A portable neurofeedback device for treating chronic subjective tinnitus: Feasibility and results of a pilot study

Robin Guillard, Marie-Josée Fraysse, Renaud Simeon, Thomas Cervoni, Jonathan Schmutz, Bastien Piedfort, Victor Ferat, Marco Congedo, Alain Londero

A portable neurofeedback device for treating chronic subjective tinnitus: feasibility and results of a pilot study.

Robin GUILLARD*[a], Marie-Josée FRAYSSE[b], Renaud SIMEON[c], Thomas CERVONI[a], Jonathan SCHMUTZ[a], Bastien PIEDFORT[a], Victor FERAT[a], Marco CONGEDO[d], Alain LONDERO[e]

(a) Zeta Technologies, 11 rue Carnot, Le Kremlin-Bicêtre, France

* Corresponding author, robin.guillard@gmail.com

(b) Service d'Otologie et d'Otoneurologie, Hopital Purpan, Place du Docteur Baylac, Toulouse, France

(c) Service ORL APHP Rothschild, 5 rue Santerre, 75012 PARIS

(d) Univ. Grenoble Alpes, CNRS, Grenoble INP, GIPSA-lab, Grenoble, France

(e) Alain LONDERO (MD). Hôpital Européen G. Pompidou, Service ORL et CCF, 20 rue Leblanc, Paris (France)

Abstract:

Background: Several clinical studies have shown that neurofeedback (NFB) has the potential to significantly improve the quality of life of patients complaining of chronic subjective tinnitus. Yet the clinical applicability of such a therapeutic approach in the everyday practice has not been tested so far.

Objective: This study aims at investigating the feasibility and efficacy of a semi-automated NFB intervention by means of a portable device that eventually could be used by the patients at home on an everyday basis. The duration of setup procedures is minimized through the use of a dry-electrodes ElectroEncephaloGraphy (EEG) headset and an automated user-interface.

Methods: We conducted a pilot clinical study (non-controlled, single arm, NCT03773926). According to a predetermined power calculation, a homogeneous population of 33 subjects with strict inclusion criteria was enrolled. After inclusion, all patients underwent 10 NFB sessions lasting 50 minutes each, over a period of 5 weeks and a 3-month follow-up period. According to previous studies, the NFB training aimed at increasing the alpha-band power (8-12 Hz) in the EEG power spectrum on the averaged signal of leads FC1, FC2, F3 and F4. Tinnitus Handicap Inventory (THI) was used as a primary outcome measure. Secondary outcome measures were the visual analog scales (VAS) and the change of the alpha-band power within sessions and across training. Time points of assessment were before intervention (T1), after intervention (T2) and at the 3-month follow-up (T3).

Results: Patient exhibited a clinically significant decrease of the THI score, with a 23% decrease (N=28) on average between T1 and T2 and a 31% decrease (N=25) between T1 and T3. A significant increase of the alpha-band power within sessions was observed. No significant increase of the alpha-band power across sessions was observed.

For the 19 subjects where sufficient data were exploitable, a significant correlation was found between the evolution of the alpha-band training across sessions and the evolution of the THI between T1 and T2.

The sessions were well tolerated and no adverse effect was reported.

Conclusion: This study suggests that neurofeedback has potential to suit everyday clinical practice with the goal to significantly reduce tinnitus intrusiveness. The merits and limitations of this NFB procedure are discussed, especially with respect to the choice of EEG electrodes to ensure a good signal quality.

Conditions : Tinnitus, hearing loss

Key-words : *neurofeedback; EEG; tinnitus; hearing loss; cognitive strategies;*

Introduction:

Subjective tinnitus (ST) is a phantom auditory percept that cannot be linked to any objective physical sound source (Baguley and al, 2013). ST is a frequent and widespread symptom experienced by approximately 10 % of the world population (Mc Cormack and al 2016). ST intrusiveness varies among patients but, for some individuals, it may lead to severe impairment of quality of life (Langguth and al 2011). ST perception is assumed to be the consequence of multiple and intertwined dysfunctions of the peripheral and central auditory systems, which have been only partially elucidated to date (Langguth and al 2013). Indeed, the neural substrate of ST remains elusive even if many pathophysiological theoretical models have been proposed. For instance, it has been proposed that tinnitus emerges because of a failure of the noise cancelling system of limbic regions on the auditory central system when the auditory pathways are damaged (Rauschecker and al 2010). Conversely, the model of adaptation of central gain proposed by Norena (2011) assumes that tinnitus may emerge from an increased neural activity in the central auditory system because of a reduced input coming from the auditory nerve. Similarly, the thalamocortical dysrhythmia model firstly described by Llinas and al 1999 and then substantiated by De Ridder and al in 2015 also assumes that the ST percept may be linked to a sustained maladaptive spontaneous brain activity triggered by a peripheral hearing loss. According to this model the resting-state α band (8-12 Hz) activity shifts to a lower-frequency θ band (4-8 Hz) activity coupled with a cortical high-frequency ɣ (35.5-45 Hz) oscillations. This hypothesis has been later corroborated by Zobay and al (2015) and Vanneste and al (2018). This model has also found support in a magnetoencephalographic (MEG) study that showed a significantly increased power of temporal cortices in the δ and ɣ bands and, conversely, a decreased power in the α band (Weisz 2005), however this abnormal pattern has not been confirmed by other studies (Adjamian and al 2009, Pierzycki and al 2016). More recent data suggest that the abnormal coupling in the auditory cortices may be specifically linked to ST, but not to other neurologic analogous conditions such as pain, where other cortical regions are

involved (Vanneste 2018a). This rhythmical imbalance in the temporal cortices constitutes a therapeutic target and a variety of neuromodulation techniques have been tested aiming at restoring a normal pattern of cortical activity (Vanneste and al 2012), including neurofeedback (NFB).

NFB can be defined as a therapeutic intervention where electrophysiological brain activity is noninvasively recorded and then analyzed in real-time in order to extract features from its measurement, e.g., frequency band power (Hammond 2011). These features are displayed to the subject in order to enable an interaction. A reward (visual, auditory, and/or tactile or a combination of them) is given according to the increase of the target activity. NFB is assumed to be a learning/conditioning process by which patients are driven to modulate their brain activity towards a predefined pattern (Sterman and Egner 2006). NFB has been proposed since the late 60s (Wyrwicka and Sterman 1968, Sterman and Friar 1972) and is currently a validated therapeutic option for attention deficit hyperactivity disorder (Arns and al 2009, Gevensleben and al 2009). The use of NFB for ST is more recent and has repeatedly demonstrated the capability of reducing ST related intrusiveness (Dohrmann and al 2007, Crocetti and al 2011, Hartmann and al 2013, Vanneste and al 2018b, Guntensperger and al 2019). The main aim of the studies from Dohrmann, Crocetti and Güntensperger was both to improve the α band power and decrease the δ band power within the auditory cortices, which Lorenz (2007) supposed to be best monitored at electrodes positions F3, F4, Fc1 and Fc2 in the 10-10 electrode position system (Jurcak and al, 2006). The choice of this α/δ metric is mainly supported by MEG data showing that both the decrease of α power and the increase of δ correlated with ST percept (Weisz et al 2005). The α/δ ratio can be optimized by an increase of the α band, a decrease in the δ band, or both. Nonetheless, preliminary analysis on data collected in a recent study (Güntensperger and al 2019) suggested that the clinical outcome mainly correlates with the α band change. Therefore, we have chosen to study a NFB protocol aimed at increasing the α band power activity extracted from these EEG electrodes.

