UNIVERSITAT DE BARCELONA

Contribution of the subcortical auditory pathway to the perception and processing of sounds

Natàlia Gorina Careta

Contribution of the subcortical auditory pathway to the perception and processing of sounds Natàlia Gorina Careta

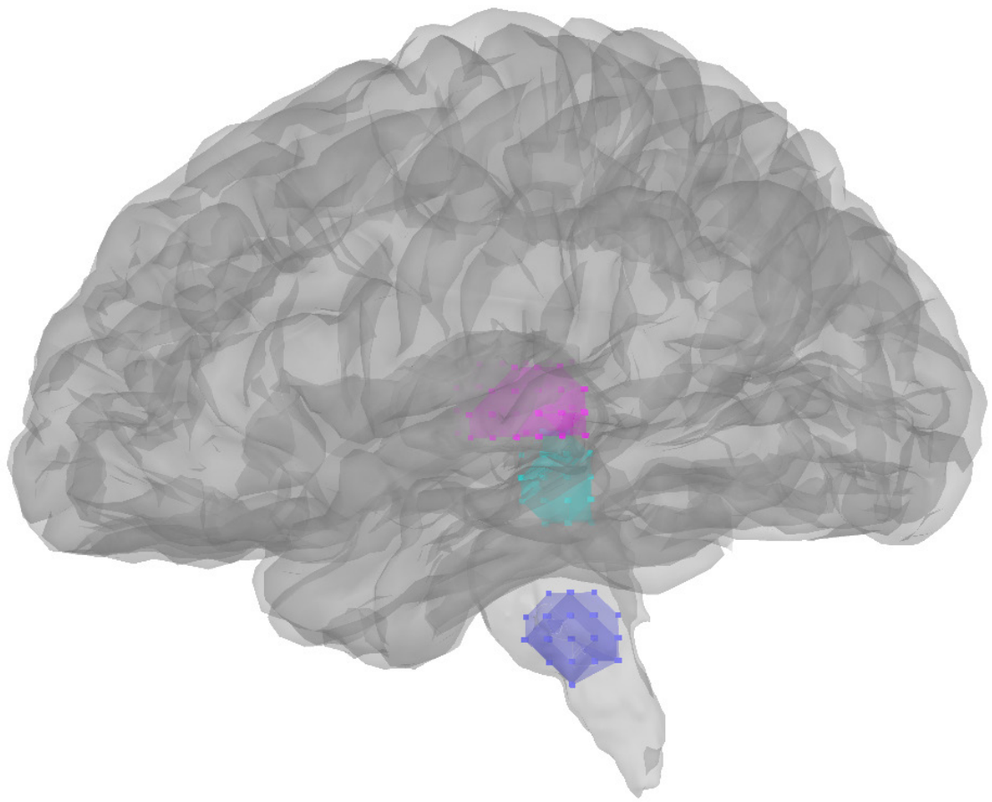
Universitat de Barcelona

Departament de Psicologia Clínica i Psicobiologia
Facultat de Psicologia
Universitat de Barcelona

Contribution of the subcortical auditory pathway to the perception and processing of sounds

Natàlia Gorina Careta

Supervised by Dr. Carles Escera



Universitat de Barcelona

Departament de Psicologia Clínica i Psicobiologia
Facultat de Psicologia
Universitat de Barcelona

UNIVERSITAT DE BARCELONA

Departament de Psicologia Clínica i Psicobiologia
Facultat de Psicologia
Universitat de Barcelona

Contribution of the subcortical auditory pathway to the perception and processing of sounds

Thesis presented by

Natàlia Gorina Careta

To obtain the

Grau de Doctora per la Universitat de Barcelona

In accordance with the requirements for the

International PhD Diploma

Programa de Doctorat en Cervell, Cognició i Conducta

Supervised by Dr. Carles Escera

Barcelona, January 2019

Institut de Neurociències
UNIVERSITAT DE BARCELONA

Brainlab — cognitive neuroscience research group
grup de recerca en neurociència cognitiva

Per als meus pares i les meves germanes, per tot el suport incondicional.

I per l'Alex, per tots els somriures diaris i no dubtar mai de mi.

ACKNOWLEDGEMENTS · AGRAÏMENTS

This work would not have been possible without the help of many people. Durant els quatre anys que ha durat aquesta tesis, ni per un moment vaig pensar que la part més difícil de tota serien els agraïments. Hi ha moltíssima gent que ha col·laborat i ha contribuït en aquesta tesis, així que espero poder expressar en un espai tant petit la profunditat dels meus agraïments.

En primer lloc al Brainlab – Grup de Recerca en Neurociència Cognitiva. Vosaltres heu estat el cos i anima del que està escrit i res, absolutament cap part d'aquesta tesis, hauria estat possible sense el vostre suport. I amb el Brainlab va el seu boss, el director d'aquesta tesis. Carles, totes les gràcies que et pugui donar es queden curtes. Gràcies per donar-li la oportunitat a una persona de biomedicina que es va plantar al teu despatx dient que volia fer ciència. Gràcies per ensenyar-me tot el que sé de psicologia i neurociència cognitiva i sobretot, gracies per deixar-me créixer, per escoltar i tenir en compte la meva opinió, per deixar-me equivocar i ensenyar-me a pensar críticament. Gràcies per totes les oportunitats i el suport incondicional que m'has donat al llarg de tots els anys que ha durat aquest procés. Gràcies per ser persona abans que boss i donar-me suport no només en la part científica, sinó en tota la part personal.

A la segunda al mando, Iria. Volviste "a lo Alemán" y solo puedo darte las gracias por ello. Gracias por el apoyo incondicional en todo y por estar allí en todo momento. Eres la primera persona a la que veo por las mañanas, ya trabajando siempre, y no podría decirte el ejemplo de constancia y dedicación que eres para mí. Ah sí, gracias también por dejarme protocolizarlo todo y mandar un poquito, ¡en el fondo me encanta!

To Kasia. My mentor, my co-supervisor, my friend, my all. Many people has contributed to this thesis, but you took me when I was just a master student willing to learn everything and turned me into a PhD student. You teached me how to run experiments, how to program and basically all the scientific

practical basis I have so far. Without you this wouldn't have been possible, so thank you my spring flower, I love you to the moon and back.

Al meu duo de Maripilis, Marta i Cristina. Totes les gràcies que us pugui donar no seran suficients per expressar tot el que vull dir-vos. Gràcies per cuidar-me i mimar-me. Gràcies per deixar-me expressar tot el que sento, per ser les meves dues constants dins de tot el caos de la tesis. Gràcies per estar absolutament sempre allà i deixar de banda tot per ajudar-me o simplement per fer un cafè. Gràcies per no ser ciència a totes hores i per distreure'm tant en els bons com en els mals moments. Y a mí dúo le falta la nota masculina, que eres tu Paco. Gracias por todo todísimo, por aguantar mi mal y mi buen humor, por los mil abrazos de soporte incondicional de cada día, por dejar que te distraiga y distraerme cada mañana para desayunar y comer. Gracias por estar al momento allí, por no matarme con mis llamadas los lunes a las 9 con problemas ni cuando te vine a ver con 3 TB de datos para colocar en un servidor. Gracias también por todo lo que me has enseñado, por tener una paciencia infinita y, en resumen, por ser tú. Als tres, us estimo i us adoro, i res no hauria estat el mateix sense vosaltres.

Al Jordi. Quan vaig començar es parlava de tu amb grans paraules i jo només sabia que a Nova York hi havia una persona brillant que havia pensat la idea per l'estudi que firma l'inici d'aquesta tesis. Vas tornar i vas complir, de sobres, totes les expectatives. Ets l'exemple de científic, amb un coneixement impressionant i una gran capacitat de veure més enllà. Però sobretot, el més important és que saps transmetre-ho. Gràcies per ser la nota crítica (sempre positiva) i ajudar-me a millorar, per fer-me veure els punts febles i els forts de tot i confiar en mi, considerant-me un punt fort del laboratori. Gràcies per col·laborar en tots i cadascun dels estudis i per ajudar-me en tots els anàlisis quan ja no podia veure més enllà.

I per moltes persones que m'hagin inspirat, aquesta tesis també està dedicada a vosaltres, Fran, Tere i Vittoria, que heu sigut els meus companys durant tot el procés. Franciiiiscu, et vaig veure entrar com un pollet i ara que t'escric aquestes paraules estàs a Estats Units aconseguint el post-doc que sempre has volgut. Gràcies per confiar en mi i per fer-me riure a totes hores. Ets la constància i el treball personificat. Tere, a tu que t'he de dir. Has estat la meva primera estudiant de màster a la que supervisar i t'has convertit en companya de despatx i sobretot, amiga. Estic orgullosíssima de tu i de veure on has arribat, i et dono gràcies per preocupar-te cada dia per mi, per alegrar-te amb cada p significativa i per posar una nota de bogeria i improvisació a la meva vida protocol·litzada. Ets com una granada de mà, petita i explosiva, i només et puc donar les gràcies per això. Y Vittoria... El yang de mi ying. Eres mi opuesto y creo que durante todos los años hemos estado de acuerdo unas dos (¿tres?) veces. Gracias por darme una oportunidad de mejorar mis argumentos para convencerte. Pero sobretodo, gracias por todo el apoyo que me has dado y por hacerme ver que siempre hay otra manera de ver las cosas.

Mil gràcies també a tots els estudiants de doctorat passats i presents, sense el suport dels quals no estaria aquí. Marc, Lenka i Míriam, vosaltres heu estat el meu exemple a seguir en aquest procés. Sou les tesis que he vist escriure i defensar, sou les persones que quan assistia als "novelty meetings" i no entenia res discutíeu les idees sent experts i us admirava , i sobretot, sou les persones que heu estat allà quan m'entraven dubtes de "i això com ho faig" o "i ara que faré". Gràcies. I als nous, Marta, Nadia y Sonia. Habéis entrado por la puerta grande y con aires nuevos. Gracias por convertiros en parte de esta gran familia tan grande y por estar allí en todo momento durante estos últimos meses. Gracias por los cafés y las comidas que me distraían de reescribir la misma línea, por ser aire fresco y motivación. Espero que yo os pueda ayudar a vosotras tanto como las personas que están en estos agradecimientos me han ayudado a mí. Y a tots els que han estat membres del Brainlab, vosaltres heu fet que el Brainlab es converteixi en la familia que es ara.

Però aquesta tesis va més enllà de la recerca, així que també he de donar les gràcies a totes les persones del Departament de Psicologia Clínica i Psicobiologia per compartir la docència amb mi. En primer lloc al Marc Via. Tu m'has ensenyat a ensenyar. M'has ensenyat com transmetre els coneixements i has estat la nota "bio" dins de tota la bogeria "psico". Gràcies per donar-me la oportunitat de seguir en una poyata i de donar les classes amb les que realment em sento còmode. Gràcies per deixar-me motivar durant les classes i ensenyar-me el camí per ser una bona docent. I les gràcies també van per tu Imma, que ets la mamà del departament. Gràcies per recolzar-me quan era una simple estudiant de màster i enviar-me mil i una convocatòries de beques perquè em pogués quedar. Gràcies per tot l'amor i suport en tots els moments. Gràcies a la Mª Àngels, a la Roser, a la Maria, a l'Ana, a la Montse... tenir tantes dones referents en un mateix departament és estar en un departament privilegiat.

Unas gracias especiales van para mi homologo, Jonatan. Te conocí aceptando la misma beca y hemos hecho este camino juntos. Gracias por todo. Por comprenderme en cada momento, por todas las charlas terapéuticas, por ser un soporte incondicional fuera del grupo, por relativizar todos los problemas. Creo que ofuscándonos juntos con los reviewers y la burocracia es menos ofuscación. Así que, pequeño, muchas gracias por todo. Gràcies també a la Júlia i la Olga, heu estat referents tant per la part de docència com per la de recerca. Espero haver seguit la vostre estela amb el nivell que heu deixat.

The group in Barcelona has been part of this thesis, but another huge thank you has to go to Jyväskylä. Piia, thank you for hosting me in your group. Thank you for supporting me, for opening me the doors to everything. Thank you for taking care both scientifically and personally when I was completely out of my comfort zone. Thank you very much for allowing me to bring a little bit of Barcelona to the cold cold Finland. A big thank you also goes to your amazing group. Jari, thank you for putting up with me during our loooong recording MEG sessions, Juho, Elisa, Xueqiao, thank you for taking care of me and including me in absolutely everything that went on. Jarmo, thank you for teaching me all the knowledge on MEG and printing my 3D brain. Many thanks also to my helpers Pyry and Iiris, thank you for letting me boss you around and being the bestest and fastest learners. *Kiitos!*

Aquesta tesis tampoc hauria estat el mateix sense tot el suport que he rebut des de fora. Aleix, gràcies per estar absolutament sempre des de fa ja moltíssims anys. Per enfadar-te amb qui em fa enfadar, indignar-te amb qui m'indigna i alegrar-te per totes i cadascuna de les coses que em passen. Gràcies per tot, no marxis mai del meu costat sisplau, la vida es més vida tenint-te amb mi. I a tu, Neus. Tot el que et pugui dir no serà res comparat amb les gràcies que et doni. Gràcies per ser el meu esmorzar, dinar o sopar setmanal, per deixar-me compartir amb tu la vida i per donar-me suport en tots i cadascun dels moments. Gràcies per fer-me quedar en el present i no deixar-me emportar per la feina. Us estimo amb bojeria als dos.

A les meves nenes boniques, Jana, Mire, Laia, Ester, Júlia i Quel. Gràcies per tot el suport tot i no entendre absolutament res del que feia. Gràcies per els mil dies esquiant, per distreure'm i portar-me a passejar i per obligar-me a disfrutar de la que és la meva passió. A tots els meus companys del CANM, sense l'esqui no tindria la estabilitat mental necessària per escriure això.

A la meva família extesa – Laura y Ranya – por ser una hermana más y darme mi primera sobrina. Laura, Vero, Andrés, Toni y Juan. Gracias por todas las cenas y vermuts y por disfrutar conmigo de todos los fines de semana. I al trio de hockey, Irene, Basti i Ferrer. Gràcies per tots els dinars i per tots els moments, per estar cada setmana a la grada i assegurar-me que "avui farà bo". A tots vosaltres, gràcies per obligar-me a sortir de festa i donar-me dies de ressaca, per carregar-me les piles i fer-me sempre enveja de que us quedeu a Puigcerdà amb la fresqueta mentre jo moro de calor a Barcelona. Gràcies per tots els sopars de fútbol, els dies de hockey i per estar sempre allà.

A la Clàudia, l'Aurora i el Guille, no sé que hauria fet a Finlàndia sense vosaltres. Gràcies per convertir-vos en la meva família allà, gracias por todas y cada una de las tardes, cenas, comidas, desayunos, meriendas y momentos. Sois mi Finlandia.

Els meus agraïments especials són per la meva família, no estaria aquí si no fos per ells. Papa, Mama, Gràcies per fer-me ser com sóc. Gràcies per ensenyar-me que és la ciència des de petita, per deixar-me experimentar i preguntar-ho tot. Gràcies per ser un suport incondicional en tots els moments de la meva vida. Per aguantar-ho tot. Us podria escriure mil pàgines però ja ho sabeu. Us estimo i no podria estar més orgullosa dels dos pares que tinc, sou el més gran exemple a seguir. A les que són el meu món, les meves dues petites. Clàudia, Laura, us ho dic sempre, però si els papes m'han ensenyat tot, vosaltres m'heu donat una lliçó de vida. Quan veu arribar vaig haver de ser més espavilada i molt millor persona, perquè tenia a dos petits monstres que em seguien els passos. Ara ja no sou petites (seguiu sent monstres) i estic orgullosíssima de les persones en les que us heu convertit. Gràcies per fer-me emprenyar i riure, per tots i cadascun dels moments. Al Joan, Carol, Maria, Miquel, Marta i Sam. Gràcies per tot. I també gràcies a tu avia. Ja feia molts anys que em tractaves de doctora i l'últim regal que em vas fer va ser un bolígraf dels bons perquè "una doctora no pot firmar amb un boli mossegat". Creu-me quan et dic que els papers que acompanyen aquesta tesis van firmats amb aquest bolígraf. A tots vosaltres, Gràcies. Us estimo.

I finalment a tu, Alex. Gràcies per ser el meu millor amic des de fa innumerables anys. Gràcies per ser el meu suport i el meu company. Gràcies per estar sempre, per alegrar-te amb totes les meves alegries i alegrar-me quan el dia em pot. En resum... gràcies per estimar-me i treure'm un somriure

a cada moment, gràcies per ser la meva persona i haver-me acompanyat en tota aquesta aventura. Gràcies per no entendre res i no deixar-me treballar els caps de setmana, per ajudar-me a trobar el balanç en la meva vida. Aquesta tesis és tant meva com teva, perquè sense tu totes i cadascuna de les paraules que venen a continuació no serien així. T'estimo, ja ho saps.

Aquesta tesis no hauria estat possible sense totes les persones que han col·laborat en els estudis i m'han cedit el seu cervell. Moltíssimes gràcies.

I finalment a tu, que estàs llegint aquestes paraules. Si estàs aquí és que t'interessa així que gràcies per dedicar-li un moment a aquesta tesis.

ABSTRACT

The auditory scene that we face during our day life is highly complex. The human auditory system is able to allow us to maintain a conversation with another person whilst ignoring the surrounding sounds but, at the same time, keeping track of what is happening to detect unexpected sounds that can be critical for survival. This suggests that whilst listening, there is an ongoing storage of information about the sounds we have already heard and how they relate to each other, thus allowing the auditory system to form expectations at different levels of complexity about what is going to come. Indeed, repetitive stimulation has been shown to reduce auditory neural activity in the human cerebral cortex and this neural activity that represents immediate or remembered features of a sensory stimulus can be used as evidence when making simple perceptual decisions. Yet, before reaching the auditory cortex, incoming auditory information is deeply processed by nuclei in the subcortical ascending auditory pathway.

In a series of three studies, we recorded the auditory frequency–following response (FFR) to study the contribution of the subcortical auditory pathway to sound encoding and processing. The FFR to periodic complex sounds provides a non-invasive measure of the neural transcription of sounds, as well as how auditory experiences transform these representations. Although it has been considered as a correlate of subcortical sound encoding, recent studies challenged this assumption, demonstrating that FFR receives major contribution from the auditory cortex.

The objective of the present PhD thesis is to investigate how stimulus statistics and temporal predictability modulate regularity encoding in the subcortical auditory pathway and how the encoding strength of sounds in this pathway influences the latter making of simple auditory perceptual decisions. Additionally, we aimed to further characterize the FFR by means of electroencephalography and magnetoencephalography to understand the role of the frequency of the eliciting stimuli and disentangle the anatomical contribution of the FFRs elicited to sounds of different frequencies.

Together our findings support the view that regularity encoding spans across the auditory hierarchy. Going a step further, temporal predictability and the frequency of the incoming stimulation also affect the subcortical sound encoding, which is reflected in the making of latter simple auditory perceptual decisions. Indeed, the frequency is a crucial parameter, as the cortical contribution to the FFR is not observable when the frequency of the sounds is around 300 Hz. Overall, we conclude that the subcortical auditory pathway has an active role in the perception and processing of the incoming sounds, consistent with the hypothesis of a distributed network for perceptual organization. Additionally, although the FFR has a multi-generator nature, it can still be used as a window into human subcortical sound encoding when using the appropriate stimulus parameters.

TABLE OF CONTENTS

FOREWORD

This thesis is presented to obtain the Degree of Doctor by the University of Barcelona (International doctor mention) and is the result of the work carried out at the Brainlab – Cognitive Neuroscience Research Group, led by Dr. Carles Escera (Excellence Research Group established by the Generalitat de Catalunya SGR2014-177 and SGR2017-974), at the department of Clinical Psychology and Psychobiology of the Faculty of Psychology, University of Barcelona (Barcelona, Spain); and the work performed in a 3-month research visit at Active Mind Lab, led by Dr. Piia Astikainen, located in the department of Psychology, University of Jyväskylä (Jyväskylä, Finland).

This work was supported by the Spanish Ministry of Economy and Knowledge: Plan Nacional de I+D+i 2013 – 2016 (PSI2015-63664-P), the grant 30/12 from the Bial Foundation (Portugal) and a grant from the University of Barcelona (Ajuts de Docència i Recerca; ADR-UB). Research was also founded by the ICREA Acadèmia distinguished professorship awarded to the PhD thesis supervisor, Dr. Carles Escera.

This thesis includes three studies that have been or are in process of being published by the PhD candidate in international scientific journals.

STUDY I

Gorina-Careta, N., Zarnowiec, K., Costa-Faidella, J., & Escera, C. (2016). Timing predictability enhances regularity encoding in the human subcortical auditory pathway. Scientific Reports, 6, 37405. http://doi.org/10.1038/srep37405

STUDY II

Gorina-Careta, N., & Escera, C. (2018). Subcortical sound encoding modulates simple auditory perceptual decisions. Scientific Reports *(under review)*

STUDY III

Gorina-Careta, N., Kurkela, J., Hämälainen, J., Astikainen, P., & Escera, C. (2018). Understanding the Frequency-Following Response and its generators to sounds of different frequencies: an MEG study. *(In preparation)*

GLOSSARY OF ABBREVIATIONS

ABR	Auditory Brainstem Response
ANOVA	Analysis of Variance
AEP	Auditory Evoked Potential
CN	Cochlear Nucleus
EEG	Electroencephalogram
EOG	Electrooculogram
Fo	Fundamental Frequency
FFR	Frequency – Following Response
FFT	Fast Fourier Transform
fMRI	Functional Magnetic Resonance Imaging
FP	Frontal Pole
HPI	Head Position Indicator
IC	Inferior Colliculus
LLR	Long Latency Response
MEG	Magnetoencephalography
MGB	Medial Geniculate Body
MLR	Middle Latency Response
MMN	Mismatch Negativity
MNE	Minimum-norm estimate modelling
OP	Occipital Pole
PAC	Primary Auditory Cortex
ROI	Region of Interest
RS	Repetition Suppression
RT	Reaction Time
SOA	Stimulus Onset Asynchrony
SSA	Stimulus Specific Adaptation

LIST OF FIGURES

CHAPTER 1

INTRODUCTION TO SOUND ENCODING AND PROCESSING ALONG THE AUDITORY HIERARCHY

Every day we deal with a continuously changing and highly complex environment and our brain has to keep track of it up to every millisecond. If I would ask you which sense you would consider the most important, most probably you will say the sight, as it allows us to move without crashing, recognize people, places and objects and basically see the beauty of everything. But consider for a moment the hearing. Thousands and thousands of sounds reach our ears and if you start listening, really listening for a second, you will realize the amount of sounds that you can perceive and the information these sounds carry, even from events that we can't see. Making sense of sound is fundamental to our everyday life. So, how does the brain disentangle what is relevant from what is irrelevant? How does the brain manage to follow a conversation and not be distracted but at the same time notice an annoying quiet mosquito near your arm?

The auditory system: a short anatomical overview

In physical terms, a sound is some pressure changes in the air or other medium (Goldstein, 2009; Schnupp et al., 2011). When the sound waves enter the ear canal and push the eardrum, these pressure changes are transmitted as vibrations through the middle ear to the cochlea – the inner

ear structure responsible for encoding the sounds as neural signals. The anatomy of the cochlea allows the conversion of the mechanical vibration into a pattern of electrical excitation that is encoded by sensory neurons. After leaving the cochlea, the auditory nerve fibers join the vestibulocochlear nerve and terminate in the cochlear nucleus (CN), where they synapse with neurons projecting to the superior olivary complex. From there, neurons project through the ascending auditory pathway via the lateral lemniscus to reach the inferior colliculus (IC) and then the medial geniculate body (MGB) of the thalamus, the last subcortical station of the auditory pathway. From the thalamus, neurons target the primary auditory cortex (PAC), where the acoustic features processed throughout the auditory pathway are integrated and inform perceptual decisions (Goldstein, 2009; Pannese et al., 2015).

This PhD thesis is composed by three studies and aims at understanding the contribution of the subcortical auditory pathway to the encoding and processing of incoming auditory stimuli in the human adult brain.

Regularity encoding and predictive coding

When listening, the sounds of the acoustic environment reach our ears in the form of an auditory stimulus that is a complex mixture of the different sound sources we can perceive in a given moment (Griffiths and Warren, 2004). In order to encode and interact with this dynamic changing environment, the auditory system needs to transform this complex time-varying acoustic waveforms into a perceptual representation of one or more auditory objects (Bizley and Cohen, 2013). Even though this acoustic information we perceive is highly dynamic, it keeps some essential invariant properties which allow the auditory system to extract and encode the relationships between discrete acoustic events, and use these short-term predictive representations to organize the acoustic background into meaningful perceptual objects (Winkler et al., 2009; McDermott et al., 2011). Therefore, auditory objects are the computational result of the auditory system's ability to detect, extract, segregate and group the spectrotemporal regularities in the acoustic environment into stable perceptual units (Bregman, 1990; Winkler et al., 2009; Schnupp et al., 2011).

Neural computations and processes that mediate auditory perceptual decisions are found in the ventral auditory pathway (for review see Bizley and Cohen, 2013), a pathway of cerebral regions that includes the PAC and its prominent connections, the middle lateral and anterolateral belt regions of the auditory cortex (Hackett, 2011). Owing to a line of neurophysiological studies in rhesus monkeys (Rauschecker and Scott, 2009; Tsunada et al., 2015; Cohen et al., 2016) and functional Magnetic Ressonance Imaging (fMRI) studies in humans (Patterson et al., 2002; Warren and Griffiths, 2003), there is a broad agreement that auditory information is organized and processed hierarchically throughout this ventral auditory pathway. Early stages in this pathway encode acoustic features relevant to stimulus identity, and become increasingly sensitive to more complex stimulus features and their relationships between the core and the belt regions of the auditory cortex. At later stages of this pathway, in the ventrolateral prefrontal cortex (vLPFC), the information extracted from the auditory stimulus informs perceptual judgements, finally leading to behavioural actions (Rauschecker and Tian, 2000; Rauschecker and Scott, 2009; Tsunada et al., 2015; Cohen et al., 2016).

Nevertheless, the auditory environment is continuously changing, so actual theories of auditory perception argue that auditory objects are not static representations stored in short-term memory but rather dynamic percepts encoded in an ongoing storage of information about the sounds we have already heard and how they relate to each other, thus allowing the auditory system to not only organize the acoustic background into meaningful percepts but also to derive predictions conveying multiple levels of complexity about future sensory events (Friston, 2005; Winkler et al., 2009) and to automatically detect deviant events which do not match such predictions (Winkler, 2008; Bendixen et al., 2012).

A broadly used approach to examine whether the spectrotemporal regularities in the acoustic environment have been encoded comes from studies on deviance detection (Escera and Malmierca, 2014; Escera et al., 2014; Malmierca et al., 2014), in which low-probability ("deviant") sounds are presented amongst high-probability ("standard") sounds. By using this approach, typically referred to as the *oddball* paradigm, and using the electroencephalogram (EEG) to record neuroelectric brain signals, it was observed that deviant sounds elicit a typical response termed mismatch negativity (MMN; Näätänen et al., 1978, 2007). The MMN is a negative long

latency range (LLR) auditory evoked potential (AEP) which peaks at about 100–250 ms after the stimulus onset and is originated in the supratemporal and prefrontal cortices (Deouell, 2007). On the other hand, regularities coming from the standard sounds are not defined by the mere stimulus repetition but by the contingency between successive discrete sounds. In fact, using complex oddball paradigms, it has been demonstrated that the encoding of auditory regularities does not occur only for simple acoustic feature repetitions (e.g., frequency, intensity, or duration; Leung et al., 2012) and complex discrete stimuli (e.g., speech sounds) (for review see Näätänen et al., 2007), but also for complex contingencies between single auditory events, such as the frequency relationship between two tones within a pair or the combination of two sound features (e.g., pitch and duration) (Paavilainen et al., 2007; Bendixen et al., 2008). In this context, MMN is frequently interpreted in terms of predictive coding, assuming that the brain does not respond passively to incoming sounds but learns the preceding sound regularities and uses that knowledge to actively predict what should be expected next (Baldeweg, 2007).

Indeed, hierarchical predictive coding models have been proposed as a general theory of perceptual inference (Friston, 2005). These models posit two functionally distinct subpopulations of neurons, one to encode the expectations of perceptual inputs and one for the prediction error. According to these models, the predictive population builds up an internal model of the regularities within the incoming stimulation in order to form relevant predictions, so that predictions at different levels of the processing hierarchy try to explain away the prediction error on preceding levels. At the same time, the predictive error population compares the incoming input to the predictions encoded by the predictive populations of neurons. The activity of the prediction error population is transmitted to the predictive population as a feedback and this error signal is used to adjust the internal model. In this context, the mismatch negativity reflects the error signal used to adjust the internal model, therefore reflecting contextual encoding (Baldeweg, 2006, 2007; Wacongne et al., 2012).

