Extracting the Cochlear Microphonic from the FFR

By P. Deltenre[1], P.Connor[1], N. Wallaert[2], and A. Calcus[1]

[1] Université Libre de Bruxelles, Belgium
[2] Ecole Normale Supérieure, Paris & CHU R. Debré, Reims; France

Introduction :

The title of this presentation should have been changed into « Trying to sort out the diverse FFR components » since we concluded that, at least in normal subjects, it is impossible to isolate the Cochlear Microphonic (CM) evoked by a given stimulus without altering it. On the other hand, data gathered from various pathological cases indicated that multiple components are likely to contribute to the scalp-recorded FFR. That the FFR could reflect the contribution of several generators located from the cochlea to the upper brainstem is not a new idea and contributes to fuel a certain amount of skepticism sometimes expressed as "The FFR is full of artefacts !"

One of the messages of this communication is that we should not be over-pessimistic…

Why did we focus on a procedure of CM extraction in the first place ?

Obviously, most research or clinical applications consider the CM as a contaminant of the FFR signal recorded to study neural encoding. It would therefore be useful to be able to isolate the CM evoked by a given stimulus in order to subtract it from the contaminated waveform.
On the other hand, as we shall see, it may be useful to scrutinize the pre-neural response for itself.

Three main strategies have been reported in the literature:

1. Simultaneous recordings along a Vertical (V: Vertex to posterior neck) and a Horizontal (H: earlobe to earlobe) axis.
2. Addition and subtraction of waveforms evoked by stimuli of opposite acoustic polarity.
3. Masking of the neural components.

Although they do contribute to disentangle the FFR components, the first two methods have their shortcomings and the third one has not been fully assessed. Figure 1 illustrates the segregation of peripheral, short latency components from central later ones according to the orientation of the recording inter-electrodes axis.
It is well known that a Horizontal recording axis (i.a. earlobe to earlobe) displays neural responses with relatively short latencies compatible with a peripheral (Cochlear Nerve) origin whereas a vertical axis (Cz to nape of the neck) records more central components arising about 2 ms later for pure tone stimuli.

Spectral amplitudes profiles according to stimulus frequency show better representation of higher frequencies in the Horizontal channel, a difference consistent with the lower frequency limit for phase-locking measured in brainstem vs auditory nerve neurones. Figure 1 illustrates these differences.

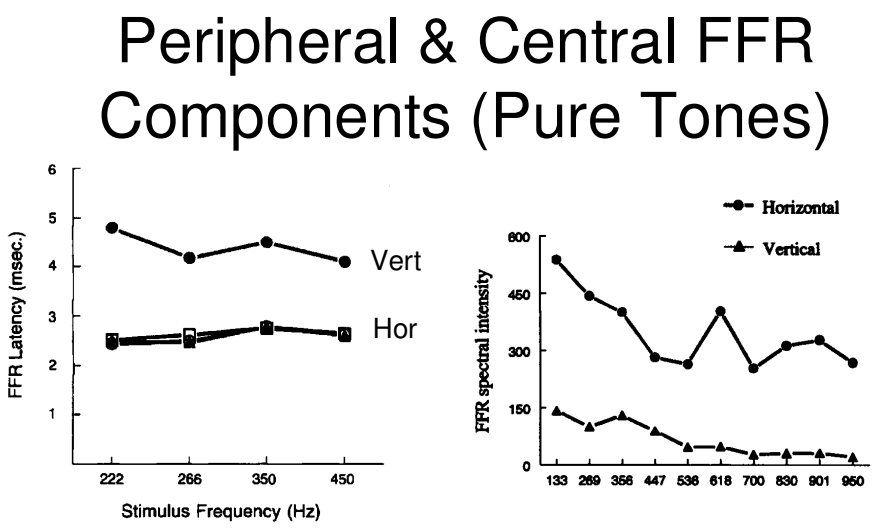


Fig. 1: *from Galbraith et al. 2000 & 2001.The average latency difference between V and H channel is about 2 ms.*

Masking of the neural response in order to isolate the CM and then subtract it from the original composite pre-neural and neural signal has been proposed as a way to record the pure neural FFR. The procedure is illustrated in Fig. 2.

Peripheral components: CM

Pure Tones

0 or 180° starting phases

Forward Masking of neural components

minus 30 dB SNR !!

CM perfectly inverted: ? =0

⇨ neural components completely masked

Subtract CM from raw FFR

!! ? raw 0 & 180° waveforms eliminates CM but distorts neural FFR (TFS)

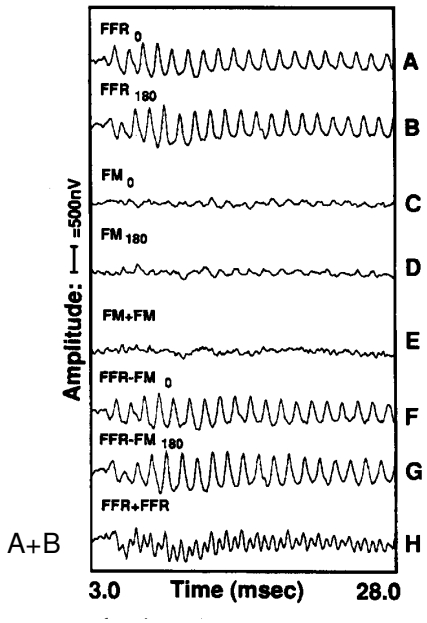


Fig. 2: *from Chimento & Schreiner 1990. A & B: FFR evoked by pure tones (800 Hz) of opposite acoustic polarities. C & D: residual CM after elimination of the neural components by a forward masker. E: Addition of the two CM waveforms shows an absence of any residual indicating complete masking of the neural components. F & G: pure neural FFRs obtained by subtracting the isolated CM. H: Simply adding the responses evoked by opposite polarity stimuli cancels the CM but distorts the FFR. Note that because of the forward masking procedure used, the authors recommend a Signal-to-Noise Ratio of -30 dB !!*

At moderate intensities at least, the CM is perfectly inverted when evoked by stimuli of opposite acoustic polarities. If we mask the neural components, we are left with the unmaskable pre-neural response that completely vanishes after addition of responses to opposite acoustic polarities. We can then subtract the isolated CM from the raw waveform to obtain the pure neural response to either one or the other stimulus polarity. The frequently used technique of adding the raw responses evoked by opposite polarities also removes the CM but distorts the neural response which is not insensitive to acoustic polarity: since it is evoked by the rarefaction phase, it is shifted by half a period.

This was for pure tone stimuli. Most current FFR studies use complex stimuli with the consequence that the number of FFR components will grow. We shall get phase-locking on both the Envelope (ENV) and Temporal Fine Structure (TFS) components of the stimulus. The ENV is much less – if at all – sensitive to acoustic polarity than TFS is.

Among the **peripheral** components we can expect the CM to include cochlear Distortion Products (DPs) which in humans are mainly the Cubic and Quadratic Distortion Tones (CDT & QDT). The ENV has been found to be 'poorly defined or non-existent' in horizontal channel recordings (Galbraith, 1994). The neural response of the Auditory nerve phase-locked on the TFS (the "auditory neurophonic") may contain the neural representation of the cochlear DPs.
The **central** components will comprise the ENV (also called the Envelope Following Response) and the TFS including the neural version of cochlear DPs.

As illustrated in Fig. 3, the cochlear DPs are by-products of the compressive non-linear amplification performed by the Outer Hair Cells (OHC). Their source has been located in the properties of the Mechano-Electrical Transduction channels (Avan et al., 2013). In humans they consist of the Cubic and Quadratic DTs.

