Auditory reverse correlation: a choose-your-own-adventure presentation
Géraldine Carranante, Léo Varnet

To cite this version:
Géraldine Carranante, Léo Varnet. Auditory reverse correlation: a choose-your-own-adventure presentation. Doctoral. Laboratoire des Systèmes Perceptifs, France. 2024. hal-04526733

HAL Id: hal-04526733
https://hal.science/hal-04526733
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.
Auditory reverse correlation: a choose-your-own-adventure presentation

Géraldine Carranante, Léo Varnet

Laboratoire des Systèmes Perceptifs, École Normale Supérieure Paris

LSP meeting, 18/01/2024
Introduction

Choose-your-own-adventure books: a concept originally proposed by Raymond Queneau ("Un conte à votre façon"), and later extended to books, card games, video games, movies... ...and now to scientific presentations!
Choose-your-own-adventure books: a concept originally proposed by Raymond Queneau ("Un conte à votre façon"), and later extended to books, card games, video games, movies... ...and now to scientific presentations!

**Rules of the game:**

- As we proceed with the presentation, we will encounter crossroads where you must choose between different paths.
- Two options represented by colors will appear on the slides: raise the colored panel corresponding to your choice.
- Majority wins.
Choose-your-own-adventure books: a concept originally proposed by Raymond Queneau ("Un conte à votre façon"), and later extended to books, card games, video games, movies... ...and now to scientific presentations!

Rules of the game:

- As we proceed with the presentation, we will encounter crossroads where you must choose between different paths.

Time to choose

- It's crystal clear, let's begin!
- I'm lost, can you repeat?
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; Varnet & Lorenzi, 2022; Osses & Varnet, 2023]

Core idea: adding random fluctuations to the stimulus and measure how they affect the participant’s responses on a trial-by-trial basis.
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; Varnet & Lorenzi, 2022; Osses & Varnet, 2023]

Time to choose

- Tell me more about previous auditory revcorr studies
- No, I just want to build my own auditory revcorr experiment
Auditory revcorr studies (full diagram on https://dbao.leo-varnet.fr/)

- **sentence recognition**
  - [Venezia et al., 2016, 2019]

- **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]
The interpretation of the resulting “kernel” varies:

- Some say it is a **direct image of our mental representations** [e.g. Ponsot et al 2018]
- Others that it is merely a **trace of the auditory processes** [Varnet & Lorenzi, 2022]
- The most radical psychophysicists argue that it should rather be seen as a **transparent descriptor of the data**, not the system [Neri, 2018]

There is also a large heterogeneity in the experimental parameters (e.g., yes/no or 2AFC? CI or Bubbles?)

We are currently preparing a review of the philosophical and practical aspects of reverse correlation.

---

**Auditory revcorr review**

**The interpretation of the resulting “kernel” varies:**
- Some say it is a **direct image of our mental representations** [e.g. Ponsot et al 2018]
- Others that it is merely a **trace of the auditory processes** [Varnet & Lorenzi, 2022]
- The most radical psychophysicists argue that it should rather be seen as a **transparent descriptor of the data**, not the system [Neri, 2018]

There is also a large heterogeneity in the experimental parameters (e.g., yes/no or 2AFC? CI or Bubbles?)

We are currently preparing a review of the philosophical and practical aspects of reverse correlation.