Another approach to investigate auditory regularity encoding is by using the so-called *roving standard* paradigm (Baldeweg et al., 2004). In this paradigm, acoustic stimuli are presented in trains, consisting of a different number of stimulus repetitions which are isochronously delivered, while a particular

stimulus feature is changed in every train. By using such a stimulus arrangement, it can be analysed how regularity encoding evolves as a function of stimulus repetition, as the first tone of a train has the role of a low-probability stimulus compared with those of the previous train (deviant stimulus), whereas the last tone of a train is a high probability stimulus within that train (standard stimulus). The use of this paradigm reveals repetition suppression (RS; Desimone, 1996; for review see Baldeweg, 2006), the attenuation of neural responses to repeated stimulation, as a potential mechanism underlying regularity encoding (Haenschel et al., 2005; Costa-Faidella et al., 2011a; Recasens et al., 2015). From the predictive coding framework, it has been suggested that RS reflects a reduction of the prediction error for expected stimuli (Grill-Spector et al., 2006; Auksztulewicz and Friston, 2015). This model emphasizes the importance of contextual factors, such as the probability of a stimulus repetition (Summerfield et al., 2011) or the temporal predictability of the upcoming stimulus (Costa-Faidella et al., 2011a) on the encoding of regularities in auditory cortical areas (Recasens et al., 2015). For example, temporal predictability of the auditory input has been shown to shape predictions in auditory cortical areas, as stimuli occurring at predictable temporal intervals advance the onset of repetition positivity, a brain potential correlate of RS, when comparing predictable to unpredictable stimulus presentations (Costa-Faidella et al., 2011a; Todorovic et al., 2011).

However, while all this research was conducted on auditory cortical responses (e.g. Deouell, 2007) and all the theoretical formulation refers to the cerebral cortex (Friston, 2005), neurophysiological investigations in humans (Costa-Faidella et al., 2011b; Chandrasekaran et al., 2012; Cacciaglia et al., 2015) and animal models (Antunes and Malmierca, 2014; Pérez-González and Malmierca, 2014) provide direct evidence that regularity encoding is a ubiquitous property of the auditory system.

Regularity encoding along the auditory hierarchy: animal and human evidence

In particular, a step forward in understanding regularity encoding along the auditory system was provided by single and multi-unit recordings in

animals, which revealed that after a few number of repetitions of a particular stimulus, neurons show a reduction of their spiking rate while maintaining almost unaffected their discharge rate when a different tone is presented. This phenomenon has been termed stimulus – specific adaptation (SSA) and has been demonstrated that is a widespread property of the auditory neurons, as it has been observed both in neurons of the PAC (Ulanovsky et al., 2003; Pérez-González et al., 2005; Taaseh et al., 2011) and in subcortical neurons from the IC (Pérez-González et al., 2005; Malmierca et al., 2009; Duque and Malmierca, 2015) and the MGB of the rat and the mouse (Anderson et al., 2009; Antunes et al., 2010).

Although it was originally proposed that subcortical SSA would emerge in the PAC and propagate to the subcortical nuclei in a top-down manner (Nelken and Ulanovsky, 2007), the genuine role of the IC and MGB in regularity encoding was demonstrated as the SSA levels and their temporal dynamics were unaffected during deactivation of the auditory cortex by cooling (Antunes and Malmierca, 2011; Anderson and Malmierca, 2013). Therefore, it was proved that neurons showing SSA integrate sensory information to create a predictive model of the stimulation, detecting deviant features in the environment. Nevertheless, according to predictive coding models, SSA could also be reflecting a reduction of the prediction error for expected stimuli, but was still uncertain if they could also account for the enhancement of responses to sensory inputs that deviate from strong predictions (Garrido et al., 2009). A recent study demonstrated that differential responses to deviant and standard tones in oddball sequences indeed reflect active predictive activity and not simply SSA in single neurons, and that this predictive activity follows a hierarchical pattern that emerges from subcortical structures to auditory cortices (Parras et al., 2017).

These findings in animal studies led to an increase in the number of studies aiming at identifying the earlier correlates of repetition effects and deviance detection in the human auditory system, especially in the range of auditory middle–latency responses (MLRs). MLRs are a series of characteristic waveforms elicited to discrete auditory stimuli in the latency of 12 to 50 ms post stimulus onset. They are labeled as No, Po, Na, Pa, Nb, and Pb and represent the earliest cortical responses to a sound. Specifically, the earliest components (No and Po) are thought to be generated in auditory thalamocortical loops (Picton, 2011) and later ones generated in PAC (Na

and Pa) or beyond (Yvert et al., 2001). By using a well-controlled oddball paradigm, significantly different MLR responses have been observed to standard and deviant sounds in latencies as short as 40 ms from sound onset for changes in simple sound features such as tone frequency or location (Althen et al., 2011; Grimm et al., 2011, 2012; Cornella et al., 2012; Grimm and Escera, 2012; López-Caballero et al., 2016). However, in contrast to what can be seen in the LLR, complex types of auditory regularities, such as feature conjunction (Althen et al., 2013) or tone alternation (Cornella et al., 2012) did not elicit a deviance-related response in MLRs, thus suggesting that the auditory system is organized in a hierarchical manner so that complex regularities require a deeper processing in higher levels of the auditory hierarchy (Grimm and Escera, 2012; Escera and Malmierca, 2014; Malmierca et al., 2014).

This hypothesis was tested in a study using magnetoencephalography (MEG), were it was observed that local regularities elicited changes in the MLR and early MMN but global regularities, which require a deeper processing, would only elicit deviant responses in the MMN range (Recasens et al., 2014). Therefore, it was confirmed that different processing stages involved in the encoding of auditory representations, and the subsequent detection of its violations engage anatomical areas which are hierarchically organized in the human auditory cortex. These different processing stages have been also functionally dissociated using EEG, as it has been observed that early thalamocortical networks of the auditory pathway are capable of encoding regularities, but encoding the deviance only takes places in higher cortical areas (Cornella et al., 2013; Aghamolaei et al., 2016; López-Caballero et al., 2016). Taken together, these findings from animal and human studies demonstrated that regularity encoding is a pervasive property of the entire auditory system, spanning from early thalamocortical structures to higher-order levels of the auditory cortex.

Nevertheless, animal studies had gone a step further, not only demonstrating the involvement of thalamic structures in regularity encoding but also showing SSA occurring in the neurons of the IC. In humans, the involvement of the subcortical stations of the auditory pathway in regularity encoding was demonstrated using event-related fMRI during a frequency oddball paradigm (Cacciaglia et al., 2015), were the encoding of regularities and deviance detection was reported to occur bilaterally on the IC, MGB and

the auditory cortical areas. However, the fMRI lacks temporal resolution to study how the subcortical auditory pathway is activated and to disclose if the hierarchy of auditory processing includes the subcortical structures or the activation seen is due to a top-down modulation which, in the predictive coding framework, would be the predictions of the upcoming stimulation propagating top-down from the auditory cortical structures.

CHAPTER 2

THE FREQUENCY – FOLLOWING RESPONSE

In humans, the contribution of the auditory subcortical stations to regularity encoding is currently studied by means of EEG and MEG, as they both provide the best temporal window into human brain function. Auditory Brainstem Responses (ABRs) are time-locked neural responses to sound that are recorded from the scalp using EEG. ABRs to periodic and complex auditory stimuli, such as speech sounds or music, have two main characteristics: a transient response and a sustained frequency - following response (FFR; Moushegian et al., 1973; Skoe and Kraus, 2010a). The transient responses are a set of fully visible responses with sharp onset generated in the first 10 ms after stimulus onset. They are characterized by fast response peaks lasting fractions of milliseconds and are evoked by brief stimulus features, such as the onset of sounds.

On the other hand, the FFR emerges at circa 7–15 ms from sound onset and reflect synchronous and sustained neural phase-locking to the spectral and temporal periodic characteristics of the eliciting acoustic signal (Skoe and Kraus, 2010a; Kraus et al., 2017) in the range of 100 to 1500Hz approximately (Galbraith et al., 2000b; Picton, 2011). It is theorized to be an aggregation of phase-locked neural activity from multiple generators within the auditory system, and it has been treated as a putative measure of subcortical auditory encoding (Chandrasekaran and Kraus, 2010; Bidelman, 2018). The FFR is highly sensitive to context-dependent contingencies (Chandrasekaran et al., 2014; Skoe et al., 2014) and to real-time statistical properties of the stimulus (Skoe and Kraus, 2010b; Escera, 2017), and it provides a non-invasive measure of the neural transcription of the sounds as well as how

short-term auditory training and auditory experiences transform their representation in a subcortico-cortical auditory network (Kraus and White-Schwoch, 2015; Kraus and Slater, 2016). When analyzed properly, the FFR becomes an objective indicator of the measurements of the fundamental acoustic features intrinsic to speech sounds, including timing (onsets), pitch (fundamental frequency, fo) and timbre (the harmonics information). In particular, it provides information about the latency and amplitude of the auditory input in the time domain; and the magnitude of the fundamental frequency and its harmonics in the frequency domain (Skoe and Kraus, 2010a; for review see Kraus et al., 2017).

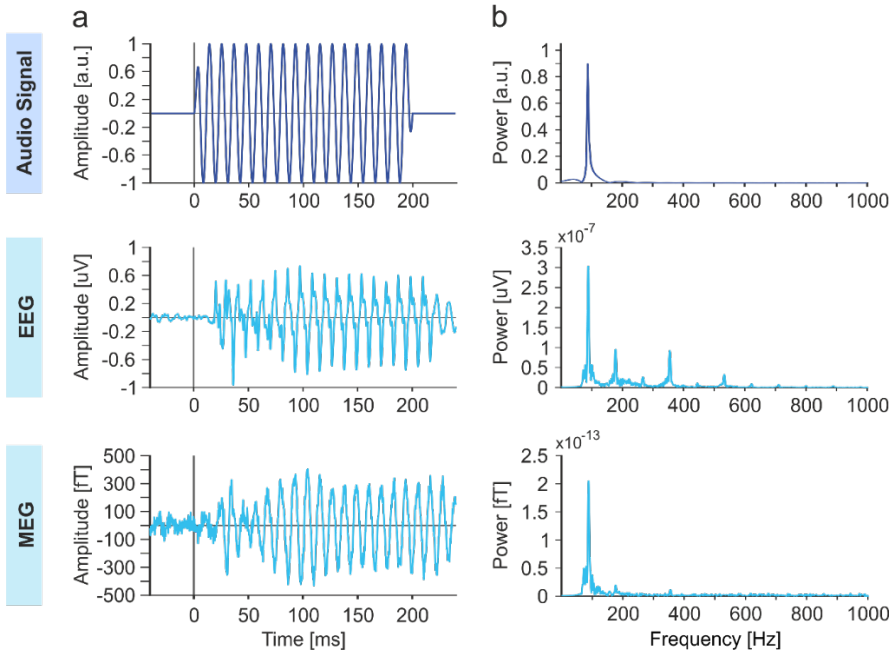


Figure 1. **Audio trace and the corresponding EEG and MEG Frequency – Following Responses.**

(a) Time course of a pure sinusoidal tone of 89 Hz and the EEG and MEG frequency – following responses recorded to the tone, displayed in the second and third lines respectively. (b) The corresponding spectra of the audio signal and the periodic portion (from 30 to 210 ms) of the recorded FFRs. FFRs are observable using both techniques and are faithfully phase-locked to the temporal (a) and spectral (b) periodic characteristics of the eliciting acoustic signal. Image: Original Figure created by the author.

Characterizing the Frequency – Following Response (FFR)

One of the mentioned features of the FFR is its sensitivity to context-dependent contingencies and to real-time statistical properties of the stimulus. A set of studies demonstrated that the FFR is able to encode for the rapid statistical features of the incoming stimulation, disclosing the encoding of regularities on the subcortical stations of the auditory hierarchy. In particular, it has been observed that when presenting in a repetitive manner a five-tone melody that features a note repetition, the amplitude of the second harmonic of the FFR is enhanced for each note between the first and second halves of the recording session. Additionally, an enhancement of these same harmonic is also observed for the note repetition within the melody, thus suggesting that the subcortical pathway can encode for both global and local statistical regularities within the ongoing stimulation and that regularity encoding mechanisms might be involved when an auditory object must be separated from background noise (Skoe and Kraus, 2010b).

Going a step further, by using an oddball paradigm with consonant-vowel stimuli, it was observed that the FFR is not only enhanced for local regularities, but also that there is a reduction of the second harmonic of the FFR response when a deviant event occurs (Slabu et al., 2012). These findings were replicated and extended, showing that pitch tracking accuracy, measured by autocorrelograms, was enhanced for standard stimuli compared to deviant and, therefore, supporting the role of the auditory brainstem in extracting statistical information from the acoustic background (Skoe et al., 2014). In addition, after a training program in which participants learned to discriminate the pitch changes that differentiated the standard and the deviant stimuli, it was observed that the relationships between deviant and standard responses varied, thus demonstrating that context-dependent contingencies and learning-dependent plasticity interact in the auditory brainstem (Skoe et al., 2014).

This neural sensitivity to stimulus statistics generalizes to more ecologically valid conditions, in which sound patterns are embedded within a single uninterrupted sequence. This was demonstrated in a number of studies in which a series of musical notes were presented in random or patterned sequences. In the patterned sequences, the occurrence of a tone predicted with high accuracy the following one. By using this paradigm,

attenuated responses were obtained for the patterned condition compared to the random one, and the more enhanced were the subcortical responses to the patterned condition to the random one, the greater was the individual capability to learn the sequence (Skoe et al., 2013). This sensitivity to stimulus statistics is biased by prior experience and the expectations arising from this experience (Skoe et al., 2015). Interestingly, this sensitivity to the contingencies of the incoming stimulation is not exclusive of adults, but is already seen in children, where an enhanced amplitude was observed in the second and fourth harmonics of the transient part of the FFR response to repetitive stimulation when compared to a variable one but only in children with good reading skills (Chandrasekaran et al., 2009). These results indicate that the human subcortical pathway of the auditory hierarchy is sensitive to ongoing stimulus context and that prior experience can modulate the responses to the incoming sounds. The first study of this PhD thesis aimed at investigating whether the context-dependent contingencies of the ongoing auditory input could modulate the stimulus statistics encoding of in the subcortical auditory pathway. Going a step further, the second study of the thesis is set to investigate how the encoding strength of sounds in this subcortical pathway influences the making of subsequent auditory simple cortical perceptual decisions.

Evidence for experience-dependent plasticity has also been provided by the results of short-term training studies, in which the FFR responses are recorded above and after a period of training (for review see Carcagno and Plack, 2017). For example, as mentioned before, the context-dependent contingencies interact with the effects of short-term training on the accuracy of the encoding of the fundamental frequency (F0) of sounds (Skoe et al., 2014) and, as a result of short-term F0 discrimination training, it has also been observed an improvement of the bilingual robustness of the subcortical temporal encoding (Carcagno and Plack, 2011). FFR plasticity has also been investigated after the training on the identification of lexical tones (Song et al., 2008; Chandrasekaran et al., 2012), as well as using general speech-in-noise training protocols (Song et al., 2012). By using this latter ones, it has been demonstrated that after training the subcortical encoding of temporal information is improved. The finding that subcortical auditory processing is not static but can be manipulated by training led to the hypothesis that sensory deficits caused by degraded sound processing could be improved by training. Indeed, it was observed that auditory training can alter the

preconscious neural encoding of complex sounds by improving the neural synchrony in the auditory brainstem in children with learning disabilities (Russo et al., 2005).

Nevertheless, the FFR is not only modulated by short-term auditory training but also by different auditory experiences, such as language experience or musical training. The first study on the influence of musical training on the FFR was conducted by Musacchia and colleagues (Musacchia et al., 2007), where they demonstrated that musicians have earlier and larger brainstem responses than non-musicians to both speech and musical stimuli presented in auditory and audiovisual conditions. Their work was extended and it was observed that musicianship enhances the FFR tracking of pitch contours (Wong et al., 2007), and that this experience-dependent plasticity of brainstem responses is shaped along the dimensions that are the most behaviorally salient for the listener. In particular, in this second study, musicians were compared to tone language speakers and it was observed that although both had stronger tracking of tone contours and musical stimuli, musicians had enhanced tracking for the musical stimuli and the tone language speakers had enhanced tracking of Chinese sounds (Bidelman et al., 2011a). Musicians also have a more robust subcortical representation of the acoustic stimulus in the presence of noise (Parbery-Clark et al., 2009) and enhanced encoding of speech syllables presented in a predictable condition relative to a variable condition than non-musicians (Parbery-Clark et al., 2011), thus leading to the hypothesis that subcortical regularity encoding is shaped by musical training and may contribute to the musicians enhanced speech-in-noise perception. Interestingly, the neural changes produced by the musical training during childhood are retained in adulthood, as the magnitude of the FFR correlates with how recently the training ceased (Skoe and Kraus, 2012).

Language experience is another factor that influences the encoding of sounds in the subcortical auditory structures. Bilingual experience enhances the neural responses to the fundamental frequency of sounds (Krizman et al., 2015; Skoe et al., 2017), as well as the subcortical representation of pitch-relevant information (Krizman et al., 2012) and neural consistency, which correlated with both a better attentional control and language proficiency (Krizman et al., 2014). In addition, long-term experience with a tone language (such as Mandarin) sharpens the tuning characteristics

of neurons along the pitch axis with enhanced sensitivity to linguistically-relevant, rapidly changing sections of pitch contours (Krishnan et al., 2008). In summary, neural encoding of sounds in the subcortical auditory pathway, studied by means of the FFR, is shaped by long-term experience with language or music, thus revealing that early sensory processing can undergo experience-dependent plasticity.

Overall, these findings establish the FFR as a stable window into neural transcription of sounds that can be obtained under passive and active listening paradigms and allow the study of how the encoding of sounds is modulated depending on context-dependent contingencies and experience-dependent plasticity. Consequently, the study of neural sound encoding using the FFRs has become a critical tool to evaluate the neural encoding of speech in clinical populations. The evoked responses for the three fundamental acoustic features intrinsic to speech sounds are shown to be inefficiently encoded, reduced or delayed in different ways for distinct clinical populations compared to typically developing controls but, overall, they all lead to a weakness in the neural processes that are important for the correct auditory processing of the auditory signal. In particular, reduced representation of the fundamental frequency and the harmonics or delayed onset of the FFRs have been observed for children with learning problems (Cunningham et al., 2001) and/or with language deficits and reading disorders such as dyslexia (Banai et al., 2005, 2009; Banai and Ahissar, 2006; Chandrasekaran et al., 2009; Anderson et al., 2010; Hornickel et al., 2012; Hornickel and Kraus, 2013). Additionally, neural synchrony (timing) and phase locking (frequency encoding) is also decreased in children with or autistic spectrum disorders (Russo et al., 2008, 2009).

Despite the abundance of FFR studies, its neural origins remain debated. Since seminal studies, the FFR has been assumed to originate from neuronal aggregates in caudal brainstem and midbrain structures, with the inferior colliculus as a major neuronal source, and has been treated as a putative measure of subcortical sound encoding. This midbrain origin is supported by the fact that the short-latency of the responses aligns with the latency of the first spikes in the IC (Langner and Schreiner, 1988) and the FFRs contained phase-locked activity up to 1500 Hz; which spans beyond the upper limit of phase-locking of cortical neurons (~100 Hz; Aiken and Picton, 2008; see next section for further details). Additionally, the cryogenic cooling of the

IC results in the disappearance of FFRs and a subsequent heating recovers the FFRs both in the colliculi and at the scalp (Marsh et al., 1970; Smith et al., 1975), and the response is eradicated with focal lesions to the IC (Sohmer et al., 1977). Nevertheless, it was suggested that a mixture of brainstem sources is indeed involved in the generation of the FFR (Chandrasekaran and Kraus, 2010; Tichko and Skoe, 2017), and this hypothesis was supported by other studies that observed weaker contributions of the IC to the FFR, with the major source on the CN (Gardi et al., 1979)accounting for an average of nearly 25% of the response amplitude. (3 or on the MGB (Weinberger et al., 1970).

However, a recent line of evidence aimed to locate the putative sources of the FFR suggested that the FFRs to an auditory stimulus of a fundamental frequency of around 100 Hz indexes the neuronal encoding of the periodic features of that sound not only in the brainstem, but also in the thalamus and mainly in the auditory cortex (Coffey et al., 2016, 2017). Yet, this cortical contribution to the FFR disappears at frequencies higher than 150 Hz (Bidelman, 2018). These findings challenge the assumption of the FFR as a correlate of subcortical sound encoding and support an emerging viewpoint in the literature that the FFR component of the ABR represents an integrated response of the entire auditory system (Kraus and White-Schwoch, 2015; Kraus and Slater, 2016). Given the multi-generator nature of the FFR and the importance of the FFR eliciting frequencies, it is highly important to understand how the frequencies are encoded throughout the auditory hierarchy to be able to disentangle the multiple sources that may be contributing to the FFR.

The frequency of sounds: a critical factor for auditory processing

The auditory system is organized tonotopically, meaning that there is a spatial distribution of neurons that respond to different frequencies, and this distribution starts at the cochlea. The cochlea is a coiled structure protected by bone and subdivided into three fluid-filled spaces by the basilar membrane. Due to its anatomy, the cochlea operates as a mechanical frequency analyzer, as different frequencies create maximal vibrations at different points along the basilar membrane (Schnupp et al., 2011).

Attached to the basilar membrane there is the primary auditory receptor structure, the organ of Corti, which is formed by the sensory receptor cells known as hair cells. Importantly, when a sound vibration propagates through, only the hair cells located at the place of maximum excitation of the basilar membrane respond and, therefore, the nerves that transmit the information from different regions of the basilar membrane already encode the frequency tonotopically. This tonotopical organization is maintained throughout the auditory pathway, from the CN to the PAC.

Additionally, as hair cells only release the neurotransmitter when depolarized, the auditory nerve fibers synchronize their spiking patterns to the temporal features of the driving stimulus by firing at a particular phase of the stimulus (Schnupp et al., 2011; Gao and Wehr, 2015). This synchronization is termed phase-locking, and it extends through all the auditory system, so that the cortical auditory system obtains information about the stimulus frequency both by the tonotopy and by the neurons phase-locking. However, the upper limit of temporal precision in phase-locked firing reduces with each ascending step in the pathway, so that the ability of neurons to follow fast modulations reduces with each ascending auditory station (Batra et al., 1989; Langner, 1992; Joris et al., 2004). Based on frequency-specific phase-locking capabilities along the auditory hierarchy, the third study of this PhD thesis aims at dissociating a hierarchy of anatomical sources contributing to the encoding of periodic stimuli of different frequencies by means of MEG, thus establishing the frequency limits by which the FFR is generated at each level of the auditory hierarchy, as well as characterizing how the FFR power is modulated as a function of the frequency of the eliciting stimulus.

Summary

Summing up, the subcortical auditory pathway, studied by means of the FFR, is sensitive to context-dependent contingencies and statistical regularities of the auditory input, and the regularity encoding mechanisms might be involved when an auditory object must be separated from background noise. This PhD thesis aimed at further studying the contribution of the subcortical pathway to the

encoding and processing of sounds, by disentangling if the context-dependent contingencies of the ongoing auditory input can modulate the encoding of stimulus statistics and if the encoding strength of the incoming sounds influences the latter making of auditory simple perceptual decisions. Furthermore, although the FFR has been used as a window into subcortical sound encoding, recent studies challenged this assumption. Based on the phase-locking capabilities of the auditory hierarchy, the third study is set to examine how the eliciting frequency modulates the FFR power and determine the anatomical contribution of the FFRs elicited to sounds of different frequencies.

CHAPTER 3

THE RESEARCH QUESTIONS

This PhD thesis aimed at examining the contribution of the subcortical auditory pathway to sound encoding and processing and to further characterize the FFR by means of EEG and MEG. Specifically, we investigated how stimulus statistics and temporal predictability modulate regularity encoding in the subcortical auditory pathway and how the encoding strength of sounds in this pathway influences the latter making of simple auditory perceptual decisions. In addition, we aimed to study how the frequency of the incoming sounds modulates the FFR power, a correlate of the encoding of these sounds, as well as the neural sources that contribute to it. This thesis is a compilation of three studies and the specific research questions of each of these are described below.

STUDY I

The first study of this PhD thesis, discussed with detail in Chapter 5, was set out to investigate whether temporal predictability in the ongoing auditory input modulates repetition suppression in subcortical stages of the auditory processing hierarchy. Previous studies proposed repetition suppression as a putative mechanism underlying regularity encoding and demonstrated that repetitive stimulation reduces auditory neural activity in animal cortical and subcortical levels and in human cerebral cortex. However, other contextual factors, such as the temporal predictability of the upcoming stimulus may influence the encoding of statistical regularities. Here we recorded the

human auditory FFR to a repeating consonant-vowel stimuli (/wa/) delivered in temporally predictable and unpredictable conditions, thus allowing us to study how the FFR is modulated both by stimulus statistics and temporal predictability. We hypothesized that FFR will be modulated both by both factors, revealing that even early neural representations of sound are sharpened by its temporal expectation of the statistical regularities.

STUDY II

The aims of the second study, presented in Chapter 6 of this PhD thesis, were to investigate the contribution of subcortical sound encoding to auditory simple cortical perceptual decisions, and to clarify a basic feature of the FFR: how its normalized power is modulated as a function of the frequency of the eliciting stimulus. Recent studies (including study I of this PhD thesis) demonstrated that the subcortical auditory system has an active role in the perception and processing of the incoming sounds, consistent with the hypothesis of a distributed network for perceptual organization. We recorded the FFRs to a set of pure tone of 20 different frequencies, along with the behavioural response times to these same sounds delivered in two blocks, one before and one after the FFR recording. By using this simple auditory reaction – time paradigm, we are able to obtain a measure of how subcortical sound encoding influences a latter perceptual encoding without the confounds of complex decision – making.

STUDY III

The third study of the present PhD thesis, which is discussed in Chapter 7, was aimed at disentangling the anatomical contribution of the FFRs elicited to sounds of different frequencies. Since seminal studies, the FFR has been considered to be a correlate of subcortical sound encoding. Yet, recent studies challenged this assumption, demonstrating that at lower frequencies (<100Hz) the FFRs reflect both cortical and subcortical activity. Based on frequency-specific phase-locking capabilities along the auditory hierarchy, in the present study we aimed to dissociate a hierarchy

of anatomical sources contributing to the encoding of periodic stimuli of different frequencies. To do so, we recorded simultaneously EEG and MEG FFRs to pure tones of 89 and 333Hz presented in a repetitive manner, thus allowing us to use distributed source modelling to analyze the contribution of the midbrain, thalamic and cortical structures to the FFR. We hypothesized that FFRs to higher frequencies would receive less cortical contribution than those to lower frequencies, hence supporting subcortical involvement for these high-frequency sounds.

CHAPTER 4

GENERAL METHODOLOGY

Participants

Studies presented in Chapters 5 and 6 were conducted at the University of Barcelona. A total of thirty students participated in study from chapter 5 (age range 19 – 27 years, mean age = 22.1, 8 males) and twelve students participated in study from chapter 6 (age range 20 – 27 years, mean age = 22.6 years, 7 males). The study discussed in Chapter 7 was hosted by the Active Mind Lab and conducted at the Jyväskylä Center for Interdisciplinary Brain Research (CIBR) from the University of Jyväskylä (Finland). A total of twenty-three students participated in the study (age range 21 – 34 years, mean age = 25.3, 4 males). All participants in the three studies reported no history of auditory, neurological or psychiatric disorders and had less than less than 4 years of musical training that ceased two or more years before the study.

All experiments were approved by the Bioethical Committee of the University of Barcelona and the Ethics Committee of the University of Jyväskylä (Study III) and were in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Written and signed informed consent was obtained from each participant before starting the corresponding experiment.

Data Acquisition

In chapters 5 and 6, FFRs were extracted from the continuous EEG recordings acquired with Neuroscan 4.4 software and Neuroscan SynAmps RT amplifier (NeuroScan, Compumedics, Charlotte, NC, USA). The EEG was recorded from 36 scalp Ag/AgCl electrodes mounted in a nylon cap (Quick-Cap; Compumedics, Charlotte, NC, USA) at the standard 10-20 system locations. Two additional electrodes were positioned at the left and the right mastoids (M1 and M2, respectively) and for study II a third additional electrode was positioned at the left earlobe (A1). The electrooculogram (EOG) was measured with two bipolar electrodes placed above and below the left eye (vertical EOG), and two horizontal electrodes placed on the outer canthi of the eyes (horizontal EOG). The ground electrode was located between Fz and FPz, and the right earlobe (A2) served as an online reference. All impedances were kept below 10 kΩ during the whole recording session and data was online bandpass–filtered from 0.05 to 3000 Hz and digitized with a sampling rate of 20 kHz.

In chapter 7, simultaneous EEG and MEG data were recorded with a 306-channel whole–head system (Elekta Neuromag® TRIUX™, Elekta Oy, Helsinki, Finland) consisting of 204 planar gradiometers and 102 magnetometers, and a compatible 64-channel EEG cap (EASYCAP GmbH, Herrsching, Germany). The EOG was measured with two bipolar electrodes placed above and below the right eye (vertical EOG), and two horizontal electrodes placed on the outer canthi of the eyes (horizontal EOG) and the ground electrode was located in the right collarbone. For the EEG recording, the right earlobe served as an online reference. Five Head Position Indicator coils (HPI-coils) were attached on top of the EEG cap; two on the forehead, two behind the ears and one on the vertex of the head. The locations of three anatomical landmarks (the nasion and left and right preauricular points) and the five HPI-coils, as well as all the locations of all the EEG electrodes and a number of additional points on the head were digitized with an Isotrak 3D digitizer (PolhemusTM, United States) before the experiment started for co-registration with the participant's anatomical MRI.

