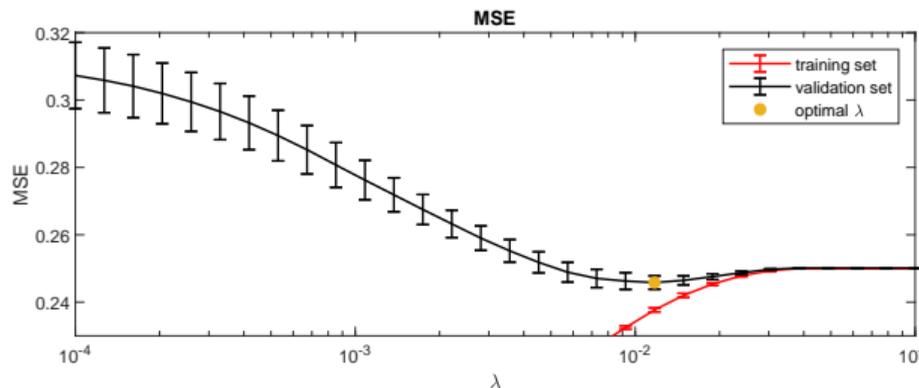
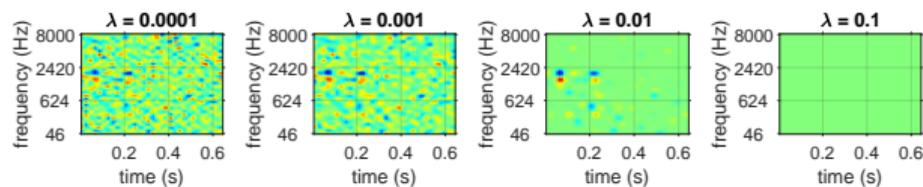
**GLM with a lasso regularization on a Gaussian basis** [Mineault et al., 2009] leads to a dramatic improvement in ACI resolution

# Lasso on Gaussian basis

Inspired by STRF estimation techniques and image denoising strategies. Reasonable assumption: the ACI is composed of a limited number of gaussian bumps.

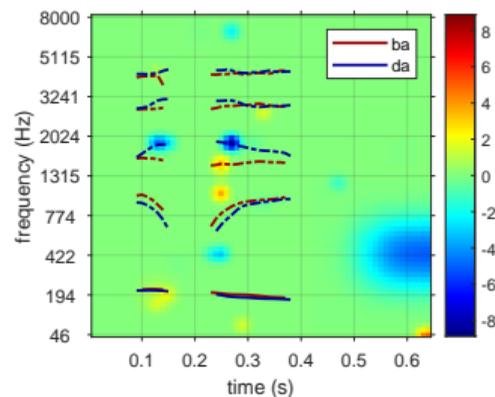
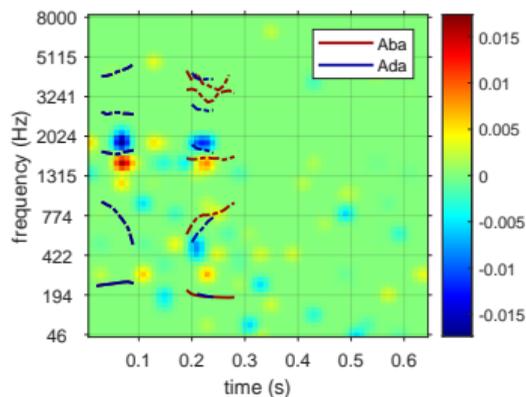


→ Finding the optimal number of bumps so that the GLM out-of-sample prediction accuracy is maximized (= the CV error is minimized).



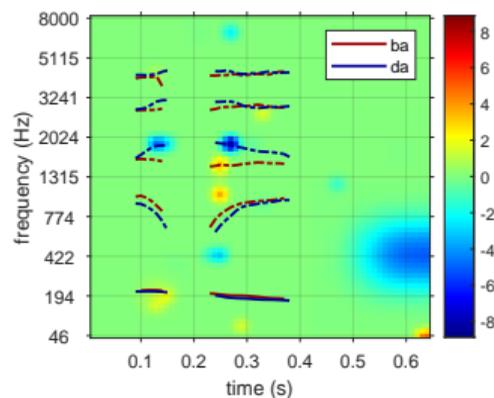
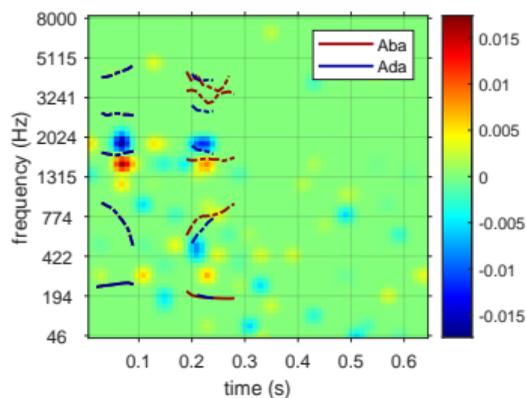
## ACIs