---

**Title** | **Authors** | **Year** | **Journal** | **Ref id** | **Auditory process Task** | **N in N stimu/categ N trials** | **N alternatives** | **N participants** | **Ci/bubble** | **Performances** | **Esti** | **Ci model**
--- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---
Spectro-Temporal Weighting of Loudness | Oberfeld, Heeren, Remi | 2012 | PLoS ONE | 499 | Loudness perception | 1 | 2 (low mean / h) 503 (for the spek 2 soft / loud) x 2.5 listeners x 21b | 16 | 0.05 | 8 CI | not reported | fixed logistic regression
Temporal loudness weights for sounds with r Ponsot, Sussini, Saint Pi | 2013 | Journal of the Ac 1175 | Loudness coding Intensity discrimi | 2 | 2 (low mean / h) 200 * 6 blocks / 2 (soft / loud) | 12 | 11 normal | CI | not reported | fixed logistic regression
Spectral features for overall level discriminativ Doherty, Lutfi, | 1996 | Journal of the Ac 622 | Loudness Coding Intensity discrs | 2 | 2 (low mean / h) 1000 | 11 normal hearing | d=1 | adjusted individual linear regression
Stimulus Features in Signal Detection | Ahumada, Lowell, | 1971 | Journal of the Ac 42 | pure-tone detect Tones in-noise de | 1 | 2 (tone absent), 50 X 32 repetitions (tone absent, 40 (exp 1), 7 (exp CI | not reported | fixed to XXX | linear regression
Time and frequency analyses of auditory sign- Ahumada, Marken, Sar | 1975 | Journal of the Ac 43 | pure-tone detect Tones in-noise de | 1 | 2 (tone absent), 400 X 8 repetitions (tone absent, | mean d=1.9 | fixed to XXX | linear regression
Psychophysical spectro-temporal receptive f Shub, Richards, | 2009 | Hearing Research 494 | pure-tone detect Tones in-noise de 1 | 2 (tone absent), 20%–22% | 6 CI | 2 for condition CI | 7 | fixed to 0 dB | Conditional again?
Identification of stimulus cues in narrow-bands Schönfelder, Wichma | 2013 | Journal of the Ac 1164 | pure-tone detect Tones in-noise de | 1 | 2 (tone absent), 40%–22% | 6 CI | 2 for condition CI | 7 | fixed to 0 dB | Conditional again?
Dynamic Reweighting of Auditory Modulation Joosten, Shamma, Lora | 2016 | PLoS One | 676 | modulation detect percept of loudness | 1 | 2 | 5.5Hz±2.8K (inre) B | 10 CI | 74±6% (inrems) fixed individually Conditional again
Human pitch detectors are tuned on a fine scale Joosten, Neri, Bar | 2012 | Biological Cybernetics | pure-tone detect Frequency discrimination | 2 | 2 (5.5Hz target) 100 Hz in 100 tria | 7±2 (and an additional) | 10 CI | 74±6%; 50% 4 individually fixed conditional average
Mechanisms of Spectrotemporal Modulation Ponsot, Varnet, Winters | 2021 | Trends in Hearing | modulation percept spectro-temporal | 2 | [fixed present, mod absent] | [modulation presented in 1st b CI] | | 10 CI | individually fixed conditional average
Probing temporal modulation detection in vel Varnet & Lorenzi | 2022 | Journal of the Acoustica modulation percept Sinusoidal AM de | 1 | 2 | 3000 | 2 | 10 CI | 7 CI | fixed 70,7% (maem adaptive) Conditional average
Speech Reductions Cause a De-Weighting of Varnet, Meunier, Ho | 2016 | Interspeech | phoneme percep consonant categor | 1 | 4 | 1000 | 2 | 7 CI | mean 60.2% and adjusted individual Conditional average

---

Choose your own adventure
Géraldine Carranante, Léo Varnet

Introduction
Auditory revcorr: a review
Decoding speech
L’amie/La mie
Gender-fair language
Abatada experiment
Stop consonants experiment
Diversity
fastACI toolbox
Auditory revcorr: technical details
Internal representation
Conceptual plurality
Conclusion
The interpretation of the resulting “kernel” varies:

- Some say it is a **direct image of our mental representations** [e.g. Ponsot et al 2018]
- Others that it is merely a **trace of the auditory processes** [Varnet & Lorenzi, 2022]
- The most radical psychophysicists argue that it should rather be seen as a **transparent descriptor of the data**, not the system [Neri, 2018]

There is also a large heterogeneity in the experimental parameters (e.g., yes/no or 2AFC? CI or Bubbles?)

We are currently preparing a review of the philosophical and practical aspects of reverse correlation.

**Time to choose**

- Yay! philosophy!
- Can you be more specific? Give us some examples of how revcorr can be applied to speech perception.
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 underlie the segmentation of the speech signal into words?
- Which acoustic cues allow listeners to differentiate one phoneme from another?
Decoding speech

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

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

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

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

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

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

Time to choose

- Let’s talk about segmentation cues!
- Let’s talk about phonetic cues!
Segmentation cues

How do listeners find the boundaries between words in continuous speech? (e.g. “I owe you a yoyo”)

- lexical cues [McClelland & Elman, 1986]
- phonotactic cues [McQueen, 1998]

... but that’s not the whole story!
Segmentation cues

We are able to discriminate between **phonemically identical sentences**

*Dis, à m’aimer, consens, va ! / Dis à mémé qu’on s’en va.*
– “Andrea c’est toi”, Boby Lapointe

*Je suis la sève héritée. Je suis lasse et vérité. Je suis la sévérité.*
– “L’Oracle”, Jean Cocteau

*Dans ces meubles laqués, rideaux et dais moroses, Danse, aime, bleu laquais, ris d’oser des mots roses.*
– “Le Coffret de santal”, Charles Cros
Segmentation cues

