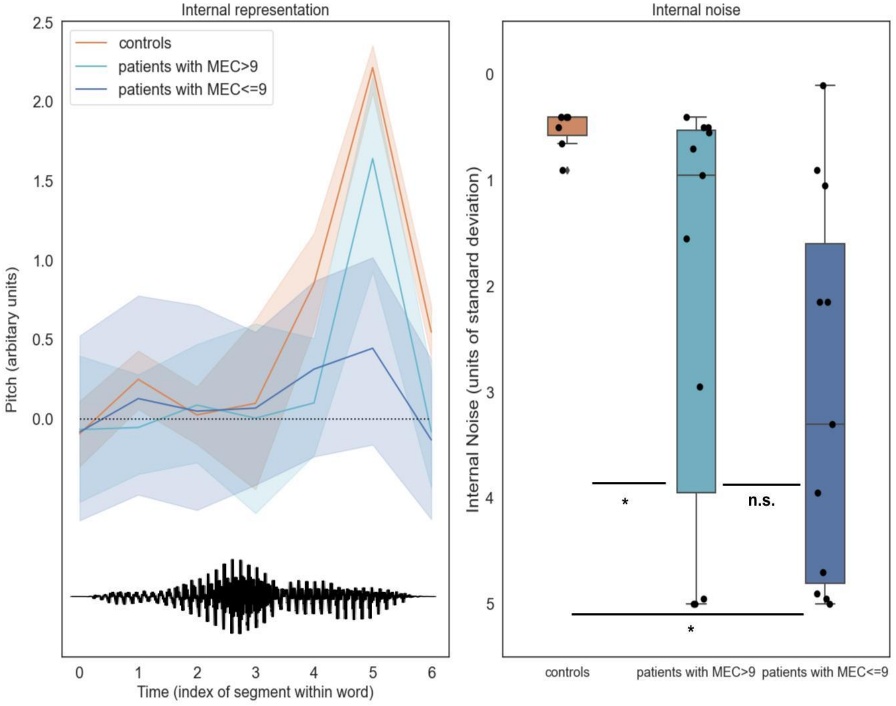
What am I working on now?

Aynaz Adl Zarrabi

Previously on Aynaz Thesis Diaries (ATD)

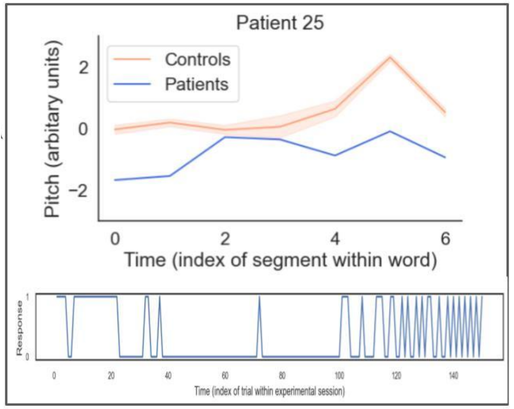
- I've been working on a reverse correlation experiment, to estimate the "internal representation" (or kernel) and "internal noise" of patients who had a right-hemisphere brain stroke, when they judge the interrogative prosody of a word.
- We have had basic results with N=22 patients, showing that representations and noise correlate with symptoms of "aprosodia".
- We have published a preprint (https://www.medrxiv.org/content/10.1101/2023.10.17.23297140v2) and are submitting the paper to journals (rejected at Brain, Stroke, etc. now trying Scientific Report)



femto-st SCIENCES & TECHNOLOGIES

What problem I have been working on this month

❖ One interesting finding from our previous results is that the patient "internal noise" seems to be an interesting biomarker.



The problem we have is that the classical way to estimate internal noise doesn't work well for these patients:

- first, it only measures internal noise in a limited number of repeated trials ("double pass") but our patient behavior is not stable in time, for instance because of perseveration.

- second, our patients have large values of internal noise (>2-3 std), and to estimate such large values, the double-pass method would need to repeat 1000s of trials, which is impractical.)

So what I'm working on now is how to estimate internal noise from continuous, non-repeated data.

