



UNIVERSITAT DE
BARCELONA

Genetic and Environmental Modulators of Neural Encoding Associated to Speech Processing and Their Relationship with Auditory Skills

Giannina Puddu Gallardo

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.



UNIVERSITAT DE
BARCELONA

Department of Clinical Psychology and Psychobiology
Faculty of Psychology

**Genetic and Environmental Modulators of Neural Encoding Associated to Speech
Processing and Their Relationship with Auditory Skills**

Thesis presented by
Giannina Puddu Gallardo

To obtain the Degree of Doctor by the University of Barcelona

In accordance with the requirements for the PhD Diploma
Doctoral Programme in Cervell, Cognició I Conducta

Supervisor
Dr. Marc Via García

Barcelona, February 2024



Institut de Neurociències
UNIVERSITAT DE BARCELONA



**EXCELENCIA
MARÍA
DE MAEZTU**

Brainlab

—
cognitive neuroscience
research group

—
grup de recerca
en neurociència cognitiva

Abstract

Individuals exhibit variability in auditory processing skills, including those related to musical or speech perception, which rely on an accurate neural representation of the spectro-temporal characteristics of the auditory signal. The Frequency-Following Response (FFR) is a neurological response originating mainly in the brainstem that synchronizes with incoming stimuli and serves as a tool to assess the tracking accuracy of the auditory system. FFR has been associated with problems in phonological awareness, reading alterations, and dyslexia. It has also been noted that this response is sensitive to environmental influences such as language experience, bilingualism, and musical training. In addition to studying environmental factors in auditory processing, genetic studies have helped to understand the contributions of genetic variants to individual differences observed in auditory skills. Twin studies have reported remarkable heritability for auditory processing, including pitch recognition, temporal processing, and dichotic word discrimination, among others. Molecular genetics approaches have investigated the association between genetic regions and behavioral phenotypes of auditory perception, such as music perception and beat synchronization. In the field of auditory responses measured by EEG, a study reported a significant association between the FFR and the serotonin uptake gene (*SLC6A4*), representing the only study relating an auditory electrophysiological response with a gene. While the study of auditory processing and perception has been focused on environmental factors or genetic contributions separately, studies that combine both approaches are scarce. Furthermore, studies that integrate measures at different levels, including genetic data, electrophysiological measures, and behavioral assessments of auditory processing in a

single study, are non-existent. This thesis includes four studies that aim to explore the genetic and environmental factors contributing to auditory processing as reflected in the FFR. It also aims to investigate the relationship between FFR and functional assessments of audiological, musical and speech perception. A total of 129 young adults volunteered as participants in an experimental session that consisted of an FFR-EEG experiment, a saliva sample collection for genetic analyses, four behavioral tests to assess speech-in-noise (SIN) perception, reading skills, musical skills, and amplitude-modulation discrimination, in addition to a language proficiency questionnaire. Study I attempted at characterizing the relationship between the FFR and processing sounds in challenging environments. Its results revealed that background noise affects the FFR by enhancing the response to the F_0 on one side and degrading the response to the formant and the neural transmission delay on the other side. Additionally, we correlated the FFR parameters with the performance in a SIN perception task, finding that only the response to the F_1 to the vowel /a/ was correlated with SIN perception. In Study II, the auditory skills measured behaviorally were correlated with four components obtained by PCA from the FFR parameters to describe functional implications of the FFR. Only the F_1 /o/ component was significantly correlated with the performance in reading skills and no other significant correlation was observed. When evaluating the relationship only between the behavioral measures, significant correlations were observed between SIN perception and reading skills. In Study III, we replicated the association between the promoter region of the SLC6A4 gene and the FFR response to the F_0 . Moreover, genome-wide analyses did not reveal significant associations after correcting for multiple comparisons, but suggestive probability values were found between FFR component and SNPs related to genes linked to different cognitive and behavioral assessments and

highly expressed in brain tissues. Based on the polygenic score analyses, we found a significant shared genetic etiology between SIN perception and ADHD. In Study IV, regarding the role of the age of bilingualism acquisition as a potential modulator of the neural response, we did not find a main effect of bilingualism across all FFR measures, nor with the behavioral assessments, indicating that by itself, this variable does not contribute to the observed differences. In summary, throughout this thesis, it was observed that the individual contribution of the different environmental and genetic effects studied has a relatively small effect size, thus highlighting the complexity underlying the in-depth study of auditory processing of sounds in our environment.