We are able to discriminate between **phonemically identical sentences** ⇒ There must be some **acoustic cues** for segmentation

\[
\text{Dis, à m’aimer, consens, va ! / Dis à mémé qu’on s’en va.} \\
\text{– “Andrea c’est toi”, Boby Lapointe}
\]

\[
\text{Je suis la sève héritée. Je suis lasse et vérité. Je suis la sévérité.} \\
\text{– “L’Oracle”, Jean Cocteau}
\]

\[
\text{Dans ces meubles laqués, rideaux et dais moroses,} \\
\text{Danse, aime, bleu laquais, ris d’oser des mots roses.} \\
\text{– “Le Coffret de santal”, Charles Cros}
\]
Segmentation cues

We are able to discriminate between **phonemically identical sentences** ⇒ There must be some **acoustic cues** for segmentation

*Dis, à m’aimer, consens, va ! / Dis à mémé qu’on s’en va.*
– “Andrea c’est toi”, Boby Lapointe

*Je suis la sève héritée. Je suis lasse et vérité. Je suis la sévérité.*
– “L’Oracle”, Jean Cocteau

*Dans ces meubles laqués, rideaux et dais moroses,*
*Danse, aime, bleu laquais, ris d’oser des mots roses.*
– “Le Coffret de santal”, Charles Cros

*C’est la mie / c’est l’amie – [Spinelli et al., 2007]*
Segmentation cues

Standard approach for identifying segmentation cues: **manipulating** acoustic content

By artificially modifying the $f_0$ on a specific phonetic segment and running a perceptual experiment, we can demonstrate that this particular information is a cue for segmentation [Spinelli, Grimault, Meunier & Welby, 2010].

**Limitations** of this hypothesis-driven approach: potential cues tested one at a time
L’amie/La mie experiment [Osses et al., 2023]

**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 et al., 2010].

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

**Stimuli:** 800 target sounds resynthesised with random \(f_0\) and segment lengths.

**Task:** Indicate whether the target was “c’est l’amie” or “c’est la mie”.

[Osses et al., 2023]
L’amie/La mie experiment \cite{Osses et al., 2023}

\begin{align*}
S_i(r_i) &= \begin{cases} 
1 ("l’amie") \\
0 ("la mie")
\end{cases} \\
\text{participant's response}
\end{align*}

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

\begin{align*}
\text{stimulus for trial } i \\
\text{(random prosody)}
\end{align*}

\begin{align*}
\text{segmentation}
\end{align*}

\begin{align*}
\text{participant's response}
\end{align*}
L’amie/La mie experiment [Osses et al., 2023]

We obtain two kernels ($f_0$ kernel and time kernel), indicating which aspects of the prosody are used as segmentation cues.
Results: effect of $f_0$ on /a/ segment

- There is a systematic effect of the random $f_0$ at 0.5 s (/a/ segment) on the response of participants.

![Graph showing the proportion of "L'amie" responses as a function of $f_0$ shift (cents).]
Results: effect of $f_0$ on /a/ segment

- There is a systematic effect of the random $f_0$ at 0.5 s (/a/ segment) on the response of participants.
- This is in line with the notion of “early $f_0$ rise” in French prosody: a high $f_0$ on the first vowel of a content word is a cue for segmentation.

[Spinelli et al. 2010; Welby, 2007]
Results: effect of $f_0$ on /a/ segment

- There is a systematic effect of the random $f_0$ at 0.5 s (/a/ segment) on the response of participants.
- This is in line with the notion of “early $f_0$ rise” in French prosody: a high $f_0$ on the first vowel of a content word is a cue for segmentation.  
  \cite{Spinelli et al. 2010, Welby, 2007}
- The strength of this relationship corresponds to the 0.5-s weight in the $f_0$ kernel.

![Graph showing proportion of "la mie" responses vs $f_0$ shift (cents)]
Results: prosody kernels

- Considerable variability at the group level (N=16)
Results: prosody kernels

- Considerable variability at the group level (N=16)
- Nevertheless, a clear prosodic pattern emerges
Results: prosody kernels

- Considerable variability at the group level (N=16)
- Nevertheless, a clear prosodic pattern emerges
- Two main segmentation cues can be identified: the relative $f_0$ of vowels /e/, /a/ and /i/ and relative duration of vowels /e/ and /i/.
- Again, consistent with previous findings obtained using hypothesis-driven methods [Spinelli et al. 2010; Welby, 2007; Shoemaker, 2014]
• Similar results were found for other pairs of ambiguous stimuli.
• Secondary cues seem to depend on the phonetical content of the targets (e.g. duration of the second syllable).
• Similar results were found for other pairs of ambiguous stimuli.
• Secondary cues seem to depend on the phonetical content of the targets (e.g. duration of the second syllable).

