Using reverse correlation to study individual perception

Léo Varnet

École Normale Supérieure Paris

IDA/LSP day, 07/11/2022

Decoding speech

“Cracking the speech code”: finding the auditory primitives of speech comprehension

Decoding speech

"Cracking the speech code": finding the auditory primitives of speech comprehension

- Which acoustic cues allow the listener to differentiate one phoneme from another?
- Which acoustic cues underlie the segmentation of the speech signal into words?

Decoding speech

"Cracking the speech code": finding the auditory primitives of speech comprehension

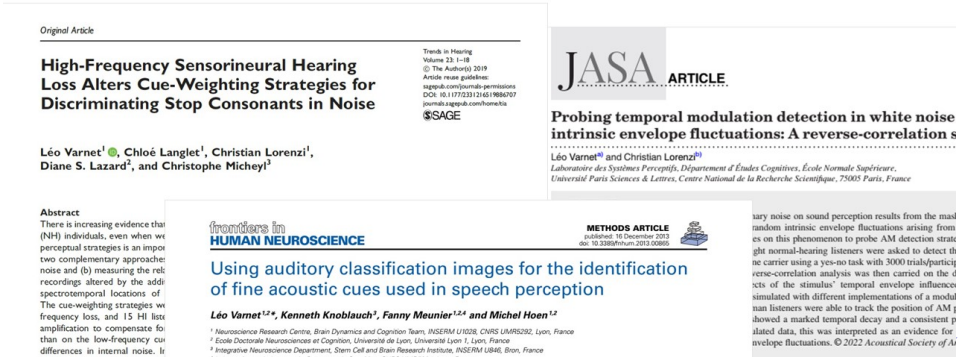
- Which acoustic cues allow the listener to differentiate one phoneme from another?
- Which acoustic cues underlie the segmentation of the speech signal into words?

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

Auditory revcorr

Reverse correlation (aka **revcorr**) is the perfect tool to reveal perceptual cues used in a psychophysical task, based on purely behavioral data... in particular for auditory categorization tasks. *[Varnet et al. 2013, 2015; Osses & Varnet, 2021; Varnet & Lorenzi, 2022]*

Core idea: adding **random fluctuations** to the stimulus and measure how they affect the participant's responses on a trial-by-trial basis.



Original Article

High-Frequency Sensorineural Hearing Loss Alters Cue-Weighting Strategies for Discriminating Stop Consonants in Noise

Trends in Hearing
Volume 23: 1–18
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2331216519886707
journals.sagepub.com/home/tia
SAGE

Léo Varnet[1], Chloé Langlet[1], Christian Lorenzi[1], Diane S. Lazard[2], and Christophe Micheyl[3]

Abstract
There is increasing evidence that (NH) individuals, even when we perceptual strategies is an impor two complementary approaches noise and (b) measuring the rel recordings altered by the addi spectrotemporal locations of The cue-weighting strategies we frequency loss, and 15 HI liste amplification to compensate fo than on the low-frequency cu differences in internal noise. In

JASA ARTICLE

Probing temporal modulation detection in white noise intrinsic envelope fluctuations: A reverse-correlation s

Léo Varnet[a] and Christian Lorenzi[b]
Laboratoire des Systèmes Perceptifs, Département d'Études Cognitives, École Normale Supérieure, Université Paris Sciences & Lettres, Centre National de la Recherche Scientifique, 75005 Paris, France

ary noise on sound perception results from the mask random intrinsic envelope fluctuations arising from es on this phenomenon to probe AM detection strate ght normal-hearing listeners were asked to detect th ne carrier using a yes-no task with 3000 trials/particip verse-correlation analysis was then carried on the d cts of the stimulus' temporal envelope influence simulated with different implementations of a modul nan listeners were able to track the position of AM p showed a marked temporal decay and a consistent p ulated data, this was interpreted as an evidence for nvelope fluctuations. © *2022 Acoustical Society of A*

frontiers in HUMAN NEUROSCIENCE

METHODS ARTICLE
published: 16 December 2013
doi: 10.3389/fnhum.2013.00865

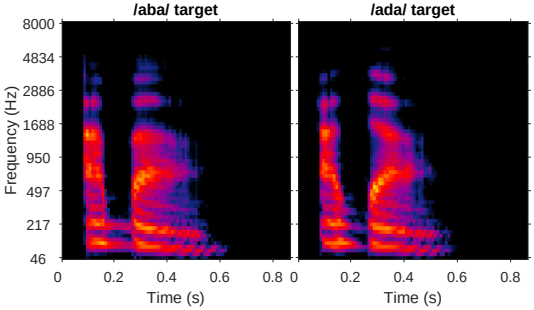
Using auditory classification images for the identification of fine acoustic cues used in speech perception

