Auditory evoked potentials

Joaquin Tomas Valderrama Valenzuela, PhD

National Acoustic Laboratories, Sydney, Australia
Macquarie University, Sydney, Australia
The HEARing CRC, Melbourne, Australia
joaquin.valderrama@nal.gov.au, joaquin.valderrama@mq.edu.au, jvalderrama@ugr.es

Diploma de Especialización Universitaria en Audiología
University of Salamanca

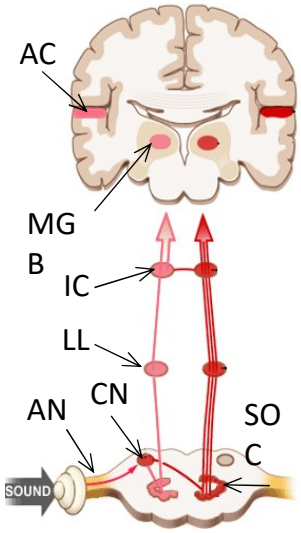
Welcome to this lecture. It is a pleasure and an honour to share with you my experience with auditory evoked potentials.

Please let me introduce myself first. I have two Bachelor of Science degrees in Telecommunication Engineering and Business Administration from the European University of Madrid (Madrid, 2008). In October 2014, I finished my PhD in the Department of Signal Theory, Telematics and Communication from the University of Granada (Granada). One part of my doctorate consisted of developing the electronics and software of an auditory evoked potentials recording system. In March 2015, I started a position as a Research Electrophysiologist in the National Acoustic Laboratories (Sydney, Australia), which is the research division of Australian Hearing. This Company was founded by the Australian Government after the II World War to evaluate and treat veterans reporting hearing difficulties, and currently, with more than 7000 audiologist, Australian Hearing is the main provider of audiology services in Australia. I also hold an *Honorary Associate* position in the Department of Linguistics at Macquarie University, while I keep a close collaboration with the University of Granada. My main research colleagues are Dr Harvey Dillon, Dr David McAlpine, Dr Ángel de la Torre, Dr Bram Van Dun, Dr Fabrice Bardy, Dr Elizabeth Beach, Dr Mridula Sharma and Dr Jaime Undurraga. If you are interested in knowing more about my work, you can visit my profile in Research Gate: https://www.researchgate.net/profile/Joaquin_Valderrama.

My intention is that the next slides give you a detailed notion about what auditory evoked potentials are, the different classes that exist, their main applications in the clinic and research, the different technologies that are available for their recording, and finally, some emerging technologies that overcome certain limitations of current systems.

Let's get started.

What are auditory evoked potentials?



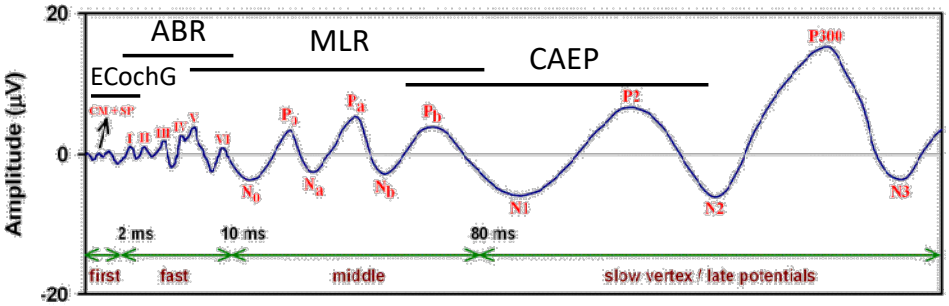


Fig 2. Different components of auditory evoked potentials. Figure adapted from: Lamoré (2009). Auditieve 'Evoked Potentials' van hogere centra. http://www.audiologieboek.nl/htm/hfd4/4-5-2.htm

Fig 1. Main neural stations of the auditory pathway: AN, auditory nerve; CN, cochlear nucleus; SOC, superior olivar complex; LL, lateral leminiscus; IC, inferior colliculus; MGB, middle geniculated body; AC, auditory cortex. Figure adapted from: Jérôme Ruel, Nuno Trigueiros-Cunha, Jean-Luc Puel. Journey into the World of Hearing.

What are auditory evoked potentials?

Auditory evoked potentials (AEPs) are voltage waves that reflect the synchronous activity of neurons in different stages of the auditory pathway in response to a sound stimulus.

Figure 1 shows a diagram with the main neural stations of the auditory pathway, from the cochlea to the auditory cortex. Figure 2 shows a diagram with different components of auditory evoked potentials. Each of these peaks and troughs reflect the neural activity of different neural stations.

For example, the cochlear microphonics (CM), the summating potential (SP) and the compound action potencial (CAP) or wave I are AEPs generated in the cochlea; waves II to VI are generated in different stages of the brainstem; waves P0 to Nb are generated in the superior portion of the brainstem and the middle geniculated body (MGB); and waves Pb to N2 are generated in the auditory cortex (AC).

Please note that the signal shown in figure 2 is not a real response, but a diagram. Currently, there is not available any technique that allows us represent all auditory evoked potentials in the same plot. AEPs are usually classified according to their latency (time from the stimulus onset to the occurrence of the peak). The main types of AEPs are: electrocochleography (ECochG), auditory brainstem responses (ABRs), middle latency responses (MLRs), and cortical auditory evoked potentials (CAEPs). Each of these potentials have their particular recording protocol. Section II in this lecture will present in detail each of these potentials, and other AEPs of similar relevance.

Why do we need to record AEPs?

To understand better our neural structures

To implement objective diagnostics

Why do we need to record AEPs?

The recording of AEPs has at least two important applications. On one hand, the evaluation of the morphology of the waves of the AEPs in response to stimuli of different nature allows investigation about the neural structures that encode and process the sounds that we perceive. Understanding these structures is fundamental to diagnose the pathologies adequately and provide efficient treatments.

On the other hand, the recording of AEPs allows an objective evaluation of the auditory pathway. The "objective evaluation" is highlighted in the previous statement to refer to diagnostic techniques that do not depend on the subjectivity of the test subjects. This way, a tonal audiometry (in which the subject responds whether or not a series of tones of different frequencies is perceived) is not considered an objective technique. On the contrary, AEPs can be used as an objective technique, which results very useful to evaluate the state of the auditory pathways in subjects without the capacity of responding, such us newborns.

Part I

Recording of auditory evoked potentials

This section describes the physiological principles that explain the generation of auditory evoked potentials, along with the most relevant parameters of an electroencephalography recording system.

What is the physical principle that explains the generation of AEPs?

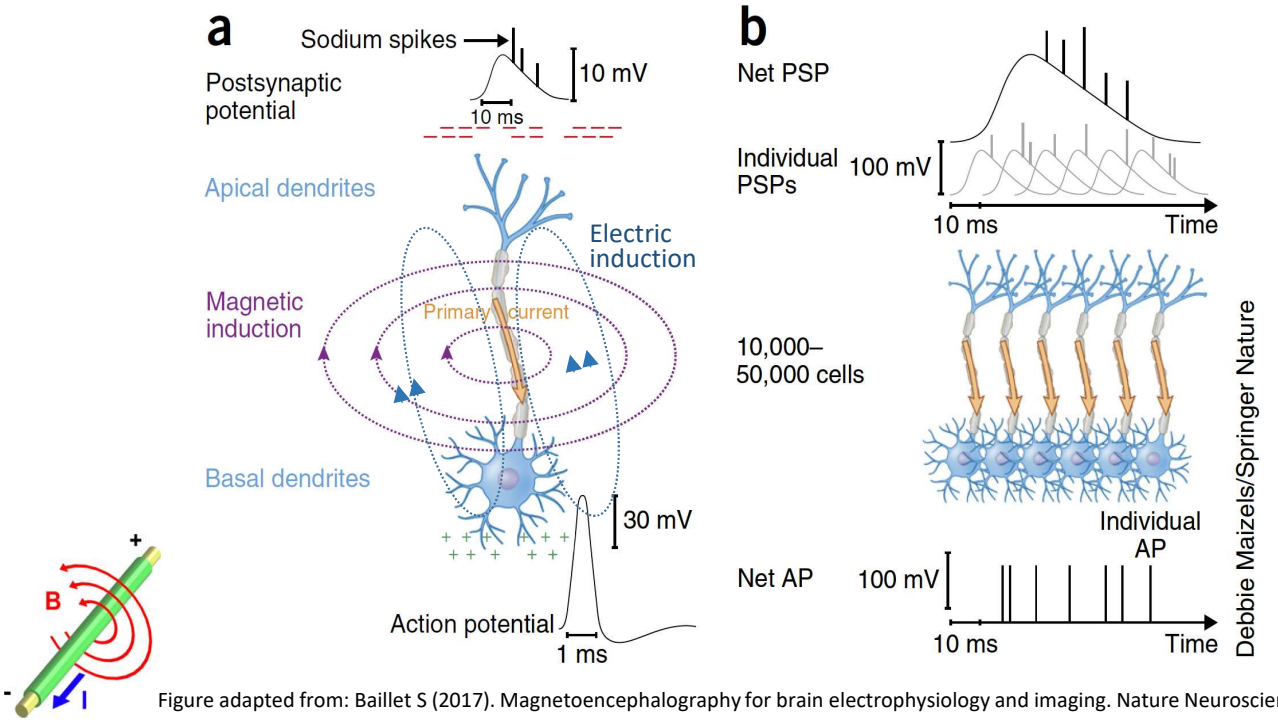


Figure adapted from: Baillet S (2017). Magnetoencephalography for brain electrophysiology and imaging. Nature Neuroscience 20, 327-339

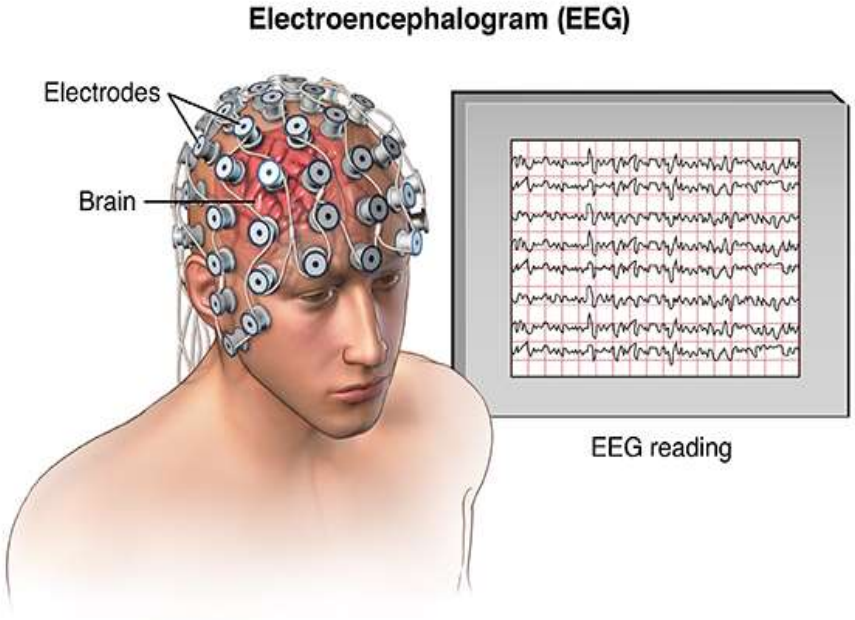
The physical principles that explain the generation of AEPs are based on **electromagnetism**, the science that describes the relation between electric and magnetic phenomena.

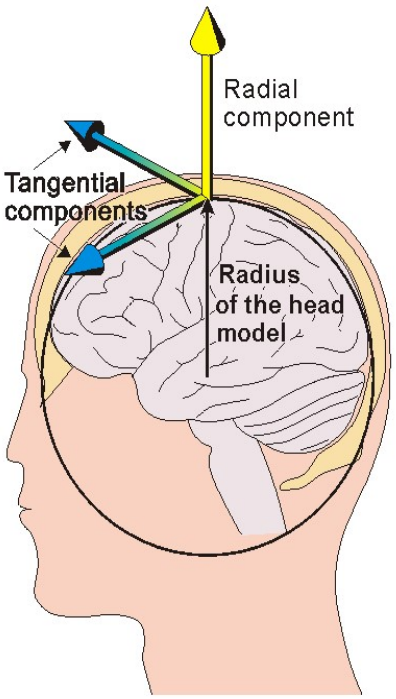
In our auditory system, when a neuron is activated, it is generated an electrical current that go across the membrane of the neuron. This flux of current is determined by two processes that affect the electrical voltage of the neuron. The **postsynaptic potential** (PSP) is a relatively slow change of voltage that is produced in the dendrites of the neuron, i.e. in the points in which the neurons make contact (synapsis) with other cells. In the synapsis, the neuron emits and receives neurotransmitter particles, that may vary the electrical voltage in the neuron. If the increase of voltage is significative, the neuron may react by depolarizing through the generation of one or more **action potentials** or **spikes**, that are rapid changes of voltage. This link shows a video from Dr Armando Hasudungan (11 minutes) that re-examines with an educational perspective the neural mechanics: https://www.youtube.com/watch?v=FVo04B0_5R4.

According to the Ampère's Law, a variable electrical current (like the one generated in the neurons) induces a variable magnetic field **B**, rotational around the conductor, as shows the diagram at the bottom left corner. On the other hand, the Faraday's Law of induction and Lenz' Law points out that any magnetic field that depend on time will evoke an electrical field **E** proportional and perpendicular to the magnetic field. Figure (a) shows a diagram of the magnetic and electric fields induced in a pyramidal neuron, along with a diagram of the action potential and the postsynaptic potential.

The electric and magnetic fields induced by a neuron are of a very small magnitude, and therefore, it is required the **synchronous activity** of a several thousands of neurons (from 10.000 to 50.000 neurons) to be able to record the neural activity with electric and magnetic sensors placed on the surface of the head. In addition, it is also required that these neurons are **spatially aligned** in order to let their individual electric and magnetic fields be added, and not be counteracted. Figure (b) shows and example of a postsynaptic potential obtained by several neurons activated synchronously and spatially aligned.

Recording of electric fields (EEG)





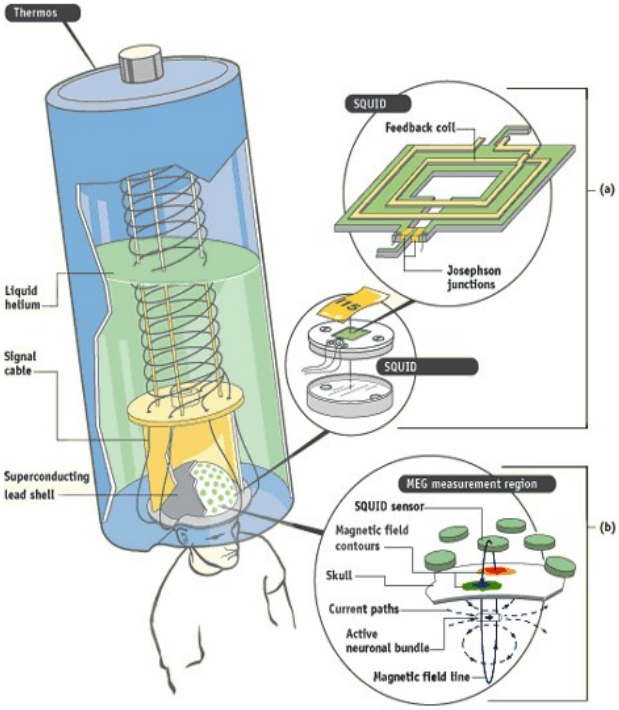
One strategy that can be followed to measure the neural activity by a non-invasive approach is to place on the head a number of sensors sensitive to magnetic and electric fields.