Time to choose
• Wait! You said “L’amie”? Why not “L’ami”? Or “L’ami·e”
• Can we move on to the study of phonetic cues?
Gender-fair language and masculine bias [Spinelli, Chevrot & Varnet, 2023]

A recent psycholinguistics study on grammatical gender, gender-fair language and perceptual biases. Collaboration with Elsa Spinelli from LPNC (Grenoble).
Gender-fair language and masculine bias \cite{Spinelli, Chevrot & Varnet, 2023}.

In languages with grammatical gender, such as French, it has been evidenced that using masculine forms as a generic reference induces a bias favoring masculine specific representations \cite{Brauer & Landry, 2008; Gygax et al., 2008; Xiao et al., 2022}.

Two main strategies to counteract these gender biases \cite{Tibblin et al., 2022}:

- **Neutralization** (e.g. gender-unmarked forms such as “l’enfant”)
- **Re-feminization** (e.g. contracted double forms such as “un-e enfant”)

Here, we explored the relative efficiency of these strategies.
Gender-fair language and masculine bias [Spinelli, Chevrot & Varnet, 2023]

Sentence evaluation paradigm: decide whether a second sentence starting with a gendered personal pronoun (“il” or “elle”) is a sensible continuation of the first sentence written in a gender-fair form (gender-unmarked form or contracted double form).

Examples of stimuli

**Gender-unmarked**: L’otage ne mange pas depuis 10 jours. Elle a perdu beaucoup de poids.

**Contracted double**: Un·e otage ne mange pas depuis 10 jours. Elle a perdu beaucoup de poids.

**Filler sentence**: Le lion est le roi de la savane. Il doit aller chercher les enfants à l’école.

Response time is measured from the onset of the second sentence.
Gender-fair language and masculine bias [Spinelli, Chevrot & Varnet, 2023]

Longer response times for sentences beginning with “elle” compared to “il” in the Gender-unmarked condition, but not in the Contracted double condition.

Gender-unmarked forms do not fully neutralize the masculine bias, while contracted double forms (here using the mid-dot) are more effective in promoting gender balanced representations.
Gender-fair language and masculine bias [Spinelli, Chevrot & Varnet, 2023]

Longer response times for sentences beginning with “elle” compared to “il” in the Gender-unmarked condition, but not in the Contracted double condition.

Time to choose

- By the way, what do we mean when we talk about “mental representations”?
- Yeah, I have the same question actually.
Aba/Ada experiment [Osses & Varnet, 2023]

**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, 2023]

**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.

---

Alejandro Osses
Aba/Ada experiment [Osses & Varnet, 2023]

One solution: using **reduced speech**

- Low-/high-pass filtered speech [Fletcher, 1922]
Aba/Ada experiment [Osses & Varnet, 2023]

One solution: using **reduced speech**

- Low-/high-pass filtered speech [Fletcher, 1922]
- Synthetic speech continuum [Liberman et al., 1954]

→ Proof that the **F2 onset** is a cue for categorizing /b/-/d/-/g/?
Aba/Ada experiment [Osses & Varnet, 2023]

One solution: using **reduced speech**

- Low-/high-pass filtered speech [Fletcher, 1922]
- Synthetic speech continuum [Liberman et al., 1954]
- 3-Dimensional Deep Search [Li & Allen, 2012], etc...
Aba/Ada experiment \cite{Osses2023} 

One solution: using **reduced speech**

- Low-/high-pass filtered speech \cite{Fletcher1922}
- Synthetic speech continuum \cite{Liberman1954}
- 3-Dimensional Deep Search \cite{Li2012}, etc...

**Problem:** the speech comprehension system shows very efficient **strategy adaptation**.
Aba/Ada experiment [Osses & Varnet, 2023]
Aba/Ada experiment [Osses & Varnet, 2023]

Auditory Classification Image (ACI): time-frequency matrix of decision weights. Shows how a specific noise configuration can mislead the participant.
Results [Osses & Varnet, 2023]

The analysis successfully identified the main cue for the task (F2 onset), consistent with the phonetics literature [Liberman et al., 1954]... as well as several secondary cues (e.g., F1 onset).
Results [Osses & Varnet, 2023]

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

...as well as several **secondary cues** (e.g., F1 onset).
Stop consonants experiment [Carranante et al., in prep.]