ACIs obtained from the classification data of participant LV, for two different experiments (left: *Varnet et al. 2013*; right: 2021 pilot)



## ACIs

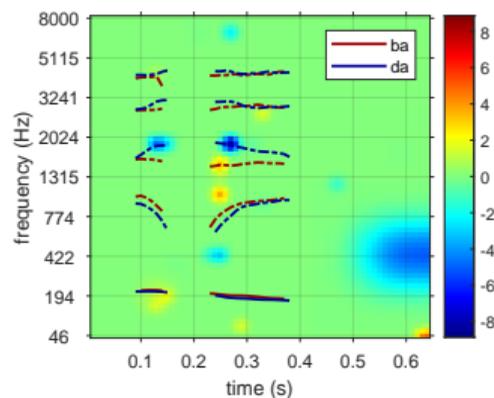
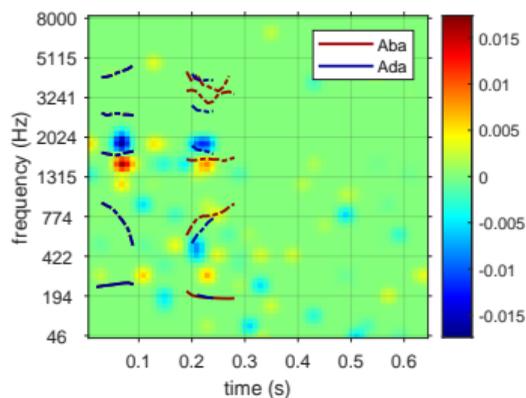
ACIs obtained from the classification data of participant LV, for two different experiments (left: *Varnet et al. 2013*; right: 2021 pilot)



- Confirms that **F2 onset is a cue** for classifying phonemes into /b/ or /d/ [*Liberman et al., 1954*]

## ACIs

ACIs obtained from the classification data of participant LV, for two different experiments (left: Varnet et al. 2013; right: 2021 pilot)



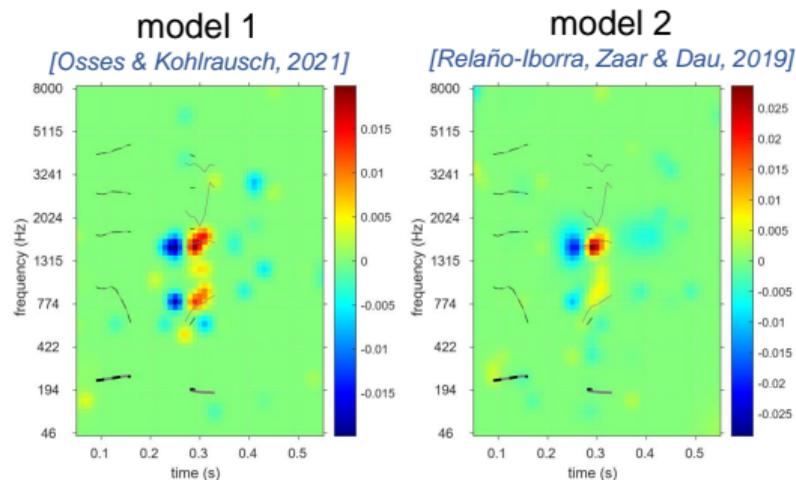
- Confirms that **F2 onset is a cue** for classifying phonemes into /b/ or /d/ [Lieberman et al., 1954]
- Other (unexpected) cues are involved in this task.