An **electroencephalogram (EEG)** signal represents the voltage difference between two electrodes placed in different positions on the surface of the head. These voltage differences are generated by variations in the **electric field** induced by the activation of the neurons. Two electrodes are needed to record one EEG channel. The electrode known as *active* (or positive) is usually placed in the area of the head that is being evaluated, while the electrode known as *reference* (or negative) is placed in one area of the head with low neural activity, like the mastoid or the ear lobule. The figure on the left shows a diagram with several EEG channels in order to monitor the neural activity in different positions of the head.

The main **advantages** of EEG are: (1) it presents a high temporal resolution (we can take measures within a very short time interval); (2) once the recording system is acquired, the cost of every EEG session is low (around 15 € per session); (3) the recording systems can be portable and can be used in kids and newborns.

The main **disadvantage** of EEG is that it does not allow an adequate temporal resolution. As we have mentioned in the previous slide, the orientation of the neurons determine the orientation of the magnetic and electric fields. The EEG signals are sensitive both to electric fields induced by neurons situated perpendicularly to the surface of the head (radial orientation), and to electric fields induced by neurons situated in parallel to the surface of the head (tangential orientation, see right figure). In addition, the EEG signals can measure activity of the neurons situated several centimetres from the sensor, which is of interest to record, for example, activity of the neurons from the brainstem. However, these characteristics make that EEG signals are not ideal to determine with precision the areas of the cortex activated under certain processes. Another disadvantage of EEG signals is that electric fields suffer a distortion when going through different surfaces of tissue between the cortex and the electrode.

Recording of magnetic fields (MEG)



The **magnetoencephalography (MEG)** signals measure variations in the **magnetic field** induced by the activity of the neurons.

The magnetic field produced by the evoked response of the neurons is around 100 femptoTesla (fT, 10^{-15} T, equivalent to around 10-100 million times smaller than the magnetic field of the Earth). These small magnetic fields can be recorded through the SQUIDS *(Superconducting Quantum Interference Devices)*, a coil that become a superconductor at very low temperatures. The figure on the left shows a real MEG recording system, and the figure on the right shows a diagram of the technology.

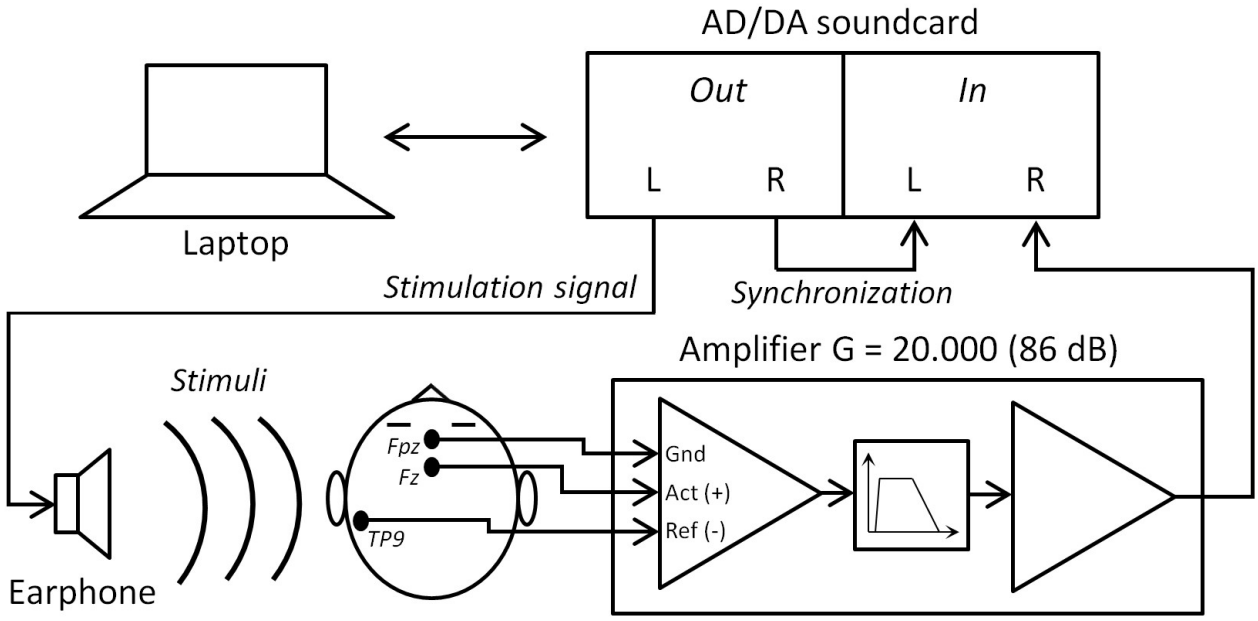
Similarly as in EEG, the MEG signals present the **advantage** of a high temporal resolution. In addition, in contrast to EEG, the main advantage of MEG is that it allows a very high spatial resolution. The SQUIDS are particularly sensitive to neurons situated near the cortex with tangential orientation. This may look like a disadvantage, but it is indeed an advantage when determining with precision the neural groups activated in certain activities (spatial resolution). Furthermore, the magnetic fields are less distorted than the electric fields when going through the different surfaces of tissue that recoat the cortex.

The main **disadvantage** of MEG is its high price. In order to make the coats become superconductors it is necessary to immerse them in helium and keep them at temperatures near the absolute zero (-269 ^{o}C or 4 Kelvin). Additionally, since SQUIDS are so sensitive to magnetic fields (even the magnetic field from the Earth can contaminate the signals), the recording sessions must be carried out in a booth with magnetic shielding, which is very expensive to build. These factors make that the cost of each recording session is around 500 €.

Other disadvantages of MEG is that the recording system is not portable and it may cause claustrophobe if the subject is particularly sensitive to close and small rooms. In addition, it is usually recommended to obtain a functional magnetic resonance imaging (fMRI) to know with precision the morphology of the brain, which significantly increases the complexity, duration and cost of the measure.

The simultaneous recording of EEG and MEG signals is technologically possible.

Hardware elements of an AEP recording system



Valderrama et al. (2014). A flexible and inexpensive high-performance auditory evoked response recording system appropriate for research purposes. Biomedical Engineering 59, 447-459.

This section will present the different *hardware* and *software* modules necessary to record AEPs with a **single-channel EEG recording system**.

The process is controlled by a **computer**, which will (a) generate the stimulation sequences according to the parameters specified by the user, and (b) process the recorded signals to obtain the AEPs.

The computer is connected to an **analogue-to-digital/digital-to-analogue (AD/DA) converter**, which connects the real world (the world that we live is analogue) with the digital world (the world of the computers). If you wonder what is the difference between an analogue and a digital system, an analogue signal is that one in which by choosing any two time-values (no matter how close they are), there will always be a value in between the chosen ones. A high-quality soundcard is generally used as AD/DA converter.

The output 1 (marked with L in the diagram) of the AD/DA converter is connected to a **transducer**, an earphone that will generate the stimulus. The neural activity in response to the stimulus will induce an electric field, which will generate a voltage difference between two **electrodes** situated in different positions on the head. The diagram of the figure shows an active electrode (situated in Fz), a reference electrode (situated in TP9) and a ground electrode (in FPz). The voltage-signal recorded by the electrodes has a very low amplitude, so the electrodes are connected to an **amplifier**, which consists of three stages: differential amplification, analogue filtering, and final amplification. The amplified signal is the EEG, which will be converted to a digital signal in the AD/DA converter, and will be processed in the computer by a number of ***software* processes**.

The output 2 of the AD/DA (marked with R in the diagram) will be connected to its input 1 (marked with L). This channel is used for the synchronization signal, which is necessary to know in which time instants the stimuli were presented (triggers).

Each of these modules deserve a more detailed study. Next, we will analyse the main parameters involved in the AD/DA converter, the transducer, the electrodes setup, the amplifier, and the software processes necessary to obtain the AEPs.

The AD/DA converter. Step 1 - Sampling

Figure 1

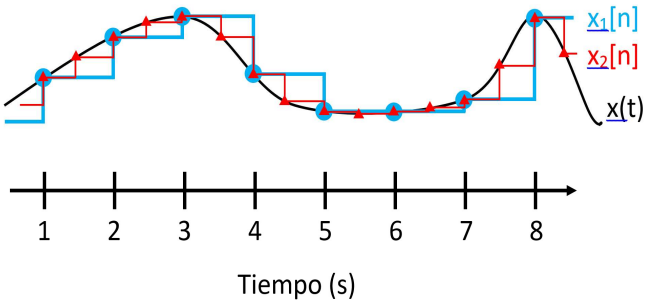
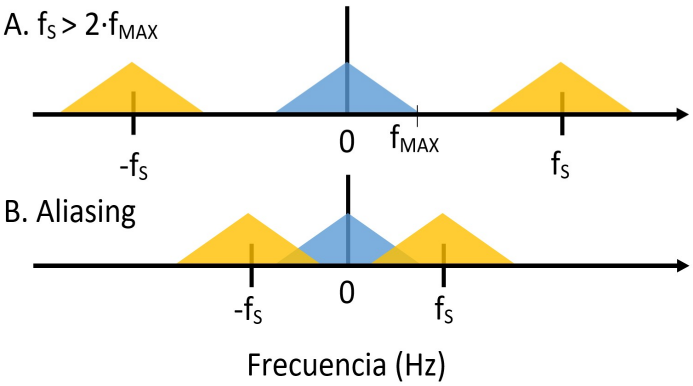


Figure 2

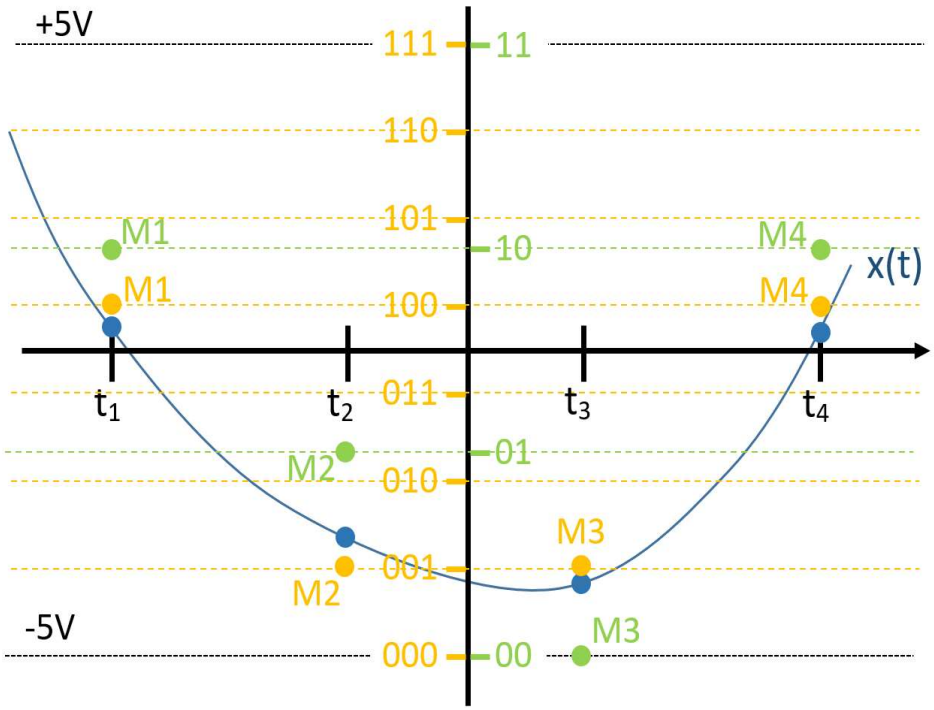


There are two main parameters that must be configured in the analogue/digital conversion: (1) the sampling rate, i.e. the time interval in which the signal is sampled; and (2) the number of bits that we use to encode each sample (quantization).

Figure 1 shows an example of an analogue signal x(t), and two digital signals $x_1[n]$ and $x_2[n]$. In $x_1[n]$, a sample is taken every second (blue circles), therefore the sampling rate (f_s) is 1 Hz. The sampling period is the inverse of the sampling rate, i.e. $T_S = 1/f_S$. In this example, the sampling period is 1 second. In the $x_2[n]$ signal, two samples per second are taken, thus fs = 2 Hz, and Ts = 0.5 s. The determination of an adequate sampling rate must consider the **sampling theorem of Nyquist-Shannon**. A consequence of this theorem is that sampling a signal with a certain sampling rate fs, the spectrum of the signal (or its frequency components, represented with a blue triangle in figure 2) is shifted towards $\pm f_S$ (yellow triangles). f_{MAX} is the highest frequency in which a signal has energy. f_{MAX} is also known as bandwidth of the signal. Figure 2A shows that to avoid that the blue and yellow triangles overlap, **fs must be greater than two-times f_{MAX}**. In other words, the bandwidth of the signal must be lower than half of the sampling rate (fs/2), which is known as *Nyquist frequency*. If this requirement is not fulfilled (figure 2B), it is produced a phenomenon known as *aliasing*, which distorts the signal.

In electrophysiology, the sampling rate must be chosen considering the frequency components of the AEPs that are being recorded. For example, the ABRs present frequency components between 100 and 3000 Hz, and therefore, the f_s must be greater than 6000 Hz, or 6 kHz. The CAEPs present frequency components between 1-100 Hz, so the f_s must be greater than 200 Hz. How much greater? Could we use a f_S = 100 kHz to record CAEPs? From the sampling theorem point of view, the signal of interest (the EEG) is perfectly well characterized using both a sampling rate of 200 Hz and 100.000 Hz. However, the use of a very high f_s presents the inconvenience of using unnecessarily more memory to store the EEG (because more samples are taken in the same time interval). In practice, f_s use to be set at a frequency near the Nyquist frequency ($2*f_{MAX}$), but with some margin. For example, f_S = 10 kHz is usually appropriate to record ABRs, and f_S = 300 Hz is adequate to record CAEPs.

The AD/DA converter. Step 2 - Quantization



The second step in the analogue-to-digital conversion is quantization. The quantization process consists of encoding the amplitude of each sample taken from a signal with a certain number of bits.

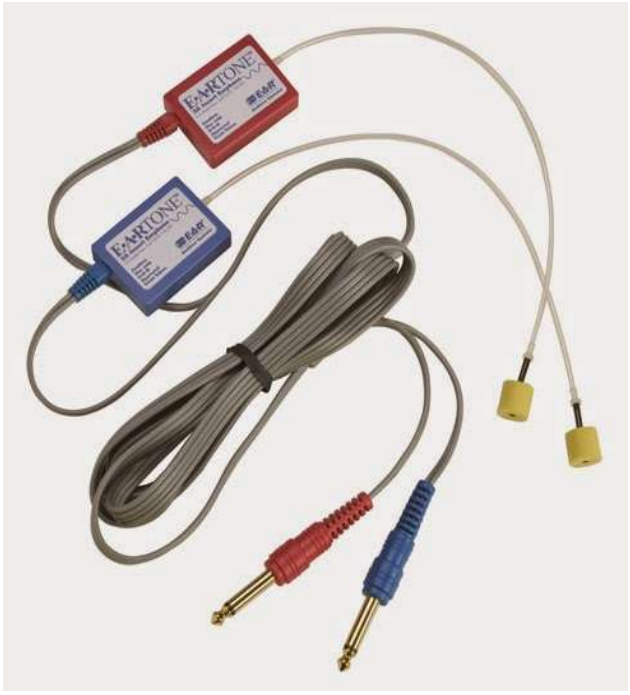
The figure in the slide shows the quantization process of the signal x(t) in 4 samples using 2 bits of quantization (green color) and 3 bits (orange color) through an AD/DA converter fed with ±5 volts. This figure shows that there are 4 possible amplitude-values that can be taken by the samples in the case of using 2 bits of quantization, and 8 possible values using 3 bits of quantization. In general, the number of possible values is 2^{bits} [2^2=4 and 2^3=8]. This figure also shows that there is a difference between the value of x(t) evaluated in the instants of the samples (blue circles) and the quantized values (orange and green circles). This difference is known as **quantization error**. This example shows that the quantization error is smaller in the case of using 3 bits, compared to using 2 bits of quantization.