- Discrimination task in noise
- 6 target sounds (plosive C in context /aCa/)
- 7 participants × 7 contrasts, 4,000 trials/condition

\[
\begin{array}{c|c|c|c|c|c|c|c|c|c|c}
\text{Labial} & \text{Dental} & \text{Velar} \\
\hline
\text{Voiced} & \text{aba} & \text{ada} & \text{aga} \\
\hline
\text{Voicing} & \text{apa} & \text{ata} & \text{aka} \\
\text{Contrasts} & \text{labial} & \text{dental} & \text{velar} \\
\end{array}
\]
Stop consonants experiment \cite{Carranante et al., in prep.}

- Discrimination task in noise
- 6 target sounds (plosive C in context /aCa/)
- 7 participants $\times$ 7 contrasts, 4,000 trials/condition

\begin{center}
\begin{tikzpicture}
  \node at (0,0) {	extbf{Place of articulation contrasts}}; \\
  \node at (0,1) {	extbf{labial}}; \node at (1,1) {	extbf{dental}}; \node at (2,1) {	extbf{velar}}; \\
  \node at (0,0.5) {aba} -- \node at (1,0.5) {ada} -- \node at (2,0.5) {aga}; \\
  \node at (0,-0.5) {apa} -- \node at (1,-0.5) {ata} -- \node at (2,-0.5) {aka}; \\
  \node at (0,1.5) {voiced}; \node at (0,0.5) {labial}; \node at (0,-0.5) {unvoiced}; \\
  \node at (1,1.5) {dental}; \node at (1,0.5) {dental}; \node at (1,-0.5) {dental}; \\
  \node at (2,1.5) {velar}; \node at (2,0.5) {velar}; \node at (2,-0.5) {velar}; \\
\end{tikzpicture}
\end{center}
Stop consonants experiment \([\text{Carranante et al., in prep.}]\)

- Discrimination task in noise
- 6 target sounds (plosive C in context /aCa/)
- 7 participants \(\times\) 7 contrasts, 4,000 trials/condition

**Place of articulation contrasts**

\[
\begin{array}{ccc}
\text{labial} & \text{dental} & \text{velar} \\
\text{voiced} & \text{aba} & \text{ada} & \text{aga} \\
\text{Voicing} & \text{apa} & \text{ata} & \text{aka} \\
\text{contrasts} & & & \\
\text{unvoiced} & & & \\
\end{array}
\]
Group-averaged ACIs

Place of articulation cues:

- F2 onset
- F1 onset
- release burst (/d/ and /t/)
- coarticulation on F1 and F2
Group-averaged ACIs

**Voicing** cues:
- voicing bar
- F1 onset
- release burst (only for /t/)

\[ \text{aga-aka} \quad \text{ada-ata} \quad \text{aba-apa} \]

These are the group-averaged ACIs. What about individual results?

Ok but the concept of "phoneme" has multiple definitions in the linguistic literature. How do you deal with that?
Group-averaged ACIs

**Voicing** cues:
- voicing bar
- F1 onset
- release burst (only for /t/)

![Graphs showing ACIs for different sounds](image)

**Time to choose**
- These are the group-averaged ACIs. What about individual results?
- Ok but the concept of “phoneme” has multiple definitions in the linguistic literature. How do you deal with that?
Individual ACIs

- Group-averaged ACIs is not the only relevant result of revcorr
- Individual ACIs show robust individual strategy (i.e. acoustic cues) used by the participant in the task
- It confirms that most participants use several cues to discriminate a single phoneme (*redundancy of acoustic cues*)
- The mean results hide some interindividual diversity
Interindividual diversity

• some participants use some acoustic cues and not others
• how could we explain this diversity when the group of participants is homogeneous? (NH participants, french locutors).
Interindividual diversity

Choose your own adventure
Géraldine Carranante, Léon Varnet

Introduction
Auditory revcorr: a review
Decoding speech
L’amie/La mie experiment
Gender-fair language
Aba/ada experiment
Stop consonants experiment
Diversity
fastACI toolbox
Auditory revcorr: technical details
Internal representation
Conceptual plurality
Conclusion
Interindividual diversity

Time to choose

- Wow! You totally convinced me, I’m ditching my past research to move to auditory revcorr experiments! How can I get started?
- Yeah, I have the same question actually
The fastACI toolbox

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

All experiments can be replicated within the toolbox and all analyses and figures can be exactly reproduced (fully reproducible workflow).

Connects with other open-source projects such as **WORLD-Straight** or the **Auditory Modelling Toolbox**.
The fastACI toolbox

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

All experiments can be replicated within the toolbox and all code is open-source.

**Time to choose**

- How does it work? I want to know more.
- Yeah yeah, cut the sales crap. I want to see what it really does.