# Modeling ACIs

**Osses & Varnet (2021, DAGA):** deriving ACIs for two models of the human auditory system

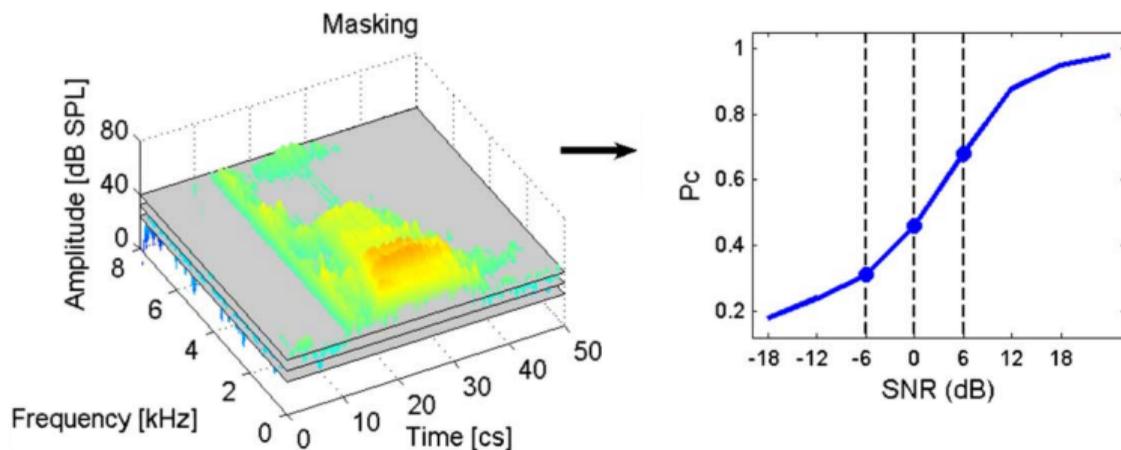
[Osses & Kohlrausch, 2021; Relaño-Iborra, Zaar & Dau, 2019]

- Similar to human participants, the two models relied on the F2 (and F1) onsets to discriminate /aba/ from /ada/
- However, they appear to use the temporal position of the formants instead of their onset frequency.



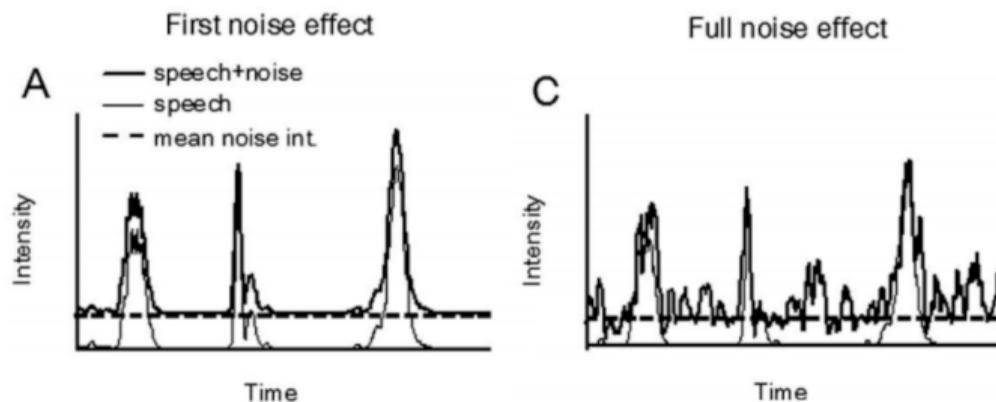
## Two effects of steady noise on phoneme perception

- Standard intelligibility models [e.g. AI, STI: Singh & Allen 2012; Houtgast et al. 2002] only consider the **macroscopic** effect of noise on speech perception, i.e., the effect of noise floor elevation (→ yields the same predictions for different realizations of the same noise type at same SNR).



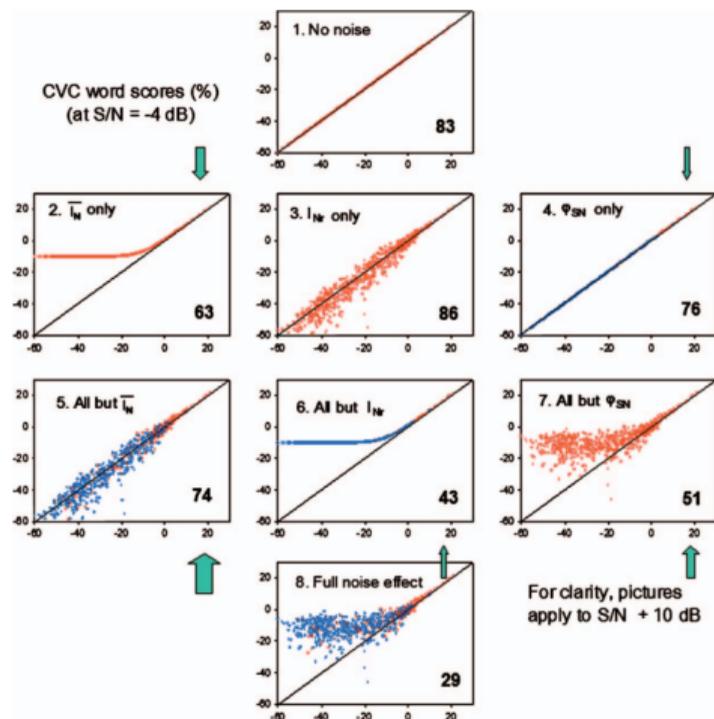
## Two effects of steady noise on phoneme perception