In particular, I'm looking at estimating kernels and internal noise as parameters of a unique model, using a LM or GLM model (which was suggested by the work of Varnet & Knoblauch)

Things I have learned (1)

I have learned that we can reformulate the estimating of kernels and internal noise as a "regression" problem.

linear observer model:
Murray 2011

$y_1 = w \cdot s_1 + z \qquad y_2 = w \cdot s_2 + z$

$r = 1\ if\ y_1 > y_2 + bias$

$y = w(s_2\_s_1) + bias$

$y = mean(s_{chosen}\_s_{not-chosen})$

LM :
w are computed as the betas of a LM
focusing on minimizing the sum of squared residuals (OLS)

$y_i = \beta_0 + \beta_1 \cdot x_{i1} + \beta_2 \cdot x_{i2} + \cdots + \beta_6 \cdot x_{i6} + \varepsilon_i$

y_i is the binary response for the i^{th} stimulus heard

β_0 is the intercept that is constant that encodes the response bias on stimulus

$\beta_1, \beta_2, \dots, \beta_6$ are the coefficients for the independent variables 'pitch 0' to 'pitch6'

ε_i is the error term for the i^{th} stimulus

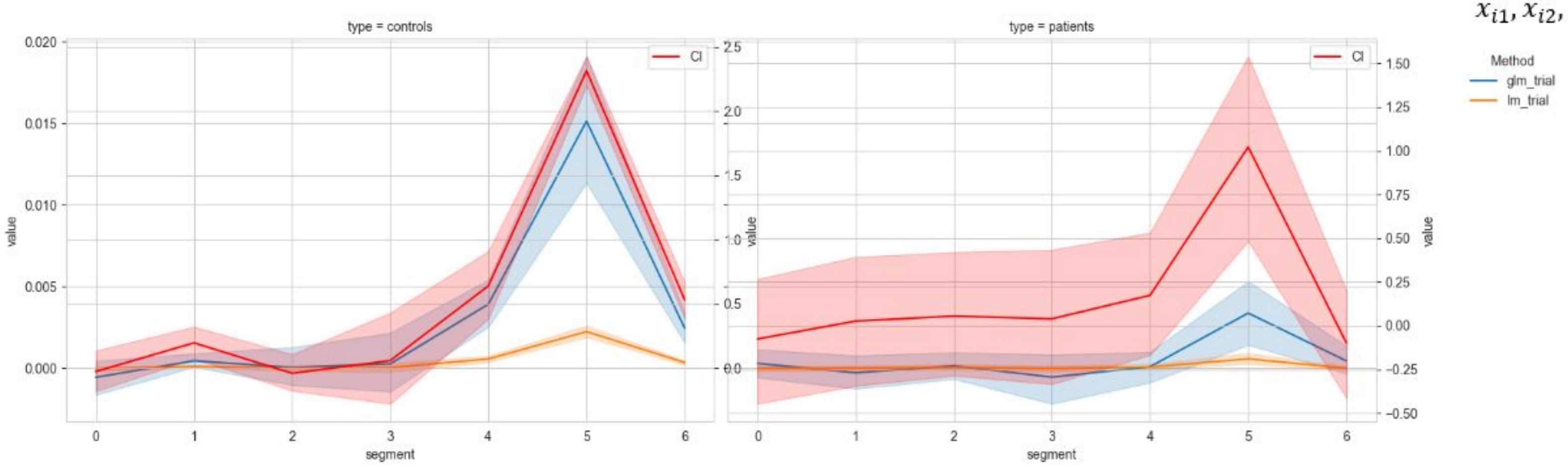
GLM:
the response is logistic
w are computed as the betas of the GLM
uses MLE to maximize the likelihood

logistic form : $P(y = 1|X) = \varphi\left(\beta_0 + \sum_{i=1}^{k} \beta_i \cdot X_i + \varepsilon_i\right)$

linear format: $(\log(\frac{y_i}{1 - y_i}) = \beta_0 + \beta_1 \cdot x_{i1} + \beta_2 \cdot x_{i2} + \cdots + \beta_6 \cdot x_{i6} + \varepsilon_i$

$\log(\frac{y_i}{1-y_i})$ is the logit (log odds) transformation of the probability y_i