<table>
<thead>
<tr>
<th>LeoVarnet</th>
<th>Adding experiment speechACI_Audika</th>
</tr>
</thead>
</table>

- **Base**: Adding option keyvaltskip_if_on_disk = 2
- **Calibration**: Saving the model calibration in the Results folder
- **Defaults**: Gammarone_pre, checking that fs element that are multiples of a pow...
- **Experiments**: Adding experiment speechACI_Audika
- **Interface**: segmentation experiment
The fastACI toolbox

The experiment is defined by a set of 4 scripts listing the experimental parameters, similar to the AFC toolbox [Ewert, 2013]. Stimuli are stored on disk or can be re-generated from their seed.

A basic text-based interface allows the participants to enter their responses. Optionally, the participant can be replaced by a model from AMT [Majdak et al, 2021].

The base functions and subfunctions from the fastACI environment ensure the data collection and analysis.
The fastACI toolbox

A few experiments are available natively:

- speechACI_Logatome-aCCa-S43M [Carranante et al., in prep]
- segmentation [Osses et al., 2023]
- speechACI_Logatome-abda-S43M [Osses & Varnet, 2023]
- modulationACI [Varnet & Lorenzi, 2022]
- speechACI_varnet2015 [Varnet & Lorenzi, 2015]
- speechACI_varnet2013 [Varnet & Lorenzi, 2013]
- ...and even ahumada1975 [Ahumada, Marken, & Sandusky, 1975]
The fastACI toolbox

A few experiments are available natively:

- speechACI_Logatome-aCCa-S43M \textit{[Carranante et al., in prep]}
- segmentation \textit{[Osses et al., 2023]}
- speechACI_Logatome-abda-S43M \textit{[Osses & Varnet, 2023]}
- modulationACI \textit{[Varnet & Lorenzi, 2022]}

Time to choose

- I’m curious about this segmentation experiment
- I’m curious about this speechACI_Logatome-abda-S43M experiment
Intrinsic fluctuations and revcorr

Part of the deleterious effect of a steady noise on speech 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]

Using the revcorr method to probe internal representations of sounds by relating random envelope fluctuations of masking noise with the response of listeners on a trial-by-trial basis.
Aba/Ada experiment [Osses & Varnet, 2023]

**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.

**Stimuli:** Target sounds in additive white noise (+ 2 non-stationary background conditions).

**SNR** adapted continuously to ensure a correct response rate of 70.7%.
Aba/Ada experiment [Osses & Varnet, 2023]

\[ r_i = \begin{cases} \text{"aba"} & \text{if participant's response is phoneme categorisation /aba/} \\ \text{"ada"} & \text{if participant's response is phoneme categorisation /ada/} \end{cases} \]

Auditory Classification Image (ACI): time-frequency matrix of decision weights. Shows how a specific noise configuration can mislead the participant.
Aba/Ada experiment [Osses & Varnet, 2023]

Auditory Classification Image (ACI): time-frequency matrix of decision weights. Shows how a specific noise configuration can mislead the participant.
Aba/Ada experiment [Osses & Varnet, 2023]

Three different types of noise:
- white noise
- bump noise (white noise with spectrotemporal regions of enhanced energy, randomly distributed)
- MPS noise (white noise low-pass filtered in the modulation domain)

Each participant performs 4,000 categorizations in each of the 3 noise conditions (12,000 trials ≈ 4h30)
Fitting procedure

Effects of type of noise:
- The final images are similar overall, with some minor differences
- The two types of noises with enhanced envelope fluctuation converge more rapidly
  \[ \rightarrow \text{Towards an adaptive bump noise?} \]
Fitting procedure

Comparing 3 different fitting procedures on the same set of data (4,000 trials in the bump noise condition).

Simple correlation
[Ahumada, 1996]

GLM with L2 penalty
[Varnet et al., 2013]

GLM with L1 penalty on a Gaussian Pyramid basis
[Osses & Varnet, 2023]

\[
P(r_i = \text{“aba”}) = \Phi(N_i^T \cdot ACI + c) \tag{1}
\]
Fitting procedure

The predictive performance of the model is quite low because the GLM does not include any information about the target sound. The effect of noise fluctuations alone explains \( \approx 8\text{-}12\% \) of the participants’ responses.

⇒ The estimation of the ACI relies only on a small subset of the data.
Fitting procedure

The predictive performance of the model is quite low because the GLM does not include any information about the target sound. The effect of noise fluctuations alone explains $\approx 8-12\%$ of the participants’ responses. The estimation of the ACI relies only on a small subset of the data.