In these studies, a very large inter-individual variability in terms of response to NFB protocols has been observed. This variability can be explained by several factors: firstly, by the recruitment of inhomogeneous samples mixing different forms of ST (both in terms of ST psychoacoustics, hearing pattern and related distress), secondly by the limited guidance provided to the patients in order to optimize the cognitive strategies they have to use for reliability modifying their brain activity. This absence of guidance could have resulted in lower motivation, whereas the active participation of the subjects is key for the success of this method (Congedo, 2004). Moreover, NFB has not been tested in ST patients with devices and protocols that can be used in a day-to-day practice outside the research settings. Consequently, we hypothesize that testing on a highly selected homogeneous subgroup of ST patients an α band power increase guided training on electrodes positions F3, F4, Fc1 and Fc2, would result in an overall improved efficacy for patients' quality of life.

The aim of this exploratory study (multicentric, single arm, non-controlled) was to test, in a highly selected homogeneous subgroup of ST patients, the feasibility and efficacy of a customized semi-automated NFB intervention by means of a portable device displaying a user-friendly automated interface that eventually could be used on an everyday basis. If proven feasible, NFB for treating tinnitus would be more time-effective and could consequently be adopted more widely by clinicians.

Material and Methods

Trial design:

Reporting follows the CONSORT guidelines (Schultz, 2010). Patients were recruited in three tertiary-care French ENT departments in Toulouse (Purpan Hospital) and Paris (HEGP and Rothschild Hospitals). Unlike in previous NFB studies (Dohrmann 2007, Crocetti and al 2011, Hartmann et al 2013, Vanneste and al 2018b, Guntensperger and al 2019), an homogeneous sample of ST patients has been recruited adopting strict inclusion/exclusion criteria (i.e., ST percept, ST intrusiveness, hearing status) (Table 1). The power calculation to reach clinical significance was based on previous results of Tinnitus Handicap Inventory (THI) scores reductions of patients with an initial THI score above 40 in Güntensperger's study (2019), (n= 8, mean reduction: -11, std: 10,1) and enabled to set the number of participants to N=30 ($p<0,05$, $1-\beta>0,8$). Power analysis was based on sample size tables for clinical studies as published in Machin and al (2018).

Table 1, Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
Age above 18 years old	Pulsatile tinnitus, highly somato-sensory tinnitus
French social security affiliation	Menière's disease, Chronic or serous otitis, acoustic neuroma
Written consent to the protocol	Epilepsy
Permanent, non-fluctuating, high pitch, bilateral or assimilated tinnitus	Pregnancy or breastfeeding
Tinnitus Handicap Inventory (THI) score superior or equal to 40	Use of a sound generator during the therapy
Mean value of the Visual Analog Scales (VAS) score of at least 6 on Loudness (VAS-L) and Intrusiveness (VAS-I) scales	Tinnitus Reaction Questionnaire (TRQ) questionnaire on tinnitus associated distress above or equal to 40
Associated hearing-loss characterized by:	Antidepressant or anti-epileptic drug treatment
o Mean value of the hearing threshold loss for the 250, 500 and 1000 Hz thresholds strictly under 25 dB	Subjects under legal protection (guardianship, curatorship or judicial protection)
o At least one hearing threshold among the 2000, 4000, 6000 and 8000 Hz with at least 30 dB of hearing loss	Notable cognitive disability impeding to understand or perform the cognitive tasks
o Mean value of the hearing threshold loss for the 2000, 4000, 6000 and 8000 Hz thresholds strictly under 70 dB	Inability to wear the electrode headset of the neuro-feedback device due to superficial cranial scars or wounds

Cut-off values for the different questionnaires, scales and auditory thresholds were chosen to recruit patients with high level of intrusiveness, mild hearing loss on high frequencies (partial deafferentation), without catastrophizing reaction and to facilitate recruitment within the inclusion time frame.

Table 1 -- Inclusion and exclusion criteria.

According to the aforementioned criteria, 181 patients were screened; 32 suitable patients with chronic ST were included, one patient declared a tinnitus lateralization between the first clinical assessment (T1) and the first session of the protocol (S1) and was thus excluded, two patients declared not having enough time to complete all sessions and stopped and one patient stopped the protocol to follow a cardiovascular treatment, the patient heart condition being present and known before the start of the protocol and unrelated to it. These four patients were excluded from all analyses. Consequently, 28/32 patients completed the protocol and had their final evaluation (19 males and 9 females; mean age 60.11 years, SD 9.48) and 25/33 patients completed the 3-month follow-up visit (Table 2). All analyses including follow-up outcomes were consequently performed on these 25 patients. The study was approved by an Ethical Committee (CPP Lyon IV, ID-RCB/EudraCT ID 2018-A00604-51) and registered online (NCT03773926).

Table 2, Demographics and tinnitus characteristics of the study sample

	Mean	Std	Median	Min	Max
Age (Sex ratio M/F : 67,8%)	60.11	9.5	60	35	75
Mean hearing loss (dB)	27.22	9.0	25	8.57	42.86
Initial THI score	53.6	13.4	52	40	82
Initial VAS-L score	6.64	1.0	7	4	8
Initial VAS-I score	6.57	1.0	7	4	9
% incomodated while awake	48.51	24.2	50	20	100
Initial TRQ score	29.78	14.29	33	5	40

Std = standard deviation, Mean hearing loss was assessed by averaging 500Hz, 1kHz, 2kHz, 4kHz and 8kHz thresholds of the study sample. THI: Tinnitus Handicap inventory. VAS-L and VAS-I: Visual Analog Scales of the Loudness (VAS-L) and of the Intrusiveness (VAS-I) associated with the tinnitus percept. TRQ : Tinnitus Reaction Questionnaire.

Table 2 -- Demographics and tinnitus characteristics of the study sample.

Thirteen visits were scheduled during the trial (Figure 1). At the inclusion visit (T0), all patients were seen by an ENT medical doctor and underwent a tonal audiogram (0.25 to 8 kHz), unless they presented with a recent one. All patients filled both an online case-form (demographics, case history), validated questionnaires required for inclusion, i.e. THI, Tinnitus Reaction Questionnaire (TRQ) and Visual Analog Scale for loudness and intrusiveness (VAS-L, VAS-I) (Table 1). The questionnaires used in this study followed the guidelines of the Tinnitus Research Initiative (TRI) (Landgrebe and al 2012, Langguth and al, 2007), and of the French association of multidisciplinary tinnitus care professionals (AFREPA).