Resum

Els individus mostren variabilitat en les habilitats de processament auditiu, incloent-hi aquelles relacionades amb la percepció musical o del discurs, que depenen d'una representació neural precisa de les característiques espectro-temporals del senyal auditiu. La Resposta de Seguiment a Freqüència (FFR) és una resposta neurològica que es genera principalment al tronc cerebral, es sincronitza amb els estímuls entrants i serveix com a eina per avaluar la precisió de codificació del sistema auditiu. La FFR s'ha associat amb problemes en la consciència fonològica, alteracions en la lectura i dislèxia. També s'ha observat que aquesta resposta és sensible a influències ambientals com el bilingüisme i la formació musical. A més d'estudiar els factors ambientals en el processament auditiu, els estudis genètics han ajudat a comprendre les contribucions de les variants genètiques a les diferències individuals observades en les habilitats auditives. Els estudis bessons han reportat una heretabilitat notable per al processament auditiu, incloent-hi el reconeixement del to, el processament temporal i la discriminació de paraules, entre d'altres. Estudis de genètica molecular han investigat l'associació entre variants genètiques i fenotips conductuals de la percepció auditiva, com la percepció musical i la sincronització del ritme. En el camp de les respostes auditives mesurades per EEG, un estudi va informar d'una associació significativa entre la FFR i el gen de la recaptació de serotonina (*SLC6A4*), representant l'únic estudi que relaciona una resposta electrofisiològica auditiva amb un gen. Tot i que l'estudi del processament auditiu i de la percepció s'ha centrat en els factors ambientals o les contribucions genètiques per separat, els estudis que combinen ambdós enfocaments són escassos. A més, els estudis que integren mesures a diferents nivells, incloent-hi dades genètiques,

electrofisiològiques i conductuals del processament auditiu en un sol estudi, són inexistents. Aquesta tesi inclou quatre estudis que tenen com a objectiu explorar els factors genètics i ambientals que contribueixen al processament auditiu reflectit en la FFR. També pretén investigar la relació entre la FFR i les avaluacions funcionals de la percepció audiològica, musical i del discurs. Un total de 129 joves adults es van participar voluntàriament en una sessió experimental que consistia en un experiment d'EEG-FFR, una recollida de mostra de saliva per a anàlisis genètiques, quatre proves conductuals per avaluar la percepció de SIN, les habilitats de lectura, les habilitats musicals i la discriminació d'amplitud modulada, a més d'un qüestionari de competència lingüística. L'Estudi I va intentar caracteritzar la relació entre la FFR i el processament de sons en ambients desafiants. Els resultats van revelar que el soroll de fons afecta la FFR al millorar la resposta a la F_0 d'una banda i degradar la resposta al formant i el retard de transmissió neural de l'altra banda. A més, vam correlacionar els paràmetres de la FFR amb el rendiment en una tasca de percepció de SIN, descobrint que només la resposta a la F_1 de la vocal /a/ estava correlacionada amb la percepció de SIN. En l'Estudi II, les habilitats auditives mesurades conductualment es van correlacionar amb quatre components obtinguts per PCA dels paràmetres de la FFR per descriure les implicacions funcionals de la FFR. Només el component F_1 /o/ va correlacionar significativament amb el rendiment en habilitats de lectura i no van observar altres correlacions significatives. Quan es va avaluar la relació només entre les mesures conductuals, van observar correlacions significatives entre la percepció de SIN i les habilitats de lectura. En l'Estudi III, vam replicar l'associació entre la regió promotora del gen SLC6A4 i la resposta de la FFR a la F_0 . A més, les anàlisis genòmiques no van revelar associacions significatives després de corregir per comparacions múltiples, però es van trobar valors de probabilitat

