Australian Government
Department of Industry, Innovation and Science

Towards the recording of brainstem and cortical evoked potentials from the fine structure of natural speech

Dr Joaquin T. Valderrama [1,2,3,#], Dr Angel de la Torre[4], Dr Bram Van Dun [1,2], Prof Jose Carlos Segura[4]

[1] *National Acoustic Laboratories, Australian Hearing, Sydney, Australia*
[2] *The HEARing CRC, Melbourne, Australia*
[3] *Dept of Linguistics, Macquarie University, Sydney, Australia*
[4] *Dept of Signal Processing, Telematics and Communications, University of Granada, Spain*
[#] joaquin.Valderrama@nal.gov.au, joaquin.valderrama@mq.edu.au

IERASG 2019 / Sydney, 30th June – 4th July 2019

creating **sound value**™

www.hearingcrc.org

- Good afternoon, I am Joaquin Valderrama, from the National Acoustic Laboratories, Macquarie University, and the HEARing CRC.
- Before starting, I would like to acknowledge the contribution of my colleagues Angel de la Torre, Bram Van Dun, and Jose Carlos Segura.
- In this presentation, I will talk about the approach that we are following to obtain both brainstem and cortical evoked potentials simultaneously from the fine-structure of natural speech.

The neural response to natural speech

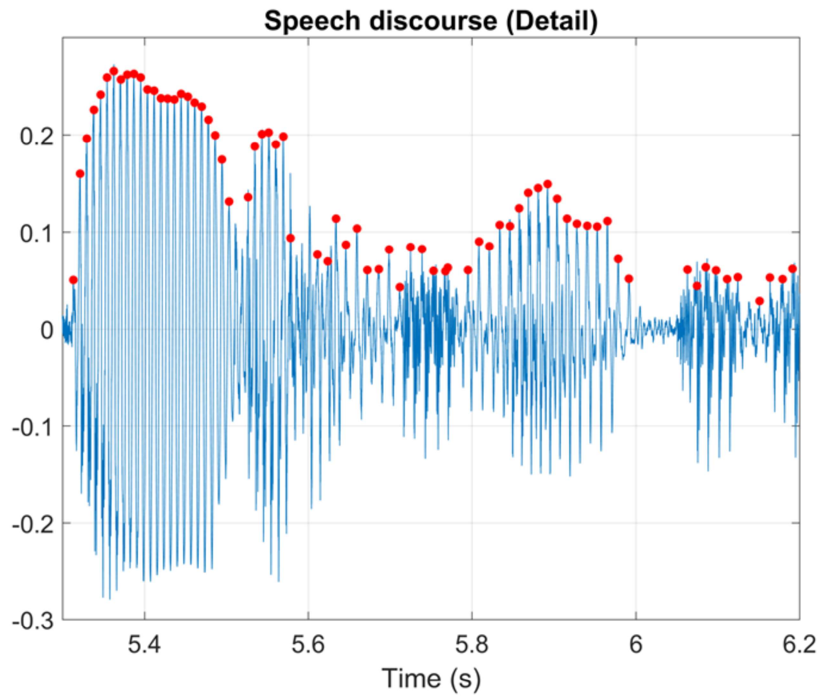
Procedure

Step 1 – Energy peaks from speech signal

Step 2 – Assume each peak elicits a response

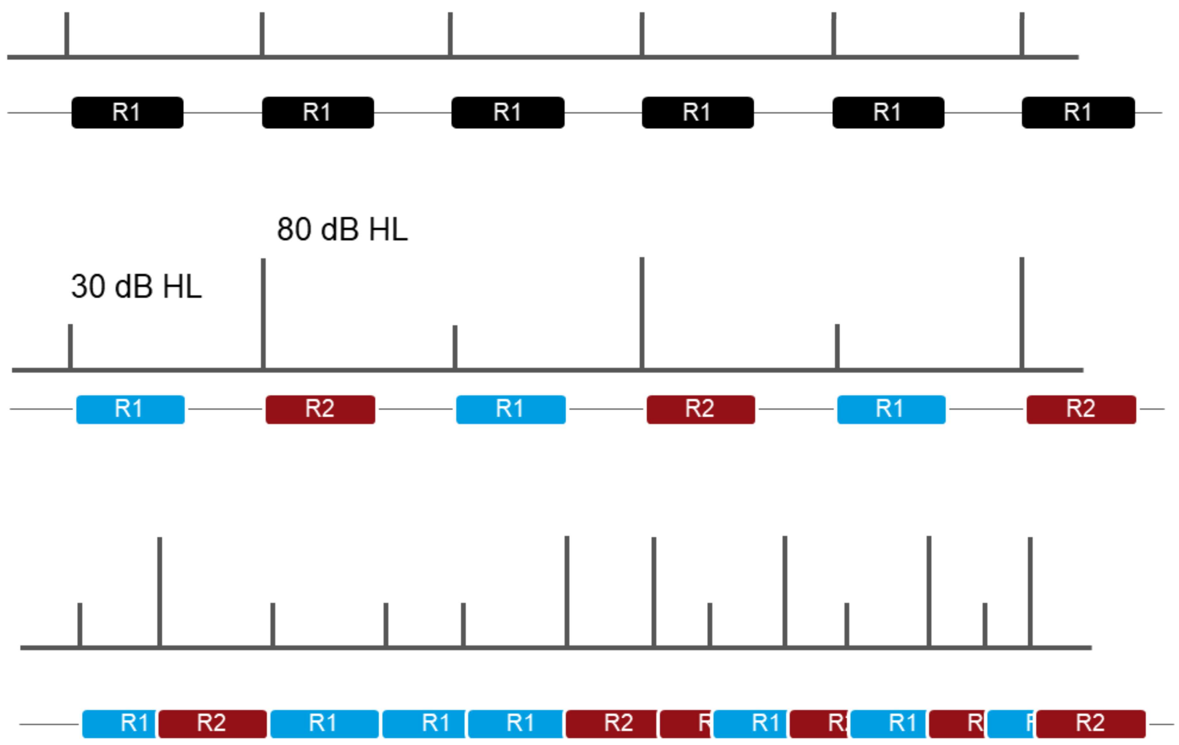
Step 3 – Deconvolve overlapping responses