Léo Varnet[1,2]*, Kenneth Knoblauch*[3]*, Fanny Meunier*[1,2,4] *and Michel Hoen*[1,2]

[1] *Neuroscience Research Centre, Brain Dynamics and Cognition Team, INSERM U1028, CNRS UMR5292, Lyon, France*
[2] *Ecole Doctorale Neurosciences et Cognition, Université de Lyon, Université Lyon 1, Lyon, France*
[3] *Integrative Neuroscience Department, Stem Cell and Brain Research Institute, INSERM U846, Bron, France*

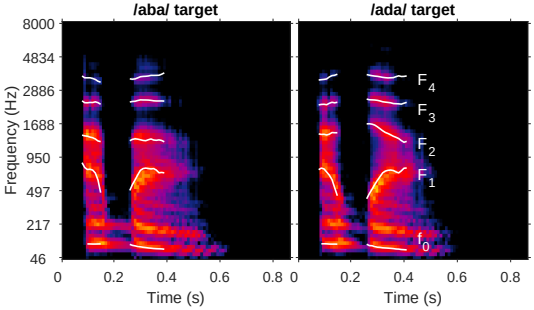
Aba/Ada experiment *[Osses & Varnet, in prep.]*

Topic: perception of stop consonants /b/ and /d/.
Targets: 2 VCV sounds (t_0=/aba/ and t_1=/ada/) from the Oldenburg Logatome Corpus *[Wesker et al., 2005]*, equalized in duration and rms.

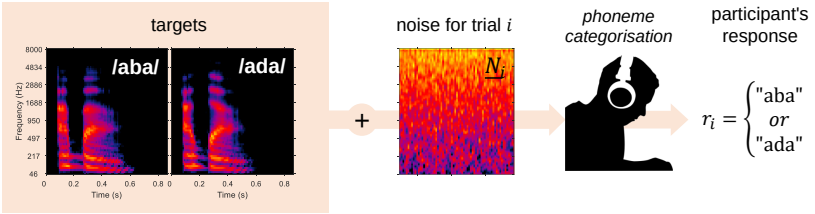


Aba/Ada experiment *[Osses & Varnet, in prep.]*

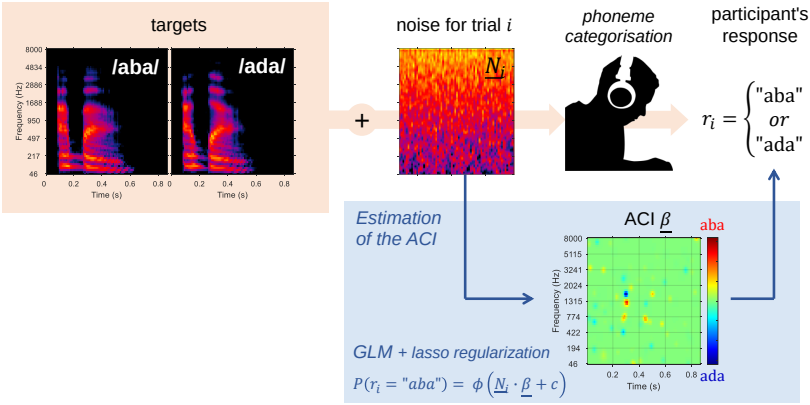
Topic: perception of stop consonants /b/ and /d/.
Targets: 2 VCV sounds (t_0=/aba/ and t_1=/ada/) from the Oldenburg Logatome Corpus *[Wesker et al., 2005]*, equalized in duration and rms.



Aba/Ada experiment *[Osses & Varnet, in prep.]*

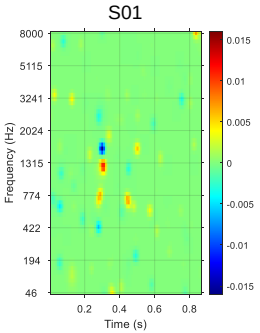


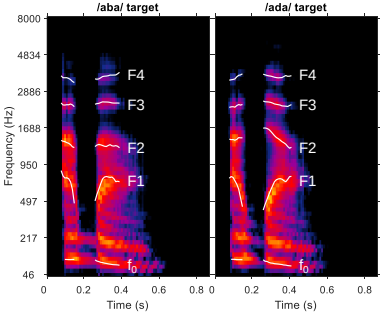
Aba/Ada experiment *[Osses & Varnet, in prep.]*



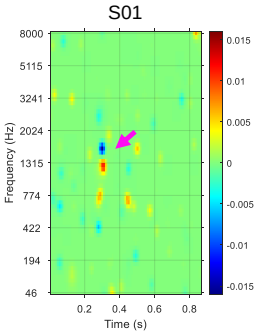
Auditory Classification Image (ACI): time-frequency matrix of decision weights. Shows how a specific noise configuration can mislead the participant.