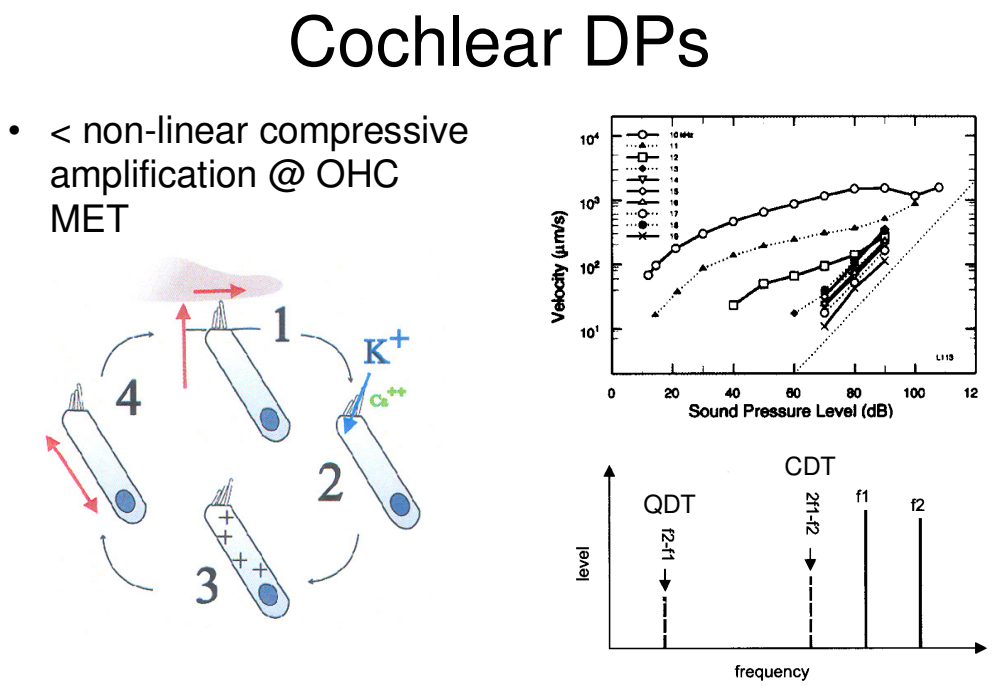


Fig. 3: *Mechano-electrical transduction followed by OHC contraction and non-linear (compressive) amplification of the vibration of the basilar membrane, with generation of CDT and QDT.*

Both DPs have been recorded in the FFRs of normal subjects. The current hypotheses about their significance is that FFR-CDT reflects neural phase-locking on cochlear DP whereas a mixed origin both cochlear and neural could cause the FFR-QDT. The next figure (from Bhagat & Champlin, 2004) illustrates the presence of Cochlear DPs in a Human FFR.

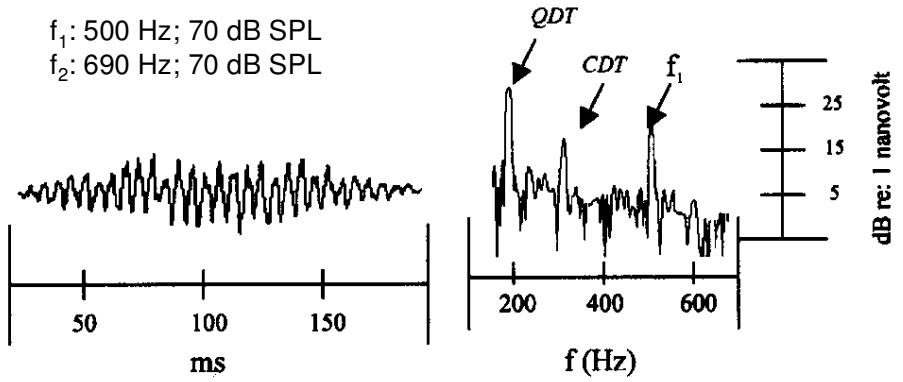


Fig. 4: *both QDT and CDT are present in the FFR recording performed in a normal-hearing adult.*

Subjects and Methods:

The results of our study have been obtained in children referred for electrophysiological evaluation of hearing under sedation. Normal hearing was defined by normal (< 20 dB) click-Auditory Brainstem Response (ABR) and Auditory Steady-State Responses (ASSR) thresholds (0.5 – 8 kHz), normal ABR Latency-Intensity function, Normal Brainstem Transmission Time (BTT) and normal DP-Oto-Acoustic Emissions (DP-OAEs).
Diverse cases of abnormal hearing like Auditory Neuropathy Spectrum Disorder (ANSD) contributed to delineate the FFR components. In most cases of ANSD, the CM is the sole recordable short-latency EP component.
Stimuli combined the third and fourth harmonic of a Missing Fundamental (MF) at 217 Hz. The overall stimulus level was 85 dB SPL delivered through an ER3-A tube-phone encased in a grounded µ-metal shield. Stimulus duration was 50 ms including 5 ms on-off $\cos^2$ ramps. Stimulation rate was 11.7 Hz. Recording was simultaneously performed on a vertical (C_z-Cv7) and horizontal ($A_{contra} - A_{ipsi}$) channel with a sampling rate of 32.768 kHz. Filter settings were band-pass 100 Hz-6 kHz, 3000 averages were collected for each

acoustic polarity (Rarefaction: R or Condensation: C, according to the polarity of the first half cycle of the stimulus).

Data were analysed after off-line addition and subtraction of the R- and C-evoked waveforms. Addition is known to enhance the ENV which is insensitive to acoustic polarity and to eliminate or markedly reduce the CM (Aiken & Picton, 2008). Subtraction will bias the higher frequency components and maximize the spectral (or TFS) response (Skoe & Kraus, 2010). Fast Fourier Transform (FFT) analysis was performed in order to evaluate frequency components and compute their Signal-to-Noise Ratio (SNR), each frequency bin being compared to the mean of the six flanking ones. Time Frequency Analysis (TFA) was performed using the Morlet wavelet transform on selected cases.

Results:

1. Normal Hearing patients

Figure 5 illustrates typical tracings obtained in a normal-hearing child referred for speech delay and unreliable behavioural testing.

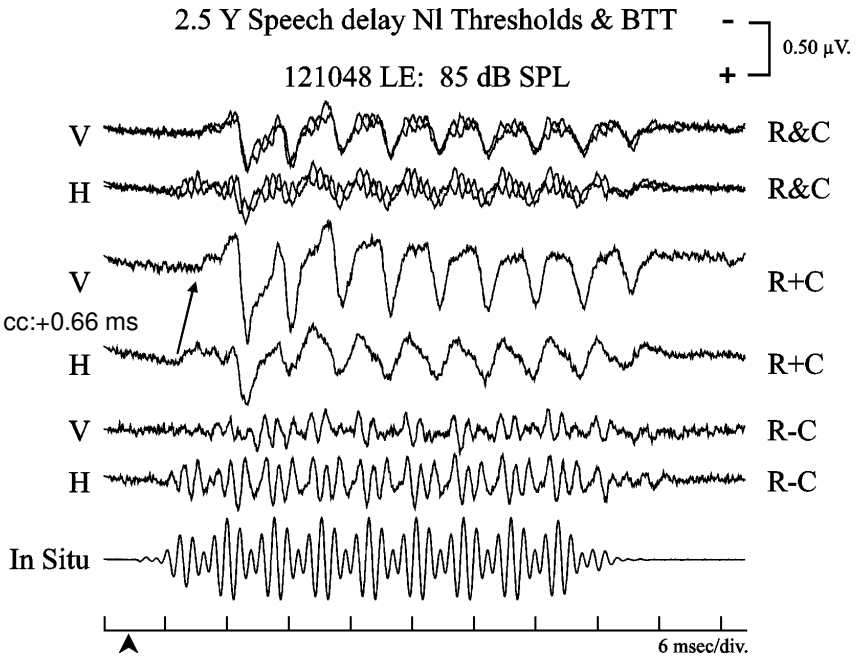


Fig. 5: *Representative FFR tracings obtained in a normal-hearing child. Addition of R- and C-evoked waveforms strongly enhances the ENV component whereas subtraction favours the TFS components. The bottom trace shows the stimulus as recorded in the external auditory canal.*

Quite typically, TFS shows higher amplitude and shorter latency in the horizontal channel whereas ENV has a larger amplitude in the vertical channel.

Figure 6 compares the relative spectral amplitudes of the four derived waveforms in the averaged spectra from 21 normal hearing subjects. The TFS shows significant spectral peaks at three frequencies: the primaries and the CDT. In the vertical channel, the TFS shows an additional peak at the frequency of QDT.
The ENV signal contains the MF/QDT, CDT and lower stimulus harmonic frequencies in both channels.