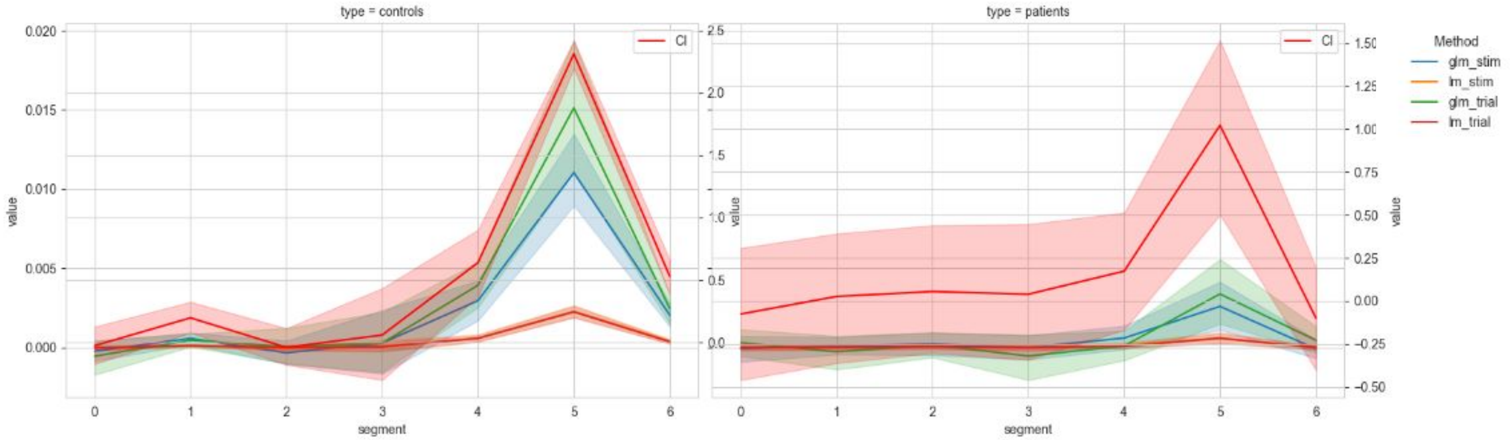
$x_{i1}, x_{i2}, \dots, x_{i6}$ are the values of the real values of pitch in 6 breakpoint (segments)

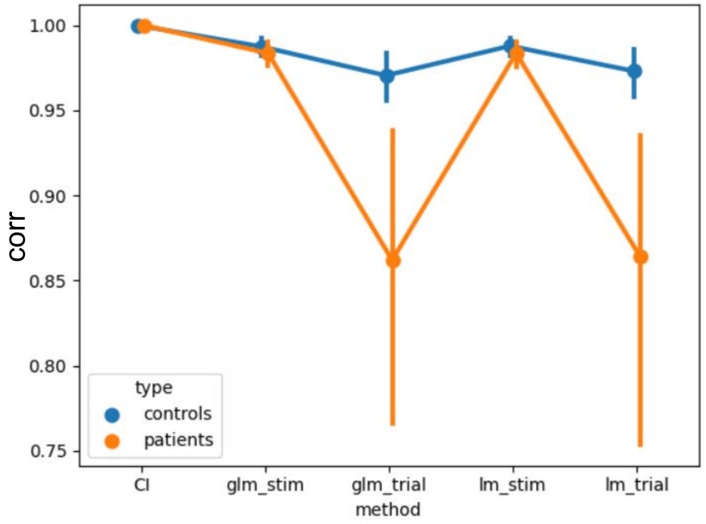


the difference between the models is the units of estimated coefficients (kernel amplitudes) for different methods across various segments.

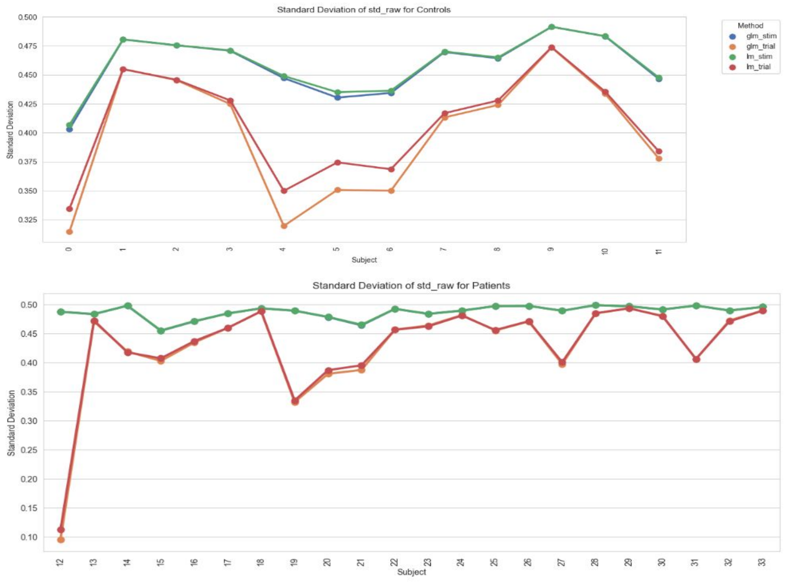
Things I have learned (2)

In more details, we can fit a LM
- on each individual sound (sound -> chosen / not)
- or each trial (difference of sounds -> chosen left or right)





The CI method assumes that this should yield the same results, but I have learned that this is not the case. While learning the LM/GLM on trials or stims is similar for controls, it can be quite different for patients, maybe because they do not evaluate both sounds in a trial independently (i.e. the assumptions of the CI model is false).