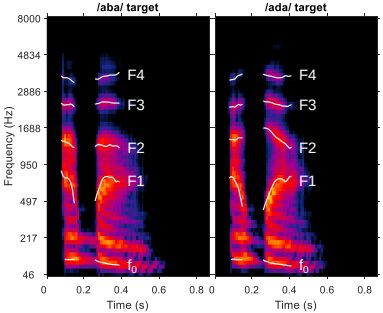
Results


S01
8000
5115
3241
2024
1315
774
422
194
46
Frequency (Hz)
0.2
0.4
0.6
0.8
Time (s)
0.015
0.01
0.005
0
-0.005
-0.01
-0.015

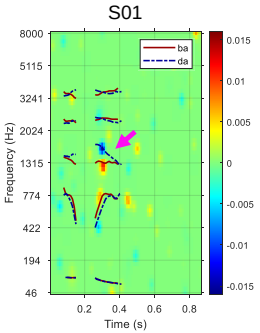

/aba/ target
/ada/ target
8000
4834
2886
1688
950
497
217
46
Frequency (Hz)
F4
F3
F2
F1
f_0
0
0.2
0.4
0.6
0.8
Time (s)

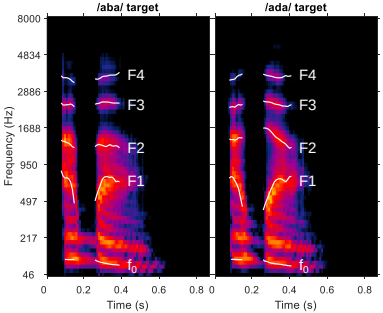
Results


S01
8000
5115
3241
2024
1315
774
422
194
46
Frequency (Hz)
0.2
0.4
0.6
0.8
Time (s)
0.015
0.01
0.005
0
-0.005
-0.01
-0.015


/aba/ target
/ada/ target
8000
4834
2886
1688
950
497
217
46
Frequency (Hz)
F4
F3
F2
F1
f_0
0
0.2
0.4
0.6
0.8
Time (s)

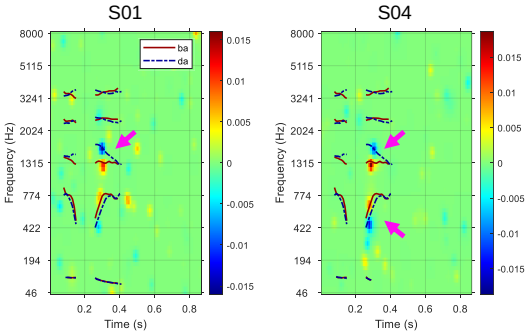
Results





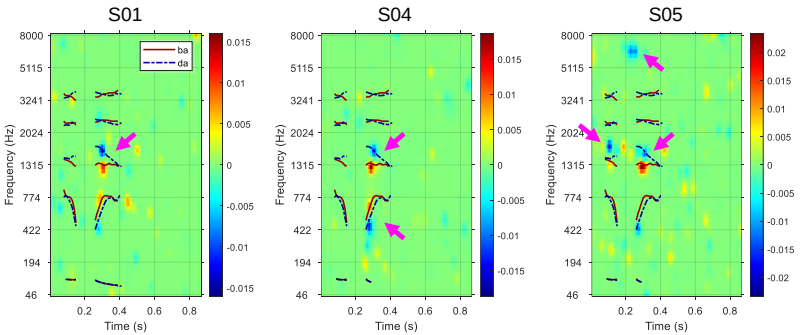
- The analysis successfully identified the **main cue** for the task (F2 onset), consistent with the phonetic literature *[Liberman, 1954]*...