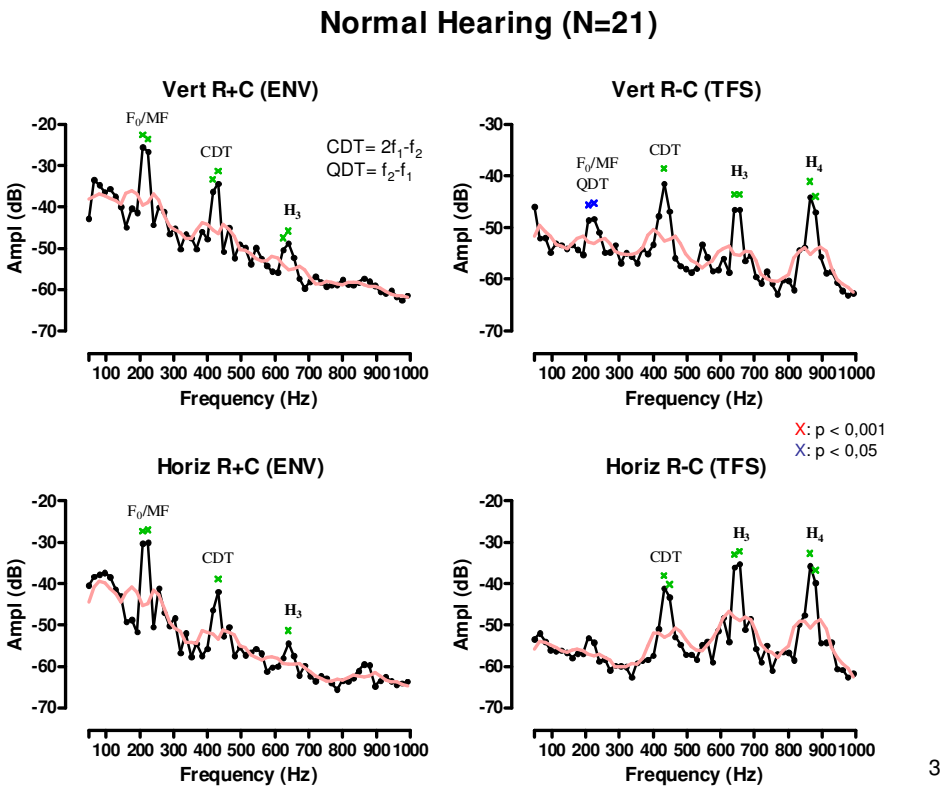


Fig. 6: *averaged spectra from 31 normal-hearing children. x = spectral bin with SNR significantly > 0 at p <0.001; x at p <0.05.*

Performing the same recordings in various instances of auditory deficits provides cues about the origins of these spectral peaks.

2. ANSD cases:

We had the opportunity to record FFRs in 6 cases of ANSD including two instances of cochlear nerve agenesis. The only recognizable activity in the representative ANSD subject illustrated in Fig.7 is a TFS pattern comprising the two stimulus primaries and the CDT, all inferred to be of pre-neural origin. Such cases led us to believe for some time

that the vertical channel was blind to pre-neural signals, but as further cases accumulated this proved to be not invariably true.

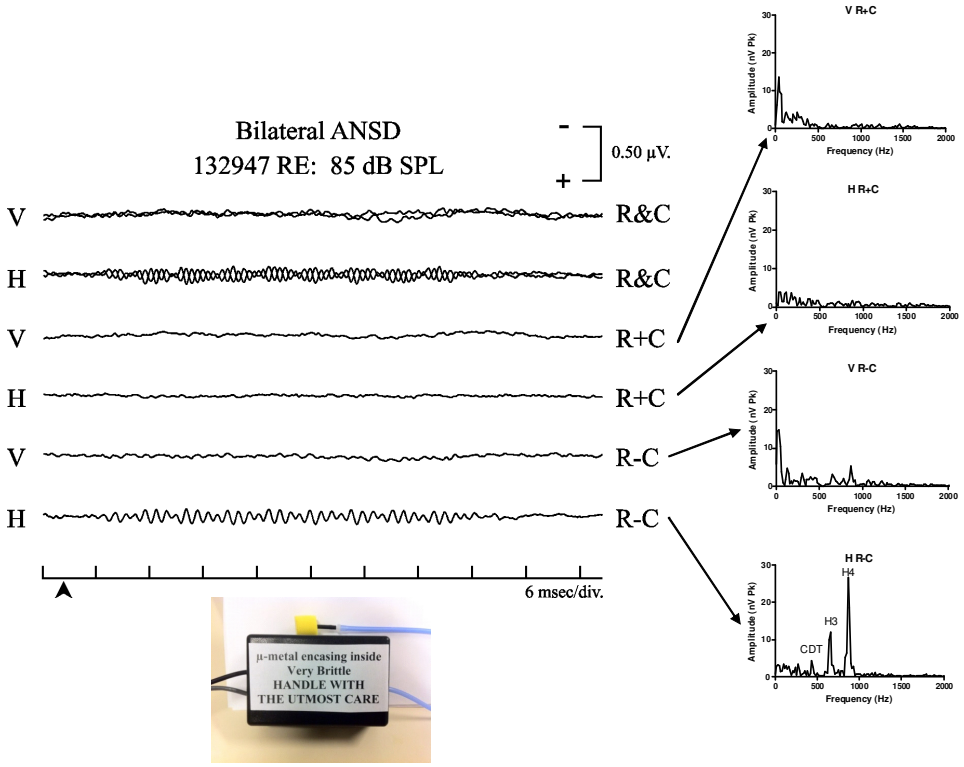


Fig. 7: *FFR recordings obtained in a representative case of ANSD. The only recognizable evoked signal is a TFS-like waveform in the horizontal channel. It contains the stimulus frequencies and their CDT. Note that clamping of the sound delivery tube was systematically applied for control measurements allowing to rule out artefactual components.*

Averaging the spectra of five ANSD patients revealed a pre-neural pattern comprising the two primaries and the CDT in the horizontal channel and a “leakage” of the higher stimulus frequency in the vertical TFS channel. There was no sign of the MSF/ENV frequency which seems therefore to be entirely dependent on neural processes.

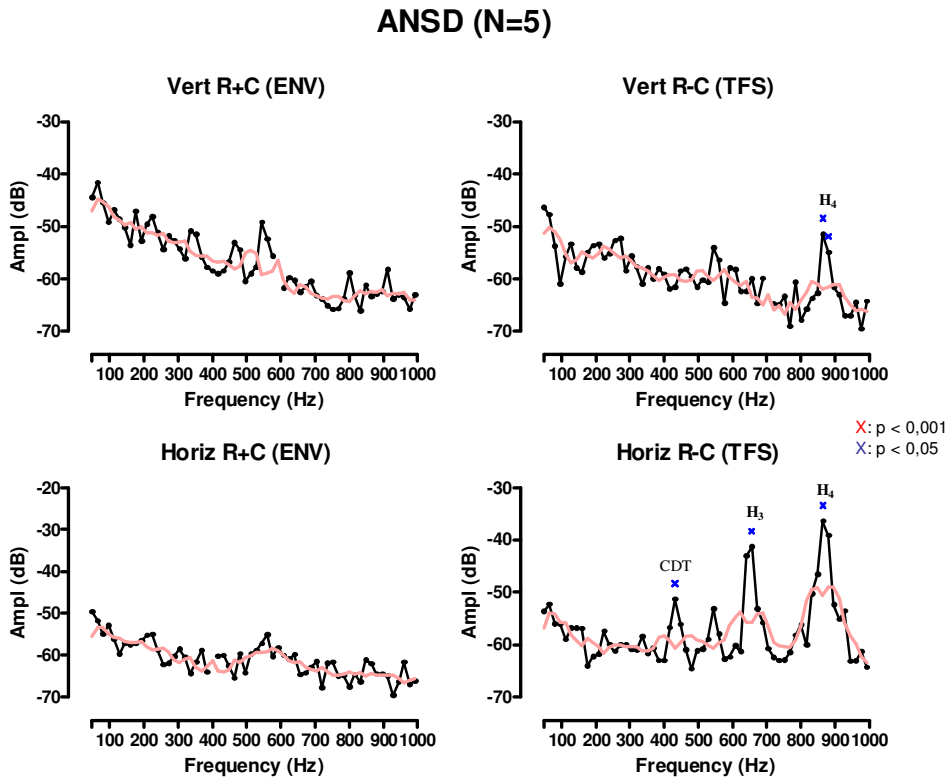


Fig. 8: *averaged spectra from five ANSD cases.*

3. Masking of the neural components in Normal Hearing Patients:

We then masked the neural responses of 21 normal hearing children in order to isolate their pre-neural components. The simultaneous masker was a Threshold Equalizing Noise (TEN) set 3 dB above the highest level needed to psycho-acoustically mask the FFR stimulus in 12 normal hearing young adults. The next figure schematizes the procedure used to define the masker level.