Individual structural magnetic resonance images (sMRI) were acquired from a private company offering MRI services (Synlab Jyväskylä, Jyväskylä, Finland). T1-weighted 3D images were collected on a GE 1.5 T (GoldSeal

Signa HDxt) MRI scanner using a standard head coil and with the following parameters: repetition time/echo time [TR/TE] = 540/10 ms, flip angle [FA] = 90°, matrix size = 256 × 256, slice thickness = 1.2 mm, sagittal orientation.

The MEG was recorded in 68° upright gantry position. All EEG impedances were kept below 10 kΩ during the whole recording session and both MEG and EEG data was online bandpass-filtered from 0.1 to 1660 Hz and digitized with a sampling rate of 5 kHz. To ensure that the participant's head position relative to the recording instrument was constant throughout the experiment, the magnetic fields produced by the HPI coils were measured before each block.

During recordings, participants sat comfortably in an electrically and acoustically shielded room (chapters 5 and 6) or in a magnetically shielded room (chapter 7) and were instructed to relax and watch a silent subtitled movie of their choice, whilst ignoring the auditory stimulation. Stimuli were generated and presented with Matlab v.2012a (chapter 5 and 6) or with Matlab v.2016a (chapter 7) (Matworks).

Data Analysis

In chapters 5 and 6, EEG data analysis was performed offline using EEGlab v.7 toolbox (Delorme and Makeig, 2004) running under Matlab v.2012a. The continuous EEG recordings extracted from Cz electrode were filtered offline with a bandpass Kaiser window FIR filter from 70 to 1500 Hz and epoched from 40 ms before the stimulus onset to 15 ms after the stimulus end. Trials with activity greater than 35 µV were removed from any further analysis and remaining epochs were baseline corrected to a 40 ms interval preceding the sound onset (Russo et al., 2008). Epochs from the different conditions of the two studies and for each participant were averaged separately.

To obtain the power spectral profile of the FFRs, Fast Fourier Transform (FFT; Cooley and Tukey, 1964) was applied to demeaned, zero-padded (1-Hz resolution) averages, windowed with a Hanning tapper. Further analysis vary depending on the specific experiment and are described in the corresponding section of the study's manuscript. Statistical analyses are performed using repeated–measures analysis of variance (rmANOVA)

running under IBM SPSS Statistics software (IBM Corporation, NY, USA). For all the analyses, Greenhouse–Geisser correction was applied when the assumption of sphericity was violated, and results were corrected using the Bonferroni correction to adjust for multiple testing. Post-hoc contrasts or pairwise comparisons were conducted when appropriate. Significance was defined for $p \leq 0.05$.

In chapter 7, continuous MEG data was pre-preprocessed off-line with the Elekta Neuromag™ MaxFilter 2.1 (Elekta Oy, Helsinki, Finland) Signal Space Separation (SSS) method (Taulu et al., 2004) to suppress external magnetic interference and remove static bad channels. MaxFilter software was also applied for head movement correction and transforming the head origin to the same position for each participant. MEG data was then imported to Brainstorm (Tadel et al., 2011) for further processing. Eye blink and heart beat artefacts were removed using Brainstorm's source signal projection (SSP) algorithm (Tesche et al., 1995; Hämäläinen, 2009) when the topography of the components matched those of ocular or cardiac origin upon visual inspection. The clean MEG recordings were bandpass filtered from 75 to 1500 Hz and epoched from -40 to 240 ms relative to stimulus onset. Epochs were baseline corrected to a 40 ms interval preceding the sound onset and averaged separately for each frequency condition, polarity of presentation and for each participant separately. Responses to alternating polarity stimuli were subtracted to maximize the response to pure tones (Aiken and Picton, 2008).

The signal source was estimated using distributed source models, which estimate the amplitude of a large number of dipoles distributed throughout the brain volume, but must be constrained by spatial priors. FreeSurfer (Fischl, 2012) was used to prepare the cortical surfaces and automatically segment subcortical structures from each subject's T1-weighted anatomical MRI scan. Anatomical data was later imported to Brainstorm, where precise co-registration of MEG and structural MRI data was accomplished. Thalamic and brainstem structures were then combined with the cortex surface to form a mixed surface/volume model with the deep brain activity (DBA) model, which included a triangulation of the cortical surface (~15,000 vertices), and brainstem and thalamus as a three-dimensional dipole grid (~18,000 points) (Attal and Schwartz, 2013). The head model was computed using the overlapping – spheres algorithm for each participant. A noise

covariance matrix, which accounts for the contaminants that remain present in the data after the preprocessing is complete, was computed from the 2-min empty – room recordings. The inverse solution was calculated on the subtracted polarities average for each subject and frequency condition using weighted minimum norm estimate (wMNE) source distribution algorithm with unconstrained source orientations using Brainstorm default parameters.

To disentangle the neural contributors of the FFR and obtain the signal originated in specific brain regions, bilateral regions of interest (ROIs) were defined in the main subcortical nuclei and cortical areas that conform the human auditory pathway (i.e. cochlear nucleus, CN; inferior colliculus, IC; medial geniculate body of the thalamus, MGB; and primary auditory cortex, PAC) as well as two control regions that are at the maximal distance from the target auditory regions: the frontal and occipital poles. As the head model used was a mixed surface/volume model, the ROIs are defined either as surfaces or volume depending on their location.

A time series of mean amplitude was then extracted for each ROI and for each of the three orientations in the unconstrained orientation source model for the FFR (30 to 210 ms from stimulus onset) and the baseline (-40 to 0 ms from stimulus onset) periods. To obtain the power spectral profile of the different extracted time series, FFT (Cooley and Tukey, 1964) was applied to zero-padded (1-Hz resolution) averages, windowed with a 5-ms raised cosine ramp. Orientations were summed in the frequency domain to obtain a single spectrum for each ROI, and posteriorly averaged to yield a final single spectrum for each bilateral pair of ROIs during the FFR and the baseline period. The mean normalized power in each ROI was computed using a 5-Hz-wide window surrounding the f0 of the presented stimuli for both the FFR and the baseline spectra, so that we calculated the increase of signal during FFR at fo over baseline for each bilateral pair. We then compared this increase in each auditory ROI to the average of the control regions and assessed statistical significance using Wilcoxon-matched pair tests. Results were corrected using the Bonferroni correction to adjust for multiple comparisons, so that significance was defined for $p \leq 0.01$ (0.05/4).

For further details in the specific methodology, refer to the corresponding chapters, where full explanation is given.

CHAPTER 5

THE ENCODING OF REGULARITIES IN THE SUBCORTICAL AUDITORY PATHWAY

In this Chapter we present the study related to the first aim of the PhD thesis: *to investigate how stimulus statistics and temporal predictability modulate regularity encoding in the subcortical auditory pathway.* This study was published in the journal *Scientific Reports* (Gorina-Careta et al., 2016).

Introduction

As discussed in Chapter 1, the encoding of regularities in the acoustic environment appears as a critical mechanism for auditory perception, as regularities shape perceptual objects in complex auditory scenes (Winkler et al., 2009; McDermott et al., 2011). Short–term predictive representations of acoustic regularities are derived from the probability of occurrence of repeating events, so that computed statistical regularities serve as a basis to automatically detect deviant events which do not match such predictions (Winkler, 2008; Bendixen et al., 2012). Regularity encoding has been inferred by studies on deviance detection (Escera et al., 2014; Malmierca et al., 2014), in which low–probability ("deviant") sounds are presented amongst high–probability ("standard") sounds. A more direct approach has been taken in studies measuring repetition suppression (RS), the attenuation of neural responses to repeated stimulation (Desimone, 1996; Baldeweg, 2006), which proposed RS as a potential mechanism underlying regularity encoding

(Costa-Faidella et al., 2011a) and therefore, sensory memory-trace formation (Haenschel et al., 2005; Recasens et al., 2015).

In the auditory modality, regularity encoding has been shown in human auditory cortex, as demonstrated by the modulation by probability of long– and middle– latency auditory evoked potentials (Haenschel et al., 2005; Sonnadara et al., 2006; Slabu et al., 2010; Costa-Faidella et al., 2011b; Grimm et al., 2011; Cornella et al., 2013; Escera and Malmierca, 2014) and functional magnetic resonance imaging (fMRI; Mutschler et al., 2010), as well as in subcortical auditory stages as revealed by fMRI (Chandrasekaran et al., 2012; Cacciaglia et al., 2015). Compelling evidence is provided by animal studies of single unit recordings, which have disclosed stimulus–specific adaptation in primary auditory cortex (Ulanovsky et al., 2003, 2004; Pérez-González et al., 2005) and in auditory subcortical stations, including the inferior colliculus (Pérez-González et al., 2005; Malmierca et al., 2009) and the medial geniculate complex of the thalamus (Anderson et al., 2009; Antunes et al., 2010).

From a predictive coding account, it has been suggested that RS reflects the correct prediction of the upcoming stimulus, that is, a reduction of the prediction error for expected stimuli. This model emphasizes the importance of contextual factors, such as the probability of a stimulus repetition (Summerfield et al., 2011), or the temporal predictability of the upcoming stimulus on the encoding of regularities. Yet, temporal predictability of the auditory input has been shown to shape predictions in auditory cortical areas, as stimuli occurring at predictable temporal intervals advance the onset of repetition positivity, a brain potential correlate of RS, when comparing predictable to unpredictable stimulus presentations (Costa-Faidella et al., 2011a; Todorovic et al., 2011).

The present study was designed to ascertain whether the modulation of RS by temporal predictability could be present in subcortical stages of auditory processing. For that aim, we measured the FFR (Skoe and Kraus, 2010a), a sustained component of the auditory brainstem potential that is phased-locked to the periodic characteristics of the eliciting stimulus. The FFR is highly sensitive to context-dependent contingencies (Chandrasekaran et al., 2009, 2014) and to real-time statistical properties of the stimulus (Skoe and Kraus, 2010a; Skoe et al., 2014), and has been used to show

regularity encoding and deviance detection in human auditory brainstem (Slabu et al., 2012; Shiga et al., 2015). Hence, we hypothesize that FFR will be modulated both by stimulus statistics and temporal predictability, revealing that even early neural representations of sound are sharpened by temporal expectation of the statistical regularities.

Methods

Participants

Thirty paid university students (aged 19–27 years, mean age = 22.1 years, 8 males, 3 left–handed) with no history of auditory, neurological or psychiatric disorders participated in the study. All participants lived in a Catalan/Spanish-speaking environment and all but two (Basque and Polish) had Catalan, Spanish or both as their mother language. Hearing thresholds were assessed with a standard pure-tone audiometry at the beginning of the experimental session using Bayerdynamic DT48-A headphones (Bayerdynamic GmbH & Co, Heilbronn, Germany). Mean hearing thresholds were below 25 dB SPL for the five test frequencies (250, 500, 1000, 2000 and 4000 Hz) in all the participants. As music experience is known to modulate the encoding of the F0 of complex sounds at the level of the brainstem (Song et al., 2011), all participants were enrolled with less than 4 years of musical training that ceased five or more years before the study. The study was approved by the Ethical Committee of the University of Barcelona and was in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Written informed consent was obtained from each participant before starting the experiment.

Stimuli and procedure

The auditory sequence was composed of a consonant–vowel (CV) syllable /wa/ (Slabu et al., 2012), generated with the Klatt speech synthesizer (Klatt, 1980). The syllable had a duration of 170 ms and a F0 of 100 Hz. Third (F_3), fourth (F_4), and fifth (F_5) formants were set at 2900, 3500 and 4900 Hz respectively. In order to elicit a large onset response, the first 5 ms of the CV syllable consisted of a rapid glide in the first (F_1; from 400 to 1700 Hz) and

second (F2; from 1700 to 1240 Hz) formants. After the initial 5 ms, there was a 50 ms transition in F1 from 125 to 800 Hz and in F2 from 571 to 1200 Hz.

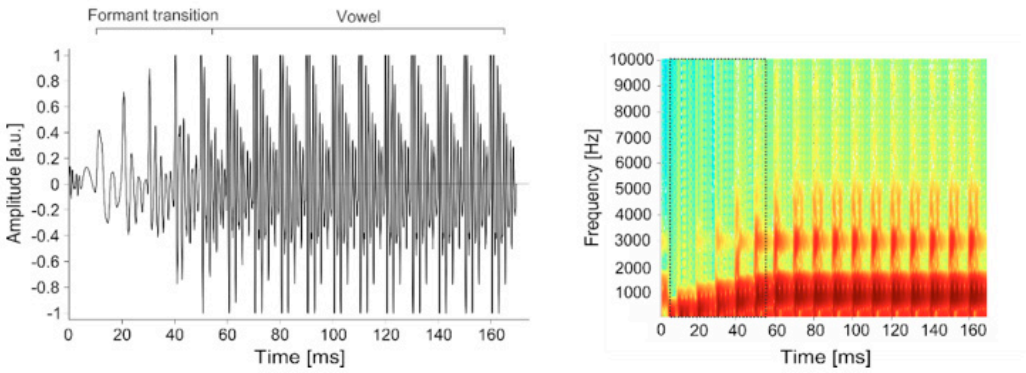


Figure 2. **Stimulus characteristics observed in the time domain and spectrogram representations**.

Participants were presented with a consonant-vowel /wa/ of 170 ms duration. As observable in the spectrogram of the left hand side of the figure, the fundamental frequency was 100 Hz, while the third, fourth, and fifth formants were set to 2900, 3500, and 4900 Hz respectively. Image: Original figure created by the author.

During the auditory stimulation with the CV syllable, a Spanish six–talker babble (four females and two males, 75 s track) was played as a background noise (10 dB SPL lower than the stimuli) in order to create a challenging listening situation (Song et al., 2011). To create the babble, speakers were recorded in a sound attenuated booth when reading in a comfortable and conversational manner semantically anomalous sentences. Tracks were acquired with 44 kHz sampling rate and 16-bit accuracy using Audacity 2.0.0 (Audacity Team® 2012). After offline root mean square amplitude normalization in Matlab v7.4 (Mathworks), all the recordings were circularly shifted and mixed together in such a way that the beginning of each speaker's track was delayed 10 s in reference to the previous speaker recording. To assure that there was no interaction between the background noise and the /wa/ stimulus, the babble was looped with no silent intervals during the experimental blocks and CV presentation was started at a random phase of the babble.

The /wa/ stimuli were presented binaurally at 75 dB SPL in alternating polarities via ER-3A ABR insert earphones (Etymotic Research, Inc., Elk Grove Village, IL-USA) in two different timing conditions: Predictable and Unpredictable. In the Predictable timing condition, stimulus onset asynchrony (SOA) was set to 366 ms. In the Unpredictable timing condition stimuli were presented with a variable SOA jittered between 183 and 549 (mean SOA 366 ms) in seven equiprobable steps of 61 ms arranged randomly. Each condition was divided into 8 blocks, each block consisting of 1001 presentations of the /wa/ stimulus. Blocks of the two conditions were presented alternately and the order was counterbalanced across participants.

During the experiment, participants sat comfortably in an electrically and acoustically shielded room and were instructed to relax and watch a silent subtitled movie of their choice, while ignoring the auditory stimulation. Pauses between blocks lasted 30 s, during which participants were allowed to move. Recording time lasted a total of two hours.

EEG recording

FFRs were extracted from the continuous EEG recording acquired with Neuroscan 4.4 software and Neuroscan SynAmps RT amplifier (NeuroScan, Compumedics, Charlotte, NC, USA). The EEG was recorded from 36 scalp Ag/AgCl electrodes mounted in a nylon cap (Quick-Cap; Compumedics, Charlotte, NC, USA) at the standard 10-20 system locations. Two additional electrodes were positioned at the left and the right mastoids (M1 and M2, respectively). The EOG was measured with two bipolar electrodes placed above and below the left eye (VEOG), and two horizontal electrodes placed on the outer canthi of the eyes (HEOG). The ground electrode was located between Fz and FPz, and the right earlobe (A2) served as an online reference. All impedances were kept below 10 kΩ during the whole recording session and data was online bandpass–filtered from 0.05 to 3000 Hz and digitized with a sampling rate of 20 kHz.

Data processing and analysis

Data analysis was performed offline using EEGlab v.7 toolbox (Delorme and Makeig, 2004) running under Matlab v.2012a. The continuous recordings extracted from the Cz electrode were filtered offline with a bandpass Kaiser

window FIR filter from 70 to 1500 Hz and epoched from 40 ms before the stimulus onset to 180 ms after the stimulus. Epochs for Predictable and Unpredictable timing conditions were sorted separately. Trials with activity greater than 35 μV were removed from any further analysis and remaining epochs were baseline corrected to a 40 ms interval preceding the sound onset (Russo et al., 2008).

Data was averaged in two different manners. To analyse the effects of temporal predictability on the FFR, epochs from each timing condition and for each participant were averaged separately (Predictable condition: mean = 7746 trials, std = 283.7; Unpredictable condition: mean = 7730 trials, std = 342.7). To analyze the effects of stimulus repetition on the FFR across time, each experimental block was divided in ten consecutive runs, each containing 100 stimulus repetitions. For each participant and condition separately, each run was averaged with the corresponding one from the other experimental blocks of the same condition. This way, we could obtain an estimation of the response based on 1000 stimulus presentations to cumulative repetition (i.e., from 1-100, 101-200, 201-300, 301-400, 401-500, 501-600, 601-700, 701-800, 801-900 and 901-1000 repetitions) for each condition separately. After artifact rejection, in the Predictable condition, 772 trials were included on average on each 100–repetition sub–average (std = 4.37), and the Unpredictable condition consisted of 770 trials per 100–repetition sub–average (std = 5.49). Responses to alternating polarity stimuli were averaged together to minimize stimulus artefact and cochlear microphonic, preserving the FFR to the stimulus envelope (Aiken and Picton, 2008).

Only the steady-state part of the FFR was analysed (65 – 180 ms), as rapid formant transitions are a perceptual challenge for the auditory system (Assmann and Summerfield, 2004). Additionally, previous studies using the consonant-vowel stimulus /da/ demonstrated that the FFR elicited by transition from the consonant to the vowel differ from the responses elicited by the steady-state vowel part of the stimulus (Chandrasekaran et al., 2009; Song et al., 2011). Therefore, as the FFR encodes better the periodic part of the stimuli, we focused on the region of the response which corresponds to the vowel steady–state part.

Fast Fourier Transform (FFT; (Cooley and Tukey, 1964) was applied to demeaned, zero-padded (1-Hz resolution) averages, windowed with a Hanning tapper. The mean response amplitude was computed using 20-Hz-wide window surrounding the F0 (90 - 110 Hz) and the subsequent five harmonics: H2, H3, H4, H5, and H6. These harmonic components were, however, not reliably present in all participants and therefore only response to the F0 was statistically analysed. Overall condition effects were assessed by means of repeated–measures ANOVA with the factor Condition (Predictable vs. Unpredictable); repetition effects in the two conditions were computed with repeated–measures ANOVAs with the factor Condition (Predictable vs. Unpredictable) and Repetition (ten 100-epoch sub–averages).

Neural Pitch Strength was quantified to analyse the magnitude of the neural phase-locking in the subcortical auditory pathway to the pitch of the stimulus waveform in both timing conditions. It was derived using a short-term autocorrelation analysis from 15 to 175 ms with 40–ms sliding window and a 1–ms step. This procedure involved cross–correlating a 40–ms frame of the response with itself and finding the height of the first peak in the autocorrelation function away from time-lag zero, which was taken as the magnitude of neural pitch strength (Boersma, 1993; Krishnan et al., 2005, 2010, Jeng et al., 2011b, 2011a). In all cases, this peak fell at a time lag of approximately 10 ms, which corresponds to the fundamental pitch period of the stimulus (i.e., frequency = 1/periodicity; e.g., 100 Hz = 1/10 ms). To account for the transmission delay of the earphones and the neural delay, the analysis bin began at 15 ms for the responses. Pitch strength values obtained from each time frame of response were Fisher-transformed and averaged, resulting in one value per each 100–epoch sub–average (ten values in total). Pitch strength on the two timing conditions was analysed with repeated–measures ANOVA with the factor Condition (Predictable vs. Unpredictable) and Repetition (ten 100-epoch sub–averages).

The Greenhouse–Geisser correction was applied when the assumption of sphericity was violated, and results were corrected using the Bonferroni correction to adjust for multiple testing. Additional Bonferroni–corrected post–hoc tests were performed to examine the direction of the effects. Significance was defined for $p \leq 0.05$.

Results

To assess temporal predictability effects on regularity encoding on the FFR, stimulus were delivered in two timing conditions. In the Predictable timing condition, stimuli were presented with a constant stimulus onset asynchrony, thus allowing a temporal prediction of the occurrence of the upcoming stimulus. In the Unpredictable timing condition, stimuli were presented with a jittered stimulus onset asynchrony so that the temporality of the upcoming stimulus could not be anticipated.

The grand–average waveforms of FFRs elicited to both Predictable and Unpredictable timing conditions are depicted in Figure 3B. As expected, the waveforms of both timing conditions resembled markedly the stimulus envelope (Figure 3A), and a small difference in the response between both timing conditions can be seen. Below we describe in detail the influence of timing predictability and the effects of repetition in these auditory subcortical responses.

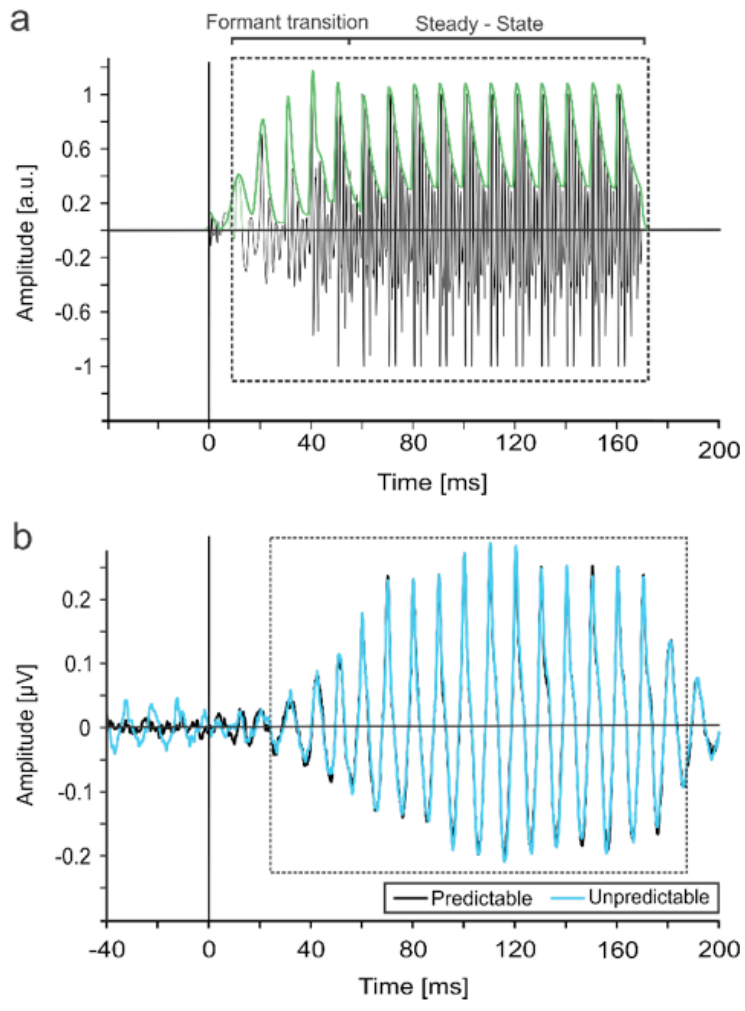


Figure 3. **Stimulus waveform and Frequency – Following Responses elicited in the two temporal conditions.**

(a) The acoustic waveform of the stimulus/wa/. The envelope of the stimulus is highlighted in green. The formant transition region and the vowel steady–state region are bracketed (a.u. = arbitrary units) (b) Grand-average FFR response recorded at Cz of all participants in the predictable (black) and unpredictable (blue) timing conditions recorded to the/wa/stimuli presented against a continuous babbling background noise. As can be seen here, the envelope of the stimulus (a, green) was preserved in the response (b) of both timing conditions. This is evidenced by the framed areas, which include the same number of cycles. Image: Figure modified from Gorina-Careta et al (2016).

When analysing the timing predictability effects of the auditory sequence on the neural response, FFRs showed a significant effect for Condition ($F(1,29) = 5.091$, $p = 0.032$, $\eta^2_{partial} = 0.149$; Figure 4A and B). The neural response to the incoming sounds had a larger amplitude when the timing was unpredictable (mean= 0.17 µV, SE = 0.08 µV) compared to when the same stimuli were presented in a predictable manner (mean= 0.16 µV, SE = 0.07 µV), thus indicating enhanced adaptation to timing–predictable repetition.

Moreover, after averaging the responses to analyse the effects of stimulus repetition across time (Figure 4C), larger FFR amplitudes were found for the Unpredictable (mean = 0.176 µV, SE = 0.014 µV) compared to Predictable timing condition (mean = 0.167 µV, SE = 0.013 µV; Condition: $F(1,29) = 5.649$, $p = 0.024$, $\eta^2_{partial} = 0.163$). Repetition effects were also statistically significant (Repetition; $F(9,261) = 3.832$, $p < 0.001$, $\eta^2_{partial} = 0.117$), indicating a decrease in the FFR amplitude as the stimulus history increased, for both timing conditions. Further post-hoc paired t-tests between repetition–averages in both conditions revealed a significant repetition effect between sub–averages 1-100 and 301-400 ($t(29) = 3.673$, $p = 0.043$), 1-100 and 401-500 ($t(29) = 5.157$, $p = 0.001$) and 1-100 and 701-800 ($t(29) = 3.609$, $p = 0.049$). There were no further significant differences in F0 amplitude between the remaining positions. The interaction between timing predictability and repetition did not reach statistical significance (Condition x Repetition: $F(9,261) = 0.684$, $p = 0.724$, $\eta^2_{partial} = 0.023$).

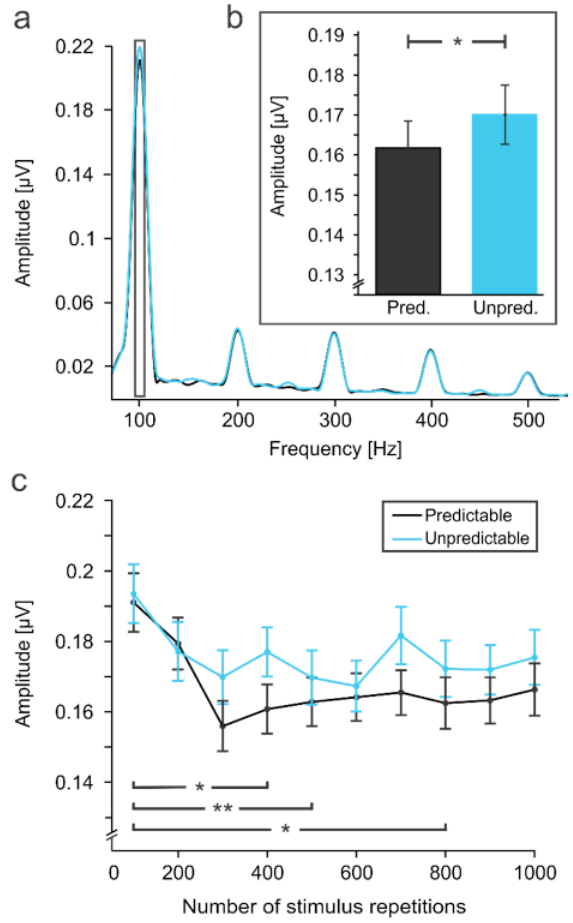


Figure 4. **FFR amplitude spectrum and mean amplitude of the fundamental frequency peak.**

(a) FFR amplitude spectrum of the steady–state part of the response in the Predictable (black) and Unpredictable (blue) timing conditions. (b) Mean amplitude of the F0 (100 Hz), computed over a 20 Hz window around the peak, is represented for both conditions. The Unpredictable timing condition yielded significantly larger amplitudes than the Predictable condition. Pred = Predictable; Unpred = Unpredictable. (c) Mean spectral amplitude of the F0 at ten consecutive 100–epoch sub–averages in both Predictable (black) and Unpredictable (blue) timing conditions. Decreased amplitude was observed in the Predictable condition compared to the Unpredictable timing condition. Also, a decrease in amplitude was observed as the number of previous repetitions increases in both timing conditions. Error bars represent ± 1 SEM. Statistically significant comparisons are marked with one ($p < 0.05$) or two ($p < 0.01$) asterisks. Image: Figure modified from Gorina-Careta et al (2016).