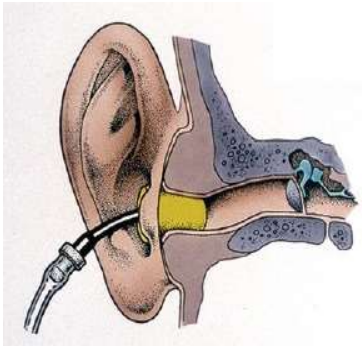
A relevant factor to bear in mind is that sampling a signal [previous slide] does not involve losing information, however, certain information is lost in the quantization process. There is a quantization error inherent to the quantization process of a sampled signal, which will be smaller as the number of quantization bits increase.

What happens if we use a large number of bits? Using a large number of bits reduce the quantization error, but increases the memory storage requirement. In the examples of the figure, 8 bits are needed to encode 4 samples with 2 bits per sample (4 samples * 2 bits/sample) and 12 bits to encode 4 samples with 3 bits per sample. If we used 100 bits to encode each sample, we would need 400 bits to encode the 4 samples. In electrophysiology it is common to use either 16, 32 or 64 bits per sample.

How much memory is needed to record an EEG of 2 minutes sampled at 25 kHz and using 64 bits per sample? The answer is 192.000.000 bits = 25.000 samples/second * 120 seconds * 64 bits/sample. 192 million bits are 24 million bytes (1 byte = 8 bits), or 24 MB (1 MB = 1.000.000 bytes).

The transducer





The transducer transforms the electric signal into a stimulus. For example, the transducer can be an earphone if the objective is to stimulate the auditory system with a sound; a bone vibrator; a cochlear implant if the stimulus is electric; or a screen if the objective is to record visual evoked potentials.

The most common transducer currently used with acoustic stimulus is the **insert earphone**. The figure on the left shows these earphones. The main feature of this type of transducer is that the earphone that generate the acoustic stimulus is spatially separated from the subject through a plastic tube, which ends in an eartip which is placed in the auditory canal of the subject (figure top right). As a standard, red color is used for the right ear and blue color for the left ear. Stimuli can be presented monaurally (in only one ear) or binaurally (in the two ears).

The main advantage of these earphones is the **spatial separation between the subject and the earphone**. The earphone consists of a coil that moves a magnet according to the flux of current that goes through the coil. Similar to the neurons, the current that goes through the coil generate magnetic and electric fields that are usually captured by the electrodes, thus contaminating the signal. This type of contamination is usually known as *stimulus artifact*. The fact of separating spatially the earphone from the electrodes significantly reduce the stimulus artifact because the electrodes are situated further from the noise source. In addition, the spatial separation also provides **temporal separation** between the generation of the stimulus (that produces a stimulus artifact) and the stimulation of the auditory system (that generate the response that we are interested to measure), and therefore, this allows to separate in the time domain the artifact from the signal of interest. The temporal separation is determined by the length of the plastic tube. Since the speed of sound on air is 343,2 m/s, and the length of a standard plastic tube is 282 mm (Etymotic ER·3A model), the temporal separation is 0,82 ms.

The stimulus artifact can also be reduced by placing these earphones inside a box with electromagnetic shielding.

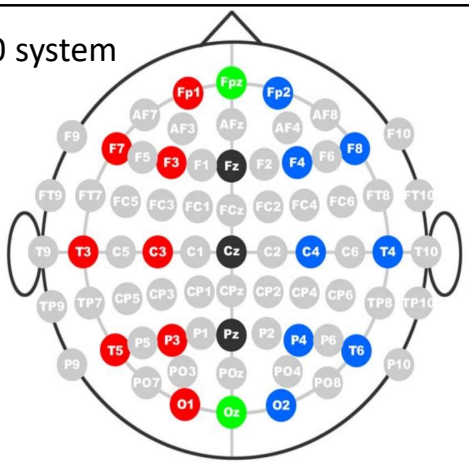
Other advantages of these earphones are that the expansive foam of the eartip provides a better acoustic insolation to the subject from possible external sounds that could interfere in the measure, the expanded foam may also prevent that part of the stimulus is transmitted to the contralateral ear by bone conduction; and furthermore, the earphones are hygienic, since the eartips are disposable.

The electrodes

1. Reusable electrodes

2. Disposable electrode

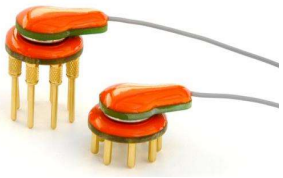
6. 10/10 system



3. Skin preparation gel

4. Conductive paste

5. Dry electrodes



The electrodes transform ionic currents (the mechanisms of conduction of bioelectric signals in tissues) in electric currents that drive the signal of interest (the evoked potentials) from the subject to the amplifier.

Because of its non-invasive nature, the most common used electrodes in electrophysiology in humans are the **surface electrodes**, that are placed on the surface of the skin. There are reusable electrodes (figure 1) and disposable electrodes (figure 2).

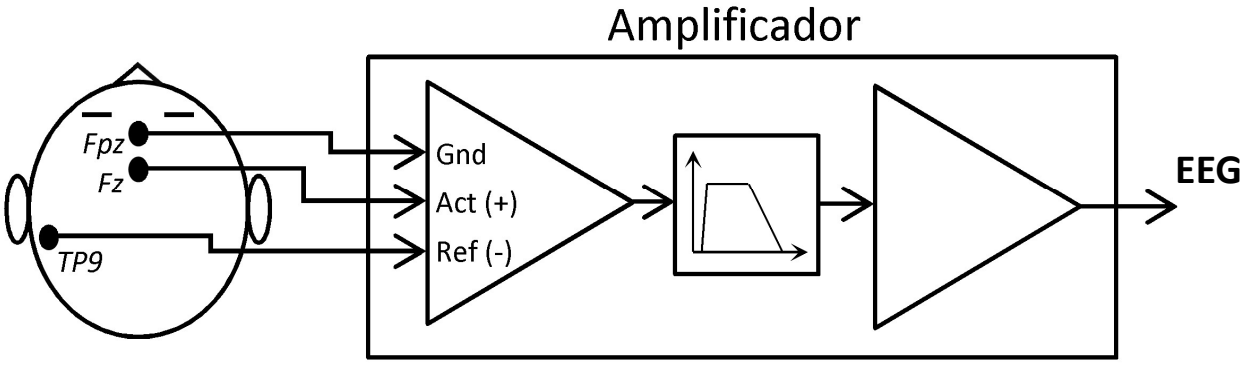
The electrodes consist of a metal body that allow conductivity, usually silver [Ag], covered with a film of gold [Au] or silver chloride [AgCl].

Since the electrodes are the first element in the recording system, it is very important that the contact impedance between the electrode and the skin is as lower as possible in order to minimize noise induced by electromagnetic sources. This contact impedance can be reduced by a soft exfoliation of the skin using a skin preparation gel (like the one of figure 3), and using conductive paste (figure 4) between the electrode and the skin. There exists a new generation of electrode that do not require this procedure to achieve good impedances, known as *dry electrodes* (figure 5), although the performance of these electrodes is pending to be evaluated by a broader range of laboratories.

The minimum configuration to record an EEG consists of 3 electrodes: an active (+), a reference (-) and a ground electrode (GND). In multichannel configurations, like in the 10/10 system (figure 6), a large number of electrodes are situated in different positions on the head. These electrodes are labelled according to the area of the head where they are situated: frontal (Fx), central (Cx), parietal (Px), occipital (Ox) and temporal (Tx). All active electrodes use to share a common reference electrode, placed in a position with low neural activity, like the mastoid or the ear lobule. These systems also use one single ground electrode.

The next slide will show why 3 electrodes are needed to record one EEG channel.

The amplifier



The objective of the amplifier is to provide to the signal of interest a sufficient gain to be able to be recorded by the AD/DA converter. The amplifier consists of 3 stages: differential amplification, analogue filtering, and final amplification.

The signal of interest (the AEPs) are recorded as the voltage difference between the active electrode (V^+, placed in Fz in the figure) and the reference (V^-, TP9 in the figure). The **differential amplification** consists of amplifying this voltage difference ($V^+ - V^-$).

The output voltage (V_{out}) of a differential amplifier depends on two different gains: (1) a differential-mode gain (Ad), that amplifies the signal of interest ($V^+ - V^-$); and (2) a common-mode gain (Ac), that amplifies the common voltage that share the two electrodes. Thus: $V_{out} = Ad\ (V^+ - V^-) + Ac\ (V^+ + V^-)/2$. A good differential amplifier is the one that presents a high Ad, and a small Ac, i.e. V_{out} depends mostly on the voltage difference between the two electrodes. The term $Ac\ (V^+ + V^-)/2$ is known as *common-mode interference*. The differential-mode gain Ad is usually moderate, with a factor around 10 or 20.

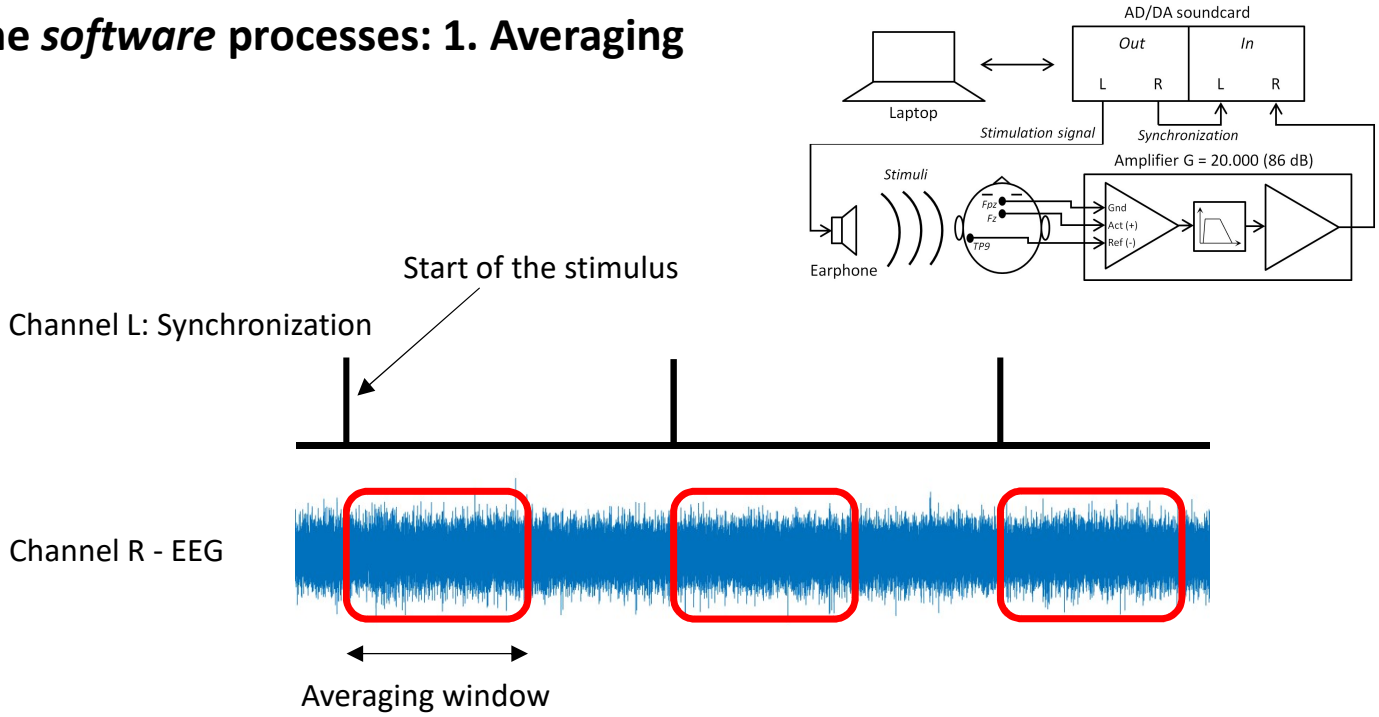
The **ground electrode** (Gnd, situated in FPz in the figure) is used to introduce some current and set the subject to the common voltage of the ground of the electronics of the amplifier, with the objective to reduce the common-mode interference. A classical electronic circuit to implement this process is the *driven right leg*.

The differential amplification stage follows an **analogue filtering**. Filtering consists of attenuating the energy of those frequency components out of the band of interest, which is determined by the type of AEP being recorded. The definition of the cut-off frequencies has been an object of debate for a long time. The current trend is to use a wide bandwidth in the analogue filters (implemented with electronic circuits), and a narrower bandwidth in the digital filters (implemented with software processes). The analogue filter must have (a) a high-pass filter, that reduces the energy of low frequency components, and especially, components near 0 Hz (DC, or *direct current*) to prevent a voltage drift that could saturate the amplifier; and (b) a low-pass filter, to attenuate components of frequency above the Nyquist frequency (fs/2) and avoid aliasing.

The last stage is **amplification**. The purpose of this stage is to provide a sufficient gain to allow the signal of interest being recorded by the AD/DA converter. This gain usually presents a factor around 500-10.000, depending on the type of AEP being recorded. The total gain of the amplifier is the

differential-mode gain (Ad) multiplied by the gain in the final stage.

The *software* processes: 1. Averaging

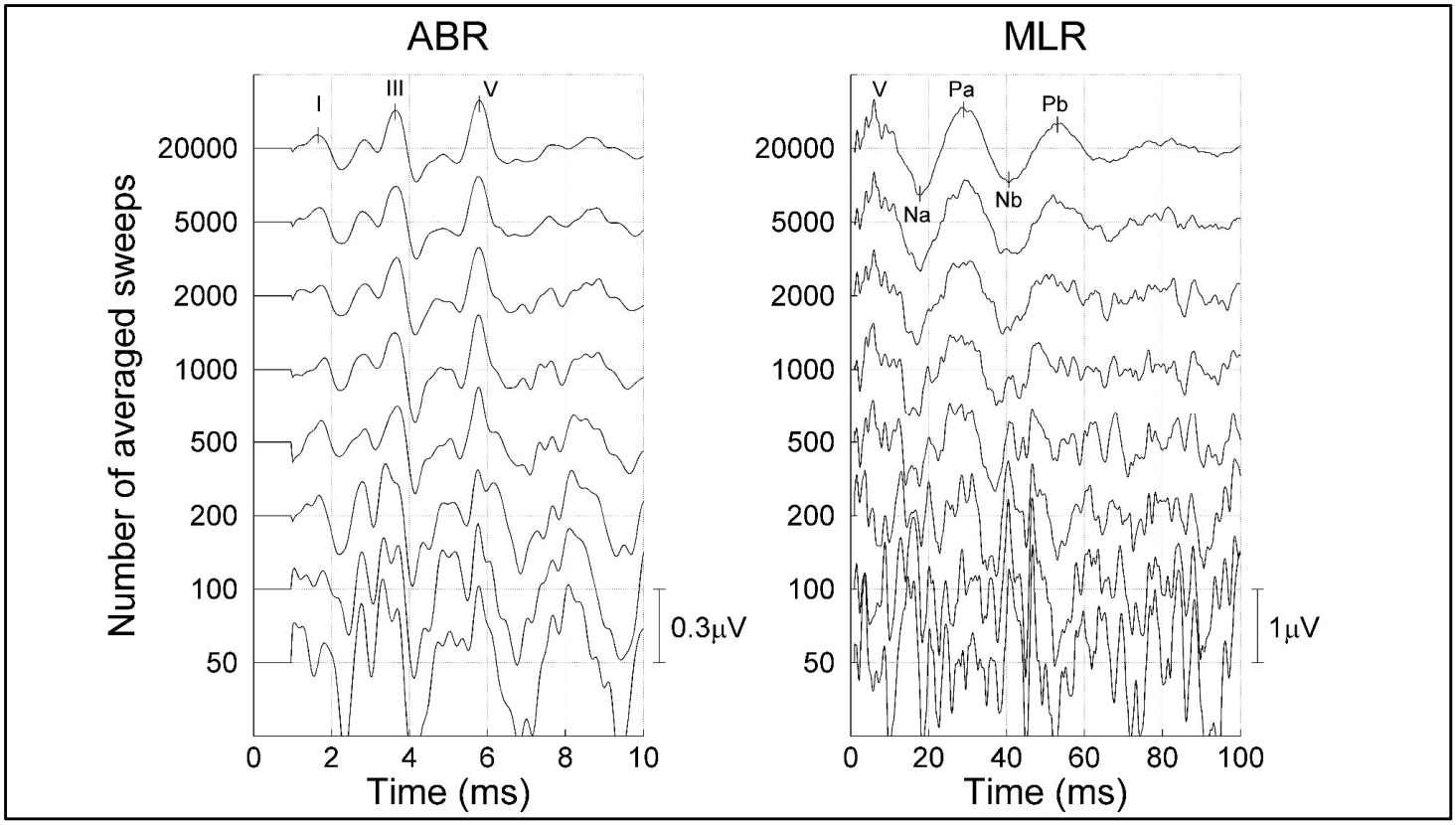


The figure on the top right corner shows the general diagram of an EEG recording system. The computer receives 2 channels of information through the two input channels of the AD/DA converter. **Channel 1 receives the synchronization signal**, that indicates the time instants in which each stimulus starts. **Channel 2 receives the amplified EEG**.