Step 4 – Latency-dependent filtering



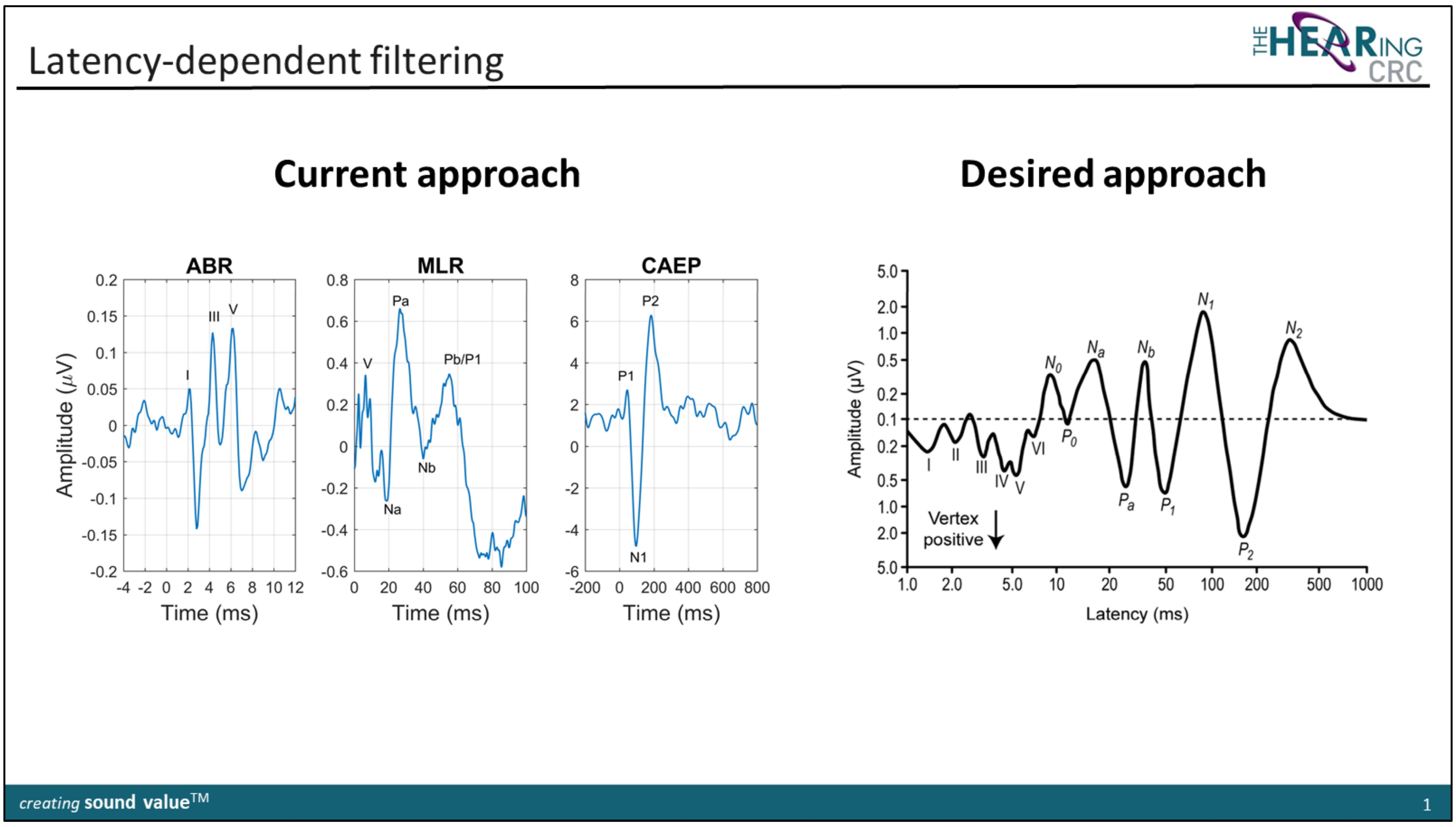
- This approach consists of a few steps.
- The first step consists of estimating the energy peaks of the speech signal. These energy peaks are associated with the glottal pulses in our vocal cords. To estimate these energy peaks, we use a pitch detection algorithm – nothing new there.
- In the second step, we assume that each energy peak will elicit a neural response from the full auditory pathway, including both brainstem and cortical responses.
- Since the duration of the neural response is much larger than the interval between pulses, there will be overlapping responses, and therefore, we will have to deconvolve (or disentangle) these overlapping responses. To do this, we will use a deconvolution algorithm known as IRSA.
- Finally, we will apply a latency-dependent filtering to represent the brainstem and cortical components in the same plot.
- I will describe more in detail steps 3 and 4.