Time to choose

- Ok that’s all well and good, but what’s the point of this research?
- Yeah I have the same question actually
Definition(s) of mental representation

- Old intuition: the mind works with ideas or images that resemble things of the world they represent. ([Aristotle, De anima]

- Computational model of the mind (XXe cent.): the mind manipulates internal states with symbolic content, that are "about" things of the world (aboutness = intentionality of the mind). ([Brentano, 1874; Fodor, 1975]

- these states are called mental or internal representations
Definition(s) of mental representation

There are several ways to understand the concept of MENTAL REPRESENTATION:

- "standard interpretation" of classical computational theories: internal representation = content of propositional attitudes.
- structural representations: scale-model of the world like "maps", in order to produce surrogate reasoning [Ramsey, 2007; Swoyer, 1991].

We defend that the relevant interpretation for revcorr studies is Ramsey’s structural representation. [Carranante and Varnet, sub.]
Defense of structural representations

CI are mostly interpreted via the linear observer model [Murray, 2011, Gosselin et al., 2003]

In this model, the perceptual system uses the *template* to "confront it" to the incoming sensory information in order to detect the signal in it. The template represents key characteristics of the signal, used for its detection.
Defense of structural representations

LOB templates show all the characteristics of a structural representation:

- They are "scale models" of the task-relevant environment (like a map)
- They are used for surrogate reasoning
- They provide scientific explanation of perceptual systems at Maar’s algorithmic level of explanation \([Maar, 1982]\)
Defense of structural representations

LOB templates show all the characteristics of a structural representation:

- They are "scale models" of the task-relevant environment (like a map)
- They are used for surrogate reasoning
- They provide scientific explanation of perceptual systems at Maar’s algorithmic level of explanation [Maar, 1982]

Time to choose

- Now that this is clear, what can you say about these representations in your revcorr psycholinguistic experiments?
- Facing several ways to understand a fundamental concept in science is frequent. Let’s explore some epistemological tools to evaluate the best version of a concept for a scientific investigation.
Some examples of conceptual plurality

Some examples from my own experience: perception [Carranante, 2020], soundscape [Grinfeder, 2022], phoneme
Some examples of conceptual plurality

Some examples from my own experience: perception [Carranante, 2020], soundscape [Grinfeder, 2022], phoneme

- "Morning star" and "Evening star" and the concept of VENUS
- The 2 concepts of PLANET in 2006, before Pluto’s classification as a DWARF PLANET [O’Madagain and Egré, 2019]
- The case of SPECIES:
  - 24 to 26 concepts of SPECIES
  - 3 main concepts: PHYLOSPECIES, BIOSPECIES, ECOSPECIES, i.e. lineage individuated by respectively phylogenetic, interbreeding and ecological approaches of species taxonomies. [Brigandt, 2010, 2012; Dupré, 1981; Ereshefsky, 1992, p.678; Grant, 1981]
- the 3 concepts of CONCEPTS [Machery, 2009]
Scientific concepts (1/3)

What is a **scientific concept**? A representational tool to talk about a category or class of things / phenomena / mechanisms, used in scientific theories and research.
Scientific concepts (2/3)

A scientific concept can be analyzed through 3 defining traits:

- **its intension** (definition, inferential role)
- **its extension** (reference)
- **its epistemic goal(s)**: the particular problem scientists are trying to solve by using this concept [Brigandt, 2012]

Exemples of epistemic goals:

- EG of NATURAL SELECTION: to account for evolutionary adaptation
- EG of EVOLUTIONARY NOVELTY: explaining the evolutionary origin of novelty.
- EG of HOMOLOGY: explaining the origin of similar morphological structures in different species. [Brigandt, 2010, 2012; Gouvea and Brigandt, 2023]
What is a **good** scientific concept?

- A functional tool, which provides good scientific explanations
- Epistemic utility theory: a good scientific belief is a belief with a good balance between plausibility and informativeness [*Huber, 2008*]
- Concept utility: Inclusiveness / Homogeneity
- A good scientific concept is a category inclusive enough to be informative; sufficiently homogeneous to support plausible inferences [*O’Madagain and Egré, 2019*]

What makes a concept **better** than another?