The figure in this slide shows in blue a segment of a real EEG. This figure shows that no clear biological response can be identified in this EEG segment. This is because when the electrodes are placed on the skin, they not only record the auditory evoked potentials, which is the signal of interest, but they also record neural activity associated to the heart beats and muscular activity, as well as they capture electromagnetic fields present in the recording booth (despite the recording sessions taking place in electromagnetic shielded booth, a complete shielding is almost impossible), and additionally, the EEG includes electric noise associated to the electronics of the amplifier and the AD/DA converter. In other words, the EEG does not only contain the signal of interest, but also **a lot of noise** of different types (electrophysiological, electromagnetic, electric, etc.).

The quality of the signal that contains the AEPs can be measured in terms of signal-to-noise ratio (SNR). The SNR is the ratio between the energy of the signal and the energy of the noise. A high SNR indicates that a specific signal has more energy from the signal of interest than the noise, and that would indicate of a signal of a high quality. **The SNR of an EEG is very low**, it is around -30 decibels (dB) in ABR signals.

The most effective technique to improve the quality of the signal of interest is **averaging**. This technique consists of the presentation of the same stimulus a certain number of repetitions and the averaging of the EEG segments corresponding to the onset of each stimulus to a specific averaging window (red rectangles in the figure). These EEG segments are known as *sweeps* or *epochs*. The length of the averaging window is set according to the duration of the AEP being recorded. For ABR signals, it is commonly used an averaging window of around 10 ms, in MLR signals the averaging window is around 100 ms, and a minimum averaging window of 300 ms is needed to record CAEP signals.



All right, averaging sweeps improves the quality of the signal, but how much?

The SNR of the signal after averaging *K* sweeps is the SNR of the raw EEG multiplied by the squared root of *K*, i.e.:

$$SNR_K = SNR_{EEG} * \sqrt{K}.$$

The SNR is usually expressed in decibels (dB) [$\mathrm{x[dB]} = 20 * log_{10}(x)$]:

$$SNR_K[dB] = SNR_{EEG}[dB] + 10 * log_{10}(K),$$

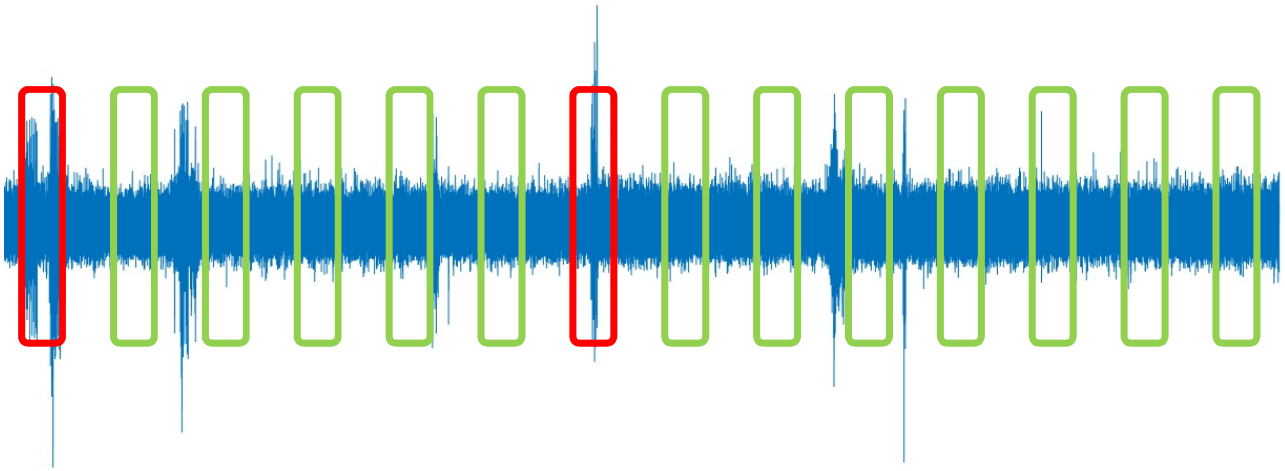
thus the result of averaging 1000 sweeps is that the SNR increases in 30 dB [$10 * log_{10}(1000)$]; and the result of averaging 2000 sweeps is that the SNR increases in 33 dB [$10 * log_{10}(2000)$]. According to this equation, every time that the number of averaged sweeps doubles, the SNR of the signal increases in 3 dB [$3\ \mathrm{dB} = \ 10 * log_{10}(2)$]. Would you be able to estimate the difference of quality between the SNR of an auditory evoked potentials obtained after averaging 20.000 responses, to those obtained with 5000 responses? Check your response in *[1].

The figure in the slide shows an example of how the quality of an ABR (left) and MLR (right) signal increases as the number of averaged sweeps increases.

The recordings obtained with 20.000 sweeps are of a very high quality, however they require a longer **recording time**. The time between the onsets of the stimuli, known as *inter-stimulus interval* (ISI), must be greater than the averaging window in order to avoid the neural evoked responses to be overlapped. In the case of ABRs, this figure shows that the averaging window used was 10 ms, which implies that the ISI must be greater than 10 ms. Supposing an ISI of 15 ms, the necessary recording time to obtain an ABR with 20.000 sweeps would be of 5 minutes (300 s = 0,015 s/stimulus * 20.000 stimuli), while the necessary time to obtain an ABR with 1000 responses is 15 seconds.

[1] 6 dB.

The *software* processes: 2. Artifact rejection



The figure in the slide shows a segment from a real EEG and the frames in which the signal of interest is contained along with noise (sweeps or epochs). This EEG visually shows that the level of the noise in certain sweeps is greater than in others.

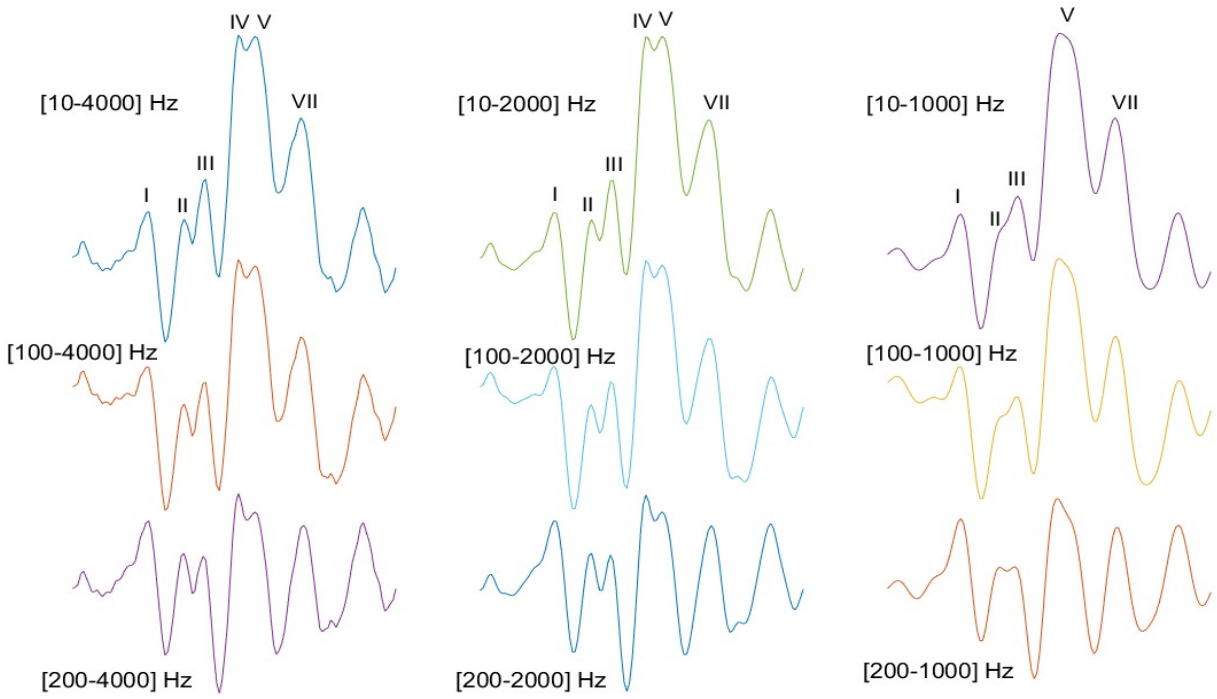
It is very common that electrophysiology tests are carried out with the subject reclining over a comfortable couch, in which they can leave the muscles of the shoulder and neck relaxed. Additionally, they are usually asked to remain as still as possible during the test to avoid contaminating the signals recorded by the electrodes. However, it is also common that the subjects move in some occasions, thus contaminating some sweeps.

The quality of the recordings can improve if **the most contaminated sweeps are identified and rejected** from averaging. The figure in the slide shows with green rectangles the less contaminated epochs, and with red rectangles the most contaminated ones.

There are several criteria that can be used to identify the most contaminated epochs. One criterion is based on the *peak value*, and other on the *root-mean-square* (RMS). Artifact rejection based on the peak value considers the maximum absolute value of each epoch, while the RMS value is an indicative of the energy of the full frame. The more energy the epoch has, the more contaminated will be.

Artifact rejection can be implemented by averaging only those epochs whose peak value or RMS value is below a certain threshold previously defined, or alternatively, by eliminating systematically a certain percentage of epochs. The first option presents the advantage of only eliminating the epochs considered contaminated, while the second option has the advantage of assuring that all recordings have been obtained by averaging the same number of epochs, and therefore, will present a similar quality.

The *software* processes: 3. Digital filtering



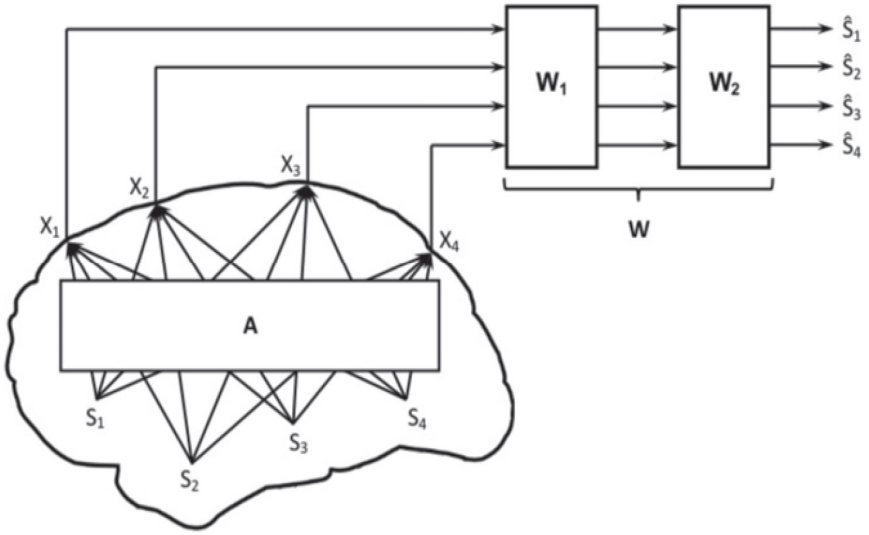
Digital filtering consists of attenuating the frequency components that are out of the band in which the signal of interest has energy in order to improve its quality. There are high-pass filters, that attenuate frequencies below a certain cut-off frequency; and low-pass filters, that attenuate frequency components above the cut-off frequency. Bandpass filters are built by both a high-pass and a low-pass filtering stages.

Slide 15 presented in the description of the amplifier that the intermediate stage consisted of analogue filtering, which is implemented with electronic circuits. The filtering described in this slide is digital filtering, which is implemented by programming codes using mathematical procedures that are out of the scope of this lesson. The aim of this slide is to show the effect on the signal of interest the application of digital filters with different cut-off frequencies.

In electrophysiology it is common to use bandpass filters with different cut-off frequencies according to the type of evoked potentials being recorded. The figure in the slide shows the same signal of auditory brainstem responses (ABRs) filtered with different cut-off frequencies. The ABR signals will be described with more detail in the next section.

In this figure, the rows show different cut-off frequencies for the high-pass filter, while the columns show different cut-off frequencies for the low-pass filter. The ABR signal on the top left corner shows an ABR processed with a filter of a wide bandwidth [10-4000] Hz. This signal shows both high-frequency and low-frequency components. This wide bandwidth makes that some segments show the signal with some ringing (high-frequency noise), but at the same time, all waves can be identified (waves I, II, III, IV, V and VII). On the other extreme, the ABR signal on the bottom right corner shows the same ABR signal filtered with a narrow bandwidth [200-1000] Hz. This signal shows that the low-frequency components are attenuated (the signal looks flatter), and the high-frequency ringing is softer, however, wave II almost disappears and waves IV and V appear as a single peak.

The *software* processes: 4. Advanced processing



Urigüen and Garcia-Zapirain (2015). EEG artifact removal – state of the art and guidelines. Journal of Neural Engineering 12, 031001 (23p).

Averaging, artifact rejection and digital filtering are techniques that are easy to implement and that can be applied to improve the quality of auditory evoked potentials recorded with a single-channel configuration, i.e. using information from a single-channel EEG. In addition to these techniques, there are other more sophisticated techniques that **take into consideration the information from several EEG channels**.

The figure of the slide shows a diagram in which there are 4 sources of neural activity situated in different points of the cortex (S_1 – S_4). This figure shows that the electrodes placed on the skin of the head are able to record the signals X_1 – X_4. This model assumes that each of the recorded signals (X_1 – X_4) contains a certain amount of energy from each of the four sources.

In matrix notation, we can call **S** to the matrix consisting of the signals generated by the four groups of neurons. The **S** matrix is unknown, and it is what we would like to estimate. We will call **X** to the matrix consisting of the activity recorded by the four electrodes. The distribution of energy of the sources S_1 – S_4 between the electrodes X_1 – X_4 is determined by the matrix **A**, i.e. **X** = **AS**.

Considering this model, the objective of these techniques consists of applying a mathematical process in which the matrix **W** is estimated, such that $\mathbf{W}^{-1}$=**A**. This way, the neural activity of the sources can be estimated ($\hat{S}_1$ - $\hat{S}_4$) using the signals recorded from the electrodes on the scalp, i.e. $\hat{\boldsymbol{S}}$ **= WX**. This model is known as *blind source separation*.