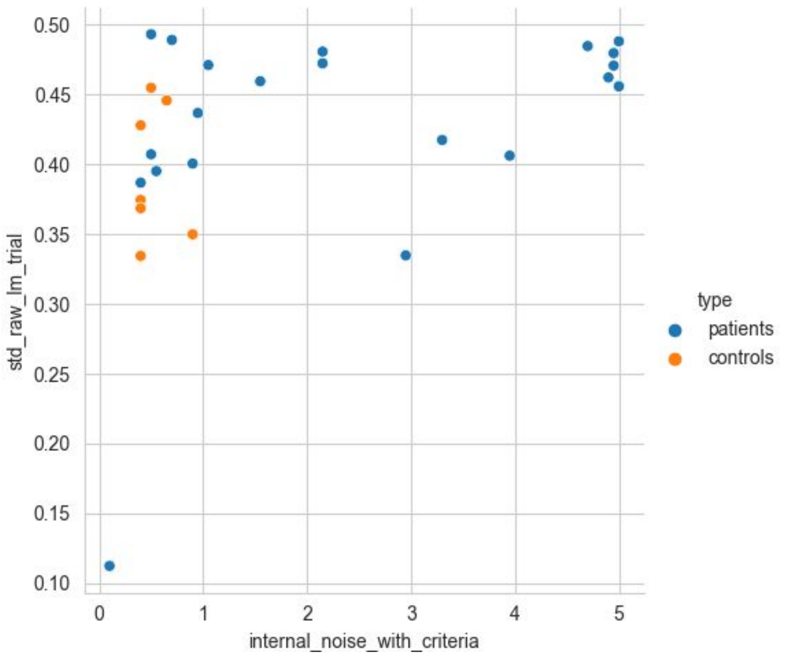
looking at the standard deviation of residuals, which shows smaller residuals for trials than for stims, regardless of lm/glm).

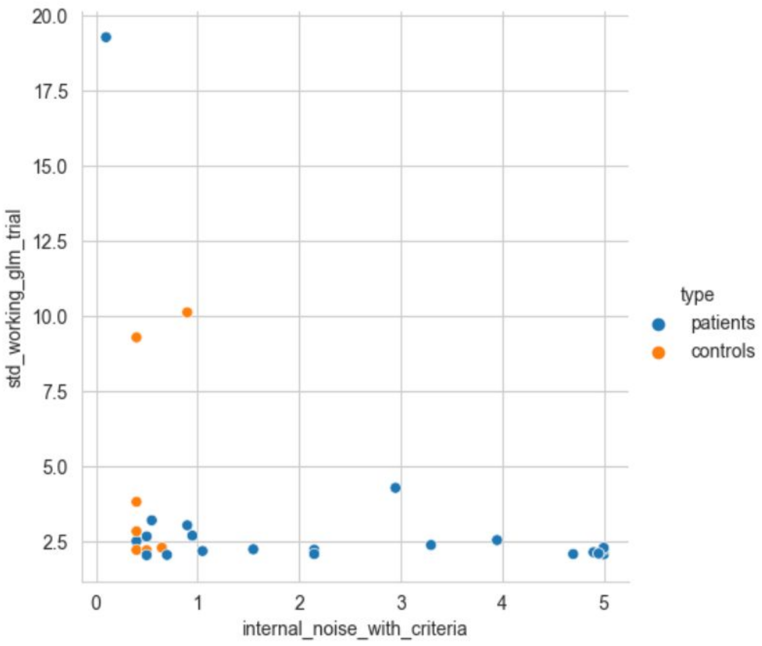
Things I have learned (3)

Our initial intuition is that the residuals of the LM/GLM could provide an estimate of internal noise.

LM: $y_i = \beta_0 + \beta_1 \cdot x_{i1} + \beta_2 \cdot x_{i2} + \cdots + \beta_6 \cdot x_{i6} + \varepsilon_i$ GLM: $\mathrm{P}(y = 1|X) = \varphi\left(\beta_0 + \sum_{i=1}^{k} \beta_i \cdot X_i + \varepsilon_i\right)$

Unfortunately, I have learned that the raw residual of the LM, or the working residual (i.e. residual in linear space, "inside" the logistic function) of the GLM do not correlate with our double-pass estimation of internal noise.





femto-st SCIENCES & TECHNOLOGIES

What is my biggest issue now?

- I have learned that LM/GLM is probably a better framework than the classification image method to estimate kernels (ex. we can do random effects on patients, i.e. LMM or GLMM).

- But my biggest issue now is how to estimate internal noise from such a model. The residual of the model does not seem to correlate with internal noise.

- What I'll be doing now is to work with simulated data, rather than real data, so I can control exactly what are the true values of internal noise, rather than trying to correlate with possibly badly-estimated values because of double-pass.