Multi-response deconvolution using IRSA



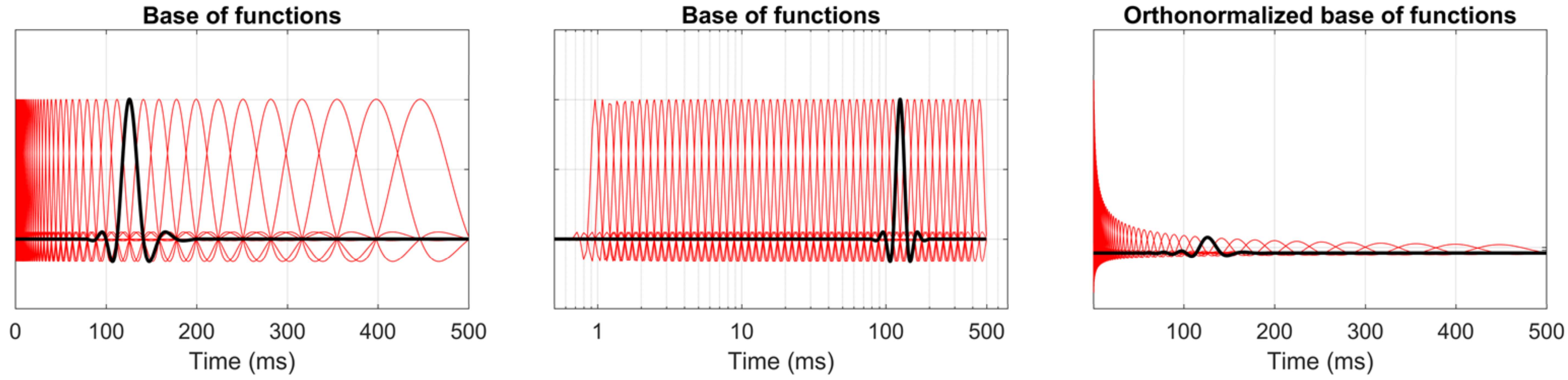
- Multi-response IRSA Iterative approach – Valderrama et al. (2014, 2016)
- Advantages – Flexible deconvolution
- Disadvantages – Takes time
- New formulation [faster]
  - Provisional patent
    Application 2019901078
  - 3 publications in progress
    stay tuned!

- The standard procedure to estimate AEPs consists of presenting one type of stimulus several times, assume that every stimulus evokes the same response (LTI assumption), and average the EEG segments containing the signal of interest.
- If rather than using the same auditory stimulus, we used different types of stimuli (like in this example – using clicks at 30 and 80 dB HL, it would be reasonable to assume that there are different types of responses.
- Multi-response deconvolution consists of estimating the different types of responses when they are overlapping.
- This multi-response deconvolution can be achieved using IRSA. For those of you interested in knowing more about this algorithm, it is fully described in these publications.
- The main advantages is that this technique allows multi-response deconvolution without imposing significant constraints in the stimulus sequence - which is of interest in many research applications.
- The principal disadvantage is that, because of its iterative approach, it requires some time to deconvolve the signals.
- As a note, we are currently working in a new formulation of the algorithm aiming to increase the processing speed amongst other advantages. There are 3 papers in preparation (one of them under review) and these new formulation is also under a provisional patent between NAL and UGR. Stay tuned!



- The latency-dependent filtering allows the representation of all components of the auditory pathway in the same plot.
- In the current approach, ABR, MLR and CAEP components are represented separately mostly because these components have energy in different frequency band, and it's difficult to find a 1 size fits all filter. Optimal filters for ABRs would attenuate CAEP, and viceversa.
- However, it just make sense to have them all represented in the same plot, as shown in this diagram.
- I presented the procedure to implement this latency-dependent filtering two years ago, in Warsaw, so I will give now just a brief summary.

Latency-dependent filtering



$$B = \begin{bmatrix} v_{1,1} & v_{1,2} & \cdots & v_{1,12500} \\ v_{2,1} & v_{2,2} & \cdots & v_{2,12500} \\ \vdots & \vdots & \ddots & \vdots \\ v_{54,1} & v_{54,2} & \cdots & v_{54,12500} \end{bmatrix}$$

- To go from the time-domain to the projected space:

  $AEP\_Projected_{(54x1)} = B_{(54x12500)}\ AEP_{(12500x1)}$

- To go from the projected space back to the time-domain:

  $AEP\_Reconstructed_{(12500x1)} = B^T\ AEP\_Projected = B^T\ B\ AEP$

creating **sound value**™ 3

- To implement this latency-dependent filtering we create a base of functions with different widths. You will notice that narrow filters in the earlier part of the response are appropriate for the higher frequency components of the ABRs, and the wider filters in the later part are more appropriate for cortical components.