- Epistemic virtues: concept utility, operationalization
- Non-epistemic virtues: intelligibility, foster social justice, simplicity, sharability, ...
Defining a situation of conceptual plurality

- A situation of conceptual plurality arises when a scientist works with a concept that they thought was univocal but discovers that there are at least 2 versions of the concept used in their field.
- $C_1$ and $C_2$ are two versions of $C$ if they differ in at least one of the 2 elements: intension ($I(C_n)$) or extension ($E(C_n)$).
The 4 strategies

- selection of the best concept (e.g. PLANET)
- unification/integration into a new concept (e.g. VENUS)
- eliminativism/replacement (e.g. CONCEPT)
- pluralism (e.g. SPECIES, PERCEPTION)
The 4 strategies

- **selection** of the best concept (e.g. PLANET)
- unification/**integration** into a new concept (e.g. VENUS)
- **eliminativism** / replacement (e.g. CONCEPT)
- **pluralism** (e.g. SPECIES, PERCEPTION)

What is the best strategy?
All 4 strategies are relevant depending on the scientific context (**contextualist thesis**).

[Carranante, in prep.]
### Decision Tree

For C_n, well-formed scientific concepts

*Initial step: I(C_1) ≠ I(C_2)*

<table>
<thead>
<tr>
<th>Solution</th>
<th>Context in which it is the <em>best</em> solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integration</td>
<td>• Compatible extension</td>
</tr>
<tr>
<td>Selection</td>
<td>• Incompatible extensions</td>
</tr>
<tr>
<td></td>
<td>• Same epistemic goals</td>
</tr>
<tr>
<td>Pluralism</td>
<td>• Incompatible extensions</td>
</tr>
<tr>
<td></td>
<td>• Different epistemic goals</td>
</tr>
<tr>
<td></td>
<td>• There is no better alternative concept</td>
</tr>
<tr>
<td>Eliminativism</td>
<td>• Different extension</td>
</tr>
<tr>
<td></td>
<td>• Different epistemic goal</td>
</tr>
<tr>
<td></td>
<td>• There are better alternative concepts</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Does C_1 and C_2 have compatible extensions?</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
</tr>
<tr>
<td>Integration</td>
</tr>
<tr>
<td>no</td>
</tr>
<tr>
<td>yes</td>
</tr>
<tr>
<td>Selection</td>
</tr>
<tr>
<td>no</td>
</tr>
<tr>
<td>yes</td>
</tr>
<tr>
<td>Pluralism</td>
</tr>
<tr>
<td>no</td>
</tr>
<tr>
<td>yes</td>
</tr>
<tr>
<td>Eliminativism</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Does C_1 and C_2 have the same epistemic goals?</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
</tr>
<tr>
<td>Integration</td>
</tr>
<tr>
<td>no</td>
</tr>
<tr>
<td>Selection</td>
</tr>
<tr>
<td>yes</td>
</tr>
<tr>
<td>Pluralism</td>
</tr>
<tr>
<td>no</td>
</tr>
<tr>
<td>Eliminativism</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Are there alternative concepts?</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
</tr>
<tr>
<td>Eliminativism</td>
</tr>
<tr>
<td>no</td>
</tr>
<tr>
<td>Pluralism</td>
</tr>
<tr>
<td>Selection</td>
</tr>
<tr>
<td>Integration</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Are they better than C_1 and C_2?</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
</tr>
<tr>
<td>Eliminativism</td>
</tr>
<tr>
<td>no</td>
</tr>
<tr>
<td>Pluralism</td>
</tr>
<tr>
<td>Selection</td>
</tr>
<tr>
<td>Integration</td>
</tr>
</tbody>
</table>
The case of phoneme

There are two main understandings of the concept PHONEME:

- The historical definition of phoneme [Jakobson, 1942], as the smallest element of language that influences the meaning of words and sentences. This is a linguistic theoretical construct. This is neither a mental object nor an acoustic object.

- A psycholinguistic definition: phonemes are mental representations, activated by some acoustic properties in specific context.
The case of phoneme

There are two main understandings of the concept PHONEME:

- The historical definition of phoneme [Jakobson, 1942], as the smallest element of language that influences the meaning of words and sentences. This is a linguistic theoretical construct. This is neither a mental object nor an acoustic object.
- A psycholinguistic definition: phonemes are mental representations, activated by some acoustic properties in specific context.

Time to choose

- How do you study phonemes with revcorr?
- I actually have the same question
Conclusion

A fruitful collaboration between psychophysics and philosophy.

A philosophy of science “in the lab”

A psychophysical approach with a special care for the definition of concepts.
Choose your own adventure

Géraldine Carranante, Léo Varnet

Introduction
Auditory revcorr: a review
Decoding speech
L’amie/La mie experiment
Gender-fair language
Aba/ada experiment
Stop consonants experiment
Diversity
Conceptual plurality
Internal representation
Auditory revcorr: technical details
fastACI toolbox
Conclusion
Thanks for your attention!

Léo Varnet, Géraldine Carranante, Alejandro Osses

Fanny Meunier, Étienne Gaudrain, Elsa Spinelli, Maria Giavazzi