A very common procedure to implement this model is the ***independent component analysis*** (ICA). This method involves a series of complex mathematical processes that separate the signals into independent components.

An advantage of separating the available EEG channels in independent components is that certain components that are contaminating the signal of interest (artifacts) can be selected, and then the signal can be reconstructed without considering these components, thus denoising the signals from contaminating sources. ICA us usually the main method to eliminate the artifact associated to blinking, which usually contaminate the frontal channels (closer to the eyes).

Because of the mathematical complexity of ICA, this method is usually implemented by functions

available in toolboxes like FieldTrip or EEGLab.

EEG recording process described in a scientific paper

2.3. Brainstem and cortical evoked response data acquisition

Participants reclined comfortably in an electro-acoustically shielded booth to facilitate recording of neurophysiologic responses. They were instructed to relax and refrain from extraneous body movement (to minimize myogenic artifacts), ignore the sounds they heard, and were allowed to sleep throughout the duration of the recording procedure (~75% fell asleep). The continuous EEG was recorded from each participant in response to binaural presentation at 80 dB SPL through magnetically shielded insert earphones (ER-3A; Etymotic Research, Elk Grove Village, IL, USA). Stimuli were presented with fixed (rarefaction) onset polarity at a repetition rate of 0.93/s. Stimulus presentation order was randomized both within and across participants. Control of the experimental protocol was accomplished by a signal generation and data acquisition system (Intelligent Hearing Systems; Miami, FL, USA) using a sampling rate of 10 kHz.

The continuous EEG was recorded differentially between Ag and AgCl scalp electrodes placed on the midline of the forehead at the hairline (~Fpz) referenced to the linked mastoids (A2/A1). Another electrode placed on the mid-forehead served as the common ground. A vertical electrode montage was selected since it provides the optimal configuration for recording brainstem responses (Galbraith et al., 2000). Inter-electrode impedances were maintained $\leq 1\ k\Omega$, amplified by 100,000, and low-pass filtered online below 5000 Hz (6 dB/octave roll-off). Raw EEGs were then divided into epochs using an analysis time window from 0 to 1075 ms (0 ms=stimulus onset). Evoked responses were obtained by time-domain averaging each epoch over the duration of the EEG recording. A total of 1000 averages were collected for each condition in blocks of 250 sweeps. Sweeps containing activity exceeding $\pm 75\ \mu V$ were rejected as artifacts and excluded from the final average. The experimental protocol took about 2 h to complete.

This slide shows a text extracted from a scientific paper [Krishnan et al. (2012). Relationship between brainstem, cortical and behavioural measures relevant to pitch salience in humans. Neuropsychologia 50, 2849-2859] that describes the EEG recording methods followed to obtain cortical and brainstem responses. After studying this section, you should be able to understand what these researchers did.

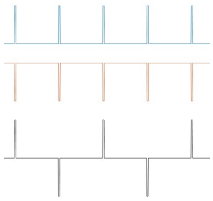
Part II

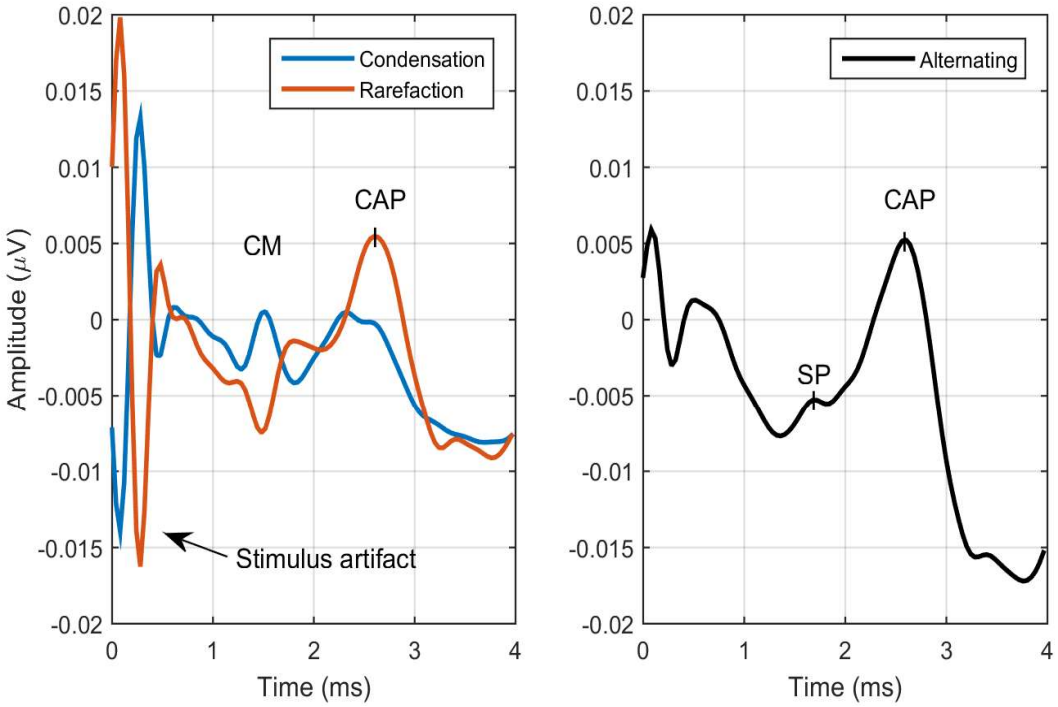
Types of Auditory Evoked Potentials

This section presents the main characteristics of some of the most relevant auditory evoked potentials. In particular I will describe the evoked potentials: (1) ECochG, (2) ABR, (3) MLR, (4) 40 Hz ASSR, (5) CAEP, (6) P300, and (7) N400.

Please consider the tables presented in this section as a guidance or recommendation. The recording and analysis parameters can always be adjusted according to the preferences of the evaluator.

1. ECochG





The signals in the slide show the main components of an **electrocochleography (ECochG)** signal: the cochlear microphonics (CM), the summating potential (SP) and the compound action potential (CAP). These components occur during the first 5 ms from the stimulus onset.

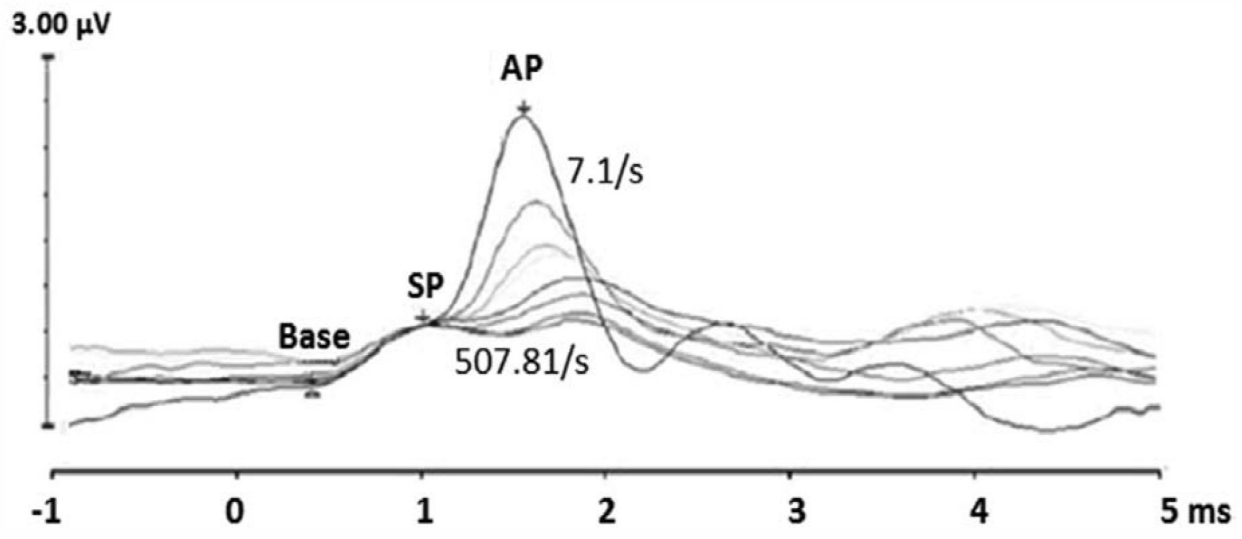
The **cochlear microphonics (CM)** are associated with the activity of the external hair cells (OHC). The CM usually present a morphology similar to the stimulus. For example, a pure tone evoke CM with a sinusoidal form with the same duration and frequency of the stimulus. The best approach to evaluate the CM is through two signals recorded in opposite polarities. The figure on the left shows two ECochG signals evoked by clicks in condensation (in blue, positive clicks) and rarefaction (in red, negative clicks). This figure shows that the CM component is sensitive to the polarity of the stimulus.

The **summating potential (SP)** is mainly associated with the activity of the inner hair cells (IHC). This potential shows the non-lineal properties of the cochlea, and can be considered as a shift in the reference voltage just before the CAP. The recording of the SP requires the presentation in alternating polarity to attenuate the component of the CM. The figure on the right shows the SP in an ECochG recording.

The **compound action potential (CAP or AP)** represents the activity of the auditory nerve. The CAP component is the same than the ABR wave I. Occasionally, ECochG recordings are presented with a negative amplitude, showing the CAP component as a negative peak. This component is effectively evoked by brief stimuli like clicks. Generally, clicks in rarefaction polarity show a CAP with a greater amplitude.

The abrupt peaks that appear at the onset of the signal (between millisecond 0 and 0.7) are the **stimulus artifact**. These voltage peaks are not a neural response, but the contamination of the signal produced by the electromagnetic fields that induce the coils of the earphones. This artifact is synchronized with the stimulus, and therefore, it cannot be reduced by averaging. The use of insert earphones allow separation in the time between the artifact and the signal of interest (description in slide 11).

Effect of stimulus rate in ECochG signals



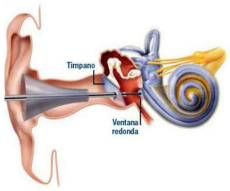
Kaf et al. (2016). Fast click rate electrocochleography and auditory brainstem response in normal hearing adults using continuous loop averaging deconvolution. Ear and Hearing 38, 244-254.

A low presentation rate (around 10 hz) enhances the recording of the CAP component. As the stimulus rate increases, the CAP component decreases as a consequence of a increase in the desynchrony of the spike firing of the neurons. At low presentation rates, most of the neurons will be ready to fire in the moment of the presentation of the stimulus, which implies that most neurons would fire in a similar time instant, and therefore, the synchronization of the firing of the neurons is high, resulting in narrow peaks of large amplitude. However, at fast stimulus rate many of the neurons would possibly be in refractory period (because of firing in the previous stimulus) and would not be ready to fire. This results into a lower neural firing synchronization, i.e. most of the neurons would not fire in the same time instant, and therefore, the peak of the CAP component would show a lower amplitude and a wider width. This effect is known as **neural adaptation**.

The figure in the slide shows the SP and AP components in ECochG signals obtained at different stimulus rates. This figure shows that the AP component is sensitive to the stimulus rate, while the SP component is almost not affected.

Stimulus	
Transducer	Insert earphones
Type	Click
Duration	0,1 ms
Polarity	Alternating or simple (R o C)
Rate	7,1 Hz
Level	70 a 90 dBnHL
Presentation	Monaural

Acquisition	
Electrodes	Depends
Filter	[10-1500] Hz
Amplification	15K – 70K
Time analysis	5 or 15 ms
Number of sweeps	50 (Promontory) o 1000 (TIPtrode)



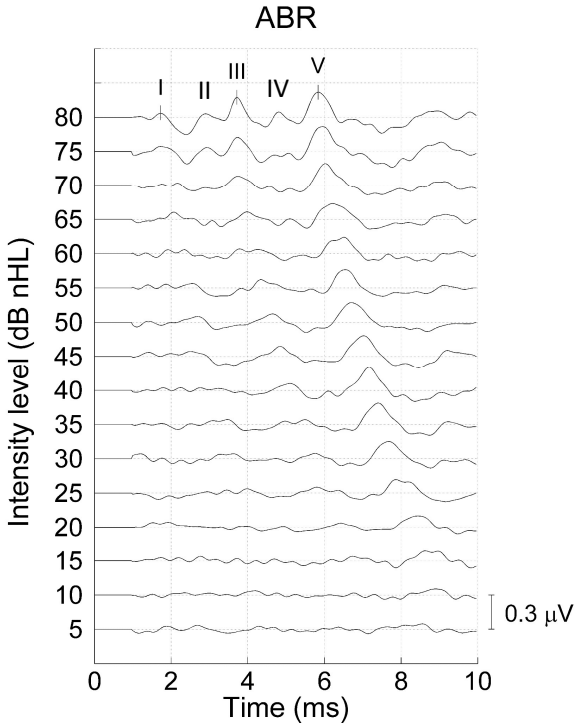
These tables summarize the main recording parameters of ECochG signals. The standard method of recording ECochG signals consists of clicks of 100 µs duration presented monaurally through insert earphones in the rarefaction (R) and condensation (C) polarities.

The recording of ECochG signals can be done by an active electrode usually placed in Fz and a reference (a) trans-tympanic electrode situated on the promontory [image on the left, this method is slightly invasive since it has to perforate the tympanic membrane]; (b) electrode situated on the tympanic membrane; (c) electrode situated on the auditory canal (TIPtrode, image on the right); or even (d) using the typical configuration of ABR signals placing a reference electrode on the mastoid.

The processing of ECochG signals usually require amplifying by a factor between 15.000 – 70.000, and filtering the signals with a filter [10-1500] Hz. The averaging window is usually around 5 ms if only the ECochG components are presented, or 15 ms if other later components (ABRs) are also shown. With a trans-tympanic electrode, averaging 50 responses gives usually a signal of sufficient quality, but with a TIPtrode it is necessary to average at least 1000 stimuli.

ECochG signals have different clinical applications. The recording of AEPs allows the diagnosis of certain auditory pathologies. For example: (a) detection of CM confirms the well-functioning of the OHC in patients that do not show optoacoustic emissions, (b) detection of SP provides evidence of the IHC activity, and (c) detection of CAP provides evidence of the integrity of IHC and the synaptic communication between IHC and the afferent fibers of the auditory nerve. The recording of ECochG is also used in the diagnosis of the **Ménière's disease** by the ratio SP/AP, characterized by a SP component significantly larger than in normal hearing subjects.

2. ABR



Valderrama et al. (2014). A flexible and inexpensive high-performance auditory evoked response recording system appropriate for research purposes. Biomedical Engineering 59, 447-459.

The **auditory brainstem response (ABR)** is without a doubt the most used AEP in the clinic. This AEP shows 7 peaks that appear within the first 10 ms from the stimulus onset. These components are labelled with Roman numerals. The main components are waves I, III and V.

The ABR **wave I** is equivalent to the CAP in ECochG signals. Wave I is generated in the cochlea and represents the activity of the afferent fibers of the auditory nerve. **Wave II** is generated in the proximal part of the auditory nerve. **Wave III** is generated in the cochlear nucleus (CN) and in the superior olivary complex (SOC). **Wave IV** is generated in the nucleus of the lateral leminiscus (LL). **Wave V** indicates activity of the ending of the LL and the inferior colliculus (IC).

These waves are usually characterized according to their *amplitude*, which is usually measured as the difference in μV between the peak and the following trough; and to their *latency*, measured in milliseconds between the occurrence of the peak and the stimulus onset.

One of the main characteristics of the ABRs is that their components are present even if the subject being tested is sleeping or under sedation, which implies that these signals do not depend on the estate of arousal of the subject, and therefore, the recording of these signals can be used as an **objective test to evaluate the auditory pathway**. Furthermore, the non-invasive nature of the recording process of these signals is appropriate for its use both in adults, children and newborns. These characteristics have promoted the use of these signals in clinical applications such as *auditory screening*; to estimate the hearing threshold; to evaluate the site of lesion of certain neural pathologies; and in intraoperative monitoring.