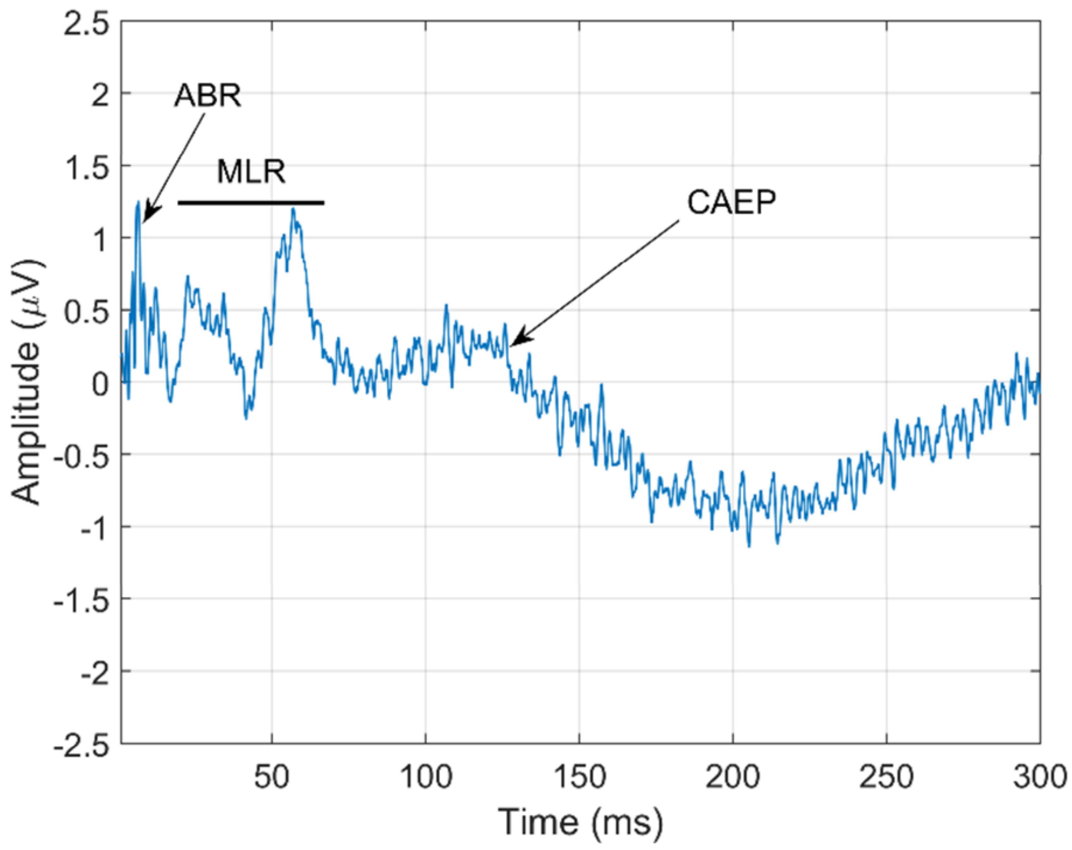
- If we represent the base of functions in the logarithmic time scale, we notice that they are uniformly distributed. In this example, there are 20 functions per decade covering 500 ms, so 54 functions.

- In order to create a base, we need to orthonormalize these functions, which is shown in this third graph.