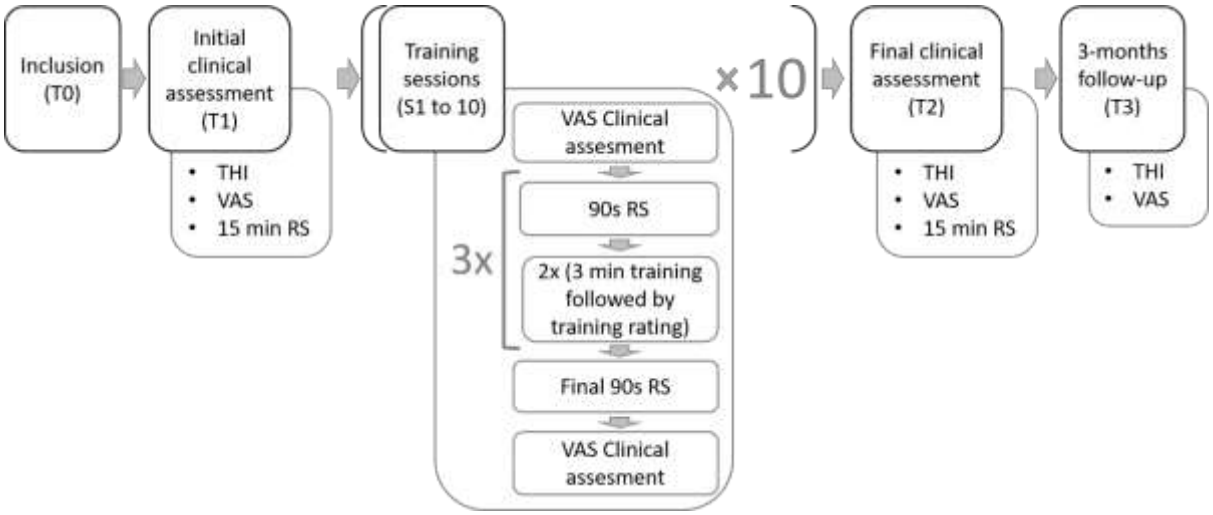

Figure 1 -- Clinical procedure of the protocol. Training ratings following the 3min trainings of S1–S10 consisted in an open question about what were the intents of the patient during the training (to control observance) and the rating on a 1–5 scale of the appreciation of the cognitive strategy used in the training.

After having received a thorough information about the potential benefits and harms of NFB all patients signed an informed consent. At the first session (T1), ST intrusiveness was reassessed and a baseline resting-state (RS) EEG measurement (7.5 min with eyes opened, 7,5 min with eyes closed) was recorded. Then 10 NFB sessions (S1 to S10) were scheduled during a 5-week period (+/- 1 week) with a mean of two sessions per week. At the end of

S10, another 15 min resting state EEG was performed. After the last NFB session, two follow-up visits were scheduled at 1 week (+/- 5 days) (T2) and at 3-month (+/- 10 days) (T3). At the end of the protocol, patients received 100 euros as a compensation and they could discuss their individual results with one of the investigators.

At T1, T2 and T3, patients completed the French version of the THI (score 0 to 100) (Ghulyan-Bédikian et al, 2010) and VAS-L and VAS-I (discreet likert-type scale, score 0 to 10). Additionally, both VAS were measured at the beginning and end of each NFB session.

The primary clinical outcome measure was the change of the THI score between T1 and T2. The secondary outcomes were the changes observed on the THI score between T1 and T3 and on VAS-L and VAS-I between T1, T2 and T3, as well as the correlation between the changes of the THI score and the trend of the α power-band measured at the beginning of each session from S1 to S10. Patients were asked by open questions at the beginning and at the end of each sessions about their well-being to identify any adverse effect.

EEG Recording:

A g-Nautilus EEG amplifier system with 32 dry g-SAHARA dry active electrodes (g.tec medical engineering GmbH, Schiedlberg, Austria) was used to record all the resting-states EEG and NFB trainings. The sampling rate for acquisition was 500 Hz. The array of 32 dry active 8-pin golden alloy electrodes were dispatched according to a subset of the 10-10 electrode position system (Jurcak and al, 2006), using positions: 'Fp1', 'Fp2', 'Af3', 'Af4', 'F7', 'F3', 'Fz', 'F4', 'F8', 'Fc5', 'Fc1', 'Fc2', 'Fc6', 'T7', 'C3', 'Cz', 'C4', 'T8', 'Cp5', 'Cp1', 'Cp2', 'Cp6', 'P7', 'P3', 'Pz', 'P4', 'P8', 'Po7', 'Po3', 'Po4', 'Po8', 'Oz', 'M1' and 'M2' . Recording was referenced against the left mastoid electrode (M1) with a ground electrode positioned at the right mastoid position (M2). As impedance value testing for determining signal quality was not possible for our system with dry electrodes, we used the FASTER method described in Nolan and al (2010) to evaluate each channel signal quality. Feedback on channels signal quality was given in real time through a graphical user-interface (GUI) to help the experimenter place the headset correctly. In order to avoid excessive movements, muscle contractions and other artifacts, patients were seated in a comfortable chair in a sound-proof and dim-light room. A custom GUI was developed (with Unity software) to guide the patient through the sessions and measurements. During the sessions, the GUI displayed easy-going step-by-step instructions to follow and allowed VAS ratings.

Neurofeedback Training and EEG online processing:

Each neurofeedback session (S1 to S10) followed the schedule described in figure 1. First, patients were asked to fill the initial VAS-L and VAS-I scales and to report any side effect that may have arisen since the last session. Then, after having equipped the patient with the EEG headset, the experimenter checked the signal quality of each channel through the GUI (which implemented the FASTER method (Nolan and al 2010)).
A resting state measurement of 90s was performed to estimate the baseline mean and standard deviation values of the α band power activity, denoted $\mu(\alpha)_{baseline}$ and $std(\alpha)_{baseline}$, respectively. During this measurement, patients were asked to sit still and fix a cross displayed in the center of the screen, avoiding unnecessary facial muscle contractions and head movements in order to minimize artifacts. Afterwards, patients accomplished two trainings sequences. In each sequence, patients were first instructed by

the GUI about the cognitive strategy they would practice in the following training and then started a 3m neurofeedback training using this cognitive strategy. The feedback was given in the form of a windmill. The speed of blades rotation was controlled by the in-line estimation of the alpha band power, denoted $\alpha_{current}$ here below,
as it follows:

$\alpha_{current} < \mu(\alpha)_{baseline} - 2 * std(\alpha)_{baseline}$: no rotation.
If $\alpha_{current}\ in\ \mu(\alpha)_{baseline} \pm 2 * std(\alpha)_{baseline}$: normal speed,
If $\alpha_{current} > \mu(\alpha)_{baseline} + 2 * std(\alpha)_{baseline}$: maximal speed.

The estimation of $\alpha_{current}$ was done on 2s signal windows, on two parallel threads with 50% overlap from the averaged power spectral density of F3, F4, Fc1 and Fc2 between 8 and 12Hz. Real-time signal pre-processing loops were implemented to minimize the risk of displaying a misleading (artifacted) feedback. Before band power extraction, each 2s signal window was first filtered between 2 and 20 Hz, then convoluted with a Hamming window to avoid border effects. FASTER method checked if some channels needed to be spline interpolated. Independent Component Analysis correction (using the FastICA algorithm as reported in Ablin and al, 2018), based on baseline data, was implemented to suppress contribution of eye blinks to the calculation of $\alpha_{current}$ .If no $\alpha_{current}$ value could be calculated, the last non-rejected $\alpha_{current}$ value was displayed, if none could be found, $\mu(\alpha)_{baseline}$ was fed back until an $\alpha_{current}$ value could be calculated again. At the end of each 3m training, in order to control and motivate observance, patients were asked to describe their mental state during the training and to rate the feasibility and comfort of each cognitive strategy tested. The same pattern RS-Tr(n)-Tr(n+1) repeated twice and the sessions ended with a RS. The whole sequence of recordings during a session can be summarized as RS1-Tr1-Tr2-RS2-Tr3-Tr4-RS3-Tr5-Tr6-RS4. Finally, patients filled their final VAS-L and VAS-I measurement to end their session. Patients were guided through all these steps by the GUI, but under experimenter supervision. Each session lasted approximately one hour.