Pitch strength values indicated a stronger phase-locking to the stimulus F0 contour when the timing was predictable (mean = 0.792, SE = 0.045) compared to when the stimuli were presented in an unpredictable manner (mean = 0.754, SE = 0.04; Condition: $F(1,29) = 8.122$, $p = 0.008$, $\eta^2_{partial} = 0.219$; Figure 5A). Furthermore, Pitch strength showed separable patterns in the two timing conditions across history of repetitions (Condition x Repetition: $F(9,261) = 2.807$, $p = 0.004$, $\eta^2_{partial} = 0.088$; Figure 5B). When stimuli occurred with an unpredictable timing, the encoding of the overall periodicity of the signal did not change as the number of repetitions increased. However, when the stimuli were presented in a predictable manner, the initial phase-locking to the stimulus was very high, but as the number of repetitions increased, the pitch strength values decreased to the same level as the unpredictable timing condition values. Further post-hoc analysis indicated that Pitch strength values differed between conditions on sub–averages ranging 1-100 ($t(29) = 2.709$, p = 0.011), 101200 ($t(29) = 4.307$, $p < 0.001$) and 401-500 ($t(29) = 2.462$, p = 0.02).

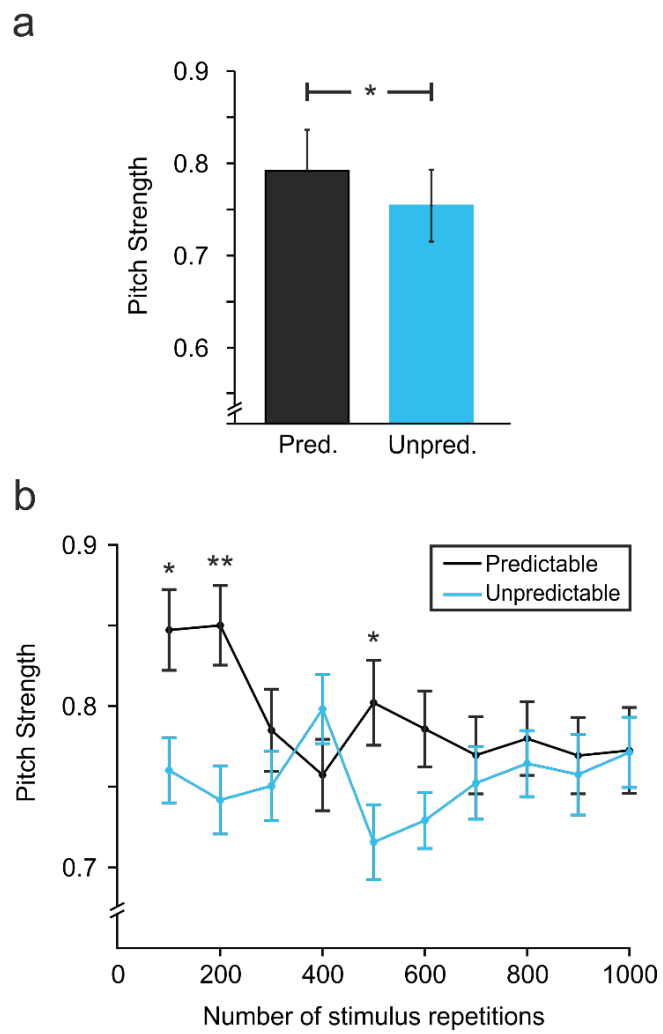


Figure 5. **Neural Pitch Strength to the pitch of the stimulus waveform in both timing conditions.**

(a) Pitch strength Fisher transformed correlation values in the Predictable (black) and Unpredictable (blue) timing conditions. Increased phase–locking to the stimulus F0 was observed on the Predictable compared to the Unpredictable timing condition. Pred = Predictable; Unpred = Unpredictable (b) Pitch strength Fisher transformed correlation values at ten consecutive 100–epoch sub–averages in both Predictable (black) and Unpredictable (blue) timing conditions. Different trends can be distinguished for both conditions as the number of repetitions increased. Error bars represent ± 1 SEM. Statistically significant comparisons are marked with one ($p < 0.05$) or two ($p < 0.01$) asterisks. Image: Figure modified from Gorina-Careta et al (2016).

Discussion

The study presented in this Chapter constitutes the first demonstration that temporal predictability enhances regularity encoding of the repetitive acoustic environment in the human auditory subcortical pathway. In particular, we have shown that the reduction of neural response caused by repetitive stimulation, although present independently of temporal aspects of the auditory input, is in fact modulated in the subcortical auditory system by the temporal predictability of the incoming stimulus. Indeed, we found a decrease in FFR amplitude when the auditory stimuli were presented with a constant presentation rate compared to when these very same stimuli were delivered at random time intervals, precluding the precise temporal anticipation of their occurrence. In addition, a general decrease on the FFR amplitude was observed as the history of stimulation increased. This effect on the FFR amplitude was clearly observed for both timing conditions, thus indicating that independently of the temporal context of the auditory stimulation, the FFR is suppressed when it faces a repetitive acoustic stimulus. Interestingly, the modulatory effects of the temporal aspects of the acoustic input on the FFR amplitude became evident only after the accumulation of 200 stimuli repetitions, when the suppression caused by the repetitions reached a plateau, causing an enhancement on the suppression when the stimuli were temporally predictable.

Our findings favour the importance of timing as a key factor in the encoding of acoustic regularities and the formation of stimulus–specific memory traces along the whole auditory hierarchy. As discussed in Chapter 1, temporal predictability of the incoming auditory stimulation has been shown to reduce the amplitude of the P50 (Schwartze et al., 2013) and N1 components (Lange, 2009; Schwartze et al., 2013) of the auditory evoked potentials, and to enhance both repetition suppression (Todorovic et al., 2011; Todorovic and de Lange, 2012) and the repetition positivity (Costa-Faidella et al., 2011a) in human auditory cortex, and has been suggested to boost the propagation of regularity encoding upstream the auditory pathway (Baldeweg, 2006; Costa-Faidella et al., 2011a). In this regard, our results expand previous findings on the role of temporal predictability on regularity encoding, by disclosing the sensitivity of the subcortical auditory pathway to temporal predictability, thus supporting

the view that the mechanisms that govern regularity encoding at cortical levels also expand to subcortical stages (Malmierca et al., 2014).

Interestingly, the effect of the temporal predictability on the subcortical auditory system that we are describing here appears as an enhancement of the repetition suppression, that is, as a pronounced reduction of the neural response to the repetitive stimulation (Desimone, 1996). Previous findings on animal studies established repetition suppression as a phenomenon that expands along the auditory hierarchy. By means of single cell recordings in anesthetized animals, it has been shown that individual neurons at both cortical (Ulanovsky et al., 2003, 2004; Pérez-González et al., 2005) and subcortical (Malmierca et al., 2009; Antunes et al., 2010; Ayala and Malmierca, 2012) levels exhibit a reduced response to a stimulus that is presented repeatedly. Repetition suppression has also been observed in the animal cortical auditory steady state responses (ASSR), as an amplitude habituation of this periodic electrical brain oscillation evoked by sinusoidally modulated acoustic stimuli (Prado-Gutierrez et al., 2015). In agreement with these animal findings at subcortical level, a recent human study described that when a stimulus feature (e.g., pitch) is repeated, the blood oxygen level–dependent (BOLD) activity can be either reduced or enhanced (Chandrasekaran et al., 2012), thus revealing that repetition suppression is a phenomenon that is not exclusive of the auditory cortex but that it can be also observed at lower stages of the auditory hierarchy. Our data confirm and expand these findings, as well as the observations from animal studies, agreeing with the emerging view that regularity encoding is a property that spans the whole auditory anatomical hierarchy, from the brainstem upwards, and in multiple temporal dimensions (Costa-Faidella et al., 2011b; Escera and Malmierca, 2014; Escera et al., 2014; Malmierca et al., 2014).

The observed sharpening of the neural representations by temporal predictability is in line with hierarchical predictive coding models (Friston, 2005; Friston and Kiebel, 2009; Wacongne et al., 2012; Wacongne, 2016). In this line, when the auditory input is temporally predictable, the input matches the prediction coming from upper levels, thus reducing the prediction error response. On the other hand, when the auditory stimulation is temporally unpredictable, there is a decrease on the prediction error due to the repetitive characteristics of the stimulation, but there is a mismatch

on the temporal expectation, leading to a repetition suppression that it is not, however, as strong as the one produced by the temporally predictable stimulation. Although the FFR has been shown to be quite insensitive to higher order perceptual processes (Bidelman et al., 2013), it is indeed modulated by stimulus regularities (Chandrasekaran et al., 2009; Skoe and Kraus, 2010b; Slabu et al., 2012; Shiga et al., 2015), which indicates that the online formation of predictive models via stimulus regularity encoding is reflected at subcortical levels despite that already established categories to interpret acoustic stimulation may not require them.

Notably, our results provide two complementary views of the effects of temporal predictability on regularity encoding in the human subcortical auditory pathway. On one side, as described above, the observed decrease on the F0 amplitude, which reflects the neural suppression underlying the encoding of regularities on the subcortical auditory pathway, as well as its modulation by the temporal predictability of the upcoming sounds. On the other side, by capitalizing on the high faithfulness of the FFR to the incoming stimulus (Skoe and Kraus, 2010a), we observed that the modulation of the early representations of regular sounds by the temporal structure of the auditory input is partially due to an increase in the robustness of the phase-locking in the auditory subcortical structures, thus indicating that the temporal predictability of the incoming stimulation increases the signal to noise ratio of the encoded repetitive stimuli. Although both findings may seem contradictory, they are, in fact, complementary, as to the periodicity of the signal contributes not only the fundamental frequency but also the whole spectral richness of the response (Schnupp et al., 2011). The increased pitch strength magnitude indicates that the response is more periodic and the phase-locking to the stimulus is more reliable (Krishnan et al., 2004; Jeng et al., 2011a), thus helping the extraction of acoustic features. As the number of temporally predictable repetitions increases, the encoding of the stimulus periodicity is reduced, revealing that whilst new predictable stimulation facilitates the neural phase-locking to the stimulus, a repeated stimulation reduces the need to represent the stimulus in a fine-grained manner. This decrease goes in parallel to the adaptation we observed on the spectral domain, where the phenomenon of repetition suppression is well described. Interestingly, the increased neural phase-locking to the incoming repetitive stimulation helps the extraction of acoustic features and aids the subcortical auditory system to better encode the upcoming

repetitive stimulation, thus making unnecessary for the auditory subcortical structures to respond strongly to the temporally predictable repetitive stimulus presentation. On the other hand, when stimuli were temporally unpredictable, there was a smaller neuronal phase-locking to the incoming stimulation but these values where stable as the stimulus history increased. Consequently, a suppression of the FFR amplitude is observed, as the stimuli are repetitive, but this suppression is reduced.

Taken together, these complementary findings led us to speculate that the temporal predictability of the upcoming stimulation may be influencing the encoding of regularities by helping the extraction of the important stimulation amongst a noisy environment. By means of this mechanism, the temporal predictability of the regular stimulation would help to extract all the features of the sounds and induce a better phase-locking of the subcortical structures to it. On the other hand, a non-temporally–predictable regular stimulation would not allow the subcortical structures to phase-lock to the auditory input as faithfully as when stimuli were predictably delivered, but as the history of stimulus presentation increases, the same early stages of the auditory hierarchy will keep extracting all the features possible from the sounds that are being presented, even if the neural response to those stimuli decreases.

Summing up, our study has shown that temporal predictability modulates the auditory FFR to a repeated stimulation, leading to enhanced repetition suppression when the incoming auditory stimuli are temporally predictable compared to when the temporality of the following sound could not be predicted. Despite this enhancement on response suppression, a temporally predictable presentation aids the encoding of the presented sounds by increasing the signal to noise ratio. Altogether, we have demonstrated that early neural representations of sounds are sharpened by the temporality of the encoded statistical regularities. Our findings add to the evidence in favour of the back–propagation hypothesis (Baldeweg, 2006), which posits that with an increasing number of stimulus repetitions, a stimulus-specific memory trace can be detected earlier on the auditory hierarchy. This hypothesis was broadened when timing was proposed to be an important variable for the formation of the aforementioned memory trace at the level of the primary auditory cortex (Costa-Faidella et al., 2011a). Crucially, our results support the view that timing is, indeed, a critical factor that affects

the formation of the stimulus-specific memory trace along the whole auditory hierarchy.

Summary

The study presented in this Chapter was designed to investigate whether temporal predictability modulates regularity encoding in the subcortical auditory pathway. By using a repetitive paradigm were consonant-vowel (/wa/) stimuli were presented in predictable and unpredictable conditions, we demonstrated that subcortical sound processing is modulated by both stimulus statistics (i.e., stimulus regularities) and temporal predictability. In particular, we observed a reduction of the neural response to repetitive stimulation in both timing conditions, thus indicating that independently of the temporal context of the auditory input, the FFR is suppressed when the acoustic stimulation is repetitive. In addition, the reduction of the neural response caused by the repetitive stimulation is enhanced in when the auditory stimulation is predictable, aiding the encoding of the presented sounds by increasing the signal to noise ratio and revealing that the subcortical auditory pathway is actively involved in the processing of incoming sounds.

The interim conclusions of this chapter are:

- Subcortical auditory pathway is sensitive to both temporal predictability and statistical regularities of the acoustic environment.
- Temporal predictability enhances regularity encoding of acoustic environment in the human subcortical auditory pathway.
- Early neural representation of sounds is sharpened by the temporality of the encoded statistical regularities.

CHAPTER 6

FROM SOUND ENCODING TO AUDITORY DECISION MAKING: THE ROLE OF THE SUBCORTICAL AUDITORY PATHWAY

In this Chapter we present the study related to the second and partially to the third aim of the PhD thesis: *How the encoding strength of sounds in the subcortical auditory pathway influences the latter making of simple auditory perceptual decisions and Characterize the FFR: how its normalized power is modulated as a function of the frequency of the eliciting stimulus*. This study is under review in the journal *Scientific Reports* (Gorina-Careta et al., under review).

Introduction

So far, we have discussed that auditory perceptual decisions rely on a complex neural processing that requires multiple computations in order to transform a time-varying acoustic waveform into a perceptual representation. In particular, it involves interpreting the incoming sensory information to detect and discriminate any auditory stimulus, and using this information to make a categorical judgement about it (Bizley and Cohen, 2013). According to evidence accumulation models (Smith and Ratcliff, 2004), simple perceptual decisions (e.g., a stimulus is present or absent) can be decomposed into three main processing stages: sensory encoding, decision formation and motor execution (Kelly and O'Connell, 2015). In this context, evidence is accumulated and integrated in favour of a particular

outcome which triggers the motor execution when reaching a threshold (Smith and Ratcliff, 2004; Gold and Shadlen, 2007; Heekeren et al., 2008; Shadlen and Kiani, 2013). In a sensory-motor task, neural activity that represents immediate or remembered features of a sensory stimulus can be used as evidence (Gold and Shadlen, 2007).

As described above, neural computations and processes that mediate auditory perceptual decisions are found in the ventral auditory pathway (for review see Bizley and Cohen, 2013), a pathway of cerebral regions that includes the PAC and its prominent connections, the middle lateral (ML) and anterolateral (AL) belt regions of the auditory cortex (Hackett, 2011). Owing to a line of neurophysiological studies in rhesus monkeys (Rauschecker and Scott, 2009; Tsunada et al., 2015; Cohen et al., 2016) and fMRI studies in humans (Patterson et al., 2002; Warren and Griffiths, 2003), there is a broad agreement that auditory information is organized and processed hierarchically throughout this ventral auditory pathway. Early stages in this pathway encode acoustic features relevant to stimulus identity, and become increasingly sensitive to more complex stimulus features and their relationships between the core and the belt regions of the auditory cortex. At later stages of this pathway, in the ventrolateral prefrontal cortex (vLPFC), the information extracted from the auditory stimulus informs perceptual judgements, finally leading to behavioural actions (Rauschecker and Tian, 2000; Rauschecker and Scott, 2009; Hackett, 2011; Tsunada et al., 2015; Cohen et al., 2016).

Yet, before reaching auditory cortex, incoming auditory stimuli are encoded in subcortical stations of the auditory pathway. A recent line of evidence has demonstrated that auditory subcortical structures are more than simple relay steps on the ascending auditory hierarchy, as they have an active role in encoding the incoming sounds (for overview see Kraus et al., 2017 and Chapters 1 and 5). In particular, evidence from stream segregation studies suggest that the subcortical auditory pathway is also involved in auditory perceptual processing (Pressnitzer et al., 2008; Shamma and Micheyl, 2010; Yamagishi et al., 2016), consistent with the hypothesis of a distributed network for perceptual organization. Thus, we hypothesize that neural processes involved in simple perceptual decisions go beyond the auditory cortex so that a subcortical contribution cannot be disregarded.

Human auditory perceptual encoding correlates of subcortical activity were recorded using EEG in the form of FFRs. Interestingly, the FFR amplitude also codes for some aspect of the auditory stimulus that correlates with simple behavioural responses to sound, as it has been demonstrated that simple motor reaction times to the auditory stimuli have a reliable correlation with FFR amplitude (Galbraith et al., 2000a). The human FFR is theorized to be an aggregation of phase-locked neural activity from multiple generators within the auditory system, and it has been treated as a putative measure of subcortical auditory encoding (Chandrasekaran and Kraus, 2010; Escera, 2017). Indeed, previous studies have observed a modulation of the FFR as a function of the stimulus frequency (Hoormann et al., 1992; Skoe and Kraus, 2012; Tichko and Skoe, 2017), confirming that whilst at lower frequencies (<100Hz) the FFRs reflect both cortical and subcortical activity (Coffey et al., 2016; Bidelman, 2018), the cortical contribution to the FFR disappears at frequencies higher than 150 Hz (Bidelman, 2018). Given the multi-generator nature of the FFR and the importance of the FFR eliciting frequencies, it is required to further characterize the FFR by confirming how the FFR power is modulated as a function of the frequency, as well as to ascertain whether auditory simple cortical perceptual decisions rely on subcortical encoding of the incoming stimuli.

Focusing on those two aims, here we recorded the FFRs to a set of 20 different tone frequencies, as well as the behavioural response times to these very same sounds presented in two separate runs: one before and one after the FFR recording. By using this simple auditory reaction – time paradigm, we are able to obtain a measure of how subcortical sound encoding influences a latter perceptual processing without the confounds of complex decision-making.

Methods

Participants

Twelve paid university students (aged 20–27 years, mean age = 22.6 years, 5 females, 1 left–handed) with no history of neurological, psychiatric or hearing impairment participated in the study. Hearing thresholds were assessed with a standard pure-tone audiometry using 5 harmonic tones of 250 Hz at the beginning of the experimental session using Bayerdynamic

DT48-A headphones (Bayerdynamic GmbH & Co, Heilbronn, Germany), with the minimum threshold requested for participation below 25 dB SPL for all tested frequencies and an interaural difference of < 10dB. As music experience is known to modulate the encoding of the fo of periodic sounds in the subcortical auditory pathway (Song et al., 2011), all participants enrolled had less than 4 years of musical training that ceased five or more years before the study. The study was approved by the Bioethical Committee of the University of Barcelona and was in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Written informed consent was obtained from each participant before starting the experiment. All registered data is available under request to the corresponding author.

Stimuli and procedure

The auditory stimuli consisted of pure sinusoidal tones of 185 ms duration, including 5 ms rise and fall times and generated with Matlab v.2012a (Matworks). Twenty different tones were used, with frequencies ranging from 121 to 577 Hz in steps of 24 Hz, and were delivered monaurally to the right ear at an intensity of ~65 dB in alternating polarities via ER-3A ABR insert earphones (Etymotic Research, Inc, Elk Grove Village, IL-USA). Sounds were calibrated to 65 dB SPL at 1 kHz using a CESVA SC102 Sound Level Meter (CESVA Instruments S.L.U., Barcelona, Spain) with frequency ponderation type A (according to IEC 61672 Class2) and temporal ponderation of 1 second. To equate the stimulus intensity across the range of frequencies used, a correction factor was added to scale the amplitude of each sine wave following the ISO 226:2003 normal equal-loudness-level contours specifications (ISO 226:2003, 2003).

The experiment consisted of a passive listening condition and a simple auditory reaction time task, both using the same stimulus frequencies. In the Passive listening condition, stimuli of all the different frequencies were randomized and presented with a variable stimulus onset asynchrony (SOA) jittered between 294 and 372 ms (mean SOA 333ms) in 2 ms equiprobable steps arranged randomly. The condition was divided in 20 blocks, each containing 1000 stimuli. In the simple auditory reaction time task, tones were delivered with a variable SOA ranging from 695 and 1303 ms (mean SOA 999 ms) in equiprobable steps of 32 ms arranged randomly.

Each stimulus was presented a total of 50 times in a randomized order across two blocks, one at the beginning of the experimental procedure and the other at the end.

During the experiment, participants sat comfortably in an electrically and acoustically shielded room, and during the listening condition were instructed to relax and watch a silent subtitled movie whilst ignoring the auditory stimulation. During the simple auditory reaction time task, participants were instructed to press the spacebar of a computer keyboard as fast as possible after hearing each sound. Pauses between blocks lasted 30 s, during which participants were allowed to move. Recording time lasted a total of two hours and thirty minutes.

Data acquisition

The EEG was continuously acquired with Neuroscan 4.4 software and Neuroscan SynAmps RT amplifier (NeuroScan, Compumedics, Charlotte, NC, USA) from the Cz scalp Ag/AgCl electrode mounted in a nylon cap (Quick-Cap, Compumedics) at the standard 10-20 system location. Three additional electrodes were positioned at the left and the right mastoids (M1 and M2, respectively) and at the left earlobe (A1). The EOG was measured with two bipolar electrodes placed above and below the left eye (VEOG), and two horizontal electrodes placed on the outer canthi of the eyes (HEOG). The ground electrode was located between Fz and FPz, and the right earlobe (A2) served as an online reference. All impedances were kept below 10 kΩ during the whole recording session and data was online bandpass–filtered from 0.05 to 3000 Hz and digitized with a sampling rate of 20 kHz.

Reaction times (RT) were obtained from the auditory reaction time blocks with Matlab by computing the time from the stimulus onset to the time of the participant's response to the stimulus. Only the spacebar presses occurring within a time window of 100 to 695 ms from stimulus onset were classified as responses to the stimuli (Chen and Sussman, 2013).

Data processing and analysis

EEG data analysis was performed offline using EEGlab v.7 toolbox (Delorme and Makeig, 2004), running under the crossplatform MATLAB

environment (The Mathworks, Inc. running under Matlab v.2012a. The continuous EEG recordings were filtered offline with a bandpass Kaiser window FIR filter from 70 to 1500 Hz and epoched from 40 ms before the stimulus onset to 200 ms after the stimulus. Trials with activity greater than 35 μV were removed from any further analysis and remaining epochs were baseline corrected to a 40 ms interval preceding the sound onset (Russo et al., 2008).

To analyse how the FFR was modulated as a function of the eliciting frequency, epochs from each stimulus frequency and for each participant were averaged separately. After artefact rejection, a mean of 976 trials were included on the average for each frequency (std = 35.4). FFRs were extracted from a window of 180 ms, starting at 10 ms to account for the neural delay required for the sound to reach subcortical structures. To obtain the power spectral profile of the FFRs to the different experimental frequencies, FFT (Cooley and Tukey, 1964) was applied to demeaned, zero-padded (1-Hz resolution) averages, windowed with a Hanning tapper.

Normalized frequency spectra were computed for each stimulus frequency by dividing the 'signal' power spectrum (i.e., the power spectrum of the neural response elicited by the tone of the frequency of interest) by the mean of the 'noise' power spectra (i.e., the mean of the power spectra of the neural responses elicited to the other tones). This was done under the assumption that the neural activity elicited by the tones whose F0 is different from the frequency of interest constitutes a measure of 'background' noise (Galbraith et al., 2000b). The mean normalized power of the FFR was computed using a 4-Hz-wide window surrounding the F0 of the presented stimuli. A repeated measures ANOVA with the factor Frequency (20 levels) was calculated on the normalized power computed for each frequency. To examine how the normalized power depended on the eliciting frequency, a linear model regression was fitted using the normalized power as the dependent variable.

For the simple auditory reaction time analysis, mean RT was calculated for each participant and for each stimulus frequency in each block separately. A response was included in the mean when it occurred between 100 and 695 ms from the offset of the auditory stimulus. The limits correspond to anticipatory responses (<100ms) and the minimum SOA (695 ms). After

discarding the extreme responses (i.e. anticipatory and delayed responses), a mean of 22.39 trials (std = 2.74) were included in the averages for each frequency in the pre–FFR block and a mean of 21.39 trials (std = 5.1) for each frequency in the post–FFR block. Reaction times were analysed with repeated measures ANOVA with the factor Frequency (20 levels) and Block (Pre– and Post–FFR). To characterize the nature of the effects, a linear regression model was fitted to the RT data for the Pre– and Post–FFR blocks, respectively.

We then computed for each stimulus frequency a circular–linear correlation between the normalized power and the RT to the sounds obtained in the two blocks (i.e., pre–FFR block and post–FFR block) for each given participant by using the two–step method described in Busch and VanRullen, 2010 Specifically, we performed a randomization test based on surrogate data, that is, under the null hypothesis that normalized power and RT are uncorrelated. In a first step, we computed the correlation value for each of the participants separately on the original data. Then, for each participant, we computed a sample of random RT with a set size identical to the number of trials in the real dataset of that given participant. These random RT where then correlated with the actual normalized power from that participant data. In this way, we obtained what would be the distribution of correlation values if we had random data from the same distribution. This first step was repeated 10,000 times per participant, thus producing for each participant a distribution of *p* values based on random RT. In a second step, one of these pseudo-p values was chosen at random for each participant and their grand-average across participants was computed. This second procedure was repeated 10,000 times also, so that a distribution of grand–average *p* values based on random data was obtained. The statistical significance of the circular–linear correlation observed in the real data was computed as the proportion of random correlations that exceeded the observed correlations. The Greenhouse–Geisser correction was applied when the assumption of sphericity was violated, and results were corrected using the Bonferroni correction to adjust for multiple testing. Significance was defined for $p \leq 0.05$.

Results

Normalized power spectra of the FFRs to the twenty experimental frequencies are depicted in Figure 6a. As expected, clear spectral peaks are observable on the fundamental frequency of each of the different experimental tone frequencies, which correspond to the FFRs elicited to the different stimuli. In addition, frequency–dependent fluctuations are evident in the FFR's power recorded across the different frequencies. When analysing the effects of Frequency on the FFR, a significant effect was found ($F(19,209) = 2.060$, $p = 0.007$, $\eta^2_{partial} = 0.158$; Figure 6b), indicating that the FFR normalized power is modulated by the eliciting frequency.

A simple linear regression model was used to predict the modulation of the normalized power based on the eliciting frequency. A significant regression equation fit was obtained ($F(1,239) = 7$, $p = 0.009$, $R^2 = 0.024$), thus revealing that the eliciting frequency is a predictor of the normalized power of the FFR, with a standardized beta of 0.169, $p = 0.009$.

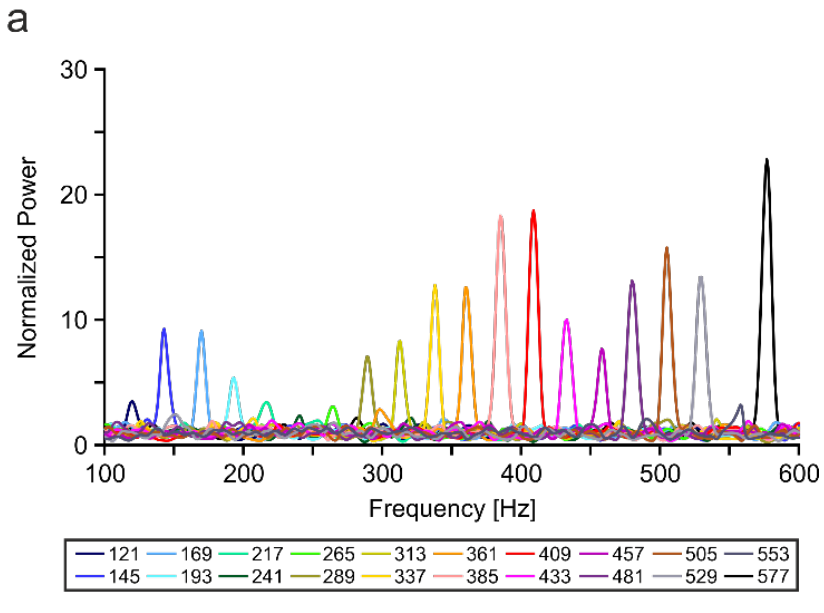


Figure 6: **FFR Normalized Power elicited to the twenty different auditory stimuli**.

(a) FFR normalized power spectra of the 20 different experimental auditory stimuli. A spectral peak can be observed for each stimulus at its respective fundamental frequency. (b) Mean spectral normalized power values for each of the stimuli fundamental frequencies. A significant modulation on the normalized power depending on the eliciting frequency is observed. In particular, there is an increase of the FFR normalized power with higher eliciting frequencies. Error bars represent ±1 SEM. Image: Original figure created by the author.