Results



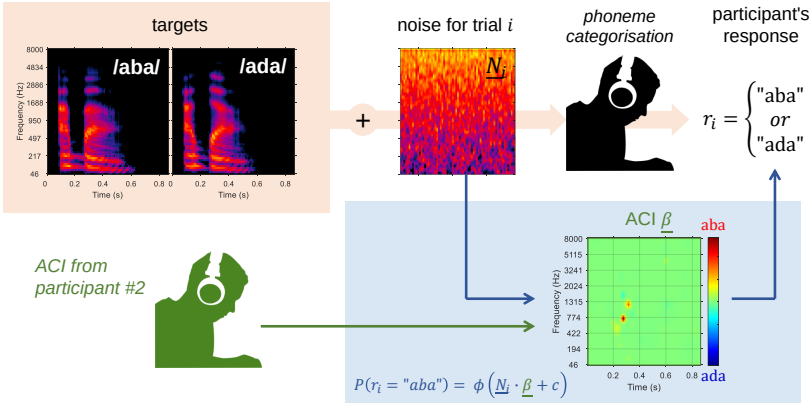
- The analysis successfully identified the **main cue** for the task (F2 onset), consistent with the phonetic literature *[Liberman, 1954]*...
- ...as well as several **secondary cues** (e.g., F1 onset).

Results



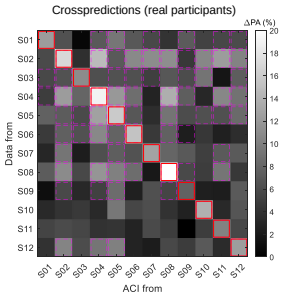
- The analysis successfully identified the **main cue** for the task (F2 onset), consistent with the phonetic literature *[Liberman, 1954]*...
- ...as well as several **secondary cues** (e.g., F1 onset).
- Contrary to our preregistered hypothesis, we observed some meaningful **interindividual variability** in the pattern of secondary cues.

Cross-predictions



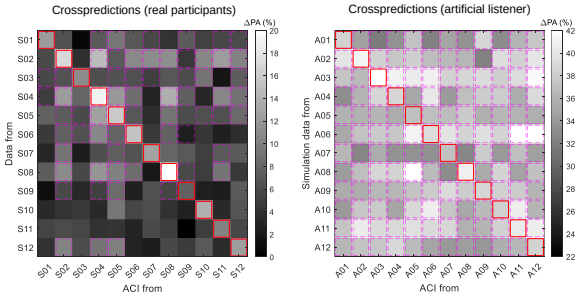
The similarity between listening strategies can be quantified by attempting to predict the responses of one participant using the ACI of another.

Cross-predictions



- Our group of normal-hearing participants shows significant heterogeneity in their listening strategies: the ACI of one participant is better at predicting new data from this participant, rather than new data from another participant.

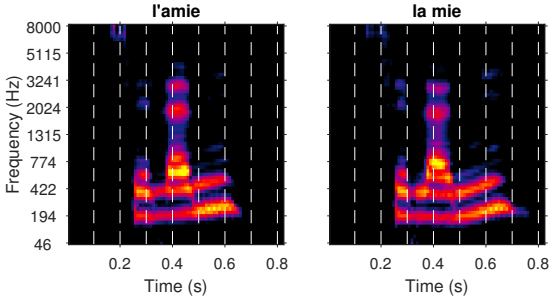
Cross-predictions



- Our group of normal-hearing participants shows significant heterogeneity in their listening strategies: the ACI of one participant is better at predicting new data from this participant, rather than new data from another participant.
- As confirmed with a simple model of the human auditory system.

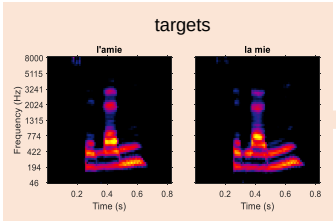
L'amie/La mie experiment *[Osses et al., in prep.]*

Topic: Acoustic cues for the segmentation of a speech sentence into words.
Targets: 2 phonetically identical sentences /selami/ (t_0="c'est l'amie" and t_1="c'est la mie"), equalized in duration and rms *[Spinelli al., 2010]*.

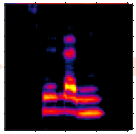


The targets were divided into 100-ms segments. Then, the f_0 trajectory in each segment was replaced by a random linear f_0 trajectory and each segment was compressed or elongated by a random amount, using WORLD *[Morise et al., 2016]*.

L'amie/La mie experiment *[Osses et al., in prep.]*



stimulus for trial i
(random prosody)

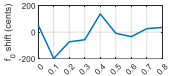


segmentation

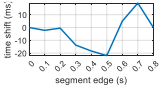
participant's response

$$r_i = \begin{cases} 1\ (\text{"}l'amie\text{"}) \\ 0\ (\text{"}la\ mie\text{"}) \end{cases}$$