Subject guidance:

In a previous study it was suggested that a better guidance may lead to improved patient adhesion and thus greater efficacy (Dohrmann 2007). Yet, according to current NFB practice, patients are usually asked to find out the best strategy by themselves, without any experimenter's suggestion. Instead, in this study eight cognitive strategies were proposed to the patients to guide them through the trainings, namely:

1. "Play a music you like in your head",
2. "Focus your attention on sounds in your surroundings",
3. "Picture yourself in a very calm and appeasing place",
4. "Remind a pleasant memory and focus on auditory details",
5. "Try to relax and meditate",
6. "Focus on the windmill and try to control it",

7. “Imagine a natural pleasant sound (sea waves, wind, bird singing…)”,
8. “Try to move the sound of tinnitus between your ears”.

The choice of these strategies was made from suggestions and feedbacks of tinnitus testers in the early technical tests of the user-interface, and in the light of the models of neurofeedback learning proposed by Sitaram and al (2016).

Each strategy was tested three times in a semi-randomized order among the 24 trainings of the four first sessions. The order was semi-randomized so that each strategy was tested at the start, middle or the end of sessions to control for the confounding factor of fatigue. At the end of S4, the four best strategies were selected for S5, S6, S7 and S8.

Ranking criteria for the strategies was estimated through a weighted average between the normalized mean value of the α ratios (α band power during a given strategy divided by the mean α band power during the preceding RS), with a weight of 2 and the normalized mean value of the appreciation ratings of a given strategy (appreciation score between 1 and 5 filled by the patient at the end of each training) with a weight of 1. The selection process was automatized by the GUI and controlled by the experimenter to ensure its validity. The same selection process was carried out at the end of S8 to choose the two best strategies for S9 and S10.

Off-line EEG data processing and analyses:

An off-line algorithmic analysis of EEG recordings was implemented with Anaconda Python 3.7 software and manually controlled (figure 2). First, an automated channel interpolation consisting in spline interpolation of the channels that exhibited abnormal spectral power, especially around 50 Hz, was performed. Then, RS(n) and two subsequent trainings Tr(n) and Tr(n+1) were concatenated to optimize the efficiency of the ICA blink rejection step and Riemannian artifact detection technique (Congedo and al 2017). Extended-infomax ICA (Lee and al 1999) blink detection was performed subsequently by identifying and suppressing the eye movement source in the ICA decomposition. Eye movements source selection from ICA sources was automatically performed by assessing the Euclidean distance between the source vectors of the mixing matrix with a predetermined model of eye blinks. The closest source was removed. Then, a threshold detection was realized with a 1s-long moving window on the signals and their derivative. Threshold peak-to-peak value for the signal was 90 μV, derivative threshold was individually assessed by an automatic detection and suppression of the tail absolute values of the derivative signals’ distribution. Then, a Riemannian artifact detection was implemented using the Riemannian potato technique (Barachant and al 2013). Finally, RS(n) and two subsequent trainings Tr(n) and Tr(n+1) were separated again for band power calculation. For band power calculation, signal was filtered on the frequency bands of interest and α band power was computed as the trace of the covariance matrices of each 2s signal window.

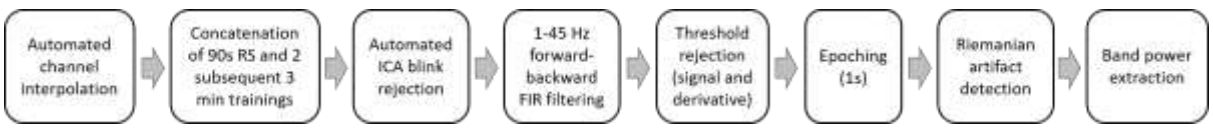

Figure 2 -- Signal processing automated pipeline.

NFB training efficiency and subjects' learning capabilities were assessed on learning criteria specifically defined for NFB (Congedo 2003, Gruzelier and al 2014). According to these criteria, "across session learning" is defined as the learning induced by the same successive trainings. This is not applicable here because multiple strategies were used. "within session learning" is defined as the ability to raise the target metric within a session. This was calculated for each patient by estimating the Spearman correlation coefficient between the sequence of $\mu(\mu(\alpha)_{RS\ or\ Tr})_{S1\rightarrow S10}$ from RS1 to RS4 and from Tr1 to Tr 6 and time. Finally, "baseline increments" are defined as the increase of the target metric across successive pre-session baselines. It was calculated for each patient by taking the Spearman correlation coefficient between the sequence of $\mu(\alpha)_{RS1}$ values from S1 to S10 and time. As it was the primary outcome measure in a recent study (Güntensperger and al 2019), we added to our analysis the "overall learning", which can be defined for each patient as the evolution of the target metric between the initial RS at T1 and the final assessment RS at T2, $\Delta_{T2-T1}(\mu(\alpha)_{RS})$. As baseline increments are supposed to be the most relevant measures for long term plasticity, correlation tests were run between baseline increments and clinical outcomes in order to verify if NFB efficiency correlates with clinical improvements.

Statistics:

We used the Python library scipy.stats functions for statistical testing. For the analysis of clinical outcomes, Wilcoxon tests were used with Cohen's d effect size measure (Cohen 1992), except for the assessment of the correlation between VAS-I and time and VAS-L and time where Spearman correlation tests were used. Bonferroni correction was applied to account for multiple comparisons on the tests performed on the clinical outcomes data.

Following Congedo (2003) and Congedo and al (2004), within learning and baseline increments learning capabilities were assessed individually to account for learning pattern variability between subjects. The Fisher's p-values combination method (Pearce, 1992) was then adopted to draw a group-scale conclusion out of individual p-value observations.

For within session learning evaluation, individual p-values were estimated for signed Spearman correlation tests between $\mu(\mu(\alpha)_{RS})_{S1\rightarrow S10}$ and time (RS1 to RS4) and between $\mu(\mu(\alpha)_{Tr})_{S1\rightarrow S10}$ and time (Tr1 to Tr6). They were then respectively combined with the Fisher's combination method. A similar procedure was applied for baseline increments, where individual Spearman correlation p-values for initial and final RS sequences were respectively combined. For overall learning, as normality test for distribution of log([mean]α) values was not significant, Wilcoxon test was performed for log([mean]α) for eyes-open RS between T1 and T2. Spearman correlation tests were used for assessing the correlation between baseline increments and THI changes.

Results:

Clinical outcome:

The clinical outcome results are displayed in Figure 3.