The analysis of the behavioural responses to the stimuli, assessed by means of the RTs (Figure 7), showed that the overall RT to the different frequencies, which ranged from 264 to 328 ms, were not significantly different in the pre– and post–FFR blocks (Block: $F(1,11) = 0.019$, $p = 0.893$, $\eta^2_{partial} = 0.002$). Nonetheless, a modulation of the RT depending on the eliciting frequency was found (Frequency: $F(19,209) = 7.188$, $p < 0.001$, $\eta^2_{partial} = 0.395$). A significant linear regression equation fit reveals that the eliciting frequency is a predictor of the reaction time to the sounds, indicating that the reaction time decreases as the frequency of the eliciting sound increases.

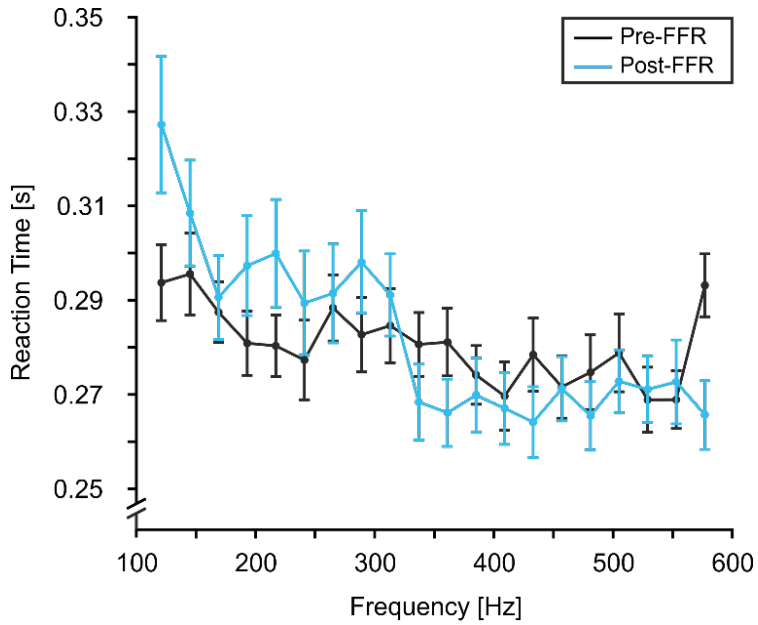


Figure 7: **Simple auditory reaction times to the different auditory stimuli.**

Simple auditory reaction times to the different stimuli in both Pre– (Black) and Post– (Blue) FFR blocks. Separable patterns can be observed in the two blocks across the range of frequencies. Whilst on the pre-FFR block the reaction times to the different stimuli were constant and not dependant on the eliciting frequency, on the post–FFR block there was a significant modulation of the reaction times depending on the frequency of the eliciting sound. Indeed, in the post–FFR block, there was a shortening of the reaction times to the sounds as the eliciting frequency increases. Error bars represent ±1 SEM. Image: Original figure created by the author.

Interestingly, our data showed separable patterns of reaction times in the pre– and post–FFR blocks across the range of frequencies (Figure 7; Block x Frequency: $F(19,209) = 1.990$, $p = 0.010$, $\eta^2_{partial} = 0.153$). To explore this interaction, a simple linear regression model was used to predict the modulation of reaction times based on the eliciting frequency depending on the block. In the pre–FFR block, the obtained linear equation fit was not significant ($F(1,239) = 1.970$, $p = 0.162$), thus indicating that in the pre– FFR block the eliciting frequency is not a predictor of the reaction times the same sounds. In contrast, a significant fit was obtained when predicting the RT in the post– FFR block based on the eliciting frequency ($F(1,239) = 11.819$, $p = 0.001$), with an R^2 of 0.043. This reveals that in the post– FFR block, the eliciting frequency becomes a predictor of the RT to the sounds, with a standardized beta of -0.218 ($p = 0.001$). Therefore, in the post– FFR block there was a shortening of the reaction times to the sounds (i.e., faster response to the auditory stimulus) as the eliciting frequency increased.

Finally, to assess the relationship between the FFR normalized power and the reaction times to the auditory stimuli depending on the eliciting frequency, Non-parametric Pearson Correlation tests were computed (Figure 8). A strong correlation was observed between the normalized power of the FFRs and the RT on the post– FFR block ($p = 0.0038$). The correlation between the normalized power and the pre– FFR block RT did not reach statistical significance ($p = 0.4738$).

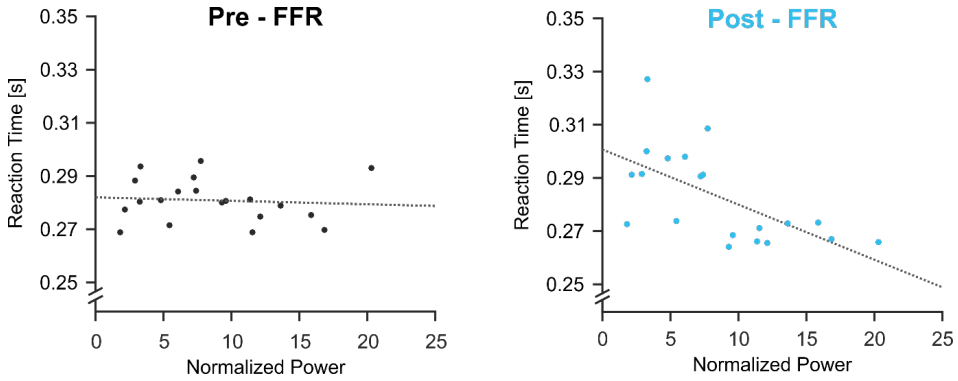


Figure 8: **Non-parametric Pearson Correlation Tests between the FFR normalized power and the RT to the same sounds.**

Scatter-plots of the non-parametric Pearson correlation values between the FFR normalized power of each of the twenty tested frequencies and the RT to the same sounds in the Pre– (Black) and Post– (Blue) FFR blocks. Whilst no significant correlation is observed between the FFR normalized power and the RTs of the pre-FFR block, after the FFR passive listening a significant correlation is obtained between the normalized power of the FFR and the reaction times to the same sounds. Indeed, the higher the normalized power was during the FFR passive listening, the shorter are the reaction times in the block post-FFR. Image: Original figure created by the author.

Discussion

The study presented in this Chapter aimed at investigating whether subcortical sound encoding modulates auditory simple cortical perceptual processing, and to clarify a basic feature of the FFR: how its normalized power is modulated as a function of the frequency of the eliciting auditory stimulus. To this end, we recorded the FFR to a set of pure tones of 20 different frequencies ranging from 121 to 577 Hz, as well as the behavioural response time to these same sounds in blocks delivered before and after the FFR recording.

We have observed a modulation of the FFR normalized power depending on the frequency of the eliciting sound, which increased as the eliciting frequency was higher and reaching a maximum ratio between 385 – 409 Hz.

Characterizing how the FFR is modulated as a function of the eliciting frequency has become of increasing relevance, since the FFR has been

adopted as a putative correlate of subcortical sound encoding. Yet, a recent line of evidence has suggested that the FFRs to an auditory stimulus of a fundamental frequency of around 100 Hz indexes the neuronal encoding of the periodic features of that sound not only in the brainstem, but also in the thalamus and in the auditory cortex (Coffey et al., 2016, 2017). Indeed, a recent study investigated whether these amplitude modulations of the FFR depending on the eliciting frequency could be the result of multiple neural generators with distinct latencies and they observed that whilst both subcortical and cortical structures are activated by low-frequency (<100 Hz) tones, the cortex becomes less sensitive to frequencies higher than 100 Hz (Tichko and Skoe, 2017). These findings challenge the assumption of the FFR as a correlate of subcortical sound encoding and support an emerging viewpoint in the literature that the FFR component of the ABR represents an integrated response of the entire auditory system (Kraus and White-Schwoch, 2015; Kraus and Slater, 2016).

Interestingly, throughout the auditory system, neurons synchronize their spiking patterns to the temporal features of the driving stimulus by firing at a particular phase of the stimulus (Schnupp et al., 2011; Gao and Wehr, 2015). However, the upper limit of temporal precision in phase-locked firing reduces with each ascending step in the pathway, so that the ability of neurons to follow fast modulations reduces with each ascending auditory station (Batra et al., 1989; Langner, 1992; Joris et al., 2004). Therefore, we hypothesized that the cortical contribution of the FFR decreases with sounds of fundamental frequencies higher than 100 Hz, thus suggesting that the FFR could still be used as a correlate of subcortical encoding at least for high frequency sounds. Recent evidence confirmed this hypothesis, where using source imaging techniques to multichannel data, demonstrated that a mixture of generators was involved in the FFR, and the relative contribution of these nuclei varied with stimulus frequency. Specifically, they observed that at lower frequencies, the PAC had a weak contribution to the FFR, but its contribution disappears at frequencies higher than 150 Hz (Bidelman, 2018).

In this framework, our results demonstrate that recording FFRs to high frequency sounds is not only feasible but also that its normalized power is more robust than that to low frequency sounds, where cortical structures may also be contributing to it. Our findings converge with previous work, where it was stablished that scalp-recorded FFRs become

increasingly stronger with frequencies up to 320 Hz, and sharply drop from approximately 440 Hz onwards (Batra et al., 1986; Hoormann et al., 1992; Skoe and Kraus, 2012). However, a number of key differences emerged from our data relative to other studies. For instance, a recent study using triangle waves to characterize the FFR found the response to be largest near 200 Hz (Tichko and Skoe, 2017). It is noteworthy to point out that the stimulus types used are different, and given that different stimuli may engage different neural mechanisms and sub-populations (Bidelman, 2015), it is probable that the differences on the FFR normalized power arise from this fact. Indeed, processing differences between the two types of stimuli start already in the cochlea. When using pure tones, only one point in the basilar membrane of the cochlea would vibrate. On the other hand, when using triangle-waves – and despite that the fundamental frequency of the sounds may be the same and, therefore, the same point in the basilar membrane will be stimulated, the sounds are more complex and the area of vibration will be broader (Schnupp et al., 2011).

Going a step further, our data demonstrates that the encoding strength in the subcortical auditory system, as indicated by the FFR normalized power, is used as evidence when making an auditory simple perceptual decision, such as deciding whether a sound was present or not. According to evidence accumulation models, there are two elements that contribute to the decision formation (or evidence accumulation process), in terms of probability theory: the priors and the evidence itself. The priors are the probability of receiving a particular stimulus before obtaining any evidence about it, and the evidence refers to the information that points to commit to a particular hypothesis (Gold and Shadlen, 2007). In this framework, cortical activity reflects the sensory evidence that is accumulated to a critical level to yield a perceptual decision (Ratcliff, 1978; Heekeren et al., 2008), and behavioural reaction times can be predicted by adding the time it takes for the evidence to reach the threshold to a non-decision time that accounts for the sensory and motor latencies (Ratcliff and McKoon, 2008).

Our results expand these models by demonstrating that subcortical sound encoding also contributes to making simple auditory perceptual decisions. In our experiment, participants had to perform a simple auditory reaction time in two separate blocks, one before and one after the FFR recording. In these blocks, they were presented with sounds of different frequencies appearing

in a randomized order. Therefore, the priors they had on those sounds (i.e., the predicted probability of receiving a particular stimulus on the upcoming trial) and the evidence they were presented with (i.e., the amount of times each sound was presented) were the same for all the stimuli and in the two blocks. Yet, we found out differential behavioural responses when comparing the reaction times in the two blocks. In the simple auditory reaction time block that took place before the FFR recording, reaction times were constant independently of the frequency of the eliciting sound. In this block, the priors and the evidence of the different sounds were the same, so it took the same time to gather the evidence favouring the detection of the sound's presence. This lead to similar reaction times to all the sounds independently of the eliciting frequency. On the other hand, in the simple auditory reaction time block that occurred after the FFR recording, a modulation of the reaction times depending on the eliciting frequency was observed, suggesting that there was an increased evidence for some sounds that lead to a faster detection of its presence. However, and similarly as the pre–FFR block, the priors and the physical evidence were the same for all the sounds, indicating that the increased evidence was obtained not by the physical stimulus presentation per se but from a better encoding of the stimulus. And indeed, we observed a correlation between the normalized power of the FFR during the passive auditory listening and the reaction times to the sounds in the post–FFR block.

Summing up, during auditory listening, increased FFR normalized power was observed for high frequency sounds, reflecting a better encoding of those sounds in the subcortical stations, which, in turn, modulates the simple perceptual decisions that took place on the post–FFR block. In particular, we observed that high frequency sounds were encoded more faithfully in the subcortical auditory system and this, in turn, provided more evidence favouring its detection, leading to faster perceptual decisions regarding these sounds. There is growing evidence that the subcortical auditory system has an important role in the perception of sounds (Pressnitzer et al., 2008; Shamma and Micheyl, 2010; Yamagishi et al., 2016) and that it encompasses more than a simple collection of relay stations in the hierarchy of auditory processing. Crucially, our findings indicate that subcortical sound encoding has an important role in the making of simple perceptual decisions, overall highlighting that evidence accumulation models are not only limited to cortical activity but that they should also take into consideration a subcortical contribution.

Summary

The study presented in this Chapter was designed to disentangle the contribution of the subcortical sound encoding to simple cortical perceptual decisions and to further characterize the FFR by observing how its normalized power is modulated as a function of the frequency of the eliciting stimulus. To do so, pure sinusoidal tones of twenty different frequencies were presented randomly in a passive listening condition and two simple auditory reaction time tasks, one before and one after the FFR recording, which allowed us to obtain the simple auditory reaction – times to the same sounds before and after a passive exposure time where the FFR was recorded. Our results showed a modulation of the FFR normalized power depending on the frequency of the eliciting sound, with increasing power as the frequency of the eliciting sound is higher. We also observed that the reaction times to the different frequencies were constant in the block before the passive listening, but were modulated depending on the frequency of the eliciting sound in the block after, with shorter reaction times to high frequency sounds. Going a step further, we observed a correlation between the normalized FFR power and the reaction times to the sounds in the block after the passive listening, thus suggesting that a better encoding of the sounds lead to a posterior faster detection of its presence and providing evidence favoring the role of subcortical sound encoding in the making of simple perceptual decisions.

The interim conclusions of this chapter are:

- The FFR normalized power, a correlate of the encoding strength of the sounds, is modulated depending on the frequency of the eliciting sound.
- The encoding strength in the subcortical auditory pathway is used as evidence when making auditory simple perceptual decisions.
- Evidence accumulation models should also take into consideration a subcortical contribution.

CHAPTER 7

NEURAL GENERATORS OF THE FREQUENCY-FOLLOWING RESPONSE

In this Chapter we present the study related to the third aim of the PhD thesis: *Characterize the FFR: disentangling the anatomical contribution of the FFRs elicited to sounds of different frequencies*. This study is still under analysis and the presented results are preliminary.

Introduction

Recapitulating on Chapter 2, the auditory Frequency – Following Response is a sustained component of the auditory brainstem potential that reflects synchronous and sustained neural phase-locking to the spectral and temporal periodic characteristics of the eliciting acoustic signal (Skoe and Kraus, 2010a; Kraus et al., 2017). in the range of 100 to 1500Hz approximately (Galbraith et al., 2000b; Picton, 2011). Recorded from the scalp with both EEG and MEG, the FFR emerges at circa 7–15 ms from sound onset and when analyzed properly, it becomes an objective indicator of the fundamental acoustic features intrinsic to speech sounds, including timing (onsets), pitch (fundamental frequency, fo) and timbre (the harmonics information). In particular, it provides information about the latency and amplitude of the auditory input in the time domain; and the magnitude of the fundamental frequency and its harmonics in the frequency domain (Skoe and Kraus, 2010a; for review see Kraus et al., 2017).

The FFR is highly sensitive to context-dependent contingencies (Slabu et al., 2012; Chandrasekaran et al., 2014; Skoe et al., 2014; Gorina-Careta et al., 2016) and to real-time statistical properties of the stimulus (Chandrasekaran et

al., 2009; Skoe and Kraus, 2010b; Skoe et al., 2013, 2015; Escera, 2017), and it provides a non - invasive measure of how short-term auditory training (Russo et al., 2005; Song et al., 2008, 2012; Carcagno and Plack, 2011; for review see Carcagno and Plack, 2017) and auditory experiences, such as language experience (Krishnan et al., 2008; Krizman et al., 2012, 2014, 2015; Skoe et al., 2017) or musical training (Musacchia et al., 2007; Wong et al., 2007; Parbery-Clark et al., 2009, 2011; Bidelman et al., 2011b; Skoe and Kraus, 2012), transform their representation in a subcortico-cortical auditory network (Kraus and White-Schwoch, 2015; Kraus and Slater, 2016). Overall, given its faithfulness in the phase-locking to the spectrotemporal detail of the incoming sounds, the FFR has been stablished as a stable window into the neural transcription of sounds that can be obtained under passive and active listening paradigms and allows the study of how the encoding of sounds is modulated depending on context-dependent contingencies and experience-dependent plasticity. Consequently, the study of neural sound encoding using the FFRs has become a critical tool to evaluate the abnormal neural encoding of speech in clinical populations (Cunningham et al., 2001; Banai et al., 2005, 2009; Banai and Ahissar, 2006; Russo et al., 2008, 2009; Chandrasekaran et al., 2009; Anderson et al., 2010; Hornickel et al., 2012; Hornickel and Kraus, 2013) that are important for the correct auditory processing of the auditory signal.

Despite the abundance of FFR studies, its neural origins remain debated. Since seminal studies, the FFR has been assumed to originate from neuronal aggregates in caudal brainstem and midbrain structures, with the inferior colliculus as a major neuronal source, and has been treated as a putative measure of subcortical sound encoding. This midbrain origin is supported by the fact that the short-latency of the responses aligns with the latency of the first spikes in the IC (Langner and Schreiner, 1988) and the FFRs contained phase-locked activity up to 1500 Hz; which spans beyond the upper limit of phase-locking of cortical neurons (~100 Hz; Aiken and Picton, 2008). Additionally, the cryogenic cooling of the IC results in disappearance of FFRs and a subsequent heating recovers the FFRs both in the colliculi and the scalp (Marsh et al., 1970; Smith et al., 1975), and the response is eradicated with focal lesions to the IC (Sohmer et al., 1977). Nevertheless, it was suggested that a mixture of brainstem sources is indeed involved in the generation of the FFR (Chandrasekaran and Kraus, 2010; Tichko and Skoe, 2017), and this hypothesis was supported by other studies that observed weaker contributions of the IC to the FFR, with the major source on the CN (Gardi et al., 1979) or on the MGB (Weinberger et al., 1970).

However, a recent line of evidence aimed to locate the putative sources of the FFR suggested that the FFRs to an auditory stimulus of a fundamental frequency of around 100 Hz indexes the neuronal encoding of the periodic features of that sound not only in the brainstem, but also in the thalamus and mainly in the auditory cortex (Coffey et al., 2016, 2017). Yet, this cortical contribution to the FFR disappears at frequencies higher than 150 Hz (Bidelman, 2018). These findings challenge the assumption of the FFR as a correlate of subcortical sound encoding and support an emerging viewpoint in the literature that the FFR component of the auditory brainstem response represents an integrated response of the entire auditory system (Kraus and White-Schwoch, 2015; Kraus and Slater, 2016).

Capitalizing on the frequency-specific phase-locking capabilities along the auditory hierarchy, where the upper limit of temporal precision in phase-locked firing reduces with each ascending step in the pathway, here we aim to determine the anatomical contribution of the FFRs elicited to sounds of different frequencies and to dissociate a hierarchy of anatomical sources contributing to the encoding of periodic stimuli of different frequencies, thus establishing the frequency limits by which the different neural generators contribute to the aggregate FFR recorded from the scalp. Considering the previous literature, we hypothesized that FFRs elicited to higher frequencies would receive less cortical contribution than those to lower frequencies, hence supporting the recent evidence of a subcortical involvement for these high-frequency sounds.

Methods

Participants

Twenty-three young adults (aged 21–34 years, mean age = 25.3 years, 4 males, 1 left–handed) with no history of auditory, neurological or psychiatric disorders participated in the study. Hearing thresholds were assessed in each ear with a standard pure-tone audiometry using 3 harmonic tones of 250 Hz at the beginning of the experimental session using SA-51 portable screening audiometer (MEDIROLL Medico Technical Ltd, Debrecen, Hungary), with the minimum threshold requested for participation below 25 dB SPL for all tested frequencies and an interaural difference of < 10dB. As music experience is known to modulate the encoding of the fundamental frequency (F0) of periodic sounds in the subcortical auditory pathway (Song

et al., 2011), all participants enrolled had less than 4 years of musical training that ceased two or more years before the study. The study was approved by the Bioethical Committee of the University of Barcelona and the Ethics Committee of the University of Jyväskylä and was in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Written and signed informed consent was obtained from each participant before starting the experiment. Pre-processed MEG and EEG data will be available under request to the corresponding author.

Stimuli and procedure

The auditory stimuli consisted of two pure sinusoidal tones of 200 ms duration, including 5 ms rise and fall times, with a frequency of 89 and 333 Hz respectively. The stimuli were delivered binaurally at an intensity of ~75 dB in alternating polarities through KAR ADU 1c audio stimulator (KAR-Audio, Unides Design Ay, Helsinki, Finland) with foam insert EAR-tips. The stimuli were generated and presented with Matlab v.2016a (Matworks). The auditory stimulation consisted in two conditions in which each of the stimuli were presented in a repetitive manner with a variable SOA jittered between 241 and 265 ms (mean SOA 253ms) arranged randomly. Each condition was divided in 16 blocks, each containing 950 trials (475 stimuli of each polarity), so that the stimuli of each frequency was presented a total of ~15200 times. All the blocks corresponding to each condition were delivered sequentially, with the order of the conditions counterbalanced across participants. Empty room activity was recorded for 2-min before each experimental session to estimate the intrinsic noise levels.

During the experiment, participants sat in an electrically, magnetically and acoustically shielded room, with their head inside the helmet-shaped magnetometer and were instructed to relax and watch a silent movie whilst ignoring the auditory stimulation. Pauses between blocks lasted 30s to allow the participants to rest, and there was a thirty-minute break between conditions were participants were allowed to move.

Data acquisition

Simultaneous magnetoencephalographic and electroencephalographic data were recorded with a 306-channel whole–head system (Elekta

Neuromag® TRIUX™, Elekta Oy, Helsinki, Finland) consisting of 204 planar gradiometers and 102 magnetometers, and a compatible 64-channel EEG cap (EASYCAP GmbH, Herrsching, Germany). The EOG was measured with two bipolar electrodes placed above and below the right eye (vertical EOG), and two horizontal electrodes placed on the outer canthi of the eyes (horizontal EOG) and the ground electrode was located in the right collarbone. For the EEG recording, the right earlobe served as an online reference. Five Head Position Indicator coils (HPI-coils) were attached on top of the EEG cap; two on the forehead, two behind the ears and one on the vertex of the head. The locations of three anatomical landmarks (the nasion and left and right preauricular points) and the five HPI-coils, as well as all the locations of all the EEG electrodes and a number of additional points on the head were digitized with an Isotrak 3D digitizer (PolhemusTM, United States) before the experiment started for co-registration with the participant's anatomical MRI. After the break between conditions, the location of five HPI-coils was re-digitized to recalculate the position of the head inside the MEG.

Individual structural magnetic resonance images (sMRI) were acquired from a private company offering MRI services (Synlab Jyväskylä, Jyväskylä, Finland). T1-weighted 3D images were collected on a GE 1.5 T (GoldSeal Signa HDxt) MRI scanner using a standard head coil and with the following parameters: repetition time/echo time [TR/TE] = 540/10 ms, flip angle [FA] = 90°, matrix size = 256 × 256, slice thickness = 1.2 mm, sagittal orientation.

The MEG was recorded in 68° upright gantry position. All EEG impedances were kept below 10 kΩ during the whole recording session and both MEG and EEG data was online bandpass-filtered from 0.1 to 1660 Hz and digitized with a sampling rate of 5 kHz. To ensure that the participant's head position relative to the recording instrument was constant throughout the experiment, the magnetic fields produced by the HPI coils were measured before each block.

Data processing and analysis

Continuous MEG data was pre-preprocessed off-line with the Elekta Neuromag™ MaxFilter 2.1 (Elekta Oy, Helsinki, Finland) Signal Space Separation (SSS) method (Taulu et al., 2004) to suppress external magnetic

interference and remove static bad channels. MaxFilter software was also applied for head movement correction and transforming the head origin to the same position for each participant. MEG data was then imported to Brainstorm (Tadel et al., 2011) for further processing. Eye blink and heart beat artefacts were removed using Brainstorm's source signal projection (SSP) algorithm (Tesche et al., 1995; Hämäläinen, 2009) when the topography of the components matched those of ocular or cardiac origin upon visual inspection. The clean MEG recordings were bandpass filtered from 75 to 1500 Hz and epoched from -40 to 240 ms relative to stimulus onset. Epochs were baseline corrected to a 40 ms interval preceding the sound onset and averaged separately for each frequency condition, polarity of presentation and for each participant separately. Responses to alternating polarity stimuli were subtracted to maximize the response to pure tones (Aiken and Picton, 2008).

The source modelling analysis for this experiment is based on the methods developed by Coffey et al., 2016, as this is the only study published so far exploring the FFR neural origins with MEG. In particular, for the present experiment, the signal source was estimated using distributed source models, which estimate the amplitude of a large number of dipoles distributed throughout the brain volume, but must be constrained by spatial priors.

FreeSurfer (Fischl, 2012) was used to prepare the cortical surfaces and automatically segment subcortical structures from each subject's T1-weighted anatomical MRI scan. Anatomical data was later imported to Brainstorm, where precise co-registration of MEG and structural MRI data was accomplished using a semiautomatic procedure. The information of the fiducial points was used for a first alignment and the digitized head shape and the scalp surface of each individual were then used to reduce the minimum distance error between them in an iterative process. Thalamic and brainstem structures were then combined with the cortex surface to form a mixed surface/volume model with the deep brain activity (DBA) model, which included a triangulation of the cortical surface (~15,000 vertices), and brainstem and thalamus as a three-dimensional dipole grid (~18,000 points) (Attal and Schwartz, 2013). The head model was computed using the overlapping – spheres algorithm for each participant. This forward model explains how neural electric currents of the source space produce magnetic fields at the external sensors with good accuracy (Huang et al., 1999). A noise covariance matrix, which accounts for the contaminants that remain

present in the data after the preprocessing is complete, was computed from the 2-min empty – room recordings. The inverse solution was calculated on the subtracted polarities average for each subject and frequency condition using wMNE source distribution algorithm with unconstrained source orientations using Brainstorm default parameters.

To disentangle the neural contributors of the FFR and obtain the signal originated in specific brain regions, bilateral regions of interest (ROIs) were defined in the main subcortical nuclei and cortical areas that conform the human auditory pathway (i.e. cochlear nucleus, CN; inferior colliculus, IC; medial geniculate body of the thalamus, MGB; and primary auditory cortex, PAC) as well as two control regions that are at the maximal distance from the target auditory regions: the frontal (FP) and occipital (OP) poles (Figure 9).

As the head model used was a mixed surface/volume model, the ROIs are defined either as surfaces or volume depending on their location.

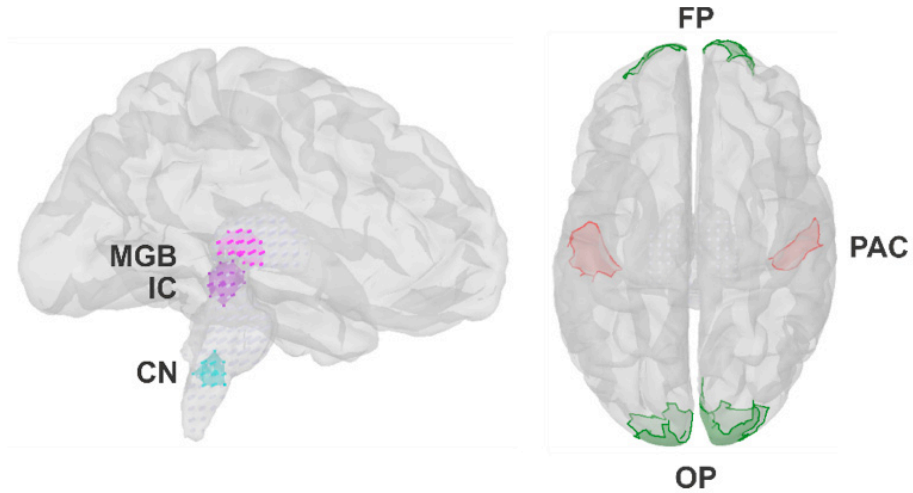


Figure 9: **Regions of Interest defined for extracting the neural contributors of the FFR**.

Bilateral regions of interest were defined on the main structures that conform the human auditory pathway and in two control regions in order to extract the neural activity in each of them and study the neural contributors to the FFRs recorded. CN = cochlear nucleus; IC = inferior colliculus; MGB = medial geniculate body of the thalamus; PAC = primary auditory cortex; FP = frontal pole; OP = occipital pole. Image: Original figure created by the author.