- Standard intelligibility models [e.g. AI, STI: Singh & Allen 2012; Houtgast et al. 2002] only consider the **macroscopic** effect of noise on speech perception, i.e., the effect of noise floor elevation (→ yields the same predictions for different realizations of the same noise type at same SNR).
- A few studies have considered a **microscopic** effect of noise: two noise realizations have different intrinsic envelop fluctuations and therefore they can have a different impact on perception [Drullman 1995a,b, 1997; Dubbelboer and Houtgast, 2007, 2008].



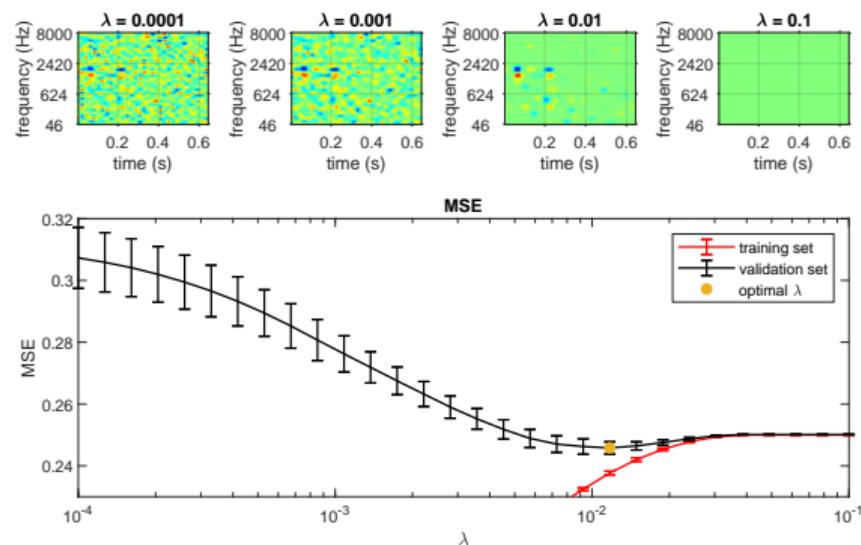
# Two effects of steady noise on phoneme perception

For CVC recognition in stationary noise, the microscopic effect of noise accounts only for a small proportion of errors (9%, vs. 54% for the full effect of noise) [Dubbelboer and Houtgast, 2007].



# Microscopic effect of noise on phoneme perception

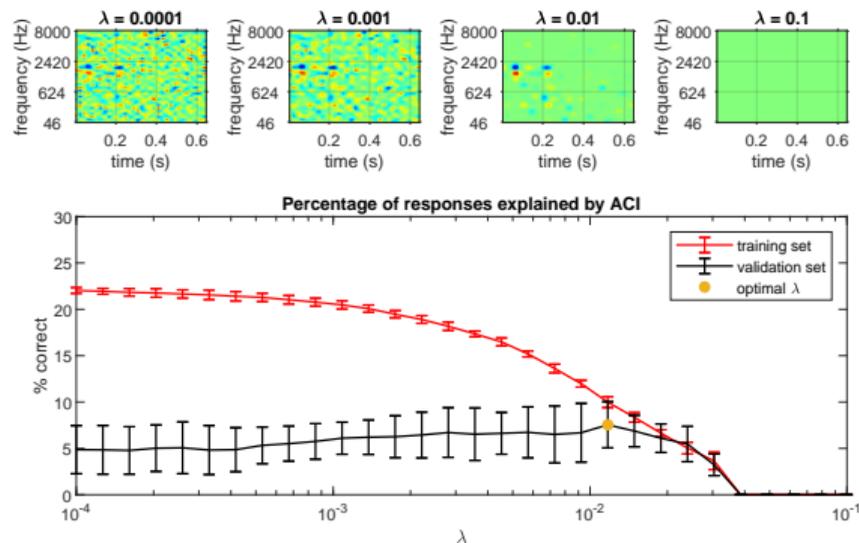
The ACI is a simple predictive model based on the microscopic effect of noise  $\rightarrow$  a way to quantify the impact of intrinsic fluctuations on phoneme categorization?