- Minimal TEN Level to mask FFR stimulus
- 12 nl young subjects
- 3-AFC procedure
- TEN Level set at highest value + 3 dB

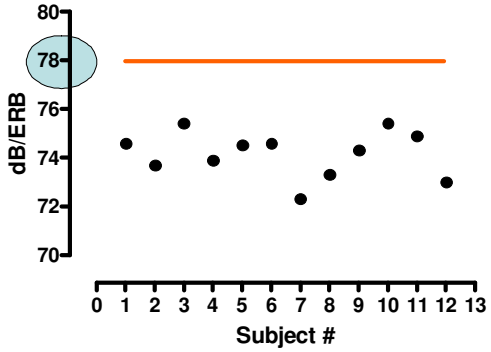


Fig. 9: P*rocedure used to define the masker level needed to mask perception of the FFR stimulus in normal hearing young listeners. It led to a TEN level of 78 dB/ERB.*

The comparison of unmasked and masked tracings in one representative subject suggests elimination of all neural components as illustrated in fig. 10.

NH Children masked by TEN

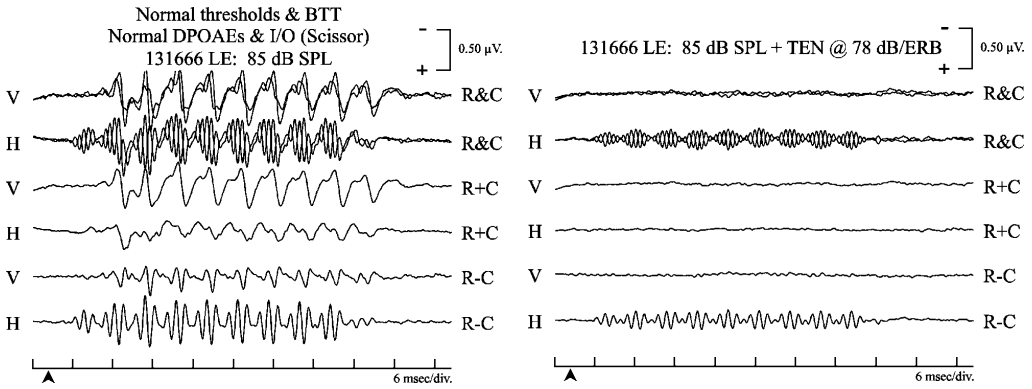


Fig. 10: *the masked waveforms obtained in normal-hearing children are quite similar to those obtained in ANSD.*

The average spectrum of the horizontal TFS confirms the pre-neural pattern observed in ANSD. However, in addition to what has been observed in ANSD, the lower stimulus frequency and the CDT are present in the TFS of the vertical channel and there is even a weak MF/QDT component in the horizontal ENV !

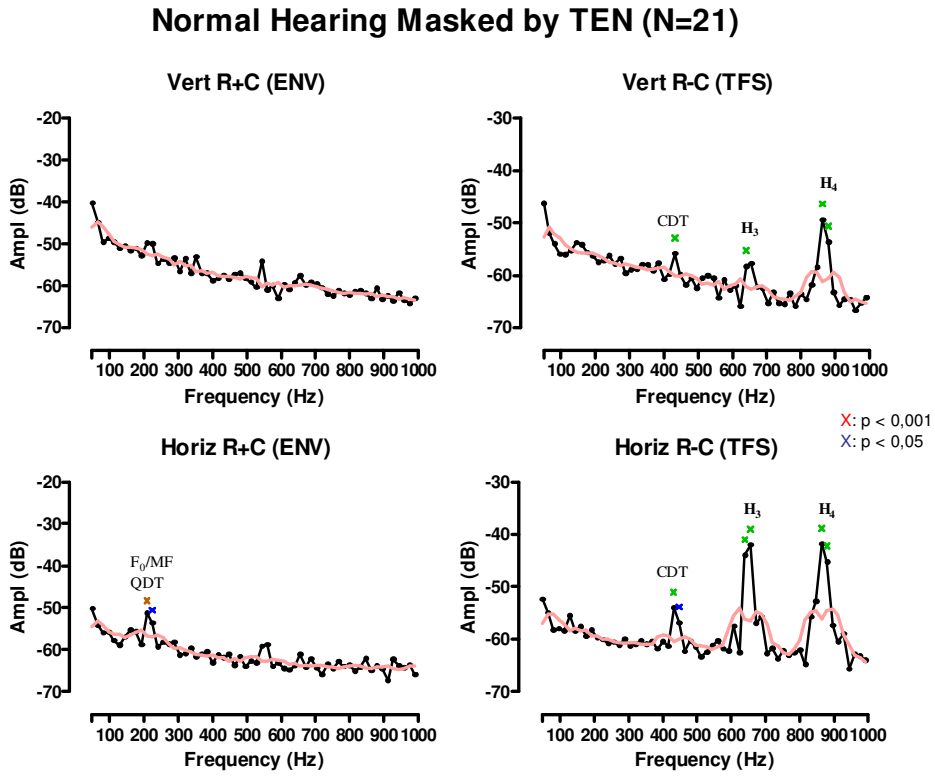


Fig. 11: *averaged spectra of the masked FFRs recorded from the 21 normal-hearing children.*

The origins of these additional peaks are currently not clear. The difference may be due to a lower residual noise for the averaged spectra of normal hearing subjects (N=21) than for ANSD cases (N=5) or to pathophysiological characteristics ANSD.

4. Masking of the isolated CM in ANSD cases:

At that stage the next logical step would have been to subtract the isolated CM from the original waveforms in order to isolate the pure neural FFR. However, the notion that masking does not affect the CM comes from studies with forward maskers (Chimento & Schreiner, 1990) whereas the CM is known to behave non-linearly and to exhibit two-tone suppression or suppressive masking. Masking cannot eliminate the CM (Aiken & Picton, 2008) but, at least when applied in simultaneous mode, can alter it. In order to verify such an effect, we compared masked and unmasked CMs in four ANSD cases (See fig.12).

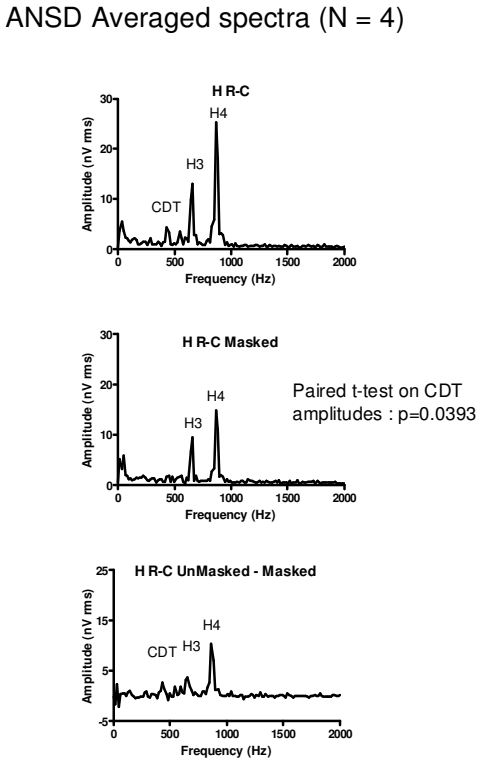


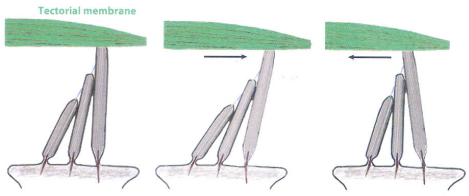
Fig. 12: *averaged spectra of the unmasked vs masked CM and of the residual waveform after subtraction of masked from unmasked tracings.*

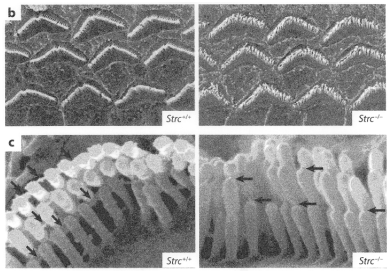
Simultaneous masking clearly reduces the spectral amplitude of the CM even in ANSD cases who are known for their absence of stapedial reflex and of efferent suppression. Therefore we interpret this effect of masking as suppressive masking of the CM.