For the surface ROIs, the right and left PAC were defined as the merged regions identified in the Destrieux Atlas (Destrieux et al., 2010) as the transverse temporal gyrus and transverse temporal sulcus (L: 7.07 cm^2 (s.d. = 1.22); R: 5.41 cm^2 (s.d. = 0.99)). The frontal poles (L: 8.45 cm^2 (s.d. = 1.50); R: 12.44 cm^2 (s.d. = 2.50)) and the occipital poles (L: 16.76 cm^2 (s.d. = 2.02); R: 24.86 cm^2 (s.d. = 3.97)) were also defined as surface ROIs. Additionally, spherical subcortical volume ROIs were grown from seeds located in the dipole grid around previously published standardized MNI coordinates corresponding to the left and right CN (MNI: [±10, -34, -45]; L: 0.49 cm^2 (s.d. = 0.02); R: 0.50 cm^2 (s.d. = 0.02)) and left and right IC (MNI: [±6, -33, -11]; L: 0.48 cm^2 (s.d. = 0.02); R: 0.48 cm^2 (s.d. = 0.03)). Additionally, ROIs capturing the activity from the thalamic MGB were defined based on the standardized MNI ([±17, -24, -2]) and covered approximately the posterior third of the thalamus (L: 1.30 cm^2 (s.d. = 0.02); R: 1.27 cm^2 (s.d. = 0.03)). A time series of mean amplitude was extracted for each ROI and for each of the three orientations in the unconstrained orientation source model for the FFR (30 to 210 ms from stimulus onset) and the baseline (-40 to 0 ms from stimulus onset) periods. To obtain the power spectral profile of the different extracted time series, Fast Fourier Transform (FFT; Cooley and Tukey, 1964) was applied to zero-padded (1-Hz resolution) averages, windowed with a 5-ms raised cosine ramp. Orientations were summed in the frequency domain to obtain a single spectrum for each ROI, and posteriorly averaged to yield a final single spectrum for each bilateral pair of ROIs during the FFR and the baseline period.

The mean normalized power in each ROI was computed using a 5-Hz-wide window surrounding the fo of the presented stimuli for both the FFR and the baseline spectra, so that we calculated the increase of signal during FFR at fo over baseline for each bilateral pair. We then compared this increase in each auditory ROI to the average of the control regions and assessed statistical significance using Wilcoxon-matched pair tests. Results were corrected using the Bonferroni correction to adjust for multiple comparisons, so that significance was defined for $p \leq 0.01$ (0.05/4).

Results

Here we present data based on a subsample of only 14 out of 21 participants, so the results in this part have to be considered preliminary. The grand–average waveforms of FFRs recorded with EEG and MEG and elicited to both frequency conditions are depicted in Figure 10, together with the corresponding spectral decompositions.

FFRs can be observed both in the time and spectral domains for both frequency conditions and using the two recording methods (EEG and MEG; Figure 10). In the time domain, although the FFRs were visible for both frequency conditions, the FFR elicited to the low frequency condition had larger amplitude compared to the ones elicited to the high frequency condition. In addition, the FFR recorded with MEG to the high frequency condition had a noisier baseline, which caused a decreased signal-to-noise ratio for the high frequency FFR compared to the low frequency FFR. In the frequency domain, both EEG and MEG showed clear peaks at the fo of the stimulus for both frequency conditions and smaller harmonic peaks (integer multiples of the fo) were observable only for the low frequency one. Consistent with what was depicted in the time domain, the ongoing noise in the whole frequency spectrum was enhanced in the high frequency condition, so that the harmonic peaks were not observable.

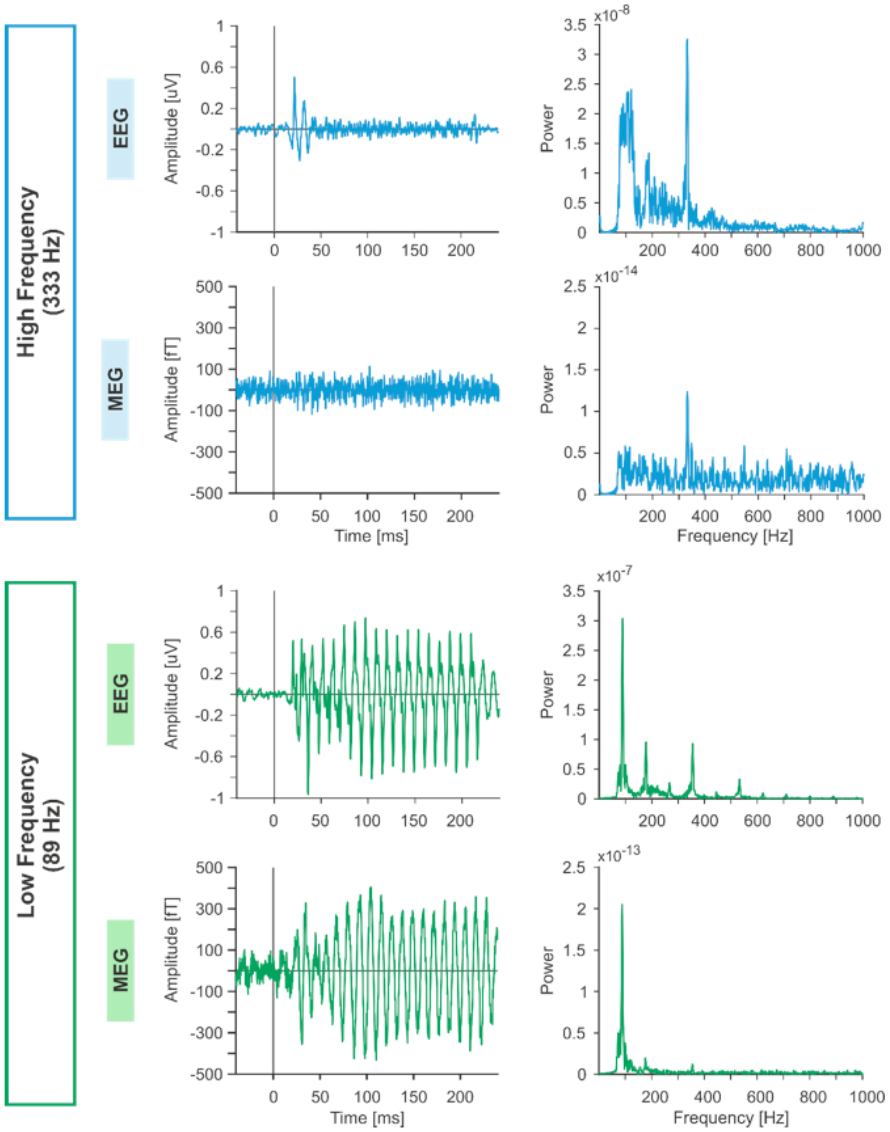


Figure 10. **Time and frequency domain representations of the EEG and MEG Frequency – Following Responses to both frequency conditions**.

Time course (first column) and spectrogram (second column) representations of the EEG and MEG recorded FFR elicited to both frequencies. FFRs are observable in the time domain using both recording techniques for the two tested frequencies. Similarly, a peak at the fundamental frequency of the eliciting stimulus is observable in the all the spectrograms. However, the amplitude of the FFRs differ depending on the frequency, being smaller and with a noisier baseline for the high frequency FFRs. The ongoing noise in the whole frequency spectrum was enhanced in the high frequency condition, so that the harmonic peaks were not observable. Image: Original figure created by the author.

To separate the contributions of subcortical and cortical FFR sources in both frequencies, we estimated the neural origin of the FFR using a minimum-norm estimate (MNE) modelling and extracted the data from bilateral pairs of regions of interest (ROIs) distributed throughout the auditory hierarchy as well as two control regions located in the frontal and occipital poles, at maximal distance from the areas of interest. For the low frequency condition, this analysis yielded strong peaks at the fo in all the subcortical auditory ROIs that were significantly larger than the signal observed in the control regions (CN: Wilcoxon – matched pair test, $Z = 3.107$, $p = 0.002$; IC and MGB: Wilcoxon – matched pair test, $Z = -3.233$, $p < 0.001$; Figure 11). A significant peak was also observed for the primary auditory cortex ROI (PAC: Wilcoxon – matched pair test, $Z = -3.296$, $p < 0.001$; Figure 11), indicating that the neural activity at 89 Hz was larger than the one in the control regions.

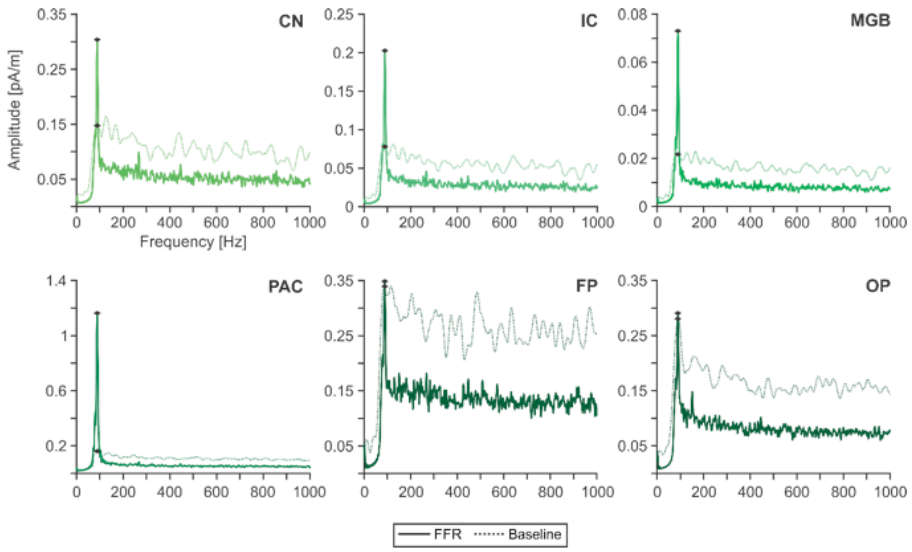


Figure 11. **ROI amplitude spectrogram during FFR and baseline for the low frequency sounds**.

Amplitude spectrograms of the time course extracted for the different auditory and control ROIs during the FFR (30 – 210 ms) and the baseline (-40 – 0) periods when stimulating with low frequency sounds. The peaks in the frequency of interest for both FFR and baseline periods are marked with a black diamond. All results are averaged across bilateral ROIs and across subjects (n = 14). Strong peaks are observable at 89 Hz, corresponding to the stimulus fo, in all the auditory ROIs as compared to the baseline period. No peaks are observed on the control areas. The signal to noise ratio obtained in both subcortical and cortical auditory ROIs was significantly larger than the one from the control regions, indicating that all the structures analysed contribute to the FFR recorded. CN = cochlear nucleus; IC = inferior colliculus; MGB = medial geniculate body of the thalamus; PAC = primary auditory cortex; FP = frontal pole; OP = occipital pole. Image: Original figure created by the author.

Regarding the results obtained for the high frequency condition, the analysis revealed peaks at the fo in the IC and MGB ROIs which showed a tendency towards significance compared to control regions, as although they were significant per se, this statistical significance vanished when correcting for multiple comparisons (IC: Wilcoxon – matched pair test, $Z = 2.291$, $p = 0.022$; MGB: Wilcoxon – matched pair test, $Z = -2.103$, $p = 0.035$; Figure 12). The neural activity recorded from the CN and PAC did not show any significant nor tending to significance peak at the fo compared to

control regions (CN: Wilcoxon – matched pair test, $Z = -0.848$, $p = 0.397$; PAC: Wilcoxon – matched pair test, $Z = 0.224$, $p = 0.221$; Figure 12).

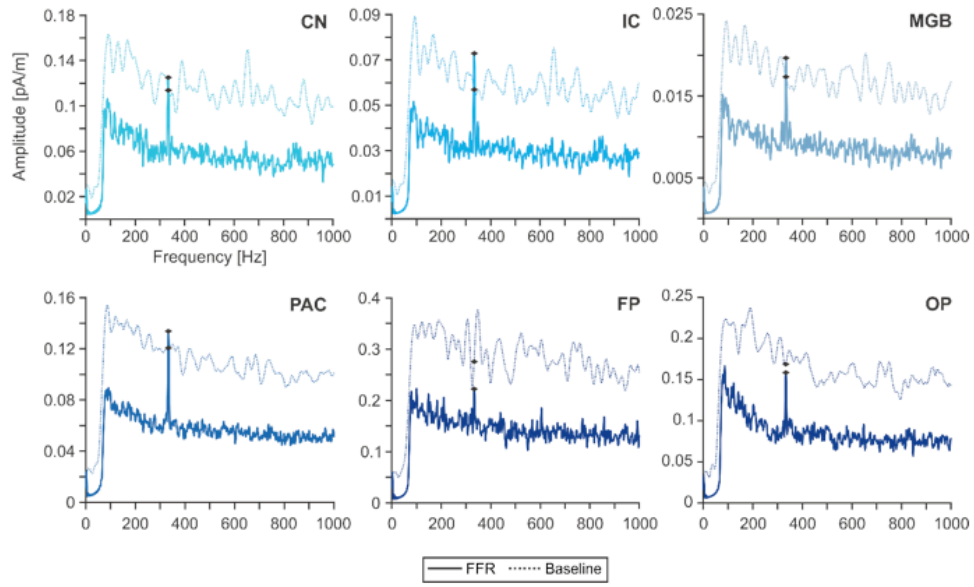


Figure 12. **ROI amplitude spectrogram during FFR and baseline for the high frequency sounds**.

Amplitude spectrograms of the time course extracted for the different auditory and control ROIs during the FFR (30 – 210 ms) and the baseline (-40 – 0) periods when stimulating with high frequency sounds. The peaks in the frequency of interest for both FFR and baseline periods are marked with a black diamond. All results are averaged across bilateral ROIs and across subjects (n = 14). Significant signal – to – noise ratio at 333 Hz is obtained only for two of the subcortical ROIs (IC and MGB) as compared to the control regions. No significant activity at the fo is obtained for the cortical ROIs, thus demonstrating that the activity recorded at the scalp only comes from IC and MGB generators. CN = cochlear nucleus; IC = inferior colliculus; MGB = medial geniculate body of the thalamus; PAC = primary auditory cortex; FP = frontal pole; OP = occipital pole. Image: Original figure created by the author.

Discussion

The present study constitutes the first demonstration, using MEG, that there is a hierarchy of anatomical sources contributing to the encoding of periodic stimuli of different frequencies in the human auditory pathway. In particular, we have observed that whilst the neural contribution of the subcortical sources is present in both stimulation frequencies (i.e. low, 89 Hz; high, 333 Hz), the cortical contribution not apparent when the eliciting frequency was high, thus favouring the hypothesis that high frequency recorded FFRs represent only subcortical activity and, therefore, it can still be considered a window into human subcortical sound encoding.

In particular, our results demonstrate that the neural contribution to the encoding of low frequency sounds is not restricted to subcortical nuclei but the FFRs recorded in the scalp represent an integrated response of different neuronal aggregates throughout the whole auditory hierarchy, including subcortical and cortical structures. By using distributed source modelling, which allows us to model and estimate the neural activity of a large number of dipoles based on spatial priors, we showed that the signal – to – noise ratio attributed to the different subcortical and cortical auditory regions is greater than the signal coming from the control regions, located in the frontal and occipital poles. On the other hand, the putative neural sources of the FFR elicited to high frequency sounds were restricted to subcortical ones, and the signal – to – noise ratio attributed to the cortical auditory regions does not differ from the signal coming from the control regions, thus meaning that no significant cortical contribution is observed.

Our findings favour the importance of frequency as a key factor when studying the encoding of sounds throughout the auditory hierarchy. As explained earlier in Chapter 2, our results go in line with previous literature which questioned the subcortical origin of the FFRs and aimed to locate the putative sources of it. Previous studies observed, using MEG, that the FFR to a speech stimulus of fundamental frequency close to 100 Hz receives major contribution from the auditory cortex (Coffey et al., 2016) and called for a re-examination of previous FFR interpretations with methods that allow source separation. This re-examination came in a recent study, where FFRs to speech sounds were recorded with EEG and source imaging techniques were applied to multichannel data (Bidelman, 2018). Indeed, and replicating

the previous MEG study, they observed that a mixture of generators was involved in the FFR, but crucially, they demonstrated that the relative contribution of these nuclei varied with stimulus frequency. Specifically, they observed that at lower frequencies the PAC had a weak contribution to the FFR, but its contribution disappears at frequencies higher than 150 Hz. Interestingly, these two studies converge and diverge at the same time, as they both observe that for low frequencies there is a cortical contribution to the FFR, but in the MEG study this contribution is the dominant on the FFR response, whilst when registering with EEG the contribution of the PAC is weak. This divergence between the two different studies could be due to the differential recording methods, as MEG is insensitive to radial sources, while EEG may reflect both radial and tangential sources, and MEG is comparatively less sensitive to deep sources.

For our study, we used simultaneous MEG and EEG recordings with the goal of finding a common point between the results observed with both techniques and correlating the MEG and EEG data with the same dataset. Unfortunately, although the EEG data that should complement the MEG findings has been recorded, only a preliminary analysis has been carried out, so no EEG results are presented in this chapter.

Nevertheless, using the same analysis methodology as the MEG study, we replicate and expand their findings, as we observe that for the low frequency pure tones, which have a frequency close to 100 Hz, there is indeed a cortical contribution to the FFRs, together with a subcortical contribution of all the main relay structures of the human auditory hierarchy (i.e. CN, IC and MGB). Going a step further, we observed that when the frequency of the incoming auditory stimulation is higher, the cortical contribution is not recorded and only a contribution from the IC and MGB can be observed. This results go in line with the ones observed by the aforementioned EEG study, confirming their claim that for high frequency sounds the cortical contribution to the FFR disappears.

Overall, what becomes clear is that the FFR has a multi-generator nature and represents an integrated response of the entire auditory system. In particular, the FFR is a composite response and the FFR at any given frequency can reflect the activity from multiple generators, which phase-lock to the incoming stimulus with different latencies (Tichko and

Skoe, 2017). Our dataset has the potential to become a critical tool to disentangle how the different structures of the auditory hierarchy interact and how depending on the eliciting frequency these structures engage and contribute to the sound encoding assessed by means of the FFR. Further latency analysis has to be carried out, which together with the correlation with the EEG data may provide a very useful insight into the FFR literature.

In this study we expand the previous findings by demonstrating, with MEG, that when increasing the frequency of the eliciting stimulation, the cortical contribution to the FFR disappears, thus allowing us to register only activity coming from the subcortical structures throughout the human auditory hierarchy.

Summary

The study presented in this chapter came after a breakthrough in the FFR literature, when it was demonstrated that at lower frequencies (<100Hz), the FFRs reflect both cortical and subcortical activity. After observing that the FFR amplitude is modulated depending on the eliciting frequency in Study II, here we aimed to dissociate a hierarchy of anatomical sources contributing to the FFRs elicited to pure tones of low (89 Hz) and high (333Hz) frequencies. To separate the contribution of the different possible generators, we used MEG and weighted minimum-norm estimate (wMNE) modelling, and defined regions of interest in the bilateral pairs of the main relays in the auditory hierarchy (i.e., Primary Auditory Cortex, Medial Geniculate Body, Inferior Colliculus and Cochlear Nucleus), plus two control regions in the frontal and occipital pole. Time course from the ROIs was extracted and their spectral amplitude computed. Our results show that for the low frequency sounds, all ROIs from the auditory hierarchy contributed to the recorded FFR as compared to the activity coming from the control regions. On the other hand, the neural structures contributing to the high frequency FFR are only the subcortical ROIs defined in the IC and MGB, and no cortical contribution has been observed. This reveals that indeed the frequency of stimulation is a key factor when studying the neural sound encoding and indicates

that with high frequency sounds, the neural contributors to the FFR are only subcortical and, therefore, the FFR can still be used as a window into subcortical sound encoding when using the appropriate stimulus parameters.

The interim conclusions of this chapter are:

- The FFR to low frequency sounds (i.e., 89 Hz) has neural contributors throughout the entire auditory hierarchy, from the CN to the PAC.
- The high frequency FFR, elicited to sounds of 333 Hz disclosed only subcortical neural generators contributing to it.
- The FFR can still be used as a window into human subcortical sound encoding when using the appropriate stimulus parameters.

CHAPTER 8

GENERAL DISCUSSION

The aim of the present PhD project was to examine the hypothesis that the subcortical auditory pathway is more than a simple collection of relay stations for sound processing and have an active role in the encoding and processing of the incoming sounds, as suggested by animal research. In particular, previous studies demonstrated that repetitive stimulation reduces auditory neural activity in animal cortical and subcortical levels and in the human cerebral cortex. In this thesis, we aimed to investigate if the human subcortical auditory pathway also plays a role in the encoding of statistical regularities (Chapter 5) and if the subcortical sound encoding contributes to the making of simple cortical perceptual decisions (Chapter 6). The contribution of the human subcortical auditory pathway to sound encoding is investigated by means of the FFR, which reflects synchronous neural phase-locking to the spectral and temporal periodic characteristics of the eliciting acoustic signal. Since seminal studies, it has been assumed that the FFR provides a window into subcortical sound processing, but some controversy about its neural origins aroused during the course of these PhD thesis. Therefore, to further understand the FFR, this PhD thesis also aimed to study how the frequency of the incoming sounds modulates the FFR strength (Chapter 6), as well as the anatomical neural sources that contribute to it (Chapter 7).

Evidence of the involvement of the subcortical auditory pathway to sound encoding and processing

One of the ultimate goals of this PhD thesis was to investigate the role of the subcortical auditory pathway to encoding and processing of sounds. As we have discussed and demonstrated throughout the different chapters of this thesis, auditory subcortical structures are more than simple relay steps on the ascending auditory hierarchy, and they have an active role in encoding the incoming sounds, which contributes to a latter simple perceptual decision making.

In particular, the results exposed in Chapter 5 show that the FFR amplitude was attenuated by repetition independently of temporal predictability, yet we observed an accentuated suppression when the incoming stimulation was temporally predictable. The importance of timing in regularity encoding has been demonstrated in previous studies, where it has been observed that temporal predictability of the incoming auditory stimulation reduces the amplitude of long latency AEPs (Lange, 2009; Schwartze et al., 2013) and enhances both repetition suppression (Todorovic et al., 2011; Todorovic and de Lange, 2012) and repetition positivity (Costa-Faidella et al., 2011a) in the human auditory cortex.

Specifically, it was observed that repetition positivity is more evident when the incoming sounds are delivered in an isochronous way and that temporal predictability enhances repetition suppression, aiding its propagation upstream the auditory pathway. Further MEG studies demonstrated that RS was modulated by the expectation of repetition of auditory events in a way that the more expected the incoming sounds were, the more its evoked response was suppressed (Todorovic et al., 2011; Todorovic and de Lange, 2012). In this regard, our findings expand the previous evidence by supporting the view that regularity encoding spans across the auditory hierarchy and point to temporal predictability as a modulatory factor of regularity encoding in early stages of the auditory pathway. This results are in line with animal studies, which by means of single cell recordings reported that individual neurons exhibit a reduced response to a repetitive stimulus both in cortical (Ulanovsky et al., 2003, 2004; Pérez-González et al., 2005) and subcortical (Malmierca et al., 2009; Antunes et al., 2010; Ayala and Malmierca, 2012) levels of the auditory hierarchy.

The results presented provide two complementary views of the effects of temporal predictability on regularity encoding in the human subcortical auditory pathway. In addition to the results described above, we observed an increase in the robustness of the phase-locking when the incoming stimulation was temporally predictable, which indicates that the response is more periodic and the phase-locking to the stimulus more reliable, thus helping the extraction of acoustic features and increasing the signal to noise ratio of the encoded repetitive stimuli. In this context, the temporal predictability would induce a better phase-locking, and therefore, enhance the adaptation to repetitive stimulation. In comparison, when the incoming stimulation is not temporally predictable, the phase-locking to the incoming acoustic stimulation is less faithful than when the stimuli are temporally predictable. Nevertheless, the stimulation is repetitive anyway, which leads to some repetition suppression although it is not as evident as when the incoming stimulation can be temporally predicted. Despite the differences observed with the different temporal conditions, we demonstrate that early neural representations of sounds are sharpened by the temporality of the encoded statistical regularities, meaning that the subcortical auditory pathway has an active role in the encoding of sounds.

Going a step further, in Chapter 6 we demonstrated that the active role in perceptual encoding of the incoming sounds of the subcortical auditory pathway is reflected in the making of latter simple cortical auditory perceptual decisions. Simple auditory perceptual decisions are thought to be mediated by neural computations occurring in the ventral auditory pathway of the cerebral cortex. Indeed, it has been shown that throughout the regions that configure this pathway, auditory information is organized and processed in a hierarchical fashion (Rauschecker and Tian, 2000; Rauschecker and Scott, 2009; Hackett, 2011; Bizley and Cohen, 2013; Tsunada et al., 2015; Cohen et al., 2016). However, as described in Chapter 5, recent evidence demonstrated that the subcortical auditory pathway is actively involved in the encoding of sounds, which lead us to the hypothesis that the processing occurring in the subcortical stages of the auditory hierarchy could influence a latter process of decision making. Indeed, as discussed in Chapter 6, we observed that our hypothesis is confirmed, as the encoding strength in the subcortical auditory system is used as evidence when deciding whether a sound is present or not. Our results disclosed a modulation of the FFR normalized power depending on the eliciting frequency, which, in turn,

correlated with faster detection of the sounds after the FFR recording, but not before. This findings led us to speculate that the increased normalized power would reflect a more faithful encoding of the sounds in the subcortical stations and this provides more evidence favouring its detection, leading to faster perceptual decisions regarding those sounds.

Although the FFR has been shown to be quite insensitive to higher order perceptual processes (Bidelman et al., 2013), as we observed in the first study of this thesis it is indeed modulated by stimulus regularities, which indicates that the online formation of predictive models via stimulus regularity encoding is reflected at subcortical levels despite that already established categories to interpret acoustic stimulation may not require them. This online formation of predictive models in the subcortical stages influences the making of simple auditory perceptual decisions, as stronger encoding of the incoming sounds is reflected into stronger predictions, which, in turn, leads to a faster detection and response to those sounds.

Interestingly, in addition to repetitive stimulation and temporal predictability, frequency has been observed to be a modulatory factor in the online formation of predictive models in the subcortical stages of the auditory pathway. This can be due to the fact that throughout the auditory system, the upper limit of temporal precision in the phase-locked firing of the neurons reduces with each ascending step in the pathway. Therefore, the modulation of the FFR depending on the frequency of the eliciting stimulus may be due to these differential phase-locking abilities, as the FFR recorded from the scalp may include activity coming from different anatomical sources.

Characterizing the FFR: a window into auditory sound encoding

The FFR has become a key tool for studying the auditory system and FFR studies have shed light on basic principles of sound processing in the brain. Characterizing the FFR is of high priority, as due to its plasticity and stability it is a sound candidate clinical tool to study individual differences and to understand communication disorders. Although the basics of measuring and interpreting the FFR are well stablished (Skoe and Kraus, 2010a; Kraus et

al., 2017), further studies are required to expand our knowledge, as there are still many methodological issues that can be technically improved to make a more powerful use of its potential. For example, a persistent challenge is the signal-to-noise problem, as the FFR needs to be the averaged response of many repetitions of the same sound, thus using long experimental times and not being suitable for studying fast sound changes which are typical from the acoustic environment that surround us. However, and despite its limitations, research on the FFR has increased exponentially during the last years, and some important contributions were made that changed all the assumptions that were the basis of the FFR before.

As discussed earlier in Chapter 2, it was assumed since seminal studies that the FFR is an aggregation of phase-locked neural activity from neuronal populations in the caudal brainstem and midbrain structures of the auditory pathway and throughout most of the studies that compose the FFR literature, this assumption was never questioned. Yet, the single study published so far using MEG to locate the sources of the FFR challenges this assumption. In the study it was observed that the FFR to a speech stimulus of fundamental frequency close to 100 Hz receives major contribution from the auditory cortex (Coffey et al., 2016), calling for a re-examination of previous FFR interpretations with methods that allow source separation. This re-examination came in a very recent study, where FFRs to speech sounds were recorded with EEG and source imaging techniques were applied to multichannel data. Indeed, they demonstrated that a mixture of generators was involved in the FFR, and the relative contribution of these nuclei varied with stimulus frequency. Specifically, they observed that at lower frequencies, the PAC had a weak contribution to the FFR, but its contribution disappears at frequencies higher than 150 Hz. This two studies converge and diverge at the same time, as they both observe that for low frequencies there is a cortical contribution to the FFR, but in the MEG study this contribution is the dominant on the FFR response, whilst when registering with EEG the contribution of the PAC is weak. Nevertheless, this two studies supposed a break-through in the FFR literature and although they did not question the basic characteristics of the FFR, they called for a re-examination of the stimulus used to measure it.

In Chapter 7, we went a step further and used simultaneous MEG and EEG recordings to try to stablish the common point between both studies

and disentangle the anatomical contributions to the FFR of sounds of two frequencies. With MEG, we were able to record the FFR to both high and low frequencies and although a peak in the fundamental frequency was observable in the spectral domain for both frequencies, the FFRs elicited to the high frequency sounds were somewhat noisier. When measuring the MEG equivalent of the EEG-FFR, we expect to be measuring distinct aspects of the same underlying phenomenon. This is because MEG is insensitive to radial sources, while EEG may reflect both radial and tangential sources, and MEG is comparatively less sensitive to deep sources. Therefore, it is completely expected that the high frequency signal was noisier, as the observed contributing sources are mainly deep ones, so that more trials are needed to have enough signal to be captured by the MEG with the same strength as the low frequency ones.