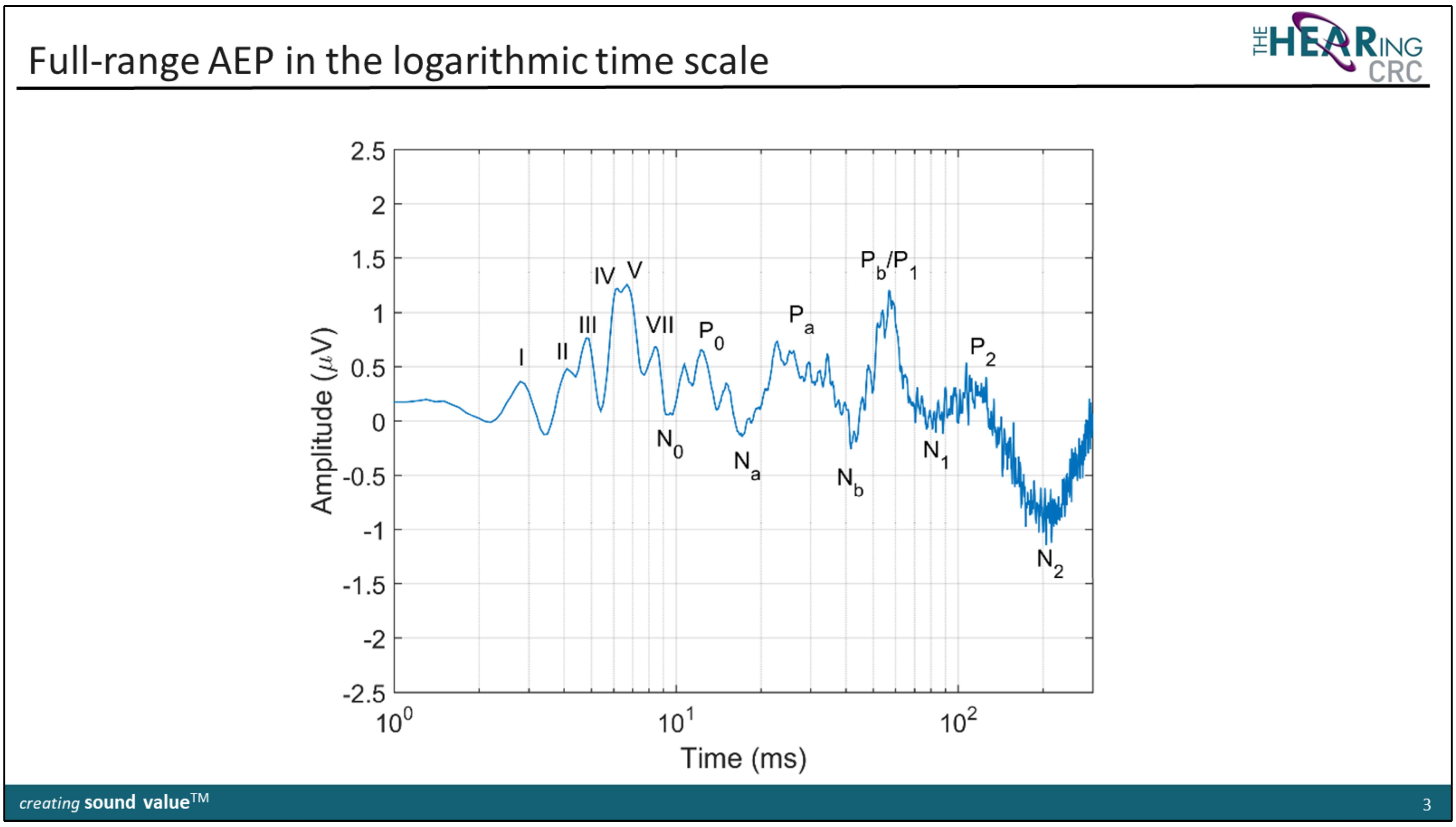
- The mathematical procedure to implement the filtering is very simple. (1) We create a matrix with the ON filters (54 rows, 12500 columns); then (2) we project our response on the reduced space by just multiplying the B matrix and the response; and (3) we reconstruct the signal by applying the transpose B matrix.

- In summary, the $B^T B$ operator applies the latency-dependent filtering. So once you have created the base of funcitons, applying the filter is a line of code.