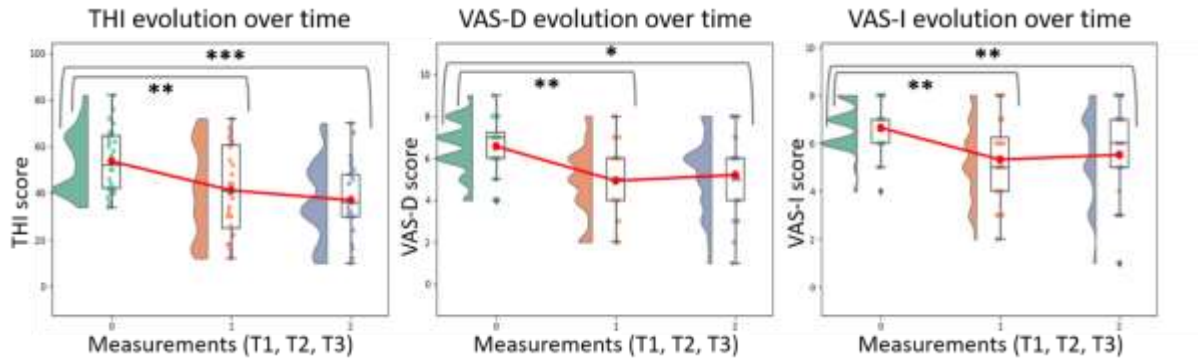

*Figure 3 -- Clinical outcomes evolution. THI scores decreased from an average of 53.64 (T1, N¼28, std.: 13.16) to an average of 41.21 (T2, N¼28, std. 19.23) to a follow-up average of 36.96 (T3, N¼25, std.: 16.13). VAS-L scores decreased from an average of 6.64 (T1, N¼28, std.: 0.97) to an average of 5.32 (T2, N¼28, std. 1.63) to a follow-up average of 5.50 (T3, N¼25, std.: 1.80). VAS-I scores decreased from an average of 6.57 (T1, N¼28, std.: 1.24) to an average of 4.93 (T2, N¼28, std. 1.67) to a follow-up average of 5.32 (T3, N¼25, std.: 1.63). Decreases of THI, VAS-L, VAS-scores were all significant between (T1) and (T2) (P <0.001). THI score evolution between (T1) and (T3) was clinically significant (P <0.01) and clinically relevant when compared to the clinical threshold of 7-point reduction (Zeman et al., 2011). *¼ P <0.001, **¼clinically significant with P <0.001.*

The primary outcome measure (THI score) showed a statistically significant decrease between T1 (M = 53.64, std = 13.16) and T2 (M = 41.21, std = 19.23), T(27) = 44.5, d= -0.75 $p < .001$, yet did not reach the clinically significant improvement of 7 points (Zeman and al, 2011) (T(27) = 138, d= - 0.33, p=0.14). THI scores continued to decrease between T2 and T3 without reaching statistical significance (T(24) = 95.0, d= -0.43 p = 0.18). However, the overall decrease of 34% of THI scores between T1 (M = 54.16, std = 13.58) and the last follow-up visit T3 (M = 36.96, std = 16.13) reached the clinical significance of 7 points, T(24) = 35.5, d = -0.83 p<0.001. Moreover, for secondary outcomes, VAS-L and VAS-I significantly decreased respectively of 17,9% and 24,2% between T1 (VAS-L: M = 6.64, std = 0.97, VAS-I: M = 6.57, std = 1.24) and T2 (VAS-L: M = 5.32, std = 1.63, VAS-I: M = 4.93, std = 1.67), VAS-L : T(27) = 35.5, d = -0.99 p < 0.001, VAS-I : T(27) = 4, d= -1.12 p < 0.001. The decrease maintained its significance for VAS-I between T1 (M = 6.64, std = 0.97) and T3 (M = 5.32, std = 1.63) for VAS-I, T(24) = 13.5, d = -1.00, p = 0.0015. The GUI enabled the patients to become autonomous after 2 or 3 sessions and patients reported that the given instructions were clear. VAS-I and VAS-L scores evaluated across sessions both exhibits significant decrease over time for initial and final assessments (N=28, all tests p<0.01) (not shown).

Neurofeedback learning evaluation:

The choice of using dry electrodes hampered the analysis of EEG recordings. EEG signal quality was very variable between subjects and between recordings. Several rejection thresholds were applied to ensure appropriate quality of the dataset. Thus a significant amount of data had to be excluded from the analysis due to poor signal quality. 786

recordings were excluded out of the 28*10*10 + 28 *2 = 2856 recordings of all sessions and 15 min RS at T1 and T2 for all 28 patients, hence the results are obtained on 72.5% of the dataset. The results of EEG learning are displayed in Figure 4. Within session learning reached statistical significance for the sequence of trainings (p=0.048) but not for the sequence of RS. Baseline increments were not significant neither for initial RS sequence nor for final RS sequences. Overall learning reached significance between T1 and T2 (p=0.048).

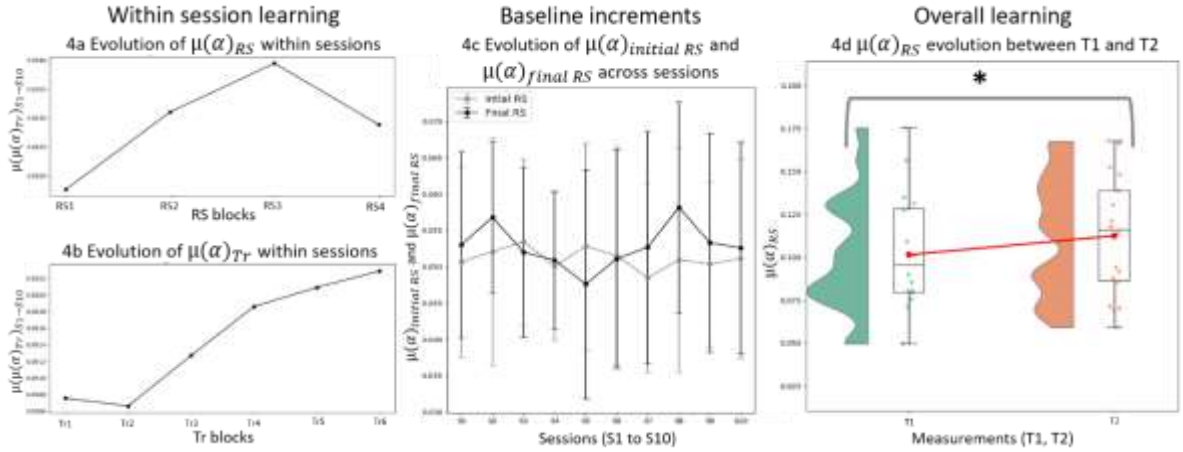

*Figure 4 -- EEG learning capabilities evaluations. From left to right: within session learning on the sequence of μ(μ(α)RS)S1!S10 values (4a) and on μ(μ(α)Tr)S1!S10 values (4b), Baseline increments (4c), assessing the accumulated learning over sessions, evaluated for the initial RS (gray) and the final RS (black). Overall learning (4d), comparing the mean band power between the 7.5min "eyes open" parts of the 15min initial and final RS in (T1) and (T2). *¼ P <0.05.*

A significant correlation was found between the individual baseline increments and THI scores between T1 and T2 (p=0.019) (Figure 5).