After computing the spectral contribution to the FFRs of different regions of interest throughout the auditory hierarchy, we confirmed that with the low frequency stimulation, both the cortical and the subcortical regions of interest were contributing to the FFR, thus replicating the results observed by Coffey and colleagues. On the other hand, when analysing the contribution of the different structures to the FFRs to the high frequency sounds, we were able to obtain a contribution from subcortical generators only, specifically from the IC and the MGB. Although the significance vanishes when correcting for multiple comparisons, the results presented in the chapter are only from a subset of participants and when including the full sample, we expect the significance to hold. This results are the first demonstration with MEG that only subcortical sources are contributing to high frequency FFRs and, therefore, that the FFRs recorded with proper stimulation can still be a window into subcortical sound encoding. Regarding the data measured with the EEG, clear FFRs were observable on the time and spectral domains, and although source imaging techniques are still to be applied, we are confident that the signal to noise ratio will be sufficient to disentangle the neural sources.

Overall, what becomes clear is that the FFR has a multi-generator nature and represents an integrated response of the entire auditory system. In particular, the FFR is a composite response and the FFR at any given frequency can reflect the activity from multiple generators, which phase-lock to the incoming stimulus with different latencies (Tichko and Skoe,

2017). Further studies are required to fully disentangle the frequency limits by which the FFR is generated at each level of the auditory hierarchy, but for low frequencies the contributors include the subcortical auditory pathway and the PAC, whilst at higher frequencies the contribution of the PAC is not observable.

Strengths and limitations

In addition to what has already been detailed throughout the different chapters of this PhD thesis, here we intend to highlight the most remarkable points of the present work, as well as the most important limitations.

Considering that the subcortical auditory pathway is composed of more than simple relay stations and that it plays a role in the encoding and processing of sounds is still a new perspective in auditory cognitive neuroscience, opposing the traditional "corticocentric" myopia that has pervaded cognitive neuroscience theory until recently (Parvizi, 2009). Although the paradigms used throughout this thesis to test the hypothesis were quite simple, they allowed us to explore some of the most basic aspects of auditory processing and set the starting point for future studies that would address this issue.

What becomes clear is that the concepts mentioned throughout this project are in the center of study in auditory neuroscience, and new features of the FFR and the subcortical auditory pathway are discovered day by day. That is the reason why, with the actual knowledge, the studies presented in chapter 5 and chapter 6 have one main limitation that is worth to be mentioned. After the recent discoveries about the origins of the FFR it becomes clear that the stimuli chosen to study how the subcortical auditory pathway processes statistical regularities and temporal predictability are not the optimal ones. In particular, we used consonant-vowel stimuli with a fundamental frequency of 100 Hz, the optimal stimuli that was used at the time the experiment was design to elicit FFRs. However, as discussed in detail in chapter 7, a major contribution of the auditory cortex is observed on the FFRs elicited to low frequency sounds, which makes us question the influence of the cortical areas on the effects of temporal predictability and statistical regularities reported. Nevertheless, our findings were recently

replicated (Bidelman and Powers, 2018) with stimulus that ensure the subcortical origin, which confirms the results and the conclusions obtained in our study.

In a similar manner, the study presented in chapter 6 shares, to some extent, the same limitation. The main feature of the experimental design of the mentioned study is the broad range of frequencies used, which allowed us to address the two hypothesis regarding the modulation of the FFR depending on the eliciting frequency and the contribution of the subcortical sound encoding to a latter perceptual decision making. By using this big range of frequencies in our auditory input, we cannot disregard a cortical contribution to the FFRs recorded to the lowest frequency sounds, up to 150 Hz approximately. On the other hand, this study has another limitation also related to the stimuli used for the paradigm. The stimuli chosen are pure tones, which are not the most ecologically valid sounds. The rationale behind the selection of this tones was to start exploring the contribution of the subcortical sound encoding to simple perceptual decision making with the most basic type of stimuli, such that the contribution observed could set a baseline for future studies and would have the less confounds possible. Therefore, future studies should extend our study by using consonant-vowel sounds or other environmental sounds, which are the ones we usually have to make perceptual decisions with.

From our point of view, another strength of this thesis lies in the different electrophysiological techniques that have been used for the study described in chapter 7, aimed at disentangling the neural generators of the FFR. Previous studies have approach the same topic with only one electrophysiological technique, which gave them some insights into the neural generators, but all of them had some limitations. In our opinion, approaching the aim both with EEG and MEG provides us the tools to join the previous research in one study, as each of the techniques has its own limitations but they complement each other. The drawback in this regard is that for a three-year thesis it is a lot of information to process, and therefore the results presented in chapter 7 are only preliminary, and further analysis has to be done to extract the most information about the FFR and its neural origins.

Nevertheless, and despite the aforementioned limitations of the studies, the studies comprised on this thesis and described and discussed throughout the chapters provide an insight in the role of the subcortical auditory pathway in the encoding and processing of sounds, as well as in the characterization of the FFR, a potential biomarker for auditory sound encoding.

CHAPTER 9

CONCLUSIONS

Based on the three studies that were conducted for this thesis, the main conclusions can be summarized as follows:

I. The early neural representation of sounds occurring in the subcortical auditory pathway is sensitive to both statistical regularities and temporal predictability of the acoustic environment. In particular, repetitive stimulation reduces the neural responses in the subcortical auditory system. This reduction, although present independently of temporal aspects of the auditory input, is further enhanced by the temporal predictability of the incoming acoustic stimulation. Overall, this reveals that temporally predictable stimulation enhances repetition suppression in the subcortical stations of the auditory hierarchy.

II. In addition to repetitive stimulation and temporal predictability, frequency has been found to modulate the online formation of predictive models in the subcortical stages of the auditory pathway. Higher frequencies are related to an enhanced signal – to –noise ratio in the FFR (a correlate of the encoding strength of the sounds) which, in turn, correlates with a faster detection of the same stimuli presented afterwards. This indicates that the encoding strength in the subcortical auditory pathway is used as evidence when making simple auditory perceptual decisions.

III. The importance of the frequency in the encoding of sounds throughout the auditory hierarchy has been highlighted by recent

studies that claimed that the FFRs elicited to stimuli below 150 Hz had clear contributors from the auditory cortex. By replicating and extending this research, we have observed that for sounds below 100 Hz both subcortical and cortical auditory stations are involved in the generation of the FFR responses. Conversely, when the frequency is around 300 Hz, only the IC and the MGB contribute to it and no activity from the cortical areas can be observed.

IV. Overall, it is concluded that the subcortical auditory pathway has an active role in the encoding of incoming auditory stimulation before the sounds reach the cortical structures. The frequency and the statistical and temporal characteristics of the auditory stimulation modulate the aggregate FFR response recorded from the scalp. Although the FFR has a multi-generator nature and represents an integrated response of the entire auditory system, by using the appropriate stimulus parameters it can still be used as a window into human subcortical sound encoding.

CHAPTER 10

RESUM EN CATALÀ (EXTENS)

Introducció

Analitzar i entendre la escena acústica en la que vivim el nostre dia a dia és un fet crucial per la nostre supervivència. Quan escoltem, els sons de l'entorn acústic arriben a les nostres orelles en forma d'un estímul auditiu que és una barreja complexa de les diferents fonts de sons que podem percebre en un moment determinat (Griffiths and Warren, 2004). Una teoria emergent en el àmbit de la neurociència cognitiva és que el sistema auditiu codifica els canvis del nostre entorn auditiu mitjançant l'extracció i la codificació de les relacions entre els diferents esdeveniments sonors i que utilitza aquestes representacions predictives de curt termini per tal d'organitzar l'entorn acústic en percepcions significatives (Winkler et al., 2009; McDermott et al., 2011; Bizley and Cohen, 2013), així com per predir esdeveniments sensorials futurs (Friston, 2005; Winkler et al., 2009) i per detectar de manera automàtica aquells esdeveniments sonors que no coincideixen amb les prediccions fetes (Winkler, 2008; Bendixen et al., 2012). Així doncs, els objectes auditius són el resultat computacional de la habilitat del sistema auditiu per detectar, extreure, segregar i agrupar les diferents regularitats espectro – temporals de l'entorn acústic en diferents unitats perceptuals estables (Bregman, 1990; Winkler et al., 2009; Schnupp et al., 2011).

Aquests processos crítics per entendre el nostre entorn auditiu es poden avaluar en el cervell humà mitjançant el registre dels potencials evocats auditius (AEP) obtinguts a partir d'un electroencefalograma (EEG). Els potencials evocats auditius humans són una sèrie de respostes cerebrals complexes provocades per estímuls auditius que, depenent de la latència post-estímul en la que ocorrin, es poden classificar en tres grups: Respostes del Tronc Encefàlic (ABR), Respostes de latència mitjana (MLR) i Respostes de latència llarga (LLR) (Luck, 2005).

Fins al moment, la codificació de les regularitats auditives s'ha estudiat de manera extensa en el còrtex humà i els processos que els causen estan clarament establerts (Haenschel et al., 2005; Baldeweg, 2006; Costa-Faidella et al., 2011a; Escera et al., 2014). Tot i així, estudis neurofisiològics realitzats en humans (Costa-Faidella et al., 2011b; Chandrasekaran et al., 2012; Cacciaglia et al., 2015) i en models animals (Pérez-González et al., 2005; Anderson et al., 2009; Malmierca et al., 2009; Antunes et al., 2010) proporcionen evidència directa que la codificació de les regularitats és una propietat ubiqua del sistema auditiu, fet que demostra que les estructures subcorticals de la via auditiva contribueixen també a aquests processos predictius i, per extensió, a la cognició auditiva.

Els processos auditius subcorticals es poden estudiar mitjançant un component de les ABR anomenat Resposta de Seguiment de Freqüència (FFR – de l'anglès Frequency - Following Response; Moushegian et al., 1973), un potencial elèctric sostingut que té la capacitat de sincronitzar-se amb la periodicitat de l'estímul auditiu, representant així les propietats temporals i espectrals específiques de la senyal auditiva (Skoe and Kraus, 2010a; Kraus et al., 2017) i que es considera una mesura de la codificació dels sons a nivell subcortical (Chandrasekaran and Kraus, 2010; Bidelman, 2018). La FFR és altament sensitiva a contingències dependents de l'estímul (Slabu et al., 2012; Chandrasekaran et al., 2014; Skoe et al., 2014) i a les propietats estadístiques de la estimulació auditiva entrant (Chandrasekaran et al., 2009; Skoe and Kraus, 2010b; Skoe et al., 2013, 2015; Escera, 2017) i ens permet obtenir una mesura no invasiva de la transcripció neural dels sons, així com de la manera en que l'entrenament auditiu a curt termini (Russo et al., 2005; Song et al., 2008, 2012; Carcagno and Plack, 2011; per una revisió veure Carcagno and Plack, 2017) i les experiències auditives, com la experiència lingüística (Krishnan et al., 2008; Krizman et al., 2012, 2014,

2015; Skoe et al., 2017) o l'entrenament musical (Musacchia et al., 2007; Wong et al., 2007; Parbery-Clark et al., 2009, 2011; Bidelman et al., 2011b; Skoe and Kraus, 2012), transformen la representació neural d'aquests en una xarxa auditiva subcortico–cortical (Kraus and White-Schwoch, 2015; Kraus and Slater, 2016).

Degut a la seva latència i a la seva amplitud, fins ara s'havia considerat que la FFR s'originava a partir d'agregats neuronals situats en diferents estructures del tronc encefàlic, essent el col·licle inferior la font neural principal. Tot i així, una recent línia d'evidència destinada a localitzar les fonts neuronals de la FFR va suggerir que la FFR a un estímul auditiu d'una freqüència fonamental d'uns 100 Hz no només representa la codificació neuronal de les característiques periòdiques d'aquest so en el tronc cerebral, sinó que també rep una important contribució de l'escorça auditiva (Coffey et al., 2016, 2017). Tanmateix, aquesta contribució cortical desapareix en freqüències superiors a 150 Hz (Bidelman, 2018). Aquests resultats posen en dubte l'assumpció que la FFR és un correlat de la codificació dels sons a nivell subcortical i donen suport a un punt de vista emergent en la literatura que proposa la FFR com una resposta integrada de tot el sistema auditiu (Kraus and White-Schwoch, 2015; Kraus and Slater, 2016).

Objectius

Aquesta tesi doctoral té com a objectiu principal examinar la contribució de la via subcortical auditiva en la codificació i processament dels sons, així com caracteritzar amb més profunditat la FFR mitjançant EEG i MEG. Més concretament, en aquesta tesi hem investigat com la estadística dels sons entrants i la predictibilitat temporal d'aquests modula la codificació de les regularitats auditives en la via subcortical auditiva i com la força d'aquesta codificació en aquesta via influencia la subseqüent presa de decisions perceptives auditives simples. A més, hem investigat com la freqüència dels sons entrants modulen la potència de la FFR, un correlat de la codificació dels sons, així com els orígens neurals que contribueixen a aquesta resposta. Aquesta tesis és una compilació de tres estudis i els objectius específics de cadascun d'ells es troben descrits a continuació.

Estudi I

El primer estudi de la tesis doctoral es va dissenyar per investigar si la predictibilitat temporal de la estimulació auditiva entrant modula la supressió per repetició en etapes subcorticals de la jerarquia de processament auditiu. Estudis previs han proposat la supressió per repetició com un mecanisme subjacent en el procés de codificació de les regularitats i han demostrat que la estimulació repetitiva redueix l'activitat neuronal auditiva en nivells corticals i subcorticals de la via auditiva animal i en nivells corticals de la via auditiva humana. Tot i així, altres factors contextuals, com la predictibilitat temporal dels estímuls pot influenciar la codificació d'aquestes regularitats estadístiques. En aquest estudi vam mesurar la FFR auditiva humana davant de la estimulació repetitiva d'un estímul consonant – vocal (/wa/) presentat de manera temporalment predictible o impredictible. Aquest disseny permet estudiar com la FFR és modulada tant per la estadística dels estímuls (és a dir, la repetició) com per la predictibilitat temporal d'aquests. La nostre hipòtesis va ser que la FFR seria modulada pels dos factors, revelant així que inclús les representacions neurals tempranes del so es troben incrementades per l'expectació temporal de les regularitats estadístiques d'aquests.

Estudi II

Els objectius del segon estudi van ser investigar la contribució de la codificació subcortical dels sons en la presa de decisions perceptives auditives simples en l'escorça cerebral i clarificar una característica bàsica de la FFR: com la seva potència es modula en funció de la freqüència de l'estímul que la produeix. Estudis recents (incloent el estudi I d'aquesta tesi doctoral) han demostrat que el sistema auditiu subcortical té un paper actiu en la percepció i el processament dels sons, consistent amb la hipòtesis de l'existència d'una xarxa distribuïda per l'organització perceptiva. En aquest estudi hem enregistrat la FFR a un grup de tons purs de 20 freqüències diferents, així com els temps de reacció a aquests mateixos sons presentats en dos blocs: un abans i un després del registre de FFR. Mitjançant l'ús d'aquest paradigma de temps de reacció auditiu simple podem obtenir una mesura de com la codificació de sons a nivell subcortical modula una codificació perceptiva subseqüent sense la confusió dels processos de presa de decisions complexes

Estudi III

Finalment, aquest tercer estudi de la tesi doctoral tenia com a objectiu entendre la contribució anatòmica a les FFRs produïdes per sons de diferents freqüències. Des dels primers estudis, la FFR s'ha considerat com un correlat de la codificació dels sons a nivell subcortical. Tot i així, estudis recents han posat en dubte aquesta suposició, demostrant que a freqüències baixes (<100Hz), les FFR reflecteixen activitat tant cortical com subcortical. Basant-nos en les capacitats de sincronització de fase específiques per freqüència al llarg de la via auditiva, aquest estudi es va dissenyar amb l'objectiu de dissociar una jerarquia d'orígens anatòmics que contribueixen a la codificació dels estímuls periòdics de diferents freqüències. Per aconseguir-ho s'han enregistrat de manera simultània amb MEG i EEG les FFRs a tons purs de 89 i 333 Hz presentats de manera repetitiva. Això ens permet utilitzar els models de generadors distribuïts per analitzar la contribució del tronc encefàlic, el tàlem i les estructures corticals a les FFRs. La nostre hipòtesis va ser que les FFRs produïdes en resposta a sons d'alta freqüència tindran una menor contribució cortical que aquelles produïdes en freqüències baixes, donant suport així a la participació de les estructures subcorticals en el processament temprà dels sons d'alta freqüència.

Estudi I

Resum (traducció de l'*abstract* original)

La codificació de les regularitats temporals és una propietat critica del sistema auditiu, ja que la representació neural a curt terme de la estadística dels sons del nostre entorn serveix per la formació d'objectes auditius i la detecció de estímuls nous potencialment rellevants. Un dels mecanismes neuronals que s'han proposat com a mecanisme subjacent en el procés de codificació de les regularitats és la supressió de la repetició, la reducció de la activitat neuronal davant la estimulació repetitiva. Tot i que s'ha demostrat que la estimulació repetitiva en si redueix l'activitat neuronal auditiva en nivells corticals i subcorticals de la via auditiva animal i en nivells corticals de la via auditiva humana, altres factors com la temporalitat dels estímuls pot influenciar la codificació d'aquestes regularitats estadístiques. En aquest estudi vam investigar si la predictibilitat de la temporalitat de la estimulació

auditiva pot modular la supressió de la repetició en nivells subcorticals de la jerarquia de processament auditiu. Amb aquest objectiu, vam mesurar la resposta de seguiment de freqüència (FFR; de l'anglès Frequency Following Response) auditiva humana davant de la estimulació repetitiva d'un estímul consonant – vocal (/wa/) presentat de manera temporalment predictible o impredictible. Vam observar una amplitud de la FFR atenuada degut a la repetició independentment de la predictibilitat temporal, però al mateix temps, vam observar una accentuació d'aquesta supressió quan la estimulació era temporalment predictible. Aquestes observacions ens permeten donar suport a la teoria que la codificació de les regularitats està present en tota la jerarquia auditiva i apunten a que la predictibilitat temporal és un factor modulador de la codificació de les regularitats en nivells primerencs de la via auditiva.

Estudi II

Resum (traducció de l'*abstract* original)

Es creu que la presa de decisions perceptives simples en el sistema auditiu està mediada pels processament neuronal que ocorre en la via auditiva ventral de l'escorça cerebral. De fet, s'ha observat que al llarg de les estructures que composen aquesta via, la informació auditiva s'organitza i es processa de manera jeràrquica. Tot i així, estudis previs han demostrat que el processament dels sons no es produeix exclusivament en regions corticals, sinó que les estructures subcorticals de la via auditiva també tenen un paper actiu en la codificació perceptiva dels sons entrants, consistent amb la hipòtesis de l'existència d'una xarxa distribuïda d'organització perceptiva. Amb l'objectiu d'avaluar la potencial contribució subcortical en la presa de decisions perceptives auditives simples, hem enregistrat la resposta de seguiment de freqüència (FFR; de l'anglès Frequency Following Response) a un grup de tons purs de 20 freqüències diferents, així com els temps de reacció a aquests mateixos sons presentats en dos blocs: un abans i un després del registre de FFR. Els nostres resultats van revelar una modulació de la potència normalitzada de la FFR depenent de la freqüència del so que es presentava que, subseqüentment, correlacionava amb una detecció més ràpida dels sons després del registre de FFR però no abans. En conjunt, els

nostres resultats donen suport a la hipòtesis que la codificació de sons a nivell subcortical té un paper important en la presa de decisions perceptives simples, indicant d'aquesta manera que els models d'acumulació d'evidència haurien de tenir en consideració una contribució subcortical.

Estudi III

Resum (traducció de l'*abstract* original)

La resposta de seguiment de freqüència auditiva (FFR; de l'anglès Frequency – Following Response) a sons periòdics complexos és una mesura no invasiva de la transcripció neural dels sons, així com de la manera en que les experiències auditives transformen aquestes representacions. Tot i que al llarg de la literatura s'ha considerat com un correlat de la codificació dels sons a nivell subcortical, l'únic estudi realitzat fins al moment que utilitza magnetoencefalografia (MEG) per localitzar l'origen neuronal de la FFR va qüestionar aquesta hipòtesis, demostrant que la FFR rep una important contribució de l'escorça auditiva. Basant-nos en les capacitats de sincronització específiques de freqüència de les diferents estructures de la jerarquia auditiva, en aquest estudi es va plantejar la hipòtesis que la FFR a freqüències altes rebrien una menor contribució de les estructures corticals en comparació a les FFRs produïdes en resposta a sons de freqüència més baixa, donant suport a la participació de les estructures subcorticals en l'origen neural de les respostes originades als sons d'alta freqüència. Amb aquest objectiu, es van registrar simultàniament amb magnetoencefalografia (MEG) i electroencefalografia (EEG) les FFRs a tons purs de 89 i 333 Hz respectivament presentats de manera repetitiva. Les FFRs produïdes tant pels sons de baixa com pels d'alta freqüència són visibles tant en els registres de MEG com en els de EEG. Utilitzant un model de localització de generadors distribuïts, hem analitzat i descrit la contribució de les estructures del tronc encefàlic i de l'escorça auditiva en les FFRs registrades. Els nostres resultats van revelar que, com es demostrava en estudis previs, la contribució a les FFRs de sons de baixa freqüència inclou totes les estructures de la via auditiva, des del nucli coclear fins a l'escorça auditiva primària. En canvi, a les FFRs produïdes pels sons d'alta freqüència només s'observa contribució provinent del col·licle inferior i del

cos geniculat medial del tàlem, i no s'observa cap contribució provinent de les estructures corticals. Aquests resultats donen suport a la hipòtesis de que augmentant la freqüència dels estímuls es pot fer decréixer, fins a desaparèixer, la contribució cortical a la resposta de seguiment de freqüència registrada degut a que les àrees corticals no tenen la capacitat de sincronitzar-se amb freqüències elevades. En conjunt, els resultats són molt rellevants per entendre la codificació dels sons al llarg de la via auditiva i ens permeten afirmar que, utilitzant estímuls amb paràmetres adequats, la resposta de seguiment de freqüència es pot seguir utilitzant com una finestra a la codificació dels sons en la via subcortical auditiva humana.

Resum dels resultats i discussió

L'objectiu d'aquesta tesis doctoral var ser examinar la hipòtesis de que la via subcortical auditiva són més que un conjunt d'estructures de pas per el processament dels sons i que tenen un paper actiu en la codificació dels sons entrants. En particular, en aquesta tesis ens vam proposar estudiar si la via subcortical auditiva humana participa en la codificació de les regularitats de l'entorn acústic (estudi I) i si aquesta codificació contribueix en la presa de decisions perceptuals auditives simples a nivell de l'escorça auditiva (Estudi II). La contribució de la via subcortical auditiva s'estudia mitjançant la resposta de seguiment de freqüència (FFR), que reflexa la activitat neuronal sincronitzada a les característiques espectrals i temporals de la senyal acústica que la causa. Al llarg de la literatura existent sempre s'ha considerat que la FFR proporciona una finestra al processament de sons a nivell subcortical, però a mitjans d'aquesta tesis, nous estudis van fer ressorgir controvèrsies respecte els seus orígens. Per això, amb l'objectiu de aprofundir en la FFR, el tercer objectiu de la tesis va ser estudiar com la freqüència dels sons modula la potència de la FFR (Estudi II) i com les diferents estructures de la via auditiva contribueixen a la seva generació (Estudi III).

Utilitzant un paradigma de repetició en el que una síl·laba es presenta de manera repetitiva de manera predictible o impredictible, els resultats de l'estudi I demostren que el processament subcortical dels sons es modula tant per les característiques estadístiques de l'estímul (es a dir,

per les regularitats auditives) com per la predictibilitat temporal. En particular, els resultats mostren una atenuació de l'amplitud de la FFR a mesura que augmenta el numero de repeticions en ambdues condicions temporals, indicant d'aquesta manera que independentment del context temporal de l'estimulació auditiva, la FFR es redueix quan els sons són repetitius. A més, aquesta reducció de la resposta neuronal causada per la estimulació repetitiva es veu incrementada quan la estimulació auditiva és temporalment predictible. Això és degut a que la predictibilitat temporal ajuda a la codificació dels sons presentats ja que incrementa la senyal de l'estímul d'interès respecte al "soroll" de fons. En conjunt, aquests resultats revelen que la via auditiva subcortical està activament involucrada en el processament dels sons entrants i que es sensitiva tant a la predictibilitat temporal com a les regularitats estadístiques de l'entorn acústic, de tal manera que la representació neural temprana dels sons és més precisa amb la temporalitat de les regularitats estadístiques codificades.

Anant un pas més enllà, en l'estudi II hem demostrat que aquest paper actiu de la via subcortical auditiva en la codificació perceptual dels sons entrants es reflexa en la presa de decisions perceptuals simples auditives a nivell de l'escorça auditiva, especialment en decidir si un so està present o no. Els nostres resultats mostren una modulació de la potència de la FFR depenent de la freqüència del so que la produeix, incrementant a mesura que la freqüència dels sons augmenta. Aquesta modulació correlaciona amb una detecció més ràpida dels sons després del temps d'escolta passiva en el qual es registrava la FFR, però no amb la detecció dels sons abans, suggerint així que una millor codificació dels sons durant el temps d'escolta passiva porta a una detecció posterior més ràpida dels sons, essent la freqüència un element modulador de la codificació crític.

Així doncs, a més de l'estimulació repetitiva i la predictibilitat temporal, s'ha observat que la freqüència és també un factor modulador en la formació del models predictius en les estacions subcorticals de la jerarquia auditiva. De fet, estudis recents han demostrat que per sons de freqüències baixes (<100 Hz) la FFR reflexa no només activitat subcortical, sinó també cortical. Després d'observar que la freqüència modula la potència de la FFR (Estudi II), l'estudi III d'aquesta tesis tenia com a objectiu dissociar la jerarquia de fonts anatòmiques que contribueixen a les FFRs produïdes per sons de diferents freqüències, en concret per tons de baixa (89 Hz) i alta (333 Hz) freqüència.

Per separar la contribució dels possibles generadors, hem registrat la FFR a aquests dos tons presentats de manera repetitiva amb MEG i hem fet una reconstrucció dels generadors neuronals mitjançant un model d'estimació de norma mínima (wMNE). Posteriorment, van definir regions d'interès bilaterals en les principals estructures de la via auditiva (còrtex auditiu primari, cos geniculat medial del tàlem, col·licle inferior i nucli coclear), així com en dues regions control localitzades en els pols frontals i occipitals. Els nostres resultats mostren que, per als sons de baixa freqüència, totes les regions d'interès de la jerarquia auditiva contribueixen a la FFR registrada en comparació amb l'activitat procedent de les regions de control, confirmant així els resultats trobats en estudis previs. En comparació, per la FFR d'alta freqüència només s'observen generadors localitzats en les estructures subcorticals, concretament en el cos geniculat medial del tàlem i en el col·licle inferior, i no s'observa cap contribució de les estructures corticals. Aquests resultats demostren que, efectivament, la freqüència dels estímuls és un factor clau en l'estudi de la codificació neuronal dels sons i indiquen que, tot i els estudis recents, si s'utilitzen els paràmetres d'estimulació adequats la resposta de seguiment de freqüència es pot seguir utilitzant com a correlat neuronal de la codificació dels sons a nivell subcortical.

Conclusions

En conjunt, a partir dels tres estudis que s'han realitzat en aquesta tesis doctoral, les principals conclusions extretes es poden resumir de la següent manera:

I. La representació neural temprana dels sons que té lloc a la via auditiva subcortical és sensible tant a la regularitat estadística com a la predictibilitat temporal de l'entorn acústic. En particular, l'estimulació repetitiva redueix les respostes neurals en el sistema auditiu subcortical. Aquesta reducció, tot i que està present independent dels aspectes temporals de l'estimulació auditiva, es veu incrementada per la predictibilitat temporal dels estímuls auditius entrants. En general, aquests resultats revelen que l'estimulació temporalment predictible millora la supressió per repetició en les estacions subcorticals de la jerarquia auditiva.

II. A més de l'estimulació repetitiva i la predictibilitat temporal, hem trobat que la freqüència també modula, de manera instantània, la formació de model predictius en les etapes subcorticals de la via auditiva. En particular, les freqüències més altes es relacionen amb una millor extracció de senyal en soroll de la FFR (un correlat del grau de codificació del sons) que, al seu torn, es correlaciona amb una detecció més ràpida dels mateixos estímuls presentats posteriorment. Això indica que el grau de codificació en la via subcortical auditiva s'utilitza com a evidència quan es prenen decisions auditives perceptives senzilles.