4. A case of suspected Stereocilin defect:

Another type of Hearing-impaired subjects is likely to improve our knowledge of the FFR components.
Stereocilin is a protein responsible for the mechanical cohesion of OHC stereocilia. It is estimated that about 5% of non syndromic mild to moderate recessive hearing loss is due to mutations of the gene encoding stereocilin.

Stereocilin





RESEARCH ARTICLE

AMERICAN JOURNAL OF medical genetics

Genome-Wide SNP Genotyping Identifies the *Stereocilin* (*STRC*) Gene as a Major Contributor to Pediatric Bilateral Sensorineural Hearing Impairment

Lauren J. Francey,[1] Laura K. Conlin,[2] Hanna E. Kadesch,[1] Dinah Clark,[1] Donna Berrodin,[1] Yi Sun,[1] Joe Glessner,[3] Hakon Hakonarson,[1,3] Chaim Jalas,[4] Chaim Landau,[4] Nancy B. Spinner,[2]

Strc mutations are may be the 2nd more frequent cause of recessive NSHL:

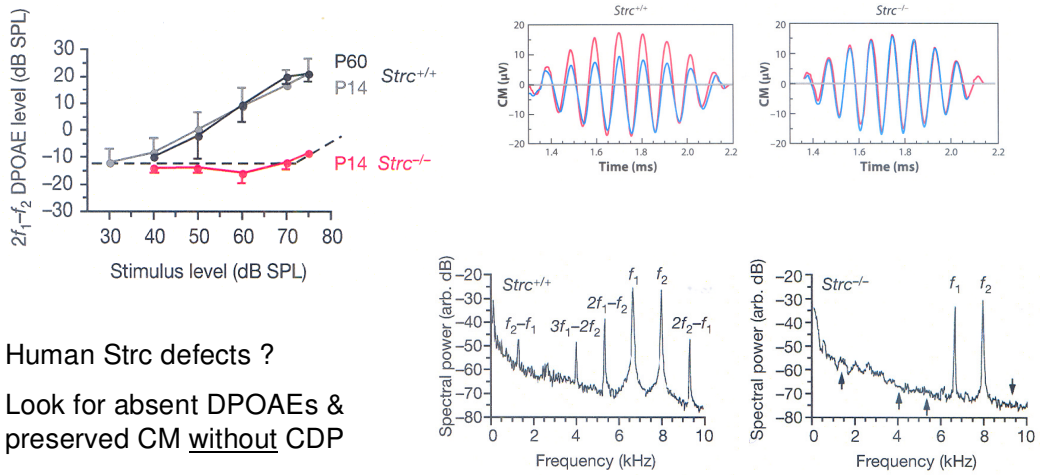
2.5 % (17/669) of NSHL

5.5% among mild-moderate HL

Fig. 13: *pictures from Avan et al. Drawing of the OHC stereocilia and their top connectors (elongated black patch between them). Photomicrograph of the stereocilia. Left hand side: normal mouse: the top connectors are indicated by the vertical arrows. Right hand side: mutant mouse with defective stereocilin: the cilia are misaligned and the top connectors are missing; the horizontal arrows indicate the tip links.*

Mutation of the gene encoding stereocilin could be a frequent cause of Human non-syndromic mild to moderate hearing loss (Francey et al., 2012). A murine model of the mutation has recently shown that without top connectors the CM looses its characteristic non-linearities as illustrated in fig. 14. In mutant mice lacking stereocilin, DP-OAEs are abolished and two-tone suppression and DP disappear from CM recordings.

Strc -/- : DPOEA abolished
CM distortions vanished



Human Strc defects ?

Look for absent DPOAEs & preserved CM without CDP

Fig. 14: *pictures from Verpy et al., 2008. Left: mutants ($Str^{-/-}$) lack DP-OAEs at post natal ages at which wild animals ($Str^{+/+}$) produce strong emissions. Right upper panels: Round-window recorded CM evoked by two pure tones creating a beat. Stimulus in red CM in blue. In the wild animal, the CM does not follow exactly the time course of the beat: its amplitude is reduced during the maximal portion of the beat. In the mutant, this two-tone interaction has gone, the CM faithfully follow the stimulus time course. Lower panels: CM spectra show that all DPs present in the wild animal are absent in the mutant.*

We recently encountered a case with an electrophysiological profile compatible with a stereocilin defect. This three-year-old girl was referred for electrophysiological testing because of a speech delay and unreliable behavioural audiometric data. She had no medical history or risk factor for hearing loss and had not benefited from neonatal hearing screening. Her electrophysiological results are illustrated in fig. 15.

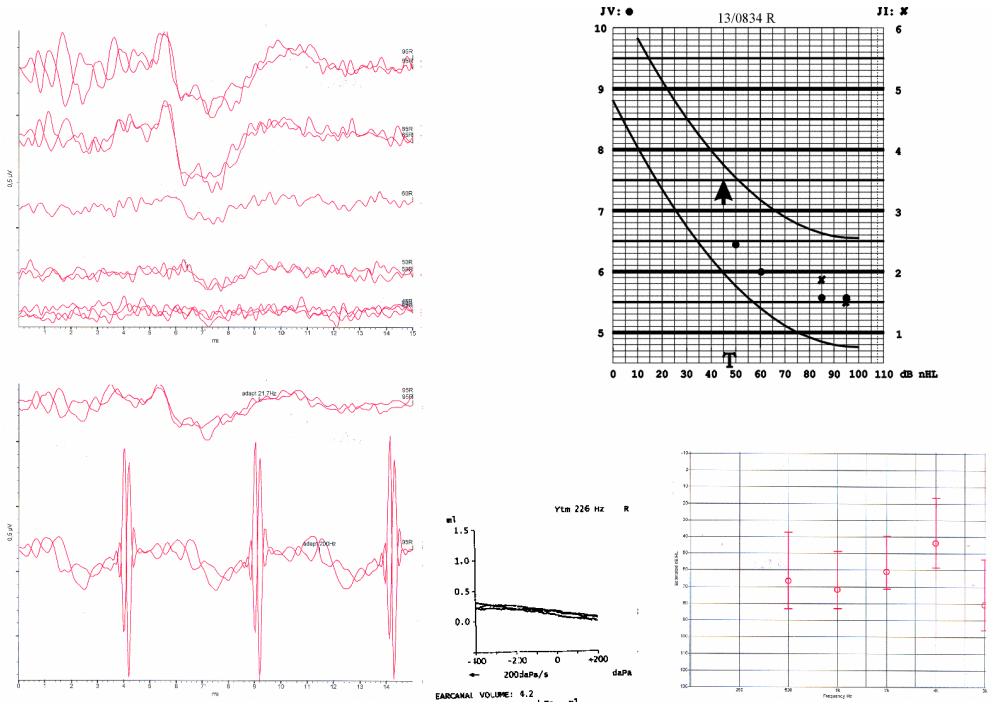


Fig. 15: *the click-evoked ABRs show a moderate (45-50 dB nHL) elevation of wave V threshold, with normal latencies for the click levels, so that the latency at threshold is abnormally short. This indicates a loss of the mechanical intra-cochlear amplification subserved by OHCs. DP-OAE are absent but this could be due to the flat tympanogram. AASRs show a moderate loss entirely coherent with the click threshold. What appears paradoxical is the presence of a clear CM of long duration suggesting normal OHC whereas cochlear amplification is defective. The pre-neural nature of the CM is attested by its absence of adaptation when the stimulation rate is raised from 21.7 to 200 Hz.*

The TFS in the horizontal channel contains the stimulus primaries only. But it is quite surprising to record an undeniable CDT from the vertical channel only !! And this CDT appears to belong to a signal with a long latency indicating a central origin.