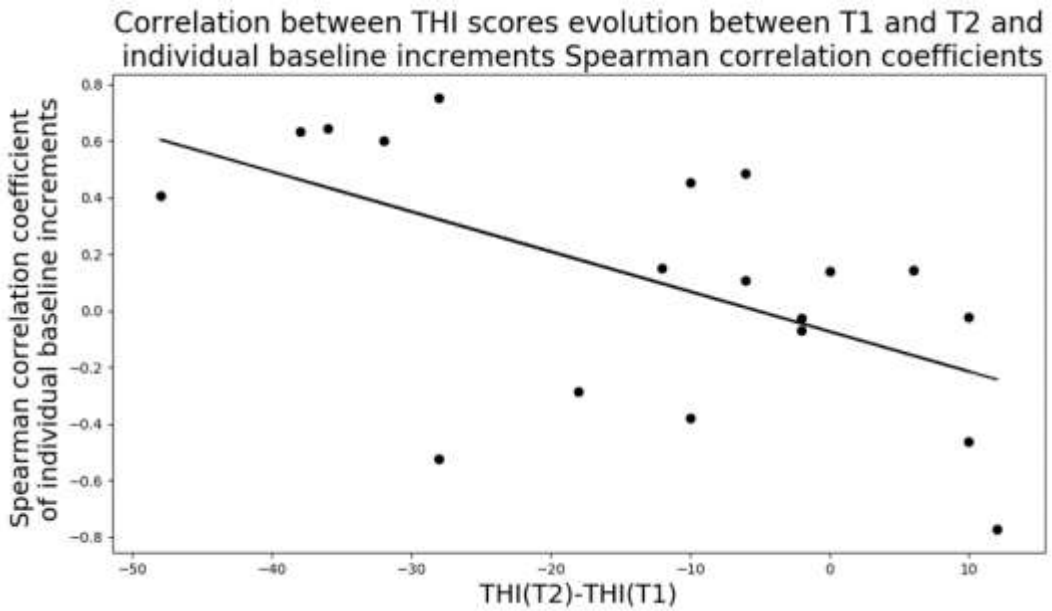

Figure 5 -- Correlation between THI score evolution and Spearman correlation coefficients of individual baseline increments correlation. Spearman correlation coefficients translates the tendency of the baseline increments across sessions, positive values representing positive learning for baseline increments across sessions, THI score reduction, represented by negative values, denotes the clinical improvement of the patient condition. The correlation was significant (P <0.05).

Discussion:

The first goal of this study was to assess if NFB can be implemented in everyday practice with minimal assistance from a clinician. The second was to verify if strict inclusion criteria and guiding for choosing cognitive strategies may improve clinical outcome.

Merits.

The present study has tested neurofeedback for treating tinnitus on a homogeneous population of patients, selected rigorously with strict inclusion criteria. Improvements over patient guidance strategies were attempted and experimental procedures were set to verify if such an approach can be implemented in everyday practice.

Limits.

This study was a single arm trial without control, and did not evaluate placebo effect. The analyses and conclusion of the training performance should be taken in account with caution as only 72,5% of the EEG recordings were exploitable, due to the choice of the use of dry EEG electrodes. Nonetheless the amount of analyzed data remains conspicuous.

NFB for ST can be implemented in a standard setting and the GUI has helped the patients to smoothly go through the NFB process. Moreover, NFB has been shown to significantly improve ST intrusiveness as measured by THI and VAS ratings. No severe or long-lasting side effects have been observed, although minor headaches and short-term increase of tinnitus were reported. The overall results are not as beneficial as expected. Even if it is easier to fit and use, we assume that the choice of a dry electrode EEG device is the main cause of such a partly positive outcome. Indeed, EEG signal quality was not optimal, hampering the potential benefits of NFB. Although the mean THI decrease between T1 and T2, and then T3 is higher than the one observed in three preceding studies (Dohrman and al 2007, Crocetti and al 2011, Güntensperger and al 2019), the overall improvement is moderate, reaching only clinical significance at T3. Only one study showed a similar positive outcome (Vanneste and al 2018b). It also included patients with high initial THI scores. Albeit not statistically significant, we also acknowledge a trend of decrease for the THI score during follow-up (T2 to T3). This has never been observed in former studies and could be interpreted as a hint for a long term NFB induced plasticity, deserving further investigation. We infer from these results that the choice of dry electrodes had only limited impact on the efficacy of the NFB approach. Although the inclusion of homogeneous samples of patients is mandatory in ST clinical trials (Londero and Hall 2017) the subgroup of patients selected here did not seem to specifically benefit more from NFB training. Indeed, the criteria we have chosen (bilateral, high pitched ST, no severe hearing loss) may have not been optimal. For

these reasons, patient subtyping, as tackled by the TRI database project (Landgrebe and al 2010), appears a central issue for the future of tinnitus research. If patients displaying other psychoacoustic properties of ST would had resulted in a better outcome remains an open question. So is the question of whether the selection of a homogeneous population driven by the etiology of ST (e.g. noise induced hearing loss, presbycusis, otosclerosis) may have led to different results. Also, the choice to only train α band power and not α/δ ratio as preceding studies, nor the θ band as current findings on thalamocortical dysrhythmia suggests, is questionable and can constitute another confounding factor.

The issue of guidance in NFB is also still controversial. Some studies reported improvement by the means of explicit cognitive strategies (Fovet and al 2015) while other studies found no differences or advocated for letting the patients find their own strategies (Hardman and al 1997, Scharnowski and al 2015). In this study, we chose to guide the patients through a relatively complex process of selection of explicit cognitive strategies. Even if the guiding helped patients to stay active and interested in the protocol, it did not dramatically improve the clinical outcome when comparing with previous studies that did not offer guiding. The analysis of RS EEG signal is debatable. Indeed, no significant baseline increments learning could be elicited at the group level. Taking into account the limits of EEG signal quality obtained with the dry electrodes we used, this does not witness in favor of a long term plasticity induced by NFB on a α band metric. Conversely, the correlation observed between the baseline increments and THI changes seems encouraging, although the design of this study cannot prove a causality link between the efficiency of NFB training and ST improvement. Indeed, a placebo effect (Hesser and al 2011) could explain both the THI changes and a decrease of anxiety that could translate in an increase of the α band power not specific to the auditory cortices (Al-Shargie and al 2016). It should be reminded here that although α decrease was found a physiological correlate of ST, this finding could not be reproduced by other studies (Pierzycki and al 2016). Nevertheless, although auditory α may not be per-se the neurophysiological correlate of ST, it is not impossible that α improvement in ST patients may have an indirect positive impact, as supported by other neuromodulation studies (Vanneste and De Ridder 2012, Londero and al 2006).

Conclusion:

This exploratory study has shown that a NFB training for treating ST in a day-to-day clinical setting is feasible and that the guidance (GUI, selection of cognitive strategies) reinforces patient compliance to the protocol. It also shed light on a key technical problem: the use of wet electrodes or other kinds of dry electrodes to ensure a proper signal quality during EEG acquisition. The inclusion criteria to be used for selecting the most responsive subgroup of ST patients is still a matter of debate. The main outcome measure (THI) did show a statistical positive result reaching a clinically significant value at the 3-month follow-up visit and a significant correlation with some RS EEG patterns recorded during the sessions. Even if these results are encouraging, the design of this pilot study (single arm, non-controlled) cannot prove the specific efficacy of NFB on alleviating ST intrusiveness. Further research (placebo-controlled study) is required to confirm or infirm these preliminary conclusions.