Full-range AEP in the linear time scale

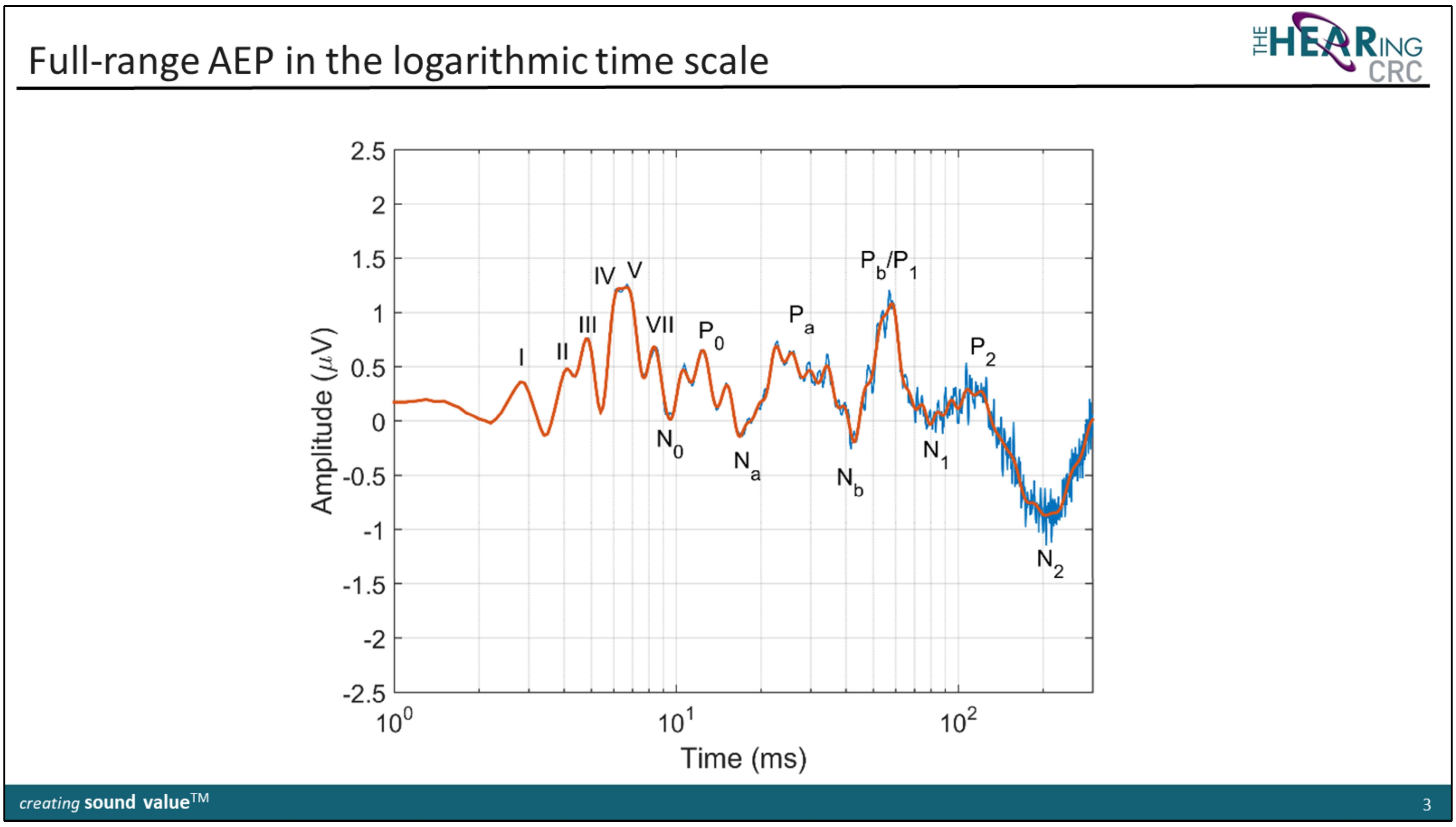


- What do we achieve with this filtering? This is an example of an AEP with components from the ABR, MLR, and CAEP.



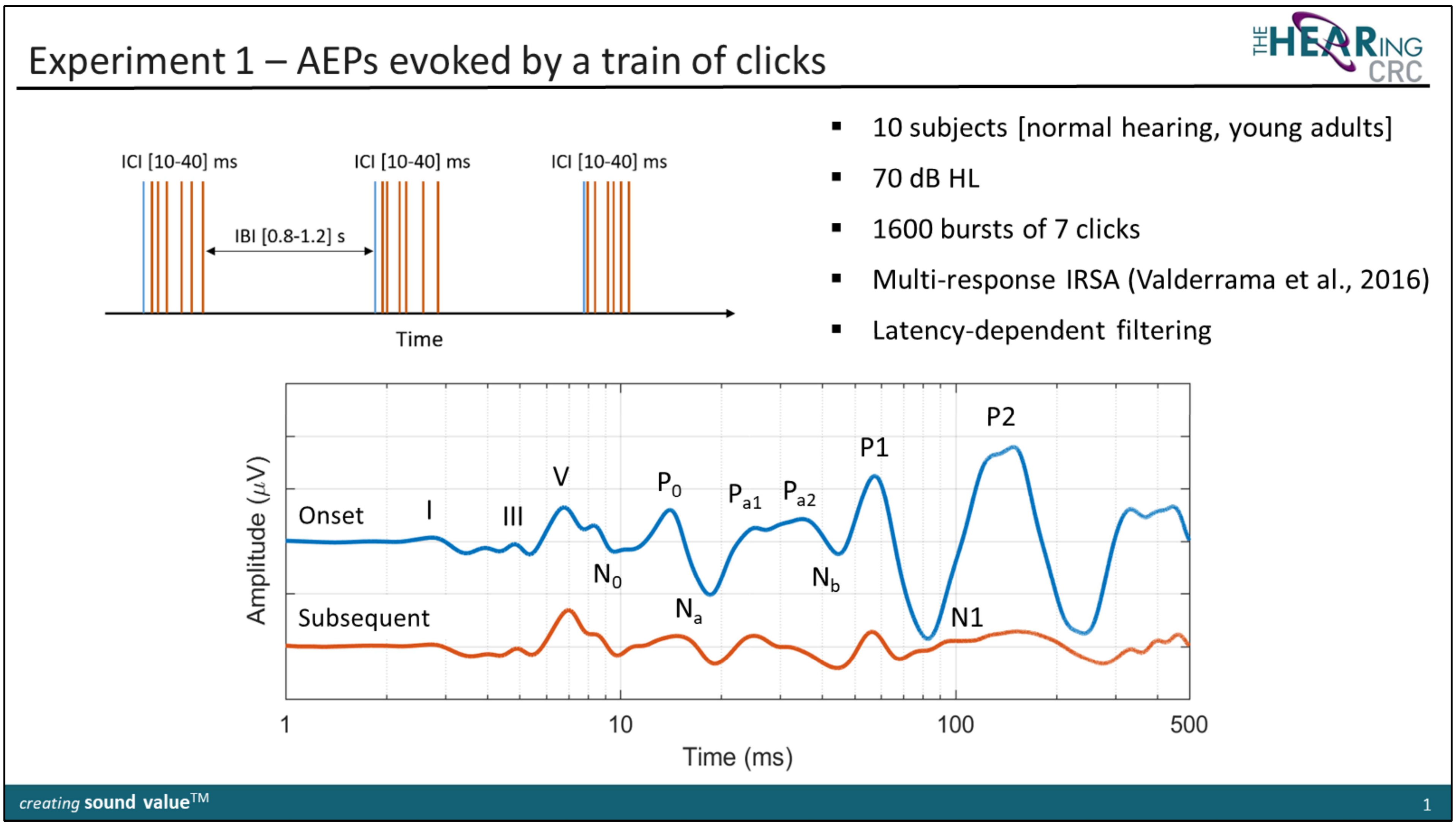
- If we represent the same signal in the logarithmic time scale, we see that this representation is appropriate to provide a full picture of the AEPs along the auditory pathway, including responses from the cochlea, brainstem and auditory cortex.