The figure in the slide shows ABR signals obtained at different levels of stimulation. This figure shows that as the level of intensity decreases, the ABR components appear with a larger latency and with a lower amplitude. Wave V is usually the most robust component. The lowest level in which wave V is identified is the threshold level. In the clinical practice, a subject is considered with normal hearing if wave V is detected at 30 dB nHL.

Stimulus	
Transducer	Insert earphones
Type	Click
Duration	0,1 ms
Polarity	Rarefaction
Rate	35 Hz, <100 Hz
Level	Variable
Presentation	Monaural

Acquisition	
Electrodes	Surface
Filters	[30-3000] Hz
Amplification	15K – 70K
Analysis window	10 ms
Number of sweeps	>1000
Sampling rate	>10.000 Hz

The recording of ABR signals usually involves the use of 100 µs duration clicks, presented monaurally through insert earphones. The rarefaction polarity is preferred to evoke larger wave I amplitudes. The stimulus rate is usually around 35 Hz.

Since the components of the ABRs appear during the first 10 ms from the stimulus onset, **the stimulus rate must be lower than 100 Hz** (100 Hz = 1/10 ms) to avoid adjacent neural responses to be overlapped.

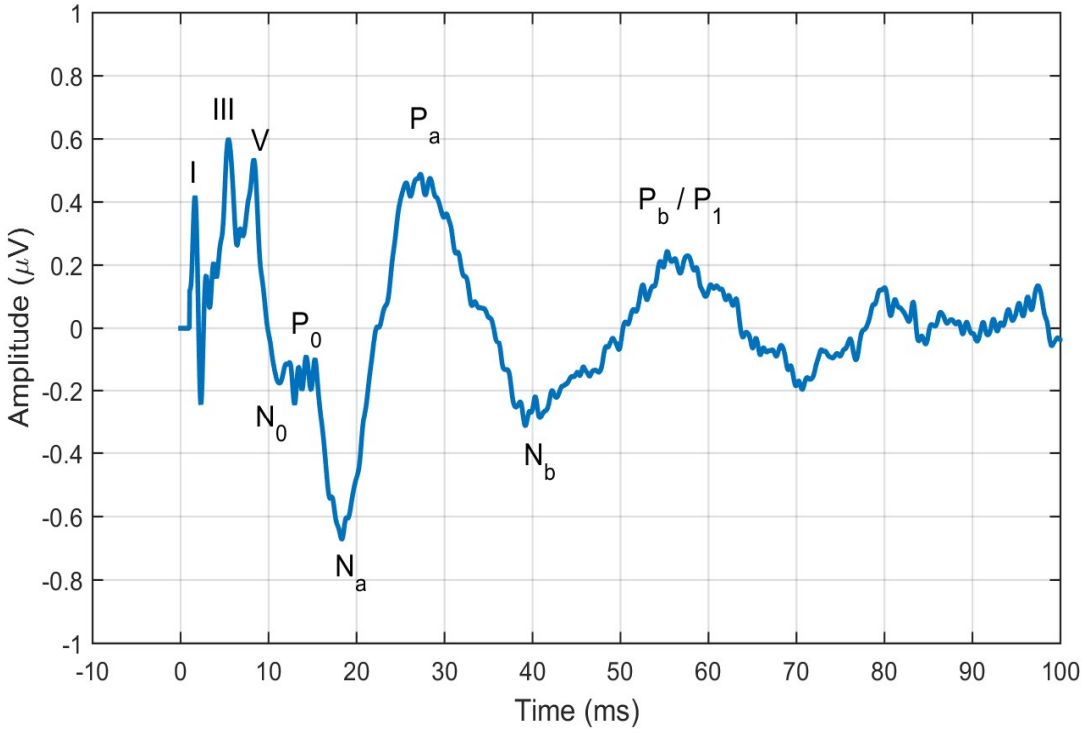
The EEG is usually recorded through surface electrodes placed in different points of the head. The optimal configuration is with an active electrode placed in middle line, in the superior part of the forehead (Fz), the reference electrode placed in the ipsilateral mastoid (right or left, same side as the ear being presented the stimulus), and a ground electrode situated in the mid-forehead.

The frequency components of ABR signals are usually in the band [30-3000] Hz. The slide 17 shows the effect of filtering the same ABR signal with different bandwidths. Since the greatest frequency component of ABRs is around 3000 Hz, the sampling frequency must be greater than 6 kHz to avoid *aliasing*. Since filters are not ideal, it is usually recommended to leave some margin and use an fs slightly greater. It is common to use a sampling rate of at least 10 kHz.

The averaging window must be at least of 10 ms, and more than 1000 responses are needed to be averaged to obtain an ABR signal of sufficient quality (see slide 15).

My personal preference in the recording of ABR signals is to use filters of a wide bandwidth and a large number of averaged responses to have ABR signals of high quality. The use of a large number of averages requires a longer recording time, which is usually feasible in adults, but more difficult in children or newborns. In situations in which the recording time is limited, one option is to increase the stimulus rate slightly to obtain the same number of averages in less recording time.

3. MLR



Middle latency responses (MLRs) present three voltage oscillations within the first 100 ms from the stimulus onset. The first components are the ABR. Wave V is usually the most prominent component. The components N_0, P_0, N_a, P_a, N_b and P_b appear after the ABR. These components are illustrated in the figure of the slide. The most robust component is N_a-P_a. The P_b component corresponds to the P_1 component of the CAEP. The generator sources of these components are mainly in the brainstem, the thalamus and the auditory cortex.

When the reference electrode is placed on the mastoid, this signal can sometimes be contaminated by an artifact associated with the post-auricular muscle, known as PAM, which can be contracted when a high-level stimulus is presented. If this artifact is present, the PAM can be reduced either by placing the reference electrode in an area slightly separated from the mastoid, and presenting the stimuli at a lower level.

The MLR waveform may vary with age, the estate of arousal of the subject, the stimulus rate, the level and the filter bandwidth. Therefore, the clinical utility of this AEP is constrained to awake and collaborative subjects. The longer latencies of these waves separate the neural responses from the artifact when using electrical stimuli in patients with cochlear implants, which is an advantage to evaluate the efficiency of the implant.

Probably the most relevant application of MLR signals is under a particular configuration in which a stimulus rate of 40 Hz is used in order to evoke an steady-state potential. This configuration is further presented in section “4. 40 Hz ASSR”.

Stimulus	
Transducer	Insert earphones
Type	Clicks or windowed tones
Duration	0,1 ms o 10 ms
Polarity	Any
Rate	<10 Hz
Level	<70 dBnHL
Presentation	Monaural

Acquisition	
Electrodes	Surface electrode
Filter	[1-1500] Hz [1-200] Hz
Amplification	15K – 70K
Analysis	100 ms
Number of responses	>1000
Sampling rate	>10.000 Hz

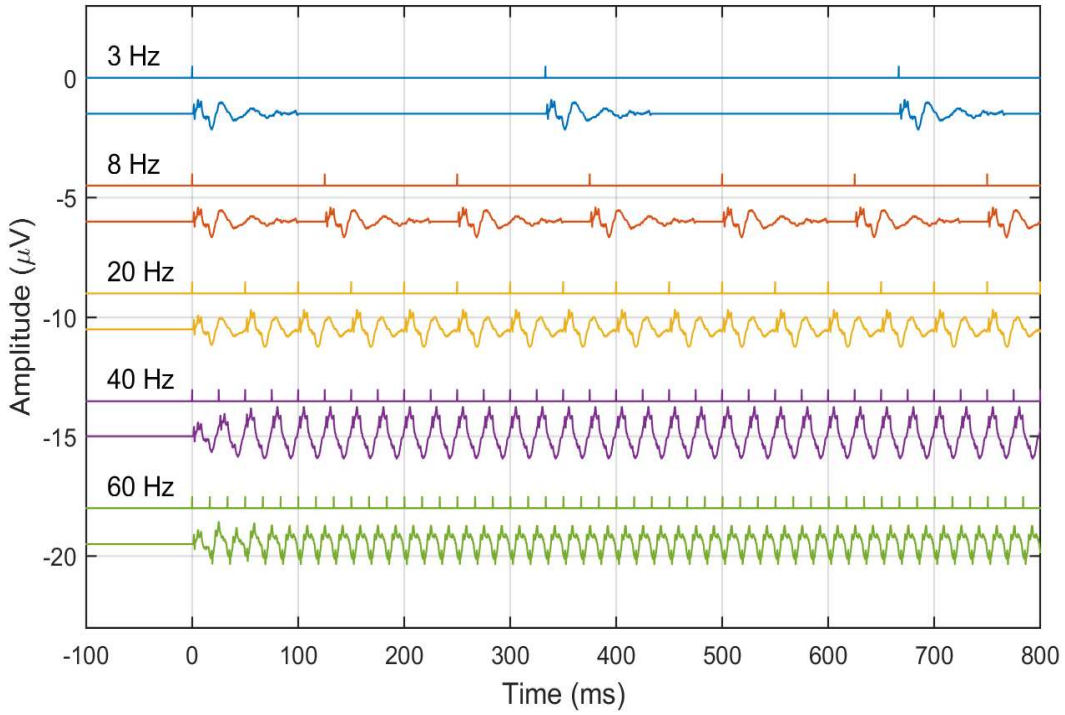
MLR signals are usually recorded with clicks of 100 µs duration or windowed tones of around 10 ms duration through insert earphones.

Considering an averaging window of 100 ms, the stimulus rate must be lower than 10 Hz to avoid overlapping responses. It is recommended a moderate level, avoiding high levels of intensity.

The minimal electrode configuration consists of an active electrode placed in Fz or Cz, a reference electrode placed on the ipsilateral mastoid, and a ground electrode placed on the forehead. Wide filters are usually recommended [1-1500] Hz to record both ABR and MLR components, or narrow filters [1-200] Hz to record only the MLR components.

It is recommended to average at least 1000 responses to obtain a signal with sufficient quality (see slide 15).

4. 40 Hz ASSR



The AEPs that we have analysed so far presented the characteristic that the interstimulus interval (ISI) [inverse of the stimulus rate] was always greater than the averaging window in oder to avoid adjacent responses to be overlapped, thus contaminating the signals.

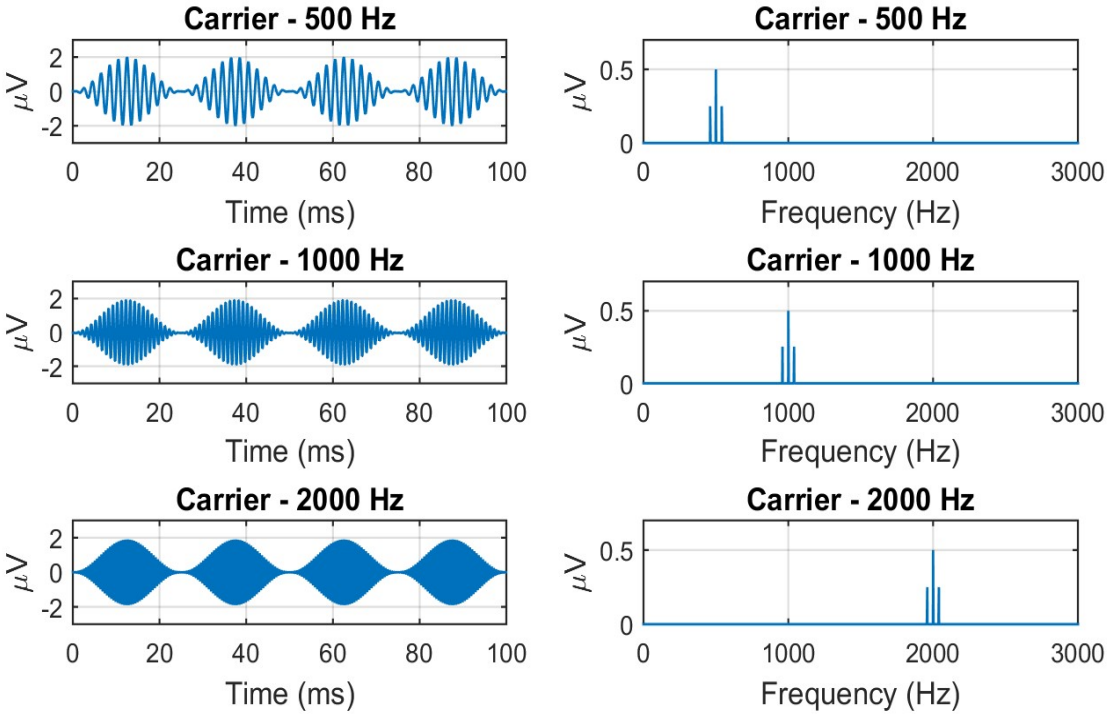
The figure in this slide shows the effect of using an ISI lower than the averaging window using synthesized signals. The blue signal shows MLR signals evoked by stimuli whose ISI is 333 ms (stimulus rate of 3 Hz, i.e. 3 stimuli per second). The top signal shows the instants in which the stimuli are presented. This example shows that the responses do not overlap because the ISI is longer than the averaging window (100 ms in MLRs).

At 8 Hz (orange signal, with ISI = 125 ms) the responses do not overlap, however, at rates greater than 10 Hz (with ISIs lower than 100 ms) the responses do overlap.

These examples show that **overlapping responses result in a periodic signal** (of a period equal to the stimulus period) in which the original response is contaminated with adjacent responses. Depending on the stimulus rate, this contamination can be constructive or destructive. This figure shows that **the stimulus rate 40 Hz produces a constructive interference**, that makes that the resulting signal presents a greater amplitude than the original one, which may facilitate neural response detection. This constructive interference is a consequence of the Na-Pa component overlapping with the Nb-Pb component of the adjacent response, resulting in a greater amplitude signal. This auditory evoked potential evoked by a stimulus rate of 40 Hz is known as **40 Hz auditory steady-state response (ASSR)**.

A relevant reference about this potential is: Bohorquez et al. (2008). Generation of the 40-Hz auditory steady-state response (ASSR) explained using convolution. Clinical Neurophysiology 119, 2598-2607.

40 Hz ASSR stimulus

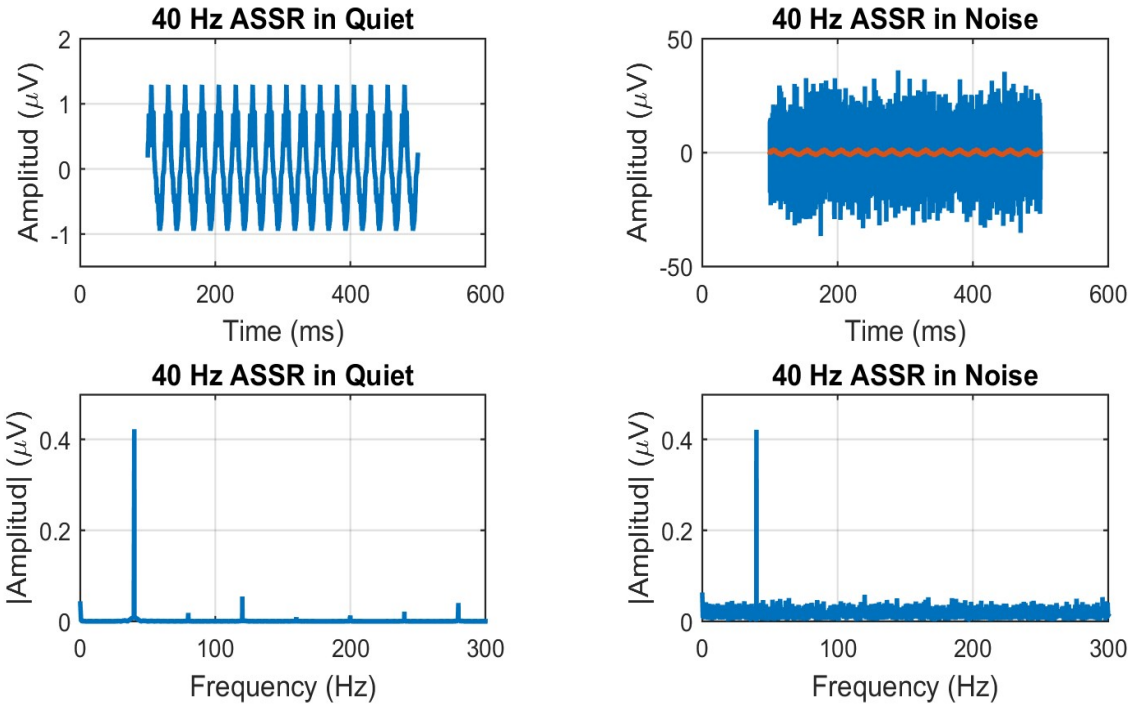


The 40 Hz ASSR can be evoked by clicks, as shown in the previous slide, or also by stimuli that present certain **frequency specificity**, i.e. stimuli in which the frequency components are controlled, stimulating therefore certain portions of the cochlea.