III. La importància de la freqüència en la codificació de sons al llarg de la jerarquia auditiva també ha estat destacada per estudis recents en els quals s'afirma que les FFR produïdes per estímuls inferiors a 150 Hz tenen clars contribuents neurals provinents de l'escorça auditiva. En aquesta tesis hem replicat i ampliat aquesta investigació, observant que per als sons inferiors a 100 Hz, tant les estacions subcorticals de la via auditiva com les corticals intervenen en la generació de la FFR. Contràriament, quan la freqüència és d'uns 300 Hz, només l'IC i el MGB contribueixen a la generació d'aquesta i no es pot observar cap contribució a la FFR provinent de les àrees corticals auditives.

IV. En general, en aquesta tesis es conclou que la via auditiva subcortical té un paper actiu en la codificació de l'estimulació auditiva abans que els sons arribin a les estructures corticals. La freqüència i les característiques estadístiques i temporals de l'estimulació auditiva modulen la resposta de seguiment de freqüència registrada amb elèctrodes situats en el cuir cabellut. Anant més enllà, tot i que la FFR està generada per múltiples generadors i representa una resposta integrada de tot el sistema auditiu, mitjançant uns paràmetres d'estímul adequats, encara es pot utilitzar com a finestra per l'estudi de la codificació dels sons a nivell subcortical en humans.

CHAPTER 11

REFERENCES

Aghamolaei M, Zarnowiec K, Grimm S, Escera C (2016) Functional dissociation between regularity encoding and deviance detection along the auditory hierarchy. Eur J Neurosci 43:529–535.

Aiken SJ, Picton TW (2008) Envelope and spectral frequency-following responses to vowel sounds. Hear Res 245:35–47.

Althen H, Grimm S, Escera C (2011) Fast detection of unexpected sound intensity decrements as revealed by human evoked potentials. PLoS One 6:e28522.

Althen H, Grimm S, Escera C (2013) Simple and complex acoustic regularities are encoded at different levels of the auditory hierarchy. Eur J Neurosci 38:3448–3455.

Anderson LA, Christianson GB, Linden JF (2009) Stimulus-specific adaptation occurs in the auditory thalamus. J Neurosci 29:7359–7363.

Anderson LA, Malmierca MS (2013) The effect of auditory cortex deactivation on stimulus-specific adaptation in the inferior colliculus of the rat. Eur J Neurosci 37:52–62.

Anderson S, Skoe E, Chandrasekaran B, Kraus N (2010) Neural timing is linked to speech perception in noise. J Neurosci 30:4922–4926.

Antunes FM, Malmierca MS (2011) Effect of Auditory Cortex Deactivation on Stimulus-Specific Adaptation in the Medial Geniculate Body. J Neurosci 31:17306–17316.

Antunes FM, Malmierca MS (2014) An overview of stimulus-specific adaptation in the auditory thalamus. Brain Topogr 27:480–499.

Antunes FM, Nelken I, Covey E, Malmierca MS (2010) Stimulus-specific adaptation in the auditory thalamus of the anesthetized rat. PLoS One 5.

Assmann P, Summerfield Q (2004) The Perception of Speech Under Adverse Conditions. In: Speech Processing in the Auditory System, pp 231–308. New York, NY: Springer New York.

Attal Y, Schwartz D (2013) Assessment of Subcortical Source Localization Using Deep Brain Activity Imaging Model with Minimum Norm Operators: A MEG Study. PLoS One 8.

Auksztulewicz R, Friston K (2015) Repetition suppression and its contextual determinants in predictive coding. Cortex:1–16.

Ayala Y, Malmierca MS (2012) Stimulus-specific adaptation and deviance detection in the inferior colliculus. Front Neural Circuits 6:89.

Baldeweg T (2006) Repetition effects to sounds: evidence for predictive coding in the auditory system. Trends Cogn Sci 10:93–94.

Baldeweg T (2007) ERP repetition effects and mismatch negativity generation: A predictive coding perspective. J Psychophysiol 21:204–213.

Baldeweg T, Klugman A, Gruzelier J, Hirsch SR (2004) Mismatch negativity potentials and cognitive impairment in schizophrenia. Schizophr Res 69:203–217.

Banai K, Ahissar M (2006) Auditory processing deficits in dyslexia: Task or stimulus related? Cereb Cortex 16:1718–1728.

Banai K, Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N (2009) Reading and subcortical auditory function. Cereb Cortex 19:2699–2707.

Banai K, Nicol T, Zecker SG, Kraus N (2005) Brainstem timing: implications for cortical processing and literacy. J Neurosci 25:9850–9857.

Batra R, Kuwada S, Maher VL (1986) The frequency-following response to continuous tones in humans. Hear Res 21:167–177.

Batra R, Kuwada S, Stanford TR (1989) Temporal coding of envelopes and their interaural delays in the inferior colliculus of the unanesthetized rabbit. J Neurophysiol 61:257–268.

Bendixen A, Prinz W, Horváth J, Trujillo-Barreto N, Schröger E (2008) Rapid extraction of auditory feature contingencies. Neuroimage 41:1111–1119.

Bendixen A, SanMiguel I, Schröger E (2012) Early electrophysiological indicators for predictive processing in audition: A review. Int J Psychophysiol 83:120–131.

Bidelman GM (2015) Multichannel recordings of the human brainstem frequency-following response: Scalp topography, source generators, and distinctions from the transient ABR. Hear Res:1–13.

Bidelman GM (2018) Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. Neuroimage 175:56–69.

Bidelman GM, Gandour JT, Krishnan A (2011a) Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. J Cogn Neurosci 23:425–434.

Bidelman GM, Krishnan A, Gandour J (2011b) Enhanced brainstem encoding predicts musicians' perceptual advantages with pitch. Eur J Neurosci 33:530–538.

Bidelman GM, Moreno S, Alain C (2013) Tracing the emergence of categorical speech perception in the human auditory system. Neuroimage 79:201–212.

Bidelman GM, Powers L (2018) Response properties of the human frequency-following response (FFR) to speech and non- speech sounds : level dependence , adaptation and phase-locking limits. Int J Audiol 0:1–8.

Bizley JK, Cohen YE (2013) The what, where and how of auditory-object perception. Nat Rev Neurosci 14:693–707.

Boersma P (1993) Accurate short-term analysus of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. IFA Proc 17:97–110.

Bregman AS (1990) Auditory scene analysis: The Perceptual Organization of Sound. Bradfort, Cambridge, MA: The MIT Press, Massachusetts Institute of Technology,Cambridge, MA.

Busch NA, VanRullen R (2010) Spontaneous EEG oscillations reveal periodic sampling of visual attention. Proc Natl Acad Sci 107:16048–16053.

Cacciaglia R, Escera C, Slabu L, Grimm S, Sanjuan A, Ventura-Campos N, Avila C (2015) Involvement of the human midbrain and thalamus in auditory deviance detection. Neuropsychologia 68:51–58.

Carcagno S, Plack CJ (2011) Subcortical plasticity following perceptual learning in a pitch discrimination task. JARO - J Assoc Res Otolaryngol 12:89–100.

Carcagno S, Plack CJ (2017) Short-Term Learning and Memory: Training and Perceptual Learning. In: The Frequency-Following Response. Springer Handbook of Auditory Research, vol 61 (Kraus N, Anderson S, White-Schwoch T, Fay RR, Popper AN, eds), pp 75–100. Springer, Cham.

Chandrasekaran B, Hornickel J, Skoe E, Nicol T, Kraus N (2009) Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia. Neuron 64:311–319.

Chandrasekaran B, Kraus N (2010) The scalp-recorded brainstem response to speech: neural origins and plasticity. Psychophysiology 47:236–246.

Chandrasekaran B, Kraus N, Wong PCM (2012) Human inferior colliculus activity relates to individual differences in spoken language learning. J Neurophysiol 107:1325–1336.

Chandrasekaran B, Skoe E, Kraus N (2014) An integrative model of subcortical auditory plasticity. Brain Topogr 27:539–552.

Chen S, Sussman E (2013) Context effects on auditory distraction. Biol Psychol 94:297–309.

Coffey EBJ, Herholz SC, Chepesiuk AMP, Baillet S, Zatorre RJ (2016) Cortical contributions to the auditory frequency-following response revealed by MEG. Nat Commun 7:11070.

Coffey EBJ, Musacchia G, Zatorre RJ (2017) Cortical Correlates of the Auditory Frequency-Following and Onset Responses: EEG and fMRI Evidence. J Neurosci 37:830–838.

Cohen YE, Bennur S, Christison-Lagay K, Gifford A, Tsunada J (2016) Functional Organization of the Ventral Auditory Pathway. Adv Exp Med Biol 894:381–388.

Cooley JW, Tukey JW (1964) An Algorithm for the Machine Calculation Complex Fourier Series. Int J Comput Math 19:297–301.

Cornella M, Leung S, Grimm S, Escera C (2012) Detection of simple and pattern regularity violations occurs at different levels of the auditory hierarchy. PLoS One 7:e43604.

Cornella M, Leung S, Grimm S, Escera C (2013) Regularity encoding and deviance detection of frequency modulated sweeps: human middle- and long-latency auditory evoked potentials. Psychophysiology 50:1275–1281.

Costa-Faidella J, Baldeweg T, Grimm S, Escera C (2011a) Interactions between "what" and "when" in the auditory system: temporal predictability enhances repetition suppression. J Neurosci 31:18590–18597.

Costa-Faidella J, Grimm S, Slabu L, Díaz-Santaella F, Escera C (2011b) Multiple time scales of adaptation in the auditory system as revealed by human evoked potentials. Psychophysiology 48:774–783.

Cunningham J, Nicol T, Zecker SG, Bradlow A, Kraus N (2001) Neurobiologic responses to speech in noise in children with learning problems : deficits and strategies for improvement. Clin Neurophysiol 112:758–767.

Delorme A, Makeig S (2004) EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21.

Deouell LY (2007) The frontal generator of the mismatch negativity revisited. J Psychophysiol 21:188–203.

Desimone R (1996) Neural mechanisms for visual memory and their role in attention. Proc Natl Acad Sci U S A 93:13494–13499.

Destrieux C, Fischl B, Dale A, Halgren E (2010) Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage 53:1–15.

Duque D, Malmierca MS (2015) Stimulus-specific adaptation in the inferior colliculus of the mouse: anesthesia and spontaneous activity effects. Brain Struct Funct 220:3385–3398.

Escera C (2017) The Role of the Auditory Brainstem in Regularity Encoding and Deviance Detection. In: The Frequency-Following Response. Springer Handbook of Auditory Research, vol 61 (Kraus N, Anderson S, White-Schwoch T, Fay RR, Popper AN, eds), pp 101–120. Springer, Cham.

Escera C, Leung S, Grimm S (2014) Deviance Detection Based on Regularity Encoding Along the Auditory Hierarchy: Electrophysiological Evidence in Humans. Brain Topogr 27:527–538.

Escera C, Malmierca MS (2014) The auditory novelty system: an attempt to integrate human and animal research. Psychophysiology 51:111–123.

Fischl B (2012) FreeSurfer. Neuroimage 62:774–781.

Friston K (2005) A theory of cortical responses. Philos Trans R Soc Lond B Biol Sci 360:815–836.

Friston K, Kiebel S (2009) Predictive coding under the free-energy principle. Philos Trans R Soc Lond B Biol Sci 364:1211–1221.

Galbraith GC, Chae BC, Cooper JR, Gindi MM, Ho TN, Kim BS, Mankowski D a, Lunde SE (2000a) Brainstem frequency-following response and simple motor reaction time. Int J Psychophysiol 36:35–44.

Galbraith GC, Threadgill MR, Hemsley J, Salour K, Songdej N, Ton J, Cheung L (2000b) Putative measure of peripheral and brainstem frequency-following in humans. Neurosci Lett 292:123–127.

Gao X, Wehr M (2015) A coding transformation for temporally structured sounds within auditory cortical neurons. Neuron 86:292–303.

Gardi J, Merzenich M, Mckean C (1979) Origins of the scalp-recorded frequency-following response in the cat. Int J Audiol 18:353–380.

Garrido MI, Kilner JM, Stephan KE, Friston KJ (2009) The mismatch negativity: A review of underlying mechanisms. Clin Neurophysiol 120:453–463.

Gold JI, Shadlen MN (2007) The Neural Basis of Decision Making. Annu Rev Neurosci 30:535–574.

Goldstein EB (2009) Sensation and Perception (Hague J-D, Perkins JA, eds)., 8th ed. Belmont, CA: Wadsworth, Cengage Learning.

Gorina-Careta N, Zarnowiec K, Costa-Faidella J, Escera C (2016) Timing predictability enhances regularity encoding in the human subcortical auditory pathway. Sci Rep 6:37405.

Griffiths TD, Warren JD (2004) The nature of auditory objects. Nat Rev Neurosci 5:887–892.

Grill-Spector K, Henson R, Martin A (2006) Repetition and the brain: neural models of stimulus-specific effects. Trends Cogn Sci 10:14–23.

Grimm S, Escera C (2012) Auditory deviance detection revisited: evidence for a hierarchical novelty system. Int J Psychophysiol 85:88–92.

Grimm S, Escera C, Slabu L, Costa-Faidella J (2011) Electrophysiological evidence for the hierarchical organization of auditory change detection in the human brain. Psychophysiology 48:377–384.

Grimm S, Recasens M, Althen H, Escera C (2012) Ultrafast tracking of sound location changes as revealed by human auditory evoked potentials. Biol Psychol 89:232–239.

Hackett TA (2011) Information flow in the auditory cortical network. Hear Res 271:133–146.

Haenschel C, Vernon DJ, Dwivedi P, Gruzelier JH, Baldeweg T (2005) Event-related brain potential correlates of human auditory sensory memory-trace formation. J Neurosci 25:10494–10501.

Hämäläinen M (2009) MNE Software User's guide v2.7. MGH/HMS/MIT Athinoula A (Martinos Cent Biomed Imaging).

Heekeren HR, Marrett S, Ungerleider LG (2008) The neural systems that mediate human perceptual decision making. Nat Rev Neurosci 9:467–479.

Hoormann J, Falkenstein M, Hohnsbein J, Blanke L (1992) The human frequency-following response (FFR): Normal variability and relation to the click-evoked brainstem response. Hear Res 59:179–188.

Hornickel J, Anderson S, Skoe E, Yi HG, Kraus N (2012) Subcortical representation of speech fine structure relates to reading ability. Neuroreport 23:6–9.

Hornickel J, Kraus N (2013) Unstable Representation of Sound: A Biological Marker of Dyslexia. J Neurosci 33:3500–3504.

Huang MX, Mosher JC, Leahy RM (1999) A sensor-weighted overlapping-sphere head model and exhaustive head model comparison for MEG. Phys Med Biol 44:423–440.

ISO 226:2003 (2003) Acoustics—Normal Equal-Loudness Contours. Geneva, Switzerland.

Jeng F-C, Chung H-K, Lin C-D, Dickman B, Hu J (2011a) Exponential modeling of human frequency-following responses to voice pitch. Int J Audiol 50:582–593.

Jeng F-C, Hu J, Dickman B, Lin C-Y, Lin C-D, Wang C-Y, Chung H-K, Li X (2011b) Evaluation of two algorithms for detecting human frequency-following responses to voice pitch. Int J Audiol 50:14–26.

Joris PX, Schreiner CE, Rees A (2004) Neural processing of amplitude-modulated sounds. Physiol Rev 84:541–577.

Kelly SP, O'Connell RG (2015) The neural processes underlying perceptual decision making in humans: Recent progress and future directions. J Physiol Paris 109:27–37.

Klatt DH (1980) Software for cascade/parallel formant synthesizer. J Acoust Soc Am 67:971–975.

Kraus N, Anderson S, White-Schwoch T (2017) The Frequency-Following Response: A Window into human communication (Kraus N, Anderson S, White-Schwoch T, Fay RR, Popper AN, eds). Cham, Switzerland: Springer International Publishing.

Kraus N, Slater J (2016) Beyond Words: How Humans Communicate Through Sound. Annu Rev Psychol 67:83–103.

Kraus N, White-Schwoch T (2015) Unraveling the Biology of Auditory Learning: A Cognitive–Sensorimotor–Reward Framework. Trends Cogn Sci 19:642–654.

Krishnan A, Bidelman GM, Gandour J (2010) Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices. Hear Res 268:60–66.

Krishnan A, Swaminathan J, Gandour JT (2008) Experience-dependent Enhancement of Linguistic Pitch Representation in the Brainstem Is Not Specific to a Speech Context. J Cogn Neurosci 21:1092–1105.

Krishnan A, XuY, Gandour J, Cariani P (2004) Human frequency-following response: representation of pitch contours in Chinese tones. Hear Res 189:1–12.

Krishnan A, Xu Y, Gandour J, Cariani P (2005) Encoding of pitch in the human brainstem is sensitive to language experience. Brain Res Cogn Brain Res 25:161–168.

Krizman J, Marian V, Shook A, Skoe E, Kraus N (2012) Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. Proc Natl Acad Sci 109:7877–7881.

Krizman J, Skoe E, Marian V, Kraus N (2014) Bilingualism increases neural response consistency and attentional control: Evidence for sensory and cognitive coupling. Brain Lang 128:34–40.

Krizman J, Slater J, Skoe E, Marian V, Kraus N (2015) Neural processing of speech in children is influenced by extent of bilingual experience. Neurosci Lett 585:48–53.

Lange K (2009) Brain correlates of early auditory processing are attenuated by expectations for time and pitch. Brain Cogn 69:127–137.

Langner G (1992) Periodicity coding in the auditory system. Hear Res 60:115–142.

Langner G, Schreiner CE (1988) Periodicity coding in the inferior colliculus of the cat. I. Neuronal Mechanisms. J Neurophysiol 60:1799–1822.

Leung S, Cornella M, Grimm S, Escera C (2012) Is fast auditory change detection feature specific? An electrophysiological study in humans. Psychophysiology 49:933–942.

López-Caballero F, Zarnowiec K, Escera C (2016) Differential deviant probability effects on two hierarchical levels of the auditory novelty system. Biol Psychol 120:1–9.

Luck SJ (2005) Introduction to the Event related potential technique, 2nd ed. The MIT Press, Massachusetts Institute of Technology,Cambridge, MA.

Malmierca MS, Cristaudo S, Pérez-González D, Covey E (2009) Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. J Neurosci 29:5483–5493.

Malmierca MS, Sanchez-vives M V, Escera C, Bendixen A (2014) Neuronal adaptation, novelty detection and regularity encoding in audition. Front Syst Neurosci 8:1–9.

Marsh JT, Worden FG, Smith JC (1970) Auditory Frequency-Following Response: Neural or Artifact? Science 169:1222–1223.

McDermott JH, Wrobleski D, Oxenham a. J (2011) Recovering sound sources from embedded repetition. Proc Natl Acad Sci 108:1188–1193.

Moushegian G, Rupert AL, Stillman RD (1973) Scalp-recorded early responses in man to frquencies in the speech range. Electroencephalogr Clin Neurophysiol 35:665–667.

Musacchia G, Sams M, Skoe E, Kraus N (2007) Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. Proc Natl Acad Sci 104:15894–15898.

Mutschler I, Wieckhorst B, Speck O, Schulze-Bonhage A, Hennig J, Seifritz E, Ball T (2010) Time scales of auditory habituation in the amygdala and cerebral cortex. Cereb Cortex 20:2531–2539.

Näätänen R, Gaillard A, Mäntysalo S (1978) Early selective-attention effect on evoked potential reinterpreted. Acta Psychol (Amst) 42:313–329.

Näätänen R, Paavilainen P, Rinne T, Alho K (2007) The mismatch negativity (MMN) in basic research of central auditory processing: a review. Clin Neurophysiol 118:2544–2590.

Nelken I, Ulanovsky N (2007) Mismatch Negativity and Stimulus-Specific Adaptation in Animal Models. J Psychophysiol 21:214–223.

Paavilainen P, Arajìrvi P, Takegata R (2007) Preattentive detection of nonsalient contingencies between auditory features. Neuroreport 18.

Pannese A, Grandjean D, Frühholz S (2015) Subcortical processing in auditory communication. Hear Res 328:67–77.

Parbery-Clark A, Skoe E, Kraus N (2009) Musical Experience Limits the Degradative Effects of Background Noise on the Neural Processing of Sound. J Neurosci 29:14100–14107.

Parbery-Clark A, Strait DL, Kraus N (2011) Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. Neuropsychologia 49:3338–3345.

Parras GG, Nieto-Diego J, Carbajal G V., Valdés-Baizabal C, Escera C, Malmierca MS (2017) Neurons along the auditory pathway exhibit a hierarchical organization of prediction error. Nat Commun 8.

Parvizi J (2009) Corticocentric myopia: old bias in new cognitive sciences. Trends Cogn Sci 13:354–359.

Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD (2002) The processing of temporal pitch and melody information in auditory cortex. Neuron 36:767–776.

Pérez-González D, Malmierca MS (2014) Adaptation in the auditory system: an overview. Front Integr Neurosci 8:1–10.

Pérez-González D, Malmierca MS, Covey E (2005) Novelty detector neurons in the mammalian auditory midbrain. Eur J Neurosci 22:2879–2885.

Picton TW (2011) Human auditory evoked potentials. San Diego: Plural Publishing, Inc.

Prado-Gutierrez P, Castro-Fariñas A, Morgado-Rodriguez L, Velarde-Reyes E, Martínez AD, Martínez-Montes E (2015) Habituation of auditory steady state responses evoked by amplitude-modulated acoustic signals in rats. Audiol Res 5:21–29.

Pressnitzer D, Sayles M, Micheyl C, Winter IM (2008) Perceptual Organization of Sound Begins in the Auditory Periphery. Curr Biol 18:1124–1128.

Ratcliff R (1978) A theory of memory retrieval. Psychol Rev 85:59–108.

Ratcliff R, McKoon G (2008) The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. Neural Comput 20:873–922.

Rauschecker JP, Scott SK (2009) Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat Neurosci 12:718–724.

Rauschecker JP, Tian B (2000) Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc Natl Acad Sci 97:11800–11806.

Recasens M, Grimm S, Wollbrink A, Pantev C, Escera C (2014) Encoding of nested levels of acoustic regularity in hierarchically organized areas of the human auditory cortex. Hum Brain Mapp 35:5701–5716.

Recasens M, Leung S, Grimm S, Nowak R, Escera C (2015) Repetition suppression and repetition enhancement underlie auditory memory-trace formation in the human brain: An MEG study. Neuroimage 108:75–86.

Russo N, Nicol T, Trommer B, Zecker S, Kraus N (2009) Brainstem transcription of speech is disrupted in children with autism spectrum disorders. Dev Sci 12:557–567.

Russo N, Nicol T, Zecker SG, Hayes E a., Kraus N (2005) Auditory training improves neural timing in the human brainstem. Behav Brain Res 156:95–103.

Russo N, Skoe E, Trommer B, Nicol T, Zecker SG, Bradlow A, Kraus N (2008) Deficient brainstem encoding of pitch in children with Autism Spectrum Disorders. Clin Neurophysiol 119:1720–1731.

Schnupp J, Nelken I, King A (2011) Auditory neuroscience: Making sense of sound, 1st ed. The MIT Press, Massachusetts Institute of Technology,Cambridge, MA.

Schwartze M, Farrugia N, Kotz SA (2013) Dissociation of formal and temporal predictability in early auditory evoked potentials. Neuropsychologia 51:320–325.

Shadlen MN, Kiani R (2013) Decision making as a window on cognition. Neuron 80:791–806.

Shamma SA, Micheyl C (2010) Behind the scenes of auditory perception. Curr Opin Neurobiol 20:361–366.

Shiga T, Althen H, Cornella M, Zarnowiec K, Yabe H, Escera C (2015) Deviance-related responses along the auditory hierarchy: Combined FFR, MLR and MMN evidence. PLoS One 10:1–14.

Skoe E, Burakiewicz E, Figueiredo M, Hardin M (2017) Basic Neural Processing of Sound in Adults is Influenced by Bilingual Experience. Neuroscience 349:278–290.

Skoe E, Chandrasekaran B, Spitzer ER, Wong PCM, Kraus N (2014) Human brainstem plasticity: the interaction of stimulus probability and auditory learning. Neurobiol Learn Mem 109:82–93.

Skoe E, Kraus N (2010a) Auditory brain stem response to complex sounds: a tutorial. Ear Hear 31:302–324.

Skoe E, Kraus N (2010b) Hearing it again and again: On-line subcortical plasticity in humans. PLoS One 5:1–9.

Skoe E, Kraus N (2012) A Little Goes a Long Way: How the Adult Brain Is Shaped by Musical Training in Childhood. J Neurosci 32:11507–11510.

Skoe E, Krizman J, Spitzer E, Kraus N (2015) Prior Experience Biases Subcortical Sensitivity to Sound Patterns. J Cogn Neurosci 27:124–140.

Skoe E, Krizman J, Spitzer ER, Kraus N (2013) The auditory brainstem is a barometer of rapid auditory learning. Neuroscience 243:104–114.

Slabu L, Escera C, Grimm S, Costa-Faidella J (2010) Early change detection in humans as revealed by auditory brainstem and middle-latency evoked potentials. Eur J Neurosci 32:859–865.

Slabu L, Grimm S, Escera C (2012) Novelty detection in the human auditory brainstem. J Neurosci 32:1447–1452.

Smith JC, Marsh JT, Brown WS (1975) Far-Field Recorded Frequency-Following Responses: Evidence for the locus of brainstem sources. Electroencephalogr Clin Neurophysiol 39:465–472.

Smith PL, Ratcliff R (2004) Psychology and neurobiology of simple decisions. Trends Neurosci 27:161–168.

Sohmer H, Pratt H, Kinarti R (1977) Sources of frequency following responses (FFR) in man. Electroencephalogr Clin Neurophysiol 42:656–664.

Song J, Skoe E, Banai K, Kraus N (2011) Perception of speech in noise: neural correlates. J Cogn Neurosci 23:2268–2279.

Song J, Skoe E, Banai K, Kraus N (2012) Training to improve hearing speech in noise: biological mechanisms. Cereb Cortex 22:1180–1190.

Song JH, Skoe E, Wong PCM, Kraus N (2008) Plasticity in the adult human auditory brainstem following short-term linguistic training. J Cogn Neurosci 20:1892–1902.

Sonnadara RR, Alain C, Trainor LJ (2006) Occasional changes in sound location enhance middle latency evoked responses. Brain Res 1076:187–192.

Summerfield C, Wyart V, Johnen VM, de Gardelle V (2011) Human Scalp Electroencephalography Reveals that Repetition Suppression Varies with Expectation. Front Hum Neurosci 5:67.

Taaseh N, Yaron A, Nelken I (2011) Stimulus-Specific Adaptation and Deviance Detection in the Rat Auditory Cortex. PLoS One 6:e23369.

Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM (2011) Brainstorm: A user-friendly application for MEG/EEG analysis. Comput Intell Neurosci 2011.

Taulu S, Kajola M, Simola J (2004) Suppression of interference and artifacts by the signal space separation method. Brain Topogr 16:269–275.

Tesche CD, Uusitalo MA, Ilmoniemi RJ, Huotilainen M, Kajola M, Salonen O (1995) Signal-space projections of MEG data characterize both distributed and well-localized neuronal sources. Electroencephalogr Clin Neurophysiol 95:189–200.

Tichko P, Skoe E (2017) Frequency-dependent fine structure in the frequency-following response: The byproduct of multiple generators. Hear Res 348:1–15.

Todorovic A, de Lange FP (2012) Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. J Neurosci 32:13389–13395.

Todorovic A, van Ede F, Maris E, de Lange FP (2011) Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: an MEG study. J Neurosci 31:9118–9123.

Tsunada J, Liu ASK, Gold JI, Cohen YE (2015) Causal contribution of primate auditory cortex to auditory perceptual decision-making. Nat Neurosci 19:135–142.

Ulanovsky N, Las L, Farkas D, Nelken I (2004) Multiple time scales of adaptation in auditory cortex neurons. J Neurosci 24:10440–10453.

Ulanovsky N, Las L, Nelken I (2003) Processing of low-probability sounds by cortical neurons. Nat Neurosci 6:391–398.

Wacongne C (2016) A predictive coding account of MMN reduction in schizophrenia. Biol Psychol 116:68–74.

Wacongne C, Changeux J-P, Dehaene S (2012) A Neuronal Model of Predictive Coding Accounting for the Mismatch Negativity. J Neurosci 32:3665–3678.

Warren JD, Griffiths TD (2003) Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. J Neurosci 23:5799–5804.

Weinberger NM, Kitzes LM, Goodman DA (1970) Some characteristics of the "auditory neurophonic." Experientia 26:46–48.

Winkler I (2008) Interpreting the Mismatch Negativity. J Psychophysiol 21:147–163.

Winkler I, Denham SL, Nelken I (2009) Modeling the auditory scene: predictive regularity representations and perceptual objects. Trends Cogn Sci 13:532–540.

Wong PCM, Skoe E, Russo N, Dees T, Kraus N (2007) Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat Neurosci 10:420–422.

Yamagishi S, Otsuka S, Furukawa S, Kashino M (2016) Subcortical correlates of auditory perceptual organization in humans. Hear Res 339:104–111.

Yvert B, Crouzeix A, Bertrand O (2001) Multiple Supratemporal Sources of Magnetic and Electric Auditory Evoked Middle Latency Components in Humans. Cereb Cortex 11:411–423.