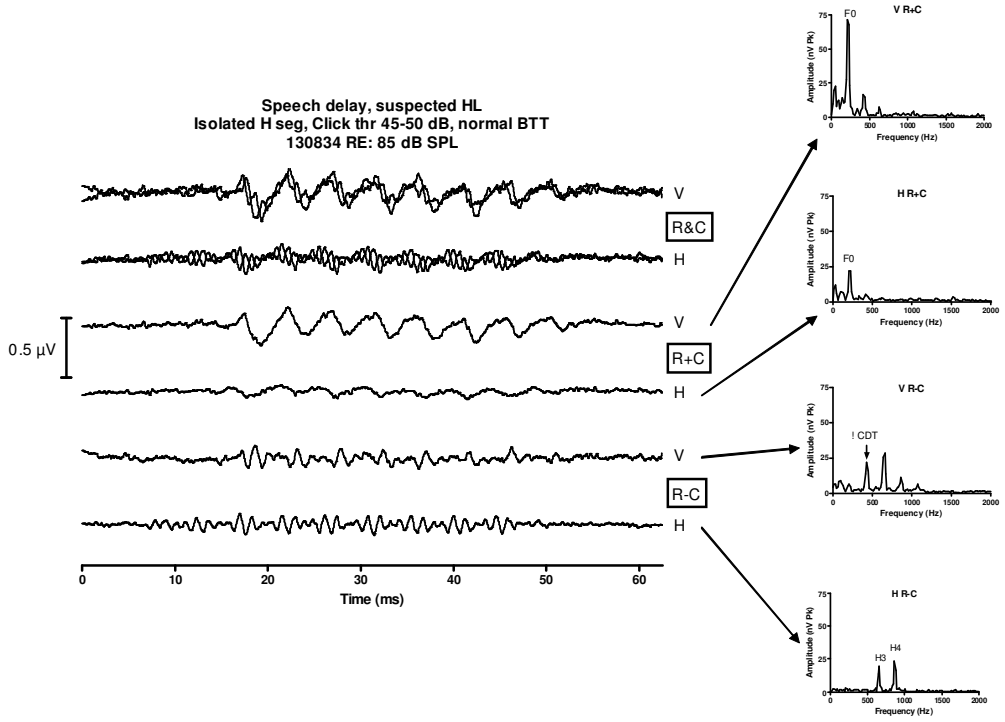


Fig. 16: *FFR waveforms and spectra showing absent CDT in the peripheral TFS signal in a case of possible stereocilin defect.*

Increasing the stimulus level by 10 dB (Fig. 17) did not reintroduce the missing CDP in the peripheral channel.

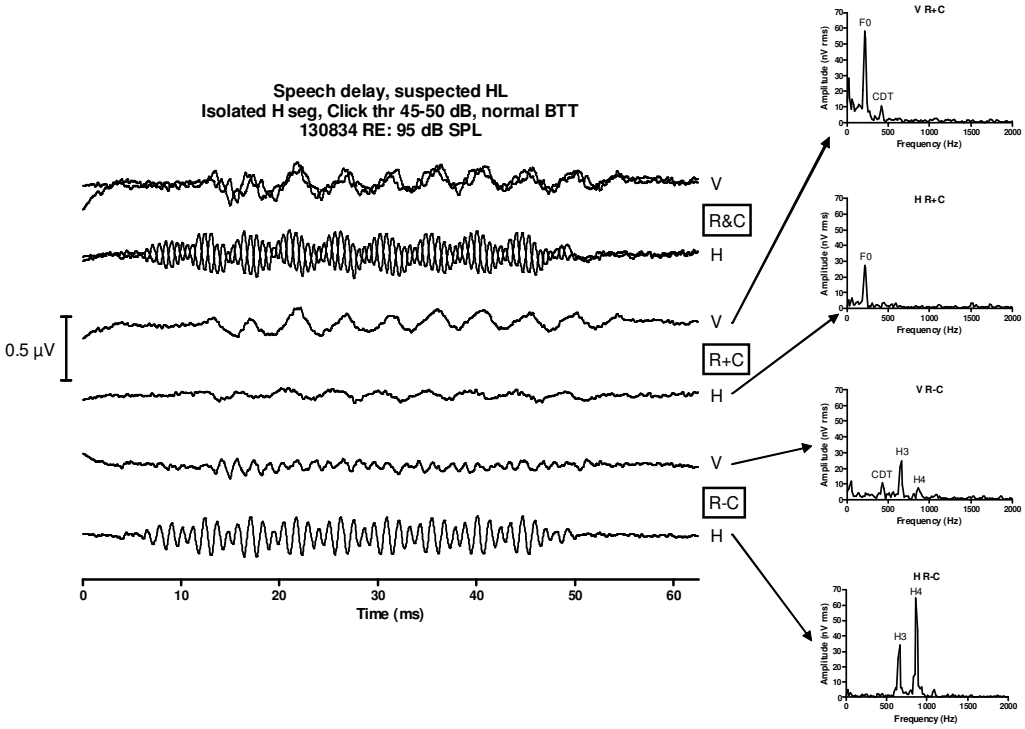


Fig. 17: *FFR waveforms and spectra showing absent CDT even for an 95 dB SPL stimulus.*

We are currently awaiting the results of genetic testing for this case whose profile is compatible with a stereocilin defect.

Since the CDT observed in the vertical channel was associated with a temporal waveform of late onset, we performed a wavelet-based TFA on the TFS signals recorded in both V and H channels in order to be confirm the latency at which the CDT appeared (Fig. 18).

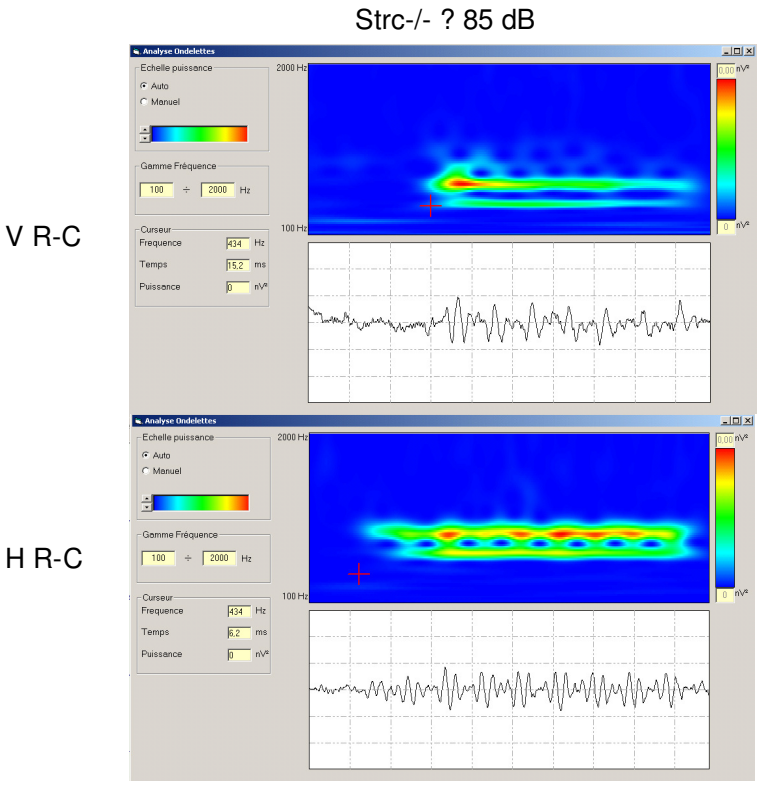


Fig. 18: *Peripheral (H) and central (V) TFS waveforms evoked by the 85 dB SPL stimulus and their TFA. In the peripheral channel, there is no energy at the frequency of the expected CDT (cross). In the central channel, the CDT frequency does not appear earlier than the visually identified temporal signal.*

Comparing (Fig. 19) the scalograms of this putative Strc $^{-/-}$ case to an ANSD one shows an obvious difference in the latency onsets of their CDP (See fig.19).