One of these options is the use a of a **sinusoidal signal modulated in amplitude by another sinusoidal signal (SAM, *sinusoidal amplitude modulation*)**. These signals consist of a pure tone of a certain frequency, known as *carrier frecuency* (fc), in which the amplitude vary as a function of another sinusoidal signal whose frequency is known as *modulating frequency* (fm). fm is usually much lower than fc. A SAM signal presents energy in the frequency of the carrier fc and in fc ± fm.

The 40 Hz ASSR can be evoked by SAM signals in which the modulating frequency is fixed at fm = 40 Hz, and the carrier frequency presents different frequencies in order to explore the ASSR response evoked by different sections of the cochlea. For example, some common carrier frequencies are fc = 500 Hz, 1000 Hz or 2000 Hz. The figure in the slide shows the representation in the time and in the frequency domains of these stimuli.

40 Hz ASSR analysis



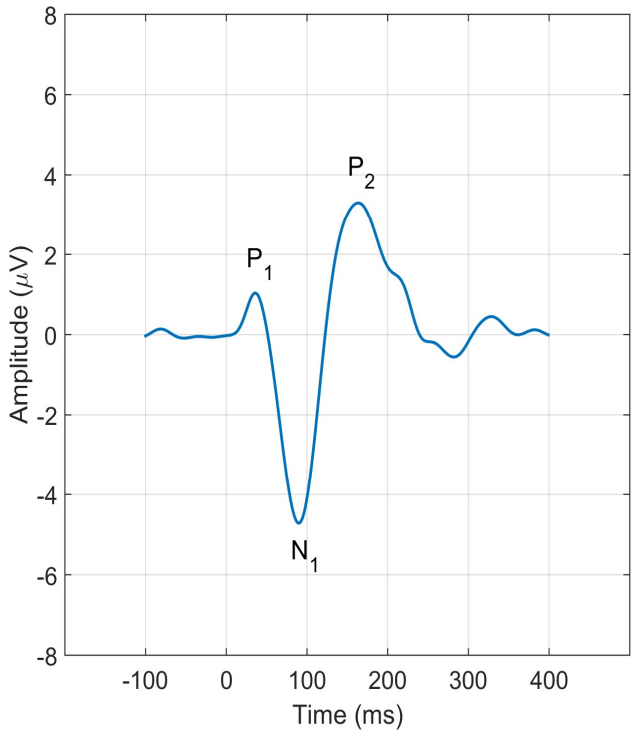
The figure on the top left corner shows a segment of an ASSR signal evoked at 40 Hz. This ASSR signal is not a real signal, this signal has been synthesized without noise to be able to study its components with more precision. This figure shows that the 40 Hz ASSR is a periodic signal, of a period equal to the stimulation period (25 ms = 1/40 Hz).

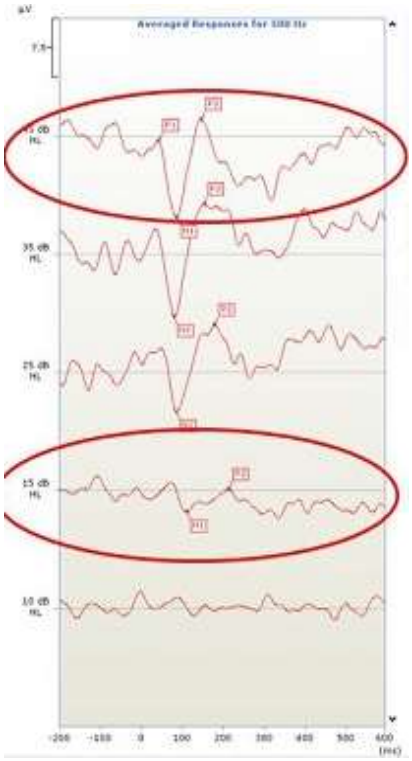
The AEPs studied so far have been analysed in the time domain, i.e. we have characterized their components as a function of the amplitude and latency of the peaks of a signal that varied as a function of time. **In the case of ASSR signals, the analysis is more effective in the frequency domain**, i.e. by analysing the energy of each frequency component in the recorded EEG.

From the field of Signal Theory, the representation in the frequency domain of any periodic signal in the time domain consists of a series of peaks of energy situated in the fundamental frequency (the inverse of the period of the signal) and its harmonics (multiples of the fundamental frequency). Therefore, since the 40 Hz ASSR are a periodic signal of a period equal to 25 ms, their representation in the frequency domain is a series of peaks of energy in multiples of 40 Hz (1/25ms): 40 Hz, 80 Hz, 120 Hz… These peaks can be observed in the figure on the bottom left corner, which shows the frequency components of the synthesized 40 Hz ASSR (no noise).

The two figures on the left column show a synthesized scenario in order to understand the representation of the signal of interest both in the time and frequency domains. However, real EEGs usually record the signal of interest along with a large amount of noise. The figure on the top right corner shows a 40 Hz ASSR signal (highlighted in red color) with noise. This figure shows that the analysis of the signal in the time domain is barely impossible. However, the representation of the same signal in the frequency domain shows a large amount of energy around the 40 Hz frequency bin, compared with the energy of the remaining components, which is an objective indicator of neural response. This figure shows that the analysis of the 40 Hz ASSR is more effective in the frequency domain.

5. CAEP





The main components of the cortical auditory evoked potentials (CAEP) are the wave complex P1-N1-P2, that appear during the first 200 ms from the stimulus onset, as shown on the figure of the left.

The P1 component is a peak generated in the primary auditory cortex, specifically in the Heschl gyrus. The N1 component has several generator sites in the primary and secondary auditory cortex. The P2 component is not as well characterised as the P1 and N1 components, however it is thought that it also has multiple generator sites in the primary and secondary auditory cortex, as well as in the ascending reticular activator system.

The morphology of these potentials can be affected by the state of arousal of the subject, by the age, gender, auditory training (musical for instance), alcohol, and by the effect of some medicines like sedatives, anaesthetics, tranquilizers and psychotherapeutic agents.

The main clinical applications of the CAEPs are in the diagnostic of sensorineural hearing loss, auditory processing disorders (APD), and in subject with learning and reading disorders. Furthermore, the CAEPs are currently being used to estimate the hearing threshold evaluating the changes in the morphology of the response obtained at different levels of stimulation, similarly as in ABR signals (figure on the right).

Stimulus	
Transducer	Insert earphones
Type	Syllables or windowed tones
Duration	Tens of ms
Polarity	Not relevant
Rate	<1,1 Hz
Level	Moderate
Presentation	Mono- / bi-aural

Acquisition	
Electrodes	Multichannel, Cz
Filters	[0,1 – 100] Hz
Amplification	10.000
Analysis	[-100 – 400] ms
Number of sweeps	150
Sampling rate (fs)	>250 Hz

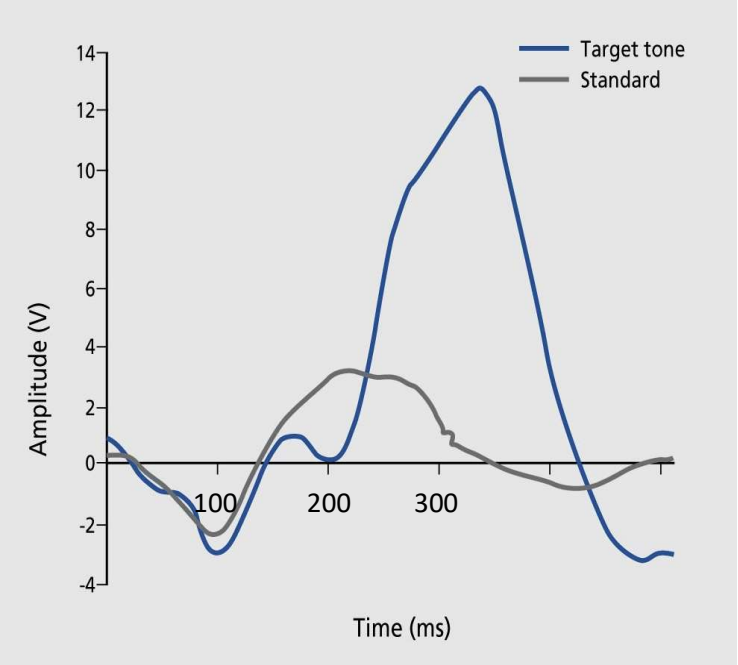
The CAEP are optimally evoked with stimuli of several tens of milliseconds, as windowed tones or syllables (/ba/, /da/, etc.). In these cases, the polarity of the stimulus is not relevant.

The stimuli can be presented both monaurally and binaurally, usually at a moderate level, leaving a long inter-stimulus interval (usually around 1 second) in order to allow most of the neurons to recover from the previous stimulus and be ready to fire with each stimulus, i.e. to reduce the effect of neuronal adaptation (described in the slide 22).

The average of 150 or 200 responses is usually enough to obtain CAEP signals of sufficient quality. The optimal electrode configuration includes an active electrode situated in Fz or Cz, a reference electrode in the ipsilateral mastoid, and a ground electrode in the forehead (FPz). It is also common to use a multichannel configuration, as the 10/10 system described in slide 12.

The EEG is usually sampled at frequencies higher than 250 Hz, it is filtered with a high-pass filter of a low cut-off frequency, e.g. 0,1 Hz, and with a low-pass filter of cut-off frequency lower than 100 Hz to avoid aliasing. The averaging window usually includes 100 ms pre-stimulus and a window of at least 400 ms from the stimulus onset.

6. P300



Boeijinga, PH (2002). Objective markers of drug effects on brain function from recordings of scalp potential in healthy volunteer. Clinical Research 4, 388-394.

So far we have studied the main auditory evoked potentials, from the cochlea to the auditory cortex. However, once these impulses reach the cortex, a number of cognitive processes *("somehow")* interpret the meaning of the message. For example, understanding a message requires (a) the development of a language [if anyone speaks to me in Chinese, at least I would not be able to understand the message]; (b) pay attention [if somebody speaks to me while I am sleeping, I would neither understand the message]; (c) inhibit perceived information that doesn't interest me [for example, the conversations in other tables in a restaurant]; etc.

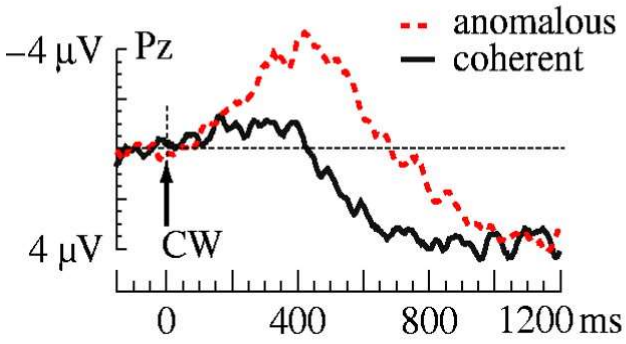
There a re a number of evoked potentials that reflect certain cognitive processes. I will present very briefly some of the most well known, the P300 and the N400 AEPs. An excellent reference about these potentials is Duncan et al. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. Clinical Neurophysiology 120, 1883-1908.

In these potentials, it is common to use a multichannel configuration (usually 64 channels, 10/10 system) situated in different positions on the head. The filters usually present a bandpass bandwidth of [0,01-100] Hz, and it is common to use a sampling rate of 250 to 1000 Hz.

The **P300 AEP** consists of a voltage peak that usually appears around 300 ms from the stimulus onset (it usually appears after the components of the CAEP). This potential is usually evoked by the *oddball* paradigm, that consists of a repetitive presentation of one stimulus that is infrequently interrupted by a different stimulus. For example, if we present the following stimulus sequence: da – da – da – da – da – da – da – da – da – BA – da – da – da… This potential is particularly sensitive to the estate of arousal of the subject, and can be affected by exercise or fatigue, by certain medicines, by age, by the intelligence coefficient, if the subject is right- or left-handed, by gender, and even by the personality of the subject.

The P300 AEP has been used in **clinical applications** to diagnose schizophrenia, humour disorders, alcohol addiction, dementia, as well as in kids with hyperactivity, attention deficits and autism.

7. N400

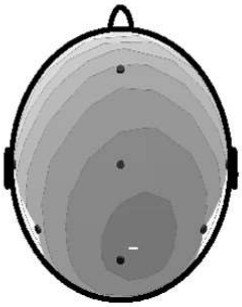


Coherent & predictable

I like coffee with milk and **sugar**

Anomalous

I like coffee with milk and **wood**



distribution of effect in 300–500 ms

Hagoort, P, van Berkum, J (2007). Beyond the sentence given. Philosofical transactions of the Royal Society B 362, 801-811.

The **N400 AEP** consists of a trough of voltage that occurs around 400 ms from the stimulus onset. In some studies, like the one of the figure, these AEPs are presented with the negative voltage in the positive abscissas axis, thus a trough is presented like a peak.

The N400 is usually recorded by the comparison of two signals evoked with sentences that are coherent and predictable against other incoherent sentences. The amplitude of the N400 is considered by many researchers as an index of the difficulty to retrieve the conceptual meaning of a specific word. These potentials can be evoked by both auditory stimuli (usually sentences presented by earphones or speakers) or visual stimuli (drawings, patterns or pictures).

The slide in the figure shows an example of a coherent and predictable sentence, like “I like coffee with milk and sugar”; and an anomalous sentence in which the last word does not fit the context: “I like coffee with milk and wood”. The signals in the slide show that evoked potentials with incongruent sentences present a deeper negativity around 400 ms from the onset of the *critical word* (CW). This potential has a broad distribution over the brain, presenting a maximum amplitude in middle line around the parietal area (Pz), as shown in the figure of the bottom right corner.

The recording of the N400 is an indicator of semantic comprehension. If a subject does not understand the meaning of the stimulus being presented, this subject would not be able to distinguish between coherent and incoherent stimuli.

The N400 has been used in patients with a broad variety of mental, neurologic and psychiatric disorders. It has also been used to examine the normal development of the language and its evolution as we get older.

Part III

Emerging techniques in the recording of AEPs

In this section, I will briefly present the main results of a research line that my group has held in collaboration with Dr Ángel de la Torre, in the Department of Signal Theory, Telematics and Communication from the University of Granada.