# Microscopic effect of noise on phoneme perception

The ACI is a simple predictive model based on the microscopic effect of noise  $\rightarrow$  a way to quantify the impact of intrinsic fluctuations on phoneme categorization?

The prediction accuracy benefit provided by the ACI, compared to a model accounting only for response bias, is  $\approx 8\%$ , consistent with the conclusions from Dubbelboer and Houtgast (my pilot data)

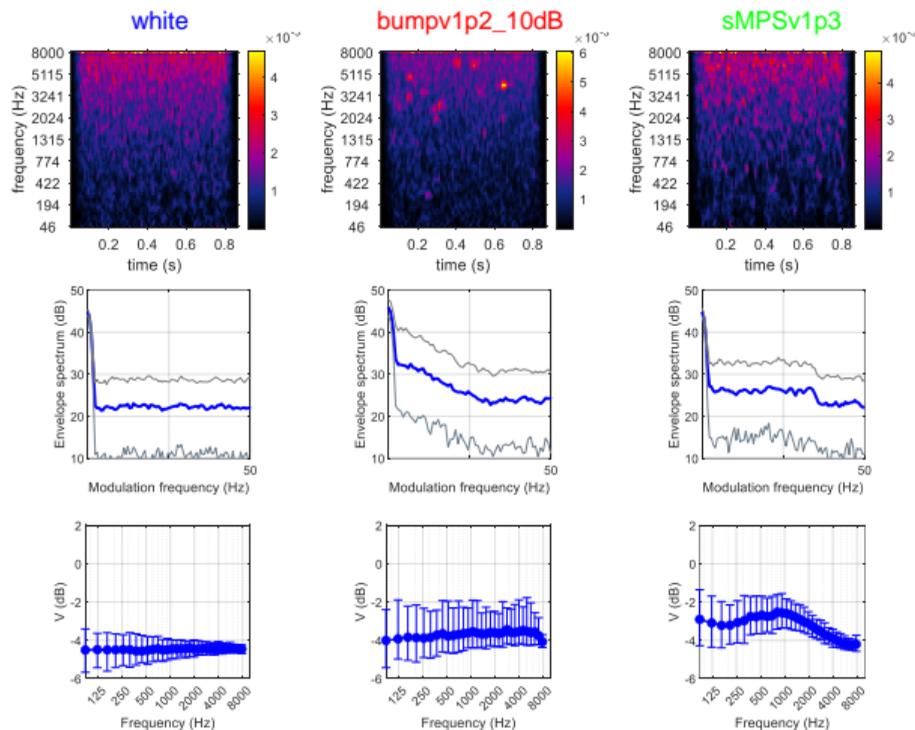


# Microscopic effect of noise on phoneme perception

Comparing the effect of 3 different types of noise:

- white noise
- bump noise (bumpv1p2\_10dB)
- noise low-pass filtered in the MPS domain (sMPSv1p3)

The two non-stationary noise induce larger random envelope fluctuations...