- HOWEVER, we also see that the cortical components are contaminated with high-frequency noise.



- The latency-dependent filtering described before (projecting over a base of functions of a reduced dimension and reconstructing back to the time domain) provides a filtering that is appropriate to represent earlier and later components in the same plot.

- In my view, this is the natural way of representing AEPs, and since we have this tool, this is the type of representation we use in most of our research.



- Before moving to speech, we wanted to do an experiment where we could train our skills, a framework where we could fine tune our research tools.

- For this purpose, we designed a test consisting of a series of bursts of clicks. [ICI] Inter-click interval; [IBI] Inter-burst interval.

- (Methods described as in text). We assumed that the onset click would elicit a neural response with a different morphology than the subsequent clicks.

- (Figure) This figure shows the grand-average of the 10 subjects.

    - The logarithmic time scale is appropriate for a full-representation of the components from the cochlea, brainstem and auditory cortex.

    - An important result is that the two responses have a different morphology.

    - The main difference is in the cortical components, which are adapted in the subsequent clicks compared to the onset click [as expected!].

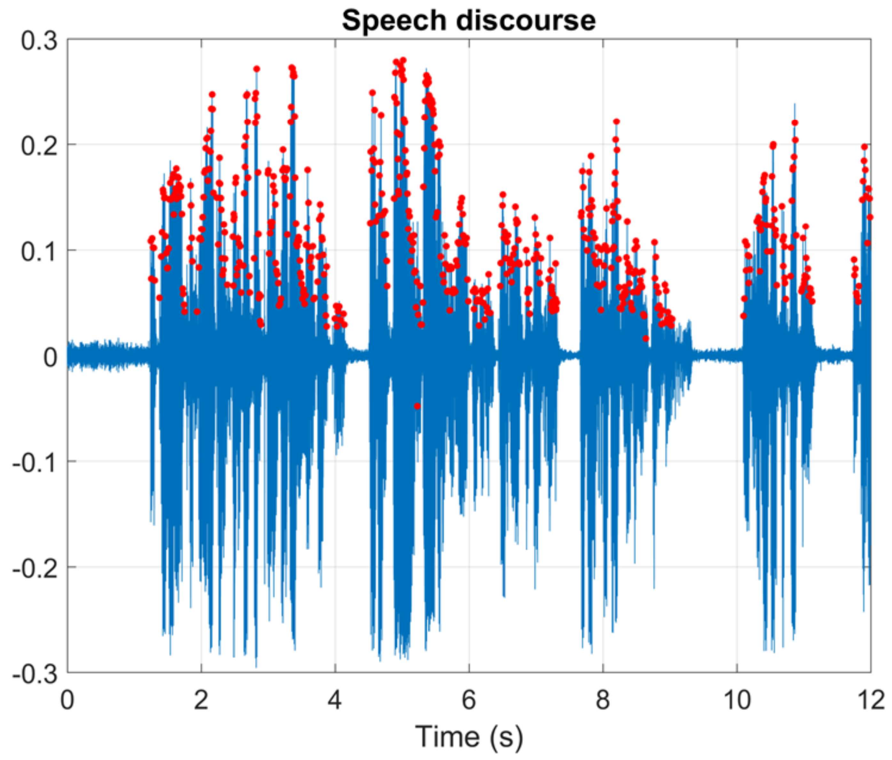
    - It is also important to note that the subsequent clicks DO EVOKE a cortical response: it’s adapted, but it’s present.

    - No particular differences in the brainstem components (as expected), as the stimulus rates are not sufficiently high to adapt these responses.

    - Main results: (1) There are two types of responses; (2) all clicks evoke all components: and (3) cortical components adapted in subsequent clicks.