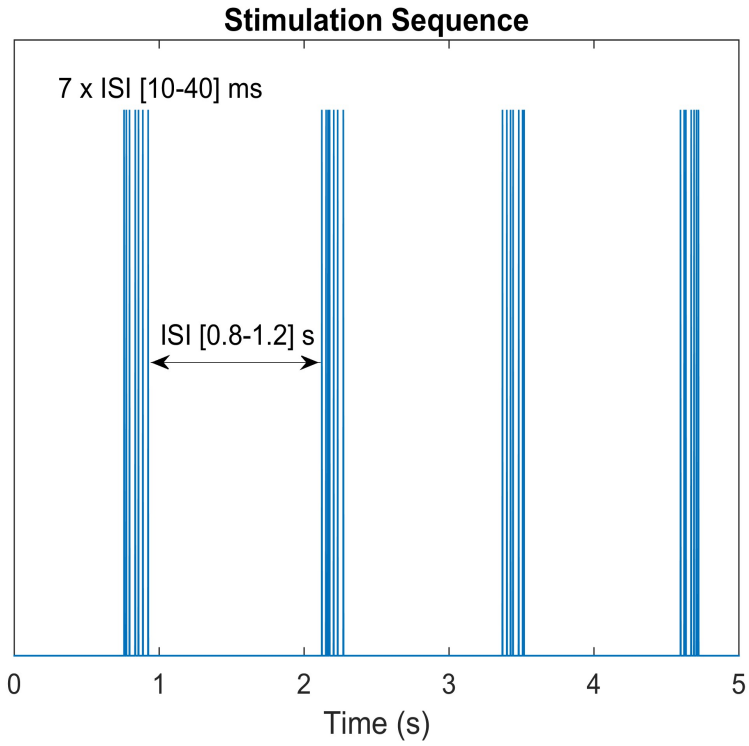
There results have been recently presented in the conference *International Evoked Response Audiometry Study Group (IERASG)*, in Warsaw (Poland), 21-25 of May of 2017, under the title "Comprehensive recording of auditory evoked potentials by projecting over a base of functions".

The objective of this research line was the development of a stimulus and an analysis procedure that would allow to **present the main auditory evoked potentials along the auditory pathway in the same figure**.

You have studied in part II of this lecture that ABR, MLR and CAEP signals are recorded with different stimulus and analysis parameters. On one hand, ABRs require the fast presentation of many brief stimuli, and the EEG is filtered with a typical bandwidth of [100-3000] Hz. On the other hand, CAEP signals are usually evoked by the presentation of longer duration stimuli at a lower rate, using filters with a bandwidth of [1-30] Hz. These differences prevent CAEP components to be shown when recording ABR signals, and vice versa.

Presenting all evoked potentials in the same figure would have the advantage of evaluating in a single test the activity of all the main stations of the auditory pathway, giving more information than the current tests.

Stimulus



The stimulus used to evoke the main components of the auditory pathway consisted of the presentation of a number of bursts of clicks, in which the *inter-stimulus interval* (ISI) between the clicks of each burst varied randomly between 10 and 40 ms. The ISI between each burst varied randomly between 0,8 and 1,2 seconds.

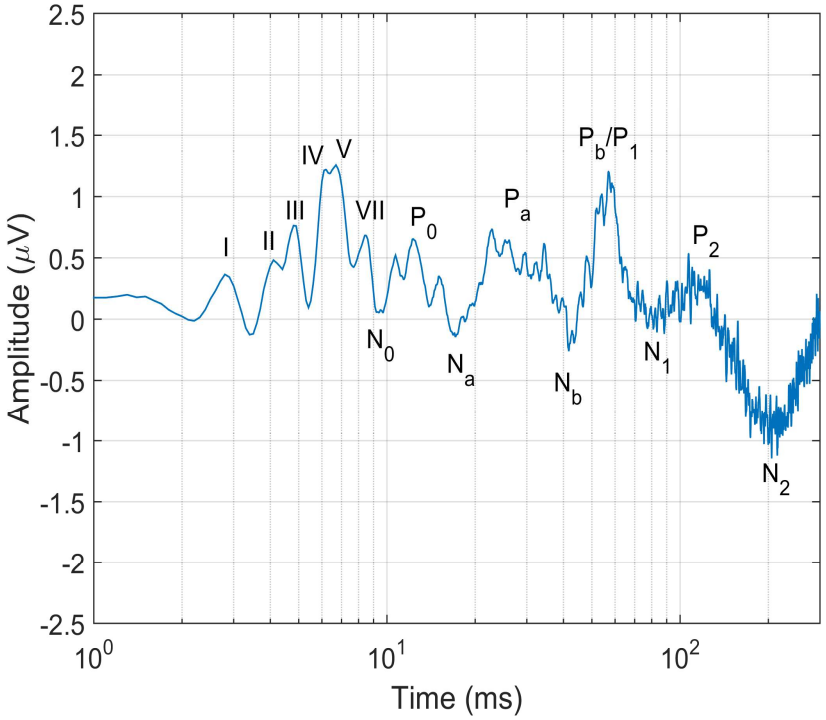
The hypothesis behind this stimulation paradigm is that each click would evoke ABR and MLR signals, while each burst would evoke CAEP components. Therefore, it was expected that only the first click of each burst would evoke CAEP compoents.

The stimulus sequence that was used in this pilot study consisted of 400 bursts of 7 clicks, presented binaurally at 95 dB peak-to-peak equivalent sound pressure level (ppeSPL) in both rarefaction and condensation polarities. The electrode montage consisted of an active electrode in FCz, two reference electrodes placed in the two mastoids, and a ground electrode placed in middle line in the forehead.

Two electroencephalograms were recorded (FCz-M1 and FCz-M2) using a sampling rate of 10 kHz, and an analogue filter with a bandpass frequency range of [1-3000] Hz. The two EEGs were averaged to obtain a single-channel EEG of FCz referenced to the combined mastoid reference, i.e.:

$$\frac{[FCz-M1]+[FCz-M2]}{2} = FCz - \frac{M1+M2}{2}.$$

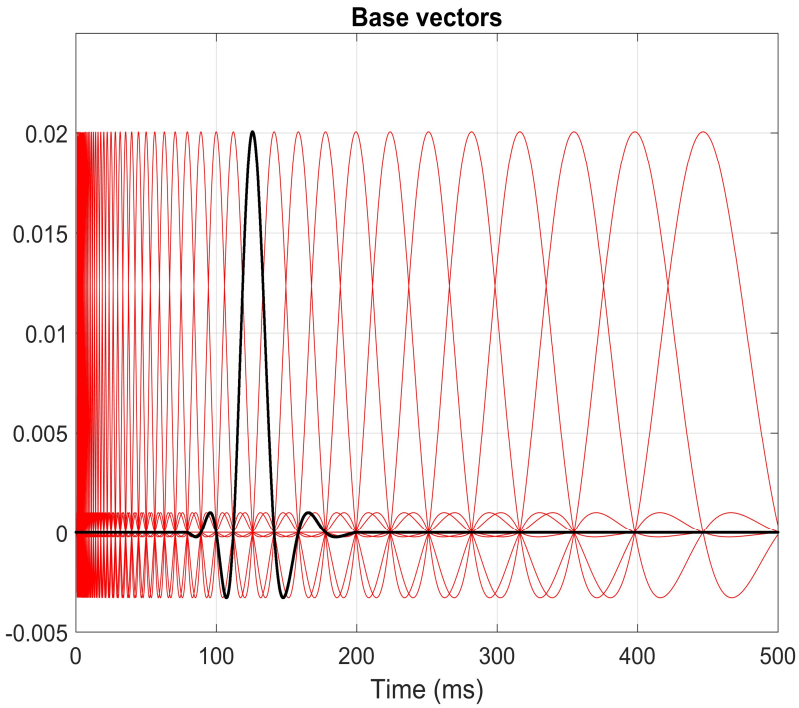
Processing – Step 1. Averaging



The figure in the slide shows the average of the EEG epochs corresponding to the first 300 ms from each click onset. Please note that the time axis is in logarithmic scale, which allows identification of all the components of the auditory pathway.

This figure shows that the proposed stimulus evoked the main components of the auditory evoked potentials, from the cochlea to the auditory cortex. However, this figure also shows that the representation of this signal in the logarithmic time-scale makes the later components to appear contaminated by high-frequency noise.

Processing – Step 2. Latency-dependant filtering

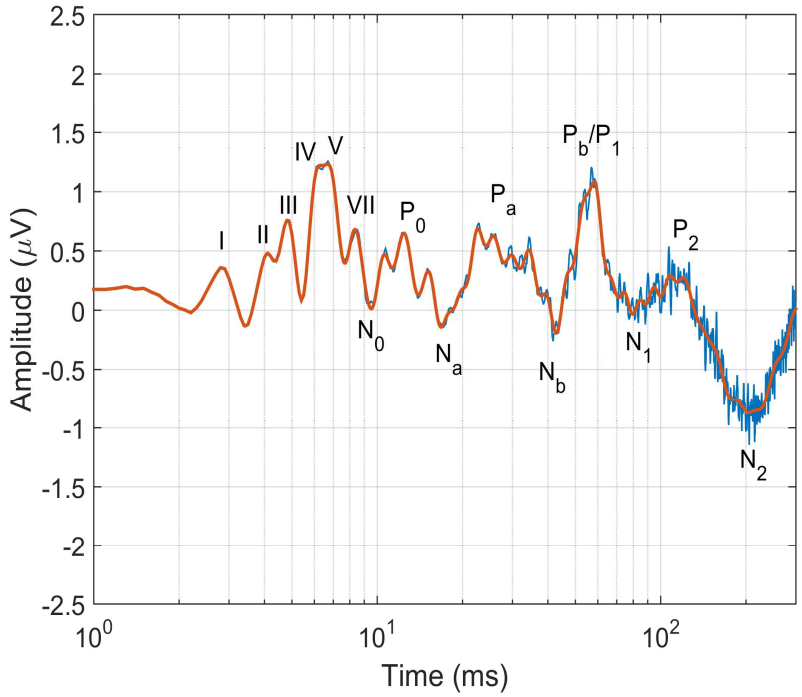


In order to represent with a higher quality the signal of the previous slide, we developed a mathematical procedure that implements a **filtering stage with cut-off frequencies dependant on the latency**, i.e. a filtering of the signal in which the frequency components being attenuated depended on the time instant of the signal.

This filtering stage is performed by projecting the signal of interest (e.g. the evoked potentials of the previous slide) over a base of functions (figure on the slide). This figure shows that the base of functions consists of narrow filters in the first milliseconds of the response, appropriate for ABR components, and that the width of the filters increases as the latency of the response increases, in an appropriate way to record MLR and CAEP components.

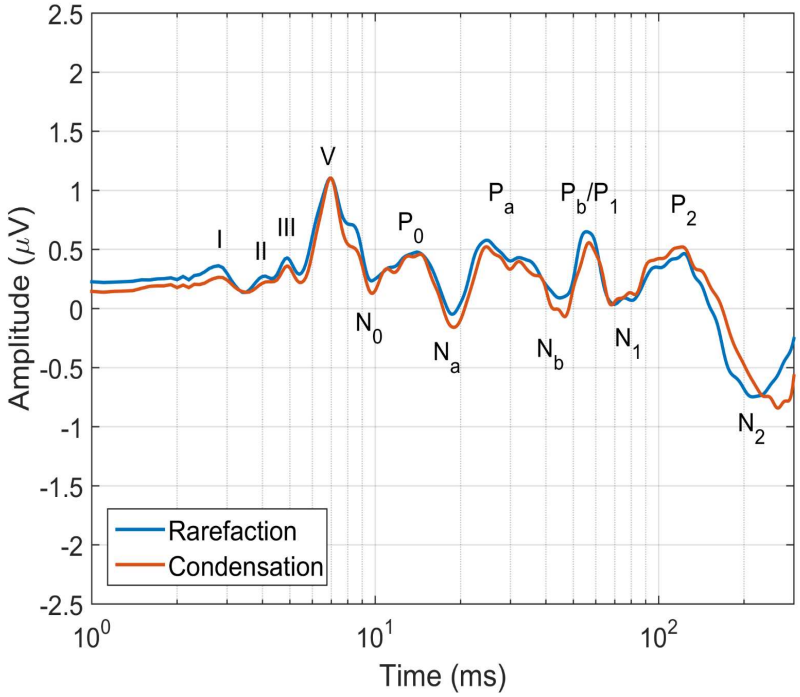
The base of functions is made of *sinc* functions uniformly distributed in the logarithmic time scale. This base of functions prevents the recording of high-frequency components in the later portions of the response.

Processing – Step 2. Latency-dependant filtering



This figure shows that projecting the signal of interest over this base of functions allows a filtering of the signal dependant on the latency (red [filtered] vs blue [original] lines in the figure), which allows an adequate representation of the activity of the main neural groups of the auditory pathway from the cochlea to the auditory cortex in the same figure.

Results – Average of 10 subjects

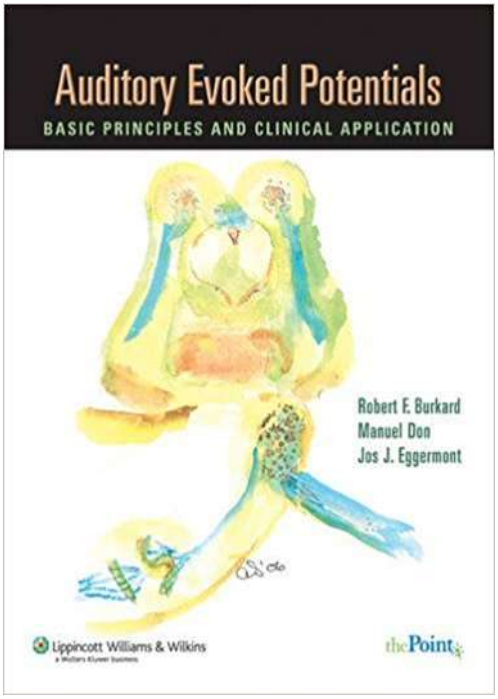


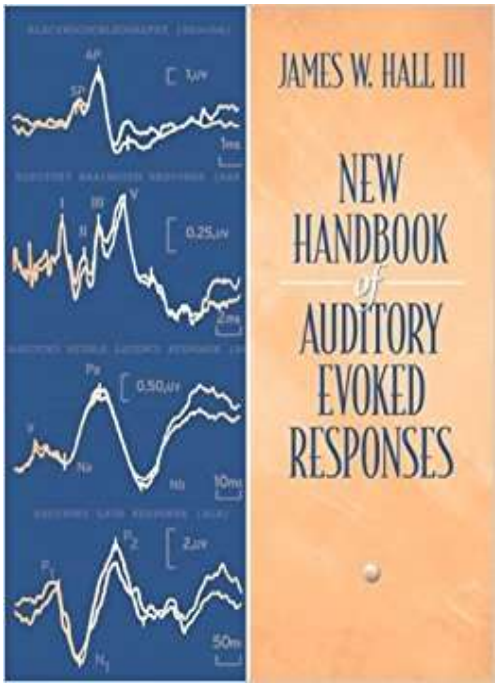
The figure in this slide shows the average of the signals obtained in 10 subjects (6 male, with an age range of [24-37] years) obtained in rarefaction (blue) and condensation (red) polarities.

This figure shows that the neural activity components in different stations of the auditory pathway, from the cochlea to the auditory cortex, are reproducible in the two stimulation polarities, which is an indication of neural response.

This emerging technique of representing the auditory evoked potentials may have potential in clinical and research applications in the future.

GOOD REFERENCES TO LEARN MORE…





I hope that you have enjoyed this lecture, and that you have acquired the essence of auditory evoked potentials, the different types that exist, the recording process, and the possibilities that they offer both in the clinic and research.

Deliberately, I leave you just a few bibliographic references, but of a very high quality. In these books you may find more information about the topics dealt in this lecture. In addition, some of the slides of this presentation include interesting references from which I obtained part of the material.

If you have any question, or should you have curiosity to know more about anything of this topic, please do not hesitate to contact me in any of my email addresses joaquin.valderrama@nal.gov.au, joaquin.valderrama@mq.edu.au or jvalderrama@ugr.es.

I look forward to hearing from you.

Best wishes,
Joaquin T. Valderrama