Acknowledgements:

This research was funded by AFREPA (French association of tinnitus healthcare professionals) and Zeta Technologies. RG, TC and JS are employees and shareholders of Zeta Technologies, and VF and BP are employees of Zeta Technologies. The authors wish to thank LEA Audika and Fondation pour l'Audition for their support in the realization of this study.

References:

Adjamian, P., Sereda, M. and Hall, D.A., 2009. The mechanisms of tinnitus: perspectives from human functional neuroimaging. *Hearing research*, *253*(1-2), pp.15-31.

Ablin P, Cardoso J, Gramfort A, 2018. Faster Independent Component Analysis by Preconditioning With Hessian Approximations. IEEE Transactions on Signal Processing 66:4040–4049

Al-Shargie, F., Kiguchi, M., Badruddin, N., Dass, S.C., Hani, A.F.M. and Tang, T.B., 2016. Mental stress assessment using simultaneous measurement of EEG and fNIRS. *Biomedical optics express*, *7*(10), pp.3882-3898.

Arns, M., De Ridder, S., Strehl, U., Breteler, M. and Coenen, A., 2009. Efficacy of neurofeedback treatment in ADHD: the effects on inattention, impulsivity and hyperactivity: a meta-analysis. *Clinical EEG and neuroscience*, *40*(3), pp.180-189.

Baguley, D., McFerran, D. and Hall, D., 2013. Tinnitus. *The Lancet*, *382*(9904), pp.1600-1607.

Barachant, A., Andreev, A. and Congedo, M., 2013, January. The Riemannian Potato: an automatic and adaptive artifact detection method for online experiments using Riemannian geometry. *TOBI Workshop IV*, Jan 2013, Sion, Switzerland. pp.19-20.

Cohen, J., 1992. Statistical power analysis. *Current directions in psychological science*, *1*(3), pp.98-101.

Congedo, M., 2003. Tomographic Neurofeedback: A new technique for the Self-Regulation of brain electrical activity. PhD Thesis, University of Tennessee, Knoxville.

Congedo M, Barachant A, Bhatia R (2017) Riemannian Geometry for EEG-based Brain-Computer Interfaces; a Primer and a Review. Brain-Computer Interfaces, 4(3), 155-174.

Congedo, M., Lubar J.,F., Joffe, D., (2004) Low-Resolution Electromagnetic Tomography Neurofeedback, IEEE Transactions on Neuronal Systems & Rehabilitation Engineering 12(4), 387-397.

Crocetti, A., Forti, S. and Del Bo, L., 2011. Neurofeedback for subjective tinnitus patients. *Auris Nasus Larynx*, *38*(6), pp.735-738.

De Ridder, D., Vanneste, S., Langguth, B. and Llinas, R., 2015. Thalamocortical dysrhythmia: a theoretical update in tinnitus. *Frontiers in neurology*, *6*, p.124.

Dohrmann, K., Weisz, N., Schlee, W., Hartmann, T. and Elbert, T., 2007. Neurofeedback for treating tinnitus. *Progress in brain research*, *166*, pp.473-554.

Fovet, T., Jardri, R. and Linden, D., 2015. Current issues in the use of fMRI-based neurofeedback to relieve psychiatric symptoms. *Current pharmaceutical design*, *21*(23), pp.3384-3394.

Gevensleben, H., Holl, B., Albrecht, B., Vogel, C., Schlamp, D., Kratz, O., Studer, P., Rothenberger, A., Moll, G.H. and Heinrich, H., 2009. Is neurofeedback an efficacious treatment for ADHD? A randomised controlled clinical trial. *Journal of Child Psychology and Psychiatry*, *50*(7), pp.780-789.

Ghulyan-Bédikian, V., Paolino, M., Giorgetti-D'Esclercs, F. and Paolino, F., 2010. Psychometric properties of a French adaptation of the Tinnitus Handicap Inventory. *L'Encephale*, *36*(5), pp.390-396.

Gruzelier, J.H., 2014. EEG-neurofeedback for optimising performance. III: a review of methodological and theoretical considerations. *Neuroscience & Biobehavioral Reviews*, *44*, pp.159-182.

Güntensperger, D., Thüring, C., Kleinjung, T., Neff, P. and Meyer, M., 2019. Investigating the Efficacy of an Individualized Alpha/Delta Neurofeedback Protocol in the Treatment of Chronic Tinnitus. *Neural plasticity*, *2019*.

Hammond, D.C., 2011. What is neurofeedback: An update. *Journal of Neurotherapy*, *15*(4), pp.305-336.

Hardman, E., Gruzelier, J., Cheesman, K., Jones, C., Liddiard, D., Schleichert, H. and Birbaumer, N., 1997. Frontal interhemispheric asymmetry: self regulation and individual differences in humans. *Neuroscience Letters*, *221*(2-3), pp.117-120.

Hartmann, T., Lorenz, I., Müller, N., Langguth, B. and Weisz, N., 2014. The effects of neurofeedback on oscillatory processes related to tinnitus. *Brain Topography*, *27*(1), pp.149-157.

Hesser, H., Weise, C., Rief, W. and Andersson, G., 2011. The effect of waiting: a meta-analysis of wait-list control groups in trials for tinnitus distress. *Journal of psychosomatic research*, *70*(4), pp.378-384.

Jurcak, V., Tsuzuki, D. and Dan, I., 2007. 10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems. *Neuroimage*, *34*(4), pp.1600-1611.

Landgrebe, M., Zeman, F., Koller, M., Eberl, Y., Mohr, M., Reiter, J., Staudinger, S., Hajak, G. and Langguth, B., 2010. The Tinnitus Research Initiative (TRI) database: a new approach for delineation of tinnitus subtypes and generation of predictors for treatment outcome. *BMC medical informatics and decision making*, *10*(1), p.42.

Landgrebe, M., Azevedo, A., Baguley, D., Bauer, C., Cacace, A., Coelho, C., Dornhoffer, J., Figueiredo, R., Flor, H., Hajak, G. and van de Heyning, P., 2012. Methodological aspects of

clinical trials in tinnitus: a proposal for an international standard. *Journal of psychosomatic research*, *73*(2), pp.112-121.

Langguth, B., Goodey, R., Azevedo, A., Bjorne, A., Cacace, A., Crocetti, A., Del Bo, L., De Ridder, D., Diges, I., Elbert, T. and Flor, H., 2007. Consensus for tinnitus patient assessment and treatment outcome measurement: Tinnitus Research Initiative meeting, Regensburg, July 2006. *Progress in brain research*, *166*, pp.525-536.

Langguth, B., Landgrebe, M., Kleinjung, T., Sand, G.P. and Hajak, G., 2011. Tinnitus and depression. *The world journal of biological psychiatry*, *12*(7), pp.489-500.

Langguth, B., Kreuzer, P.M., Kleinjung, T. and De Ridder, D., 2013. Tinnitus: causes and clinical management. *The Lancet Neurology*, *12*(9), pp.920-930.

Lee, T.W., Girolami, M., Sejnowski, T.J., 1999. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. Neural computation, 11(2), pp.417-441.