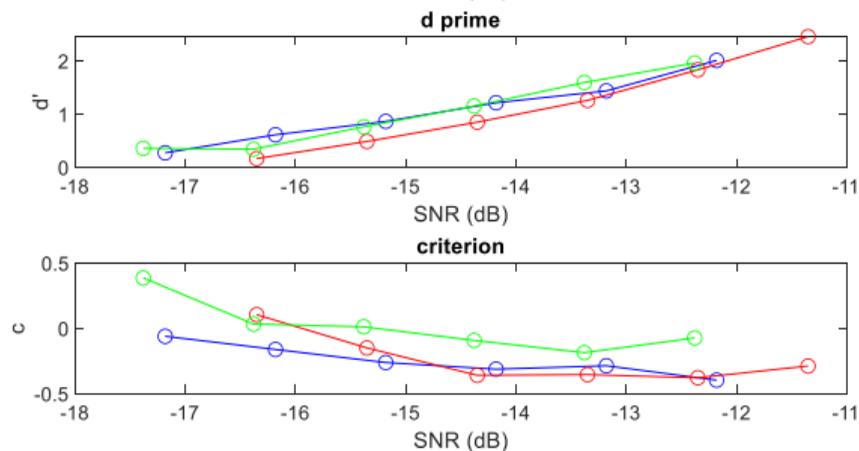
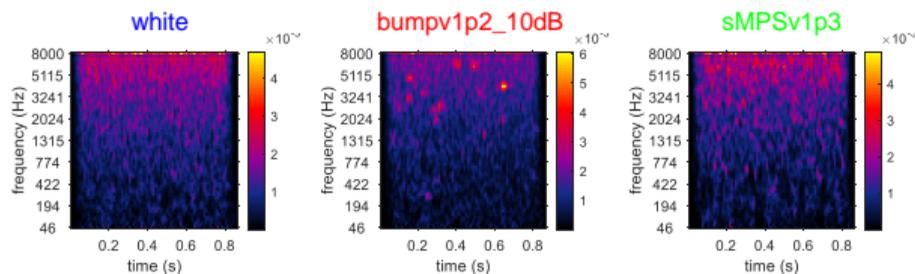


# Microscopic effect of noise on phoneme perception

Comparing the effect of 3 different types of noise:

- white noise
- bump noise (bumpv1p2\_10dB)
- noise low-pass filtered in the MPS domain (sMPSv1p3)

The two non-stationary noise induce larger random envelope fluctuations...

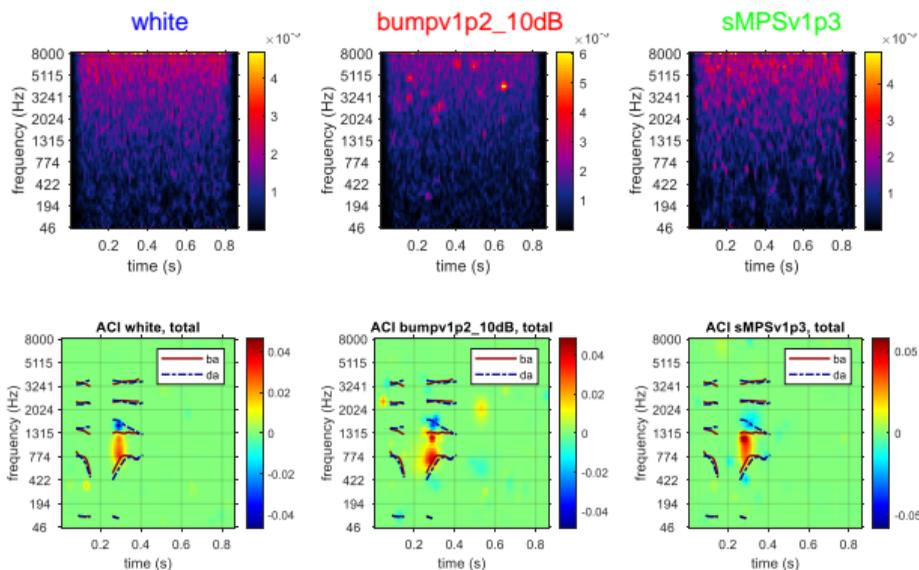


# Microscopic effect of noise on phoneme perception

Comparing the effect of 3 different types of noise:

- white noise
- bump noise (bumpv1p2\_10dB)
- noise low-pass filtered in the MPS domain (sMPSv1p3)

The two non-stationary noise induce larger random envelope fluctuations...