Experiment 2 – AEPs evoked by natural speech

THE HEARING CRC

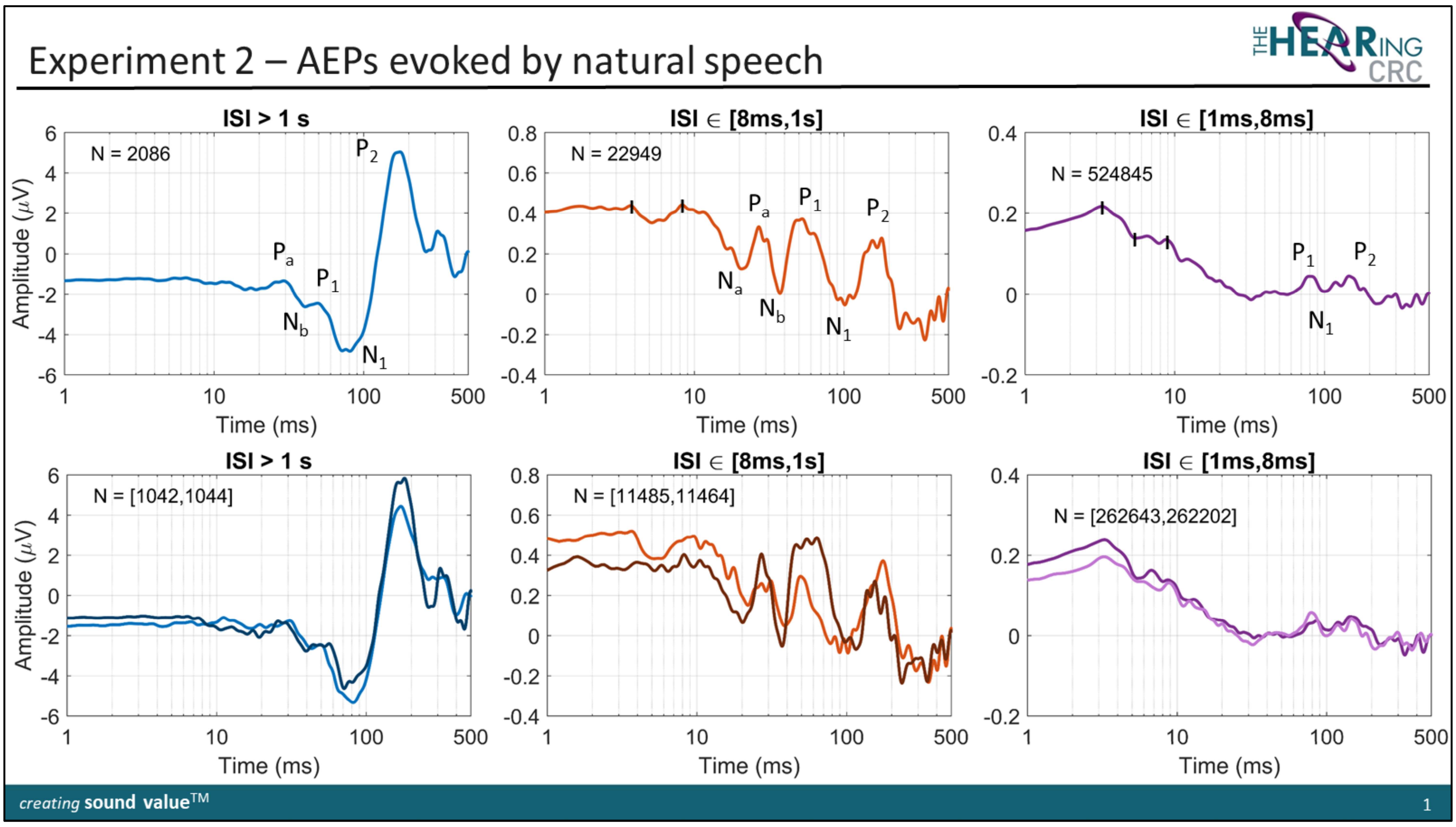


- 8 subjects [normal hearing, young adults]
- 260 short sentences

  The toddler likes the biscuits with some milk.
- Assumption – 3 categories
  - ISI > 1 s
  - ISI ϵ [8ms , 1s]
  - ISI ϵ [1ms , 8ms]
- Multi-response IRSA (Valderrama et al., 2016)
- Latency-dependent filtering

creating **sound value**™ 1

- (Methods)



- Onset of the sentence evoke large cortical components

- Onset of the syllables evoke adapted cortical components. This is consistent with what we found in the previous experiment using bursts of clicks. There are some components before 10 ms that seem to be a brainstem response to natural speech.

- The response from the glottal pulses show very adapted cortical responses, and some brainstem components.

- (Analysis of reproducibility)

- This is ongoing work. There is still much to do, but we have a procedure and we have the research tools. Next, we will continue investigating which are the optimal categories in order to model how the brainstem and the auditory cortex encode speech.

Take-home message

- The proposed methodology aims to obtain brainstem and cortical evoked activity from unprocessed natural speech.
- Results point out that natural speech evokes AEPs of different morphology, thus the assumption of a single-type neural response typically made in speech-evoked AEPs should be reconsidered.
- This tool will provide a more accurate modelling of how the auditory system encodes speech, and may facilitate objective tests of attention and language comprehension.

THANK YOU FOR YOUR ATTENTION!

ANY QUESTIONS?

Australian Government
Department of Industry, Innovation and Science

NAL NATIONAL ACOUSTIC LABORATORIES

MACQUARIE University

UNIVERSIDAD DE GRANADA

THE HEARING CRC

creating **sound value**™ 1

- Text