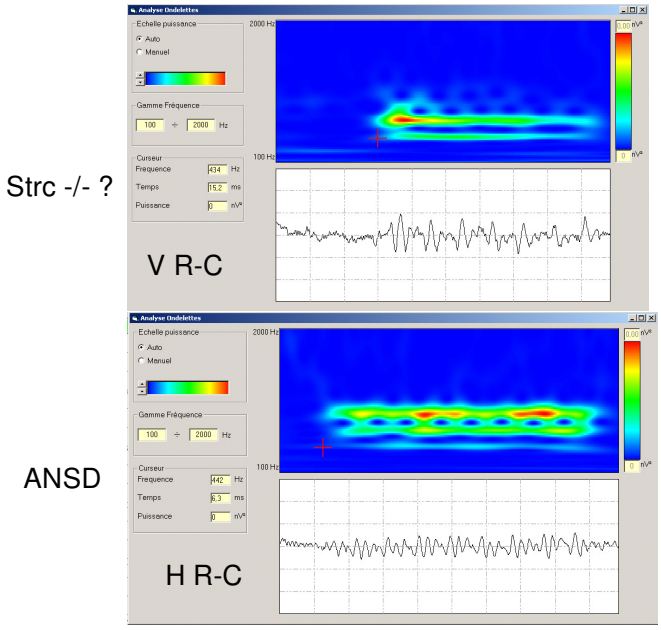


Fig. 19: *compared TFA of the CDT present in the putative Strc $^{-/-}$ case and one ANSD case.*

6. Cases of proximal auditory nerve lesions:

The next case who brought interesting information about the FFR generators was one of severe central hypomyelination (Pelizaeus-Merzbacher-like disease due to a GJA12 mutation). These subjects harbour a severely abnormal central (oligodendrocytes-dependent) myelin whereas their peripheral (Schwann cells-dependent) myelin is normal. This means that the distal, extra-meningeal portion of the cochlear nerve is normally myelinated whereas severe desynchronization occurs as soon as the nerve crosses the meningeal envelope. Fig. 20 compares the myelin MRI signals between an age-matched control and the affected child.

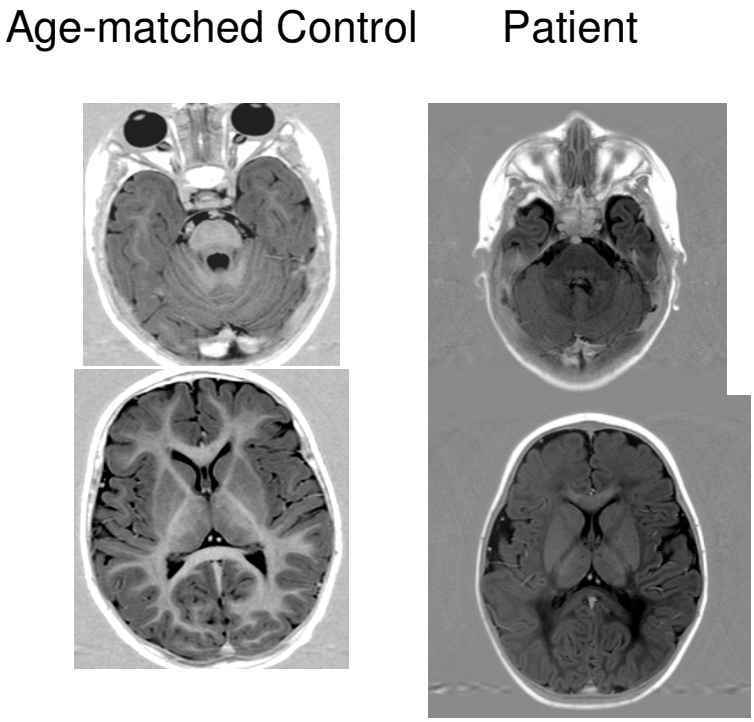


Fig. 20: *T1-weighted MRI scans showing severe hypomyelination both at brainstem and hemispheric levels.*

The next figure demonstrates normal cochlear function in this child as DP-OAEs, click-evoked ABR wave I threshold and latency-intensity function are normal.

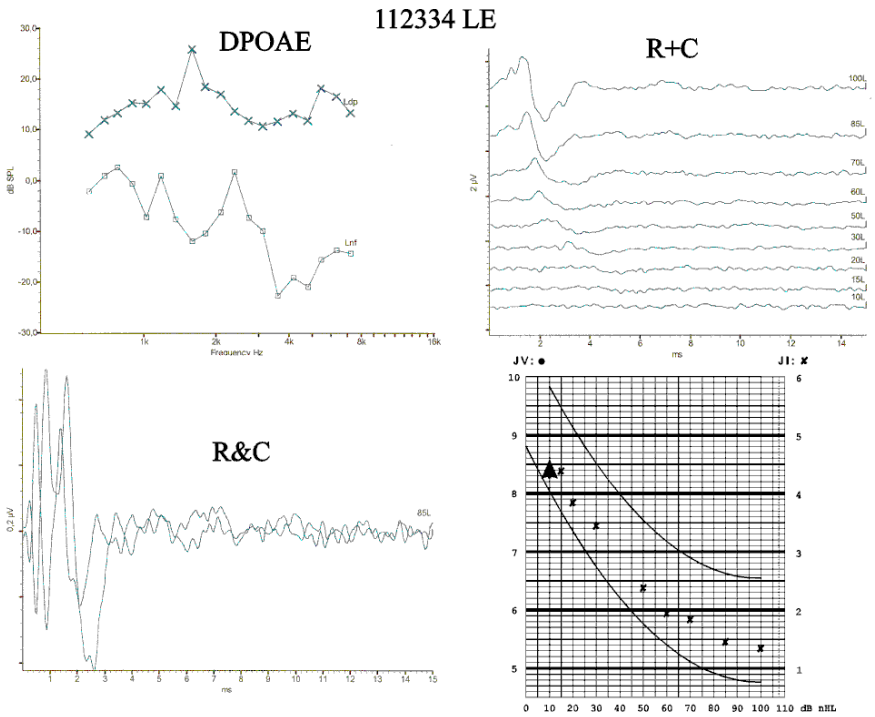


Fig. 21: *Results of physiological testing from the left ear in a case of Pelizaeus-Merzbacher-like disease. OAEs, click-evoked CM and ABR wave I are normal, but more central ABR waves are absent or very weak. The FFR recordings show prominent activity in the peripheral (H) channel and very weak components in the central (V) channel. The peripheral channel displays an undeniable ENV component.*

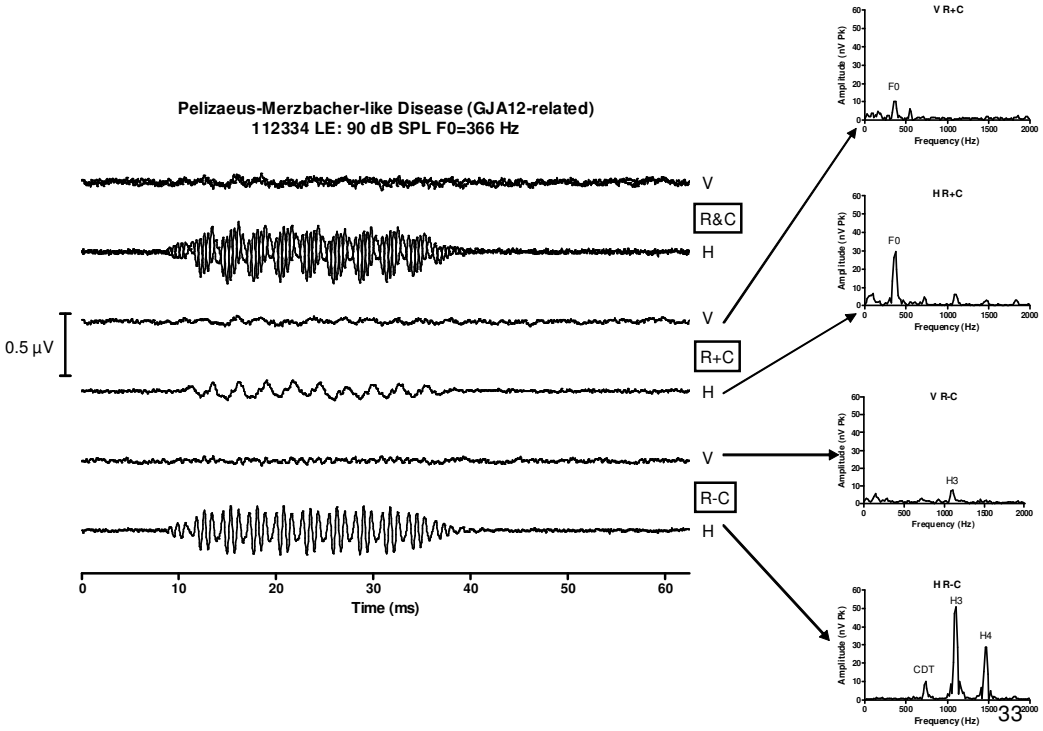


Fig. 22: *FFR components obtained from the left ear in a case of Pelizaeus-Merzbacher-like disease. Peripheral channels show a normal profile of components including an ENV whereas central channels contain very weak signals.*

The last case is one of severe tuberculous meningitis with brainstem compression and cochlear nerve entrapment at meningeal envelope crossing. Here also the DP-OAEs are normal, but as illustrated by the next figure, ABR wave I shows a moderately elevated threshold and amplitude loss.