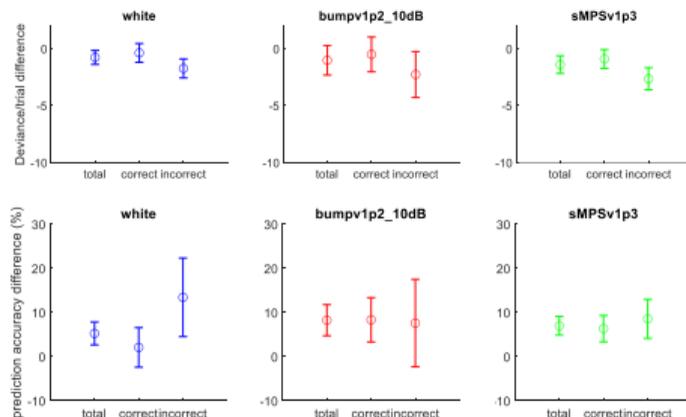
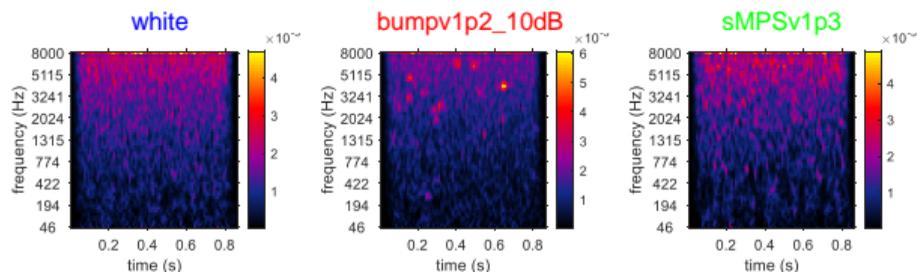


# Microscopic effect of noise on phoneme perception

Comparing the effect of 3 different types of noise:

- white noise
- bump noise (bumpv1p2\_10dB)
- noise low-pass filtered in the MPS domain (sMPSv1p3)

The two non-stationary noise induce larger random envelope fluctuations... which result in a stronger microscopic effect and therefore a **higher predictive power of the ACI** (my pilot data).



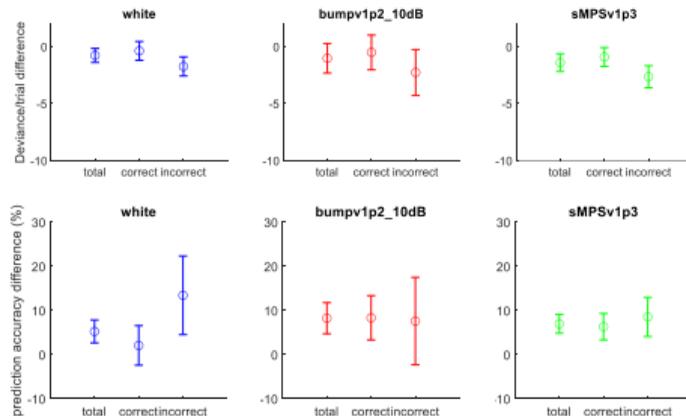
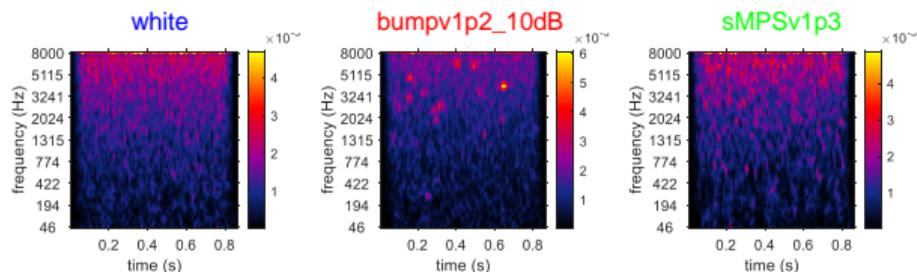
# Microscopic effect of noise on phoneme perception

Comparing the effect of 3 different types of noise:

- white noise
- bump noise (bumpv1p2\_10dB)
- noise low-pass filtered in the MPS domain (sMPSv1p3)

The two non-stationary noise induce larger random envelope fluctuations... which result in a stronger microscopic effect and therefore a **higher predictive power of the ACI** (my pilot data).

Objective of our next paper: measuring this effect on a group of  $N \approx 12$  participants (12.000 trials/participant!)



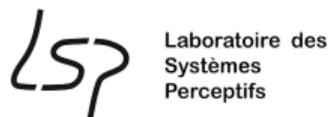
# Conclusions

- The auditory revcorr approach can reveal listening strategies in simple (e.g., AM detection) and complex (e.g., phoneme categorization) tasks...
- ...but it must be supplemented by auditory modelling (or system identification) analysis.
- However, the conclusions are based in fact on a very tenuous effect, raising epistemological and practical questions.

Objectives of the future Osses & Varnet paper:

- First (?) evidence and quantification of the microscopic effect using a 'natural' task.
- Demonstration that the relative roles of macroscopic/microscopic effect can vary depending on the type of noise.

**Any inputs?**



**Thank you for your attention!**

and thanks to **Alejandro Osses & Christian Lorenzi**