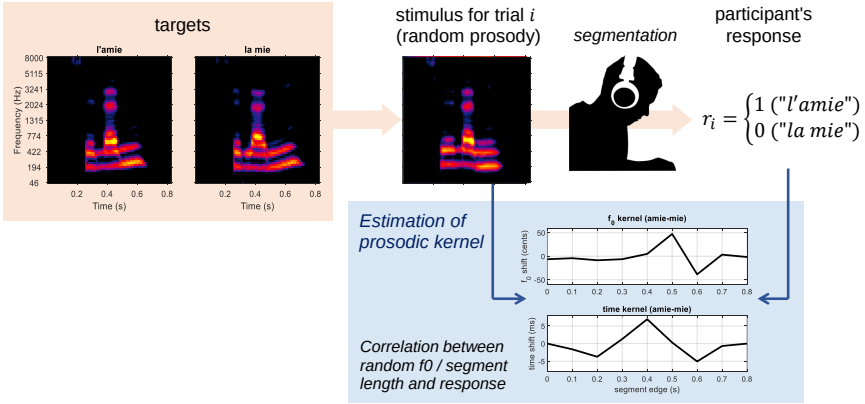
with random f0



and random segment length



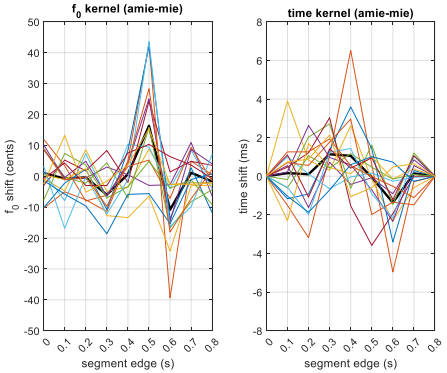
L'amie/La mie experiment *[Osses et al., in prep.]*



We obtain two kernels (f0 kernel and time kernel), indicating which aspects of the prosody are used as segmentation cues.

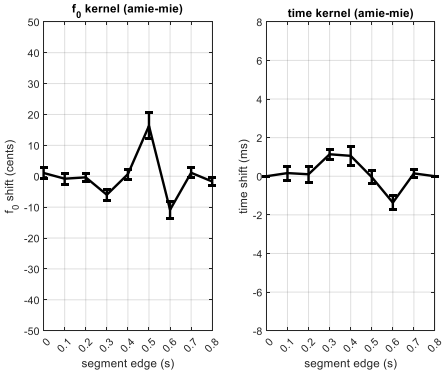
Results

- Considerable variability at the group level (N=15)



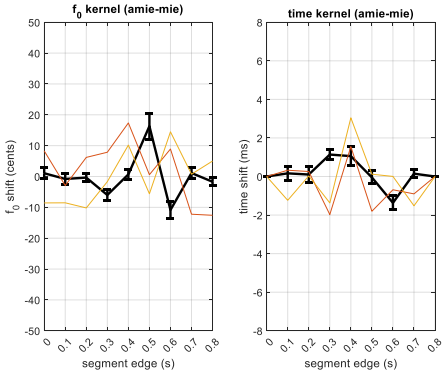
Results

- Considerable variability at the group level (N=15)
- Nevertheless, a clear prosodic pattern emerges



Results

- Considerable variability at the group level (N=15)
- Nevertheless, a clear prosodic pattern emerges
- Two dyslexic participants seem to use a different listening strategy for segmentation



Conclusions

- The **auditory revcorr approach** can reveal listening strategies and acoustic cues in psychoacoustic tasks (e.g., pitch perception) and psycholinguistic tasks (e.g., phoneme discrimination, segmentation).

Code available on GitHub as an open-source MATLAB toolbox with documentation and turnkey experiments *[Osses & Varnet, 2021]*: `https://github.com/aosses-tue/fastACI`

Conclusions

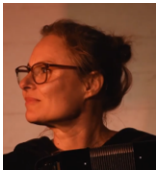
- The **auditory revcorr approach** can reveal listening strategies and acoustic cues in psychoacoustic tasks (e.g., pitch perception) and psycholinguistic tasks (e.g., phoneme discrimination, segmentation).
- It can also produce reliable results at the individual level, making it possible to explore **individual listening strategies**.
- **Heterogeneity** of individual strategies for phoneme discrimination in a normal-hearing group. Possible **difference between the strategies of** dyslexic and control participants in a segmentation task.

Code available on GitHub as an open-source MATLAB toolbox with documentation and turnkey experiments *[Osses & Varnet, 2021]*: `https://github.com/aosses-tue/fastACI`

Thanks for your attention!

And thanks to:



Alejandro **Osses**, Fanny **Meunier**, Étienne **Gaudrain**, Elsa **Spinelli**




Christian **Lorenzi**, Michel **Hoen**, Ken **Knoblauch**, Emmanuel **Ponsot**