Llinás, R.R., Ribary, U., Jeanmonod, D., Kronberg, E. and Mitra, P.P., 1999. Thalamocortical dysrhythmia: a neurological and neuropsychiatric syndrome characterized by magnetoencephalography. *Proceedings of the National Academy of Sciences*, *96*(26), pp.15222-15227.

Lorenz, I., 2007. From Signal Space to Source Space: Does Source Space Projection Improve the Neurofeedback Therapy in Chronic Tinnitus Patients?

Londero, A., Langguth, B., De Ridder, D., Bonfils, P. and Lefaucheur, J.P., 2006. Repetitive transcranial magnetic stimulation (rTMS): a new therapeutic approach in subjective tinnitus?. *Neurophysiologie Clinique/Clinical Neurophysiology*, *36*(3), pp.145-155.

Londero, A. and Hall, D.A., 2017. Call for an evidence-based consensus on outcome reporting in tinnitus intervention studies. *Frontiers in Medicine*, *4*, p.42.

Machin, D., Campbell, M.J., Tan, S.B. and Tan, S.H., 2018. *Sample sizes for clinical, laboratory and epidemiology studies*. Wiley-Blackwell.

McCormack, A., Edmondson-Jones, M., Somerset, S. and Hall, D., 2016. A systematic review of the reporting of tinnitus prevalence and severity. *Hearing research*, *337*, pp.70-79.

Nolan, H., Whelan, R. and Reilly, R.B., 2010. FASTER: fully automated statistical thresholding for EEG artifact rejection. *Journal of neuroscience methods*, *192*(1), pp.152-162.

Norena, A.J., 2011. An integrative model of tinnitus based on a central gain controlling neural sensitivity. *Neuroscience & Biobehavioral Reviews*, *35*(5), pp.1089-1109.

Pearce, S.C., 1992. Introduction to Fisher (1925) statistical methods for research workers. In *Breakthroughs in statistics* (pp. 59-65). Springer, New York, NY.

Pierzycki, R.H., McNamara, A.J., Hoare, D.J. and Hall, D.A., 2016. Whole scalp resting state EEG of oscillatory brain activity shows no parametric relationship with psychoacoustic and psychosocial assessment of tinnitus: a repeated measures study. *Hearing research*, *331*, pp.101-108.

Rauschecker, J.P., Leaver, A.M. and Mühlau, M., 2010. Tuning out the noise: limbic-auditory interactions in tinnitus. *Neuron*, *66*(6), pp.819-826.

Scharnowski, F., Veit, R., Zopf, R., Studer, P., Bock, S., Diedrichsen, J., Goebel, R., Mathiak, K., Birbaumer, N. and Weiskopf, N., 2015. Manipulating motor performance and memory through real-time fMRI neurofeedback. *Biological psychology*, *108*, pp.85-97.

Schultz, K.F., 2010. CONSORT statement: updated guidelines for reporting parallel group randomised trials. *BMC med*, *8*, pp.18-27.

Sitaram, R., Ros, T., Stoeckel, L., Haller, S., Scharnowski, F., Lewis-Peacock, J., Weiskopf, N., Blefari, M.L., Rana, M., Oblak, E. and Birbaumer, N., 2017. Closed-loop brain training: the science of neurofeedback. *Nature Reviews Neuroscience*, *18*(2), p.86.

Sterman, M.B. and Friar, L., 1972. Suppression of seizures in an epileptic following sensorimotor EEG feedback training. *Electroencephalography and clinical neurophysiology*, *33*(1), pp.89-95.

Sterman, M.B. and Egner, T., 2006. Foundation and practice of neurofeedback for the treatment of epilepsy. *Applied psychophysiology and biofeedback*, *31*(1), p.21.

Vanneste, S. and De Ridder, D., 2012. Noninvasive and invasive neuromodulation for the treatment of tinnitus: an overview. *Neuromodulation: Technology at the Neural Interface*, *15*(4), pp.350-360.

Vanneste, S., Song, J.J. and De Ridder, D., 2018a. Thalamocortical dysrhythmia detected by machine learning. *Nature communications*, *9*(1), pp.1-13.

Vanneste, S., Joos, K., Ost, J. and De Ridder, D., 2018b. Influencing connectivity and cross-frequency coupling by real-time source localized neurofeedback of the posterior cingulate cortex reduces tinnitus related distress. *Neurobiology of stress*, *8*, pp.211-224.

Weisz, N., Moratti, S., Meinzer, M., Dohrmann, K. and Elbert, T., 2005. Tinnitus perception and distress is related to abnormal spontaneous brain activity as measured by magnetoencephalography. *PLoS Med*, *2*(6), p.e153.

Welch P., 1967. The use of the fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms, IEEE Trans. Audio Electroacoust. vol. 15, pp. 70-73.

Wyrwicka, W. and Sterman, M.B., 1968. Instrumental conditioning of sensorimotor cortex EEG spindles in the waking cat. *Physiology & Behavior*, *3*(5), pp.703-707.

Figure Descriptions :

Fig 1 : Clinical procedure of the protocol. Training ratings following the 3min trainings of S1 to S10 consisted in an open question about what were the intents of the patient during the training (to control observance) and the rating on a 1 to 5 scale of the appreciation of the cognitive strategy used in the training.

Fig 2 : Signal processing automated pipeline.

*Fig 3: Clinical outcomes evolution. THI scores decreased from an average of 53.64 (T1, N=28, std: 13.16) to an average of 41.21 (T2, N=28, std 19.23) to a follow-up average of 36.96 (T3, N=25, std: 16.13). VAS-L scores decreased from an average of 6.64 (T1, N=28, std: 0.97) to an average of 5.32 (T2, N=28, std 1.63) to a follow-up average of 5.50 (T3, N=25, std: 1.80). VAS-I scores decreased from an average of 6.57 (T1, N=28, std: 1.24) to an average of 4.93 (T2, N=28, std 1.67) to a follow-up average of 5.32 (T3, N=25, std: 1.63). Decreases of THI, VAS-L, VAS- scores were all significant between (T1) and (T2) (p<0.001). THI score evolution between (T1) and (T3) was clinically significant (p<0.01) and clinically relevant when compared to the clinical threshold of 7-point reduction (Zeman and al, 2011). * = p<0.001, ** = clinically significant with p<0.001.*

Fig 4: EEG learning capabilities evaluations. From left to right: within session learning on the sequence of $\mu(\mu(\alpha)_{RS})_{S1\rightarrow S10}$ *values (4a) and on* $\mu(\mu(\alpha)_{Tr})_{S1\rightarrow S10}$ *values (4b), Baseline increments (4c), assessing the accumulated learning over sessions, evaluated for the initial RS (grey) and the final RS (black). Overall learning (4d), comparing the mean band power between the 7.5 min « eyes open » parts of the 15 min initial and final RS in (T1) and (T2). * = p<0.05*

Fig 5: Correlation between THI score evolution and Spearman correlation coefficients of individual baseline increments correlation. Spearman correlation coefficients translates the tendency of the baseline increments across sessions, positive values representing positive learning for baseline increments across sessions, THI score reduction, represented by negative values, denotes the clinical improvement of the patient condition. The correlation was significant (p < 0.05).