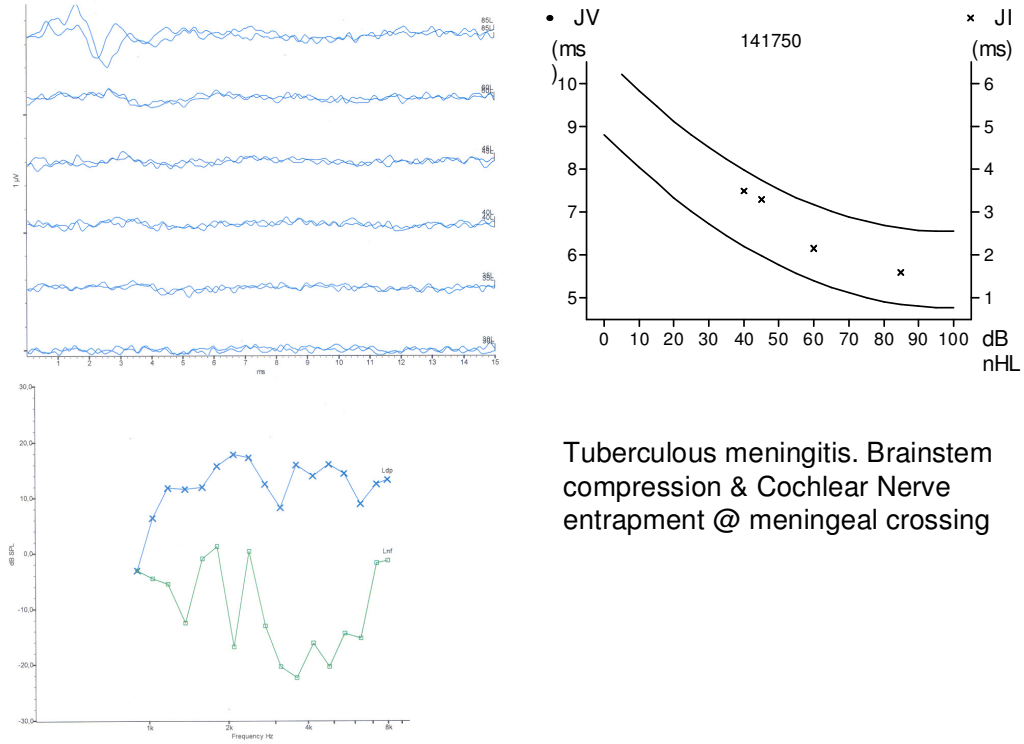


Fig. 23: *Results of physiological testing from the left ear in a case of cochlear nerve entrapment at meningeal envelope crossing.*

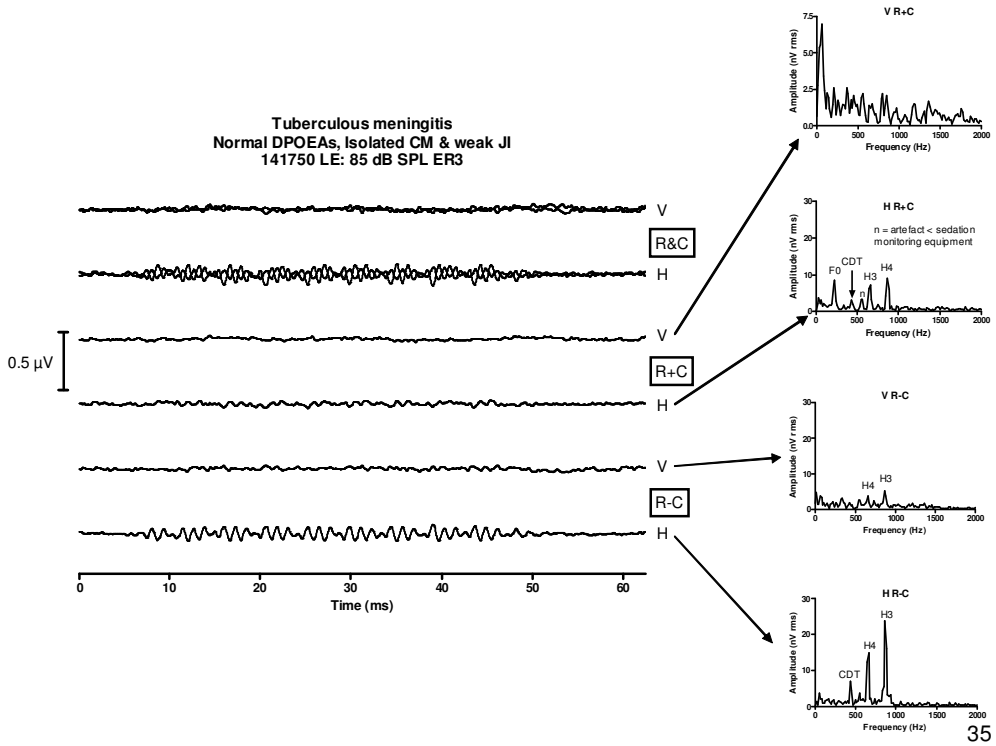


Fig. 24: *FFR components obtained from the left ear in a case of cochlear nerve entrapment. The peripheral TFS profile is normal, but the ENV component which is present in the peripheral channel only is very weak.*

Discussion:

Taken together, the above results show that the peripheral (H) TFS (R-C) channel contains a robust frequency pattern associating the two stimulus frequencies and their CDT. This pattern is recorded in Normal-Hearing subjects and ANSD cases and resist to masking of the neural components in the former subjects. This indicates a cochlear pre-neural origin possibly enriched by phase-locked neural components when neural responses are preserved. The ENV (R+C) waveforms contain the MF, the CDT and the lower stimulus frequency. They are mostly of neural origin since they are absent in ANSD and after masking (except for a weak MF component in the peripheral channel that could also be a QDT).

The idea of isolating the CM by masking in order to subtract it from the original waveforms to obtain the pure neural FFR seems unrealistic since simultaneous maskers induce suppressive masking that alters the CM. Forward masking might avoid suppressive masking, but its use in the context of Human FFR recordings is severely limited by the massively negative (-30 dB) Signal-to-Noise Ratio required to maintain the adequate level of masking over the entire stimulus duration.

Patients' data indicate the presence of an ENV component in the peripheral portion of the cochlear nerve and strongly suggest the late generation of a component at the CDT frequency by central neurones.

Although simultaneous masking alters the CM, it remains useful to reveal the spectral profile of the pre-neural components. This profile can be modified in some specific cochlear pathologies like the stereocilin defect.

Time Frequency Analysis appears to be a promising method to disentangle the various frequency components and their origins from cochlea to upper brainstem.

The systematic investigation of well-documented patients series (including those with a genetically proven molecular mechanism) should contribute to improve our understanding of the FFR content which is undeniably complex.

This complexity should not make us over-pessimistic: there is gold to be mined in the FFR data, but a lot of excavating work remains to be done.

References:

Aiken & Picton Hear Res 2008;245:35-47.
Avan et al. Physiol Rev;2013:1563-1619.
Bhagat & Champlin Hear Res ;2004 :193 :51-67
Chimento & Schreiner EEG Clin Neurophysiol 1990;75:88-96.
Francey et al. Am J Med Genet 2012 Part A 158A:298-308.
Galbraith EEG Clin Neurophysiol;1994:92:321-330.
Galbraith et al. Neurosci Lett 2000;292:123-127.
Galbraith et al. Percept Mot Skills 2001;92:99-106.
Skoe & Kraus Ear Hear 2010;3 :302-324.
Verpy et al. Nature 2008 ;456 :255-258.