suggeridors entre les components de la FFR i SNPs relacionats amb gens relacionats amb diferents fenotips cognitius i conductuals i altament expressats en teixit cerebral. Les anàlisis de puntuacions poligèniques van detectar una etiologia genètica compartida entre la percepció de SIN i el TDAH. En l'Estudi IV, pel que fa al paper de l'edat d'adquisició del bilingüisme com a modulador potencial de la resposta neural, no vam trobar un efecte principal del bilingüisme en cap mesura de la FFR ni amb les avaluacions conductuals, indicant que per si sola, aquesta variable no contribueix a les diferències observades. En resum, al llarg d'aquesta tesi s'ha observat que la contribució individual dels diferents efectes ambientals i genètics estudiats té una mida d'efecte relativament petita, destacant així la complexitat que subjau en l'estudi profund del processament auditiu dels sons del nostre entorn.

Fundings

The present project was conducted at the Brainlab – Cognitive Neuroscience Research Group (Excellence Research Group 2017SGR-974 and 2021SGR-00356, funded by the Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement de la Generalitat de Catalunya) within the Psychobiology Section of the Department of Clinical Psychology and Psychobiology, Faculty of Psychology, University of Barcelona (UB; Barcelona, Catalonia, Spain). This work was supported by an ANID PhD Scholarship awarded by the Government of Chile, recipient: Giannina Puddu-Gallardo, Scholarship ID number: 72190276, with additional funding awarded to Marc Via García within the framework of the "Estrategia Española de Ciencia y Tecnología y de Innovación" and the goals of the Programa Estatal de Generación de Conocimiento y Fortalecimiento Científico y Tecnológico del Sistema de I+D+i (PGC2018-099013-A-100), from the "Ministerio de Ciencia, Innovación y Universidades" of Spain.

INDEX OF ABBREVIATIONS AND ACRONYMS

A1	Minor Allele
ABR	Auditory Brainstem Response
AEPs	Auditory Evoked Potential
AM	Amplitude Modulated
AP	Auditory Processing
DNA	Deoxyribonucleic Acid
EEG	Electroencephalography
EOG	Electrooculogram
F ₀	Fundamental frequency
F ₁	First Formant
FDR	False-discovery rate
FFR _(ENV)	Frequency Following Response Envelope
FFR _(TFS)	Frequency Following response Temporal Fine Structure
FUMA	Functional Mapping and Annotation
GEPLEM	The Role of Genes and Epigenetic Factors in Modulating Brain Potentials Associated with Speech and Musical Processing
GWAS	Genome Wide Association
H _n	Harmonic _(number)
h ²	Heritability
Hz	Hertz
IBD	Identity by Descent
IC	Inferior Colliculus
LD	linkage disequilibrium
LEAP-Q	Language Experience and Proficiency Questionnaire

LLR	Late Latency Response
MAF	Minor Allele Frequency
MAGMA	Multi-marker Analysis of GenoMic Annotation
MEG	Magnetoencephalography
MLR	Middle Latency Response
PCA	Principal Component Analysis
PC	Principal Component
PCR	Polymerase Chain Reaction
PGS	Polygenic Scores
PRS	Polygenic Risk Scores
PROLEC-SE-R	Batería de Evaluación de los Procesos Lectores en Secundaria y Bachillerato, revisada
PROMS-S	Profile of Music Perception Skills-short version
QC	Quality Control
SIN	Speech In noise
SNP	Single Nucleotide Polymorphism
SOA	Stimulus Onset Asynchrony
SPIN	Speech in Noise Test
WF	Weber Fraction

INDEX OF FIGURES

Figure 1. Schematic representation of the human auditory neural pathway.

Figure 2. Graphical representation of a typical Auditory Evoked Potential, displaying positive and negative waves deflections across time.

Figure 3. Illustration of a typical sine wave in each frequency band.

Figure 4. Spectro-temporal characteristics of the /ba/ stimulus.

Figure 5. Examples of various types of stimuli and their corresponding FFR are depicted.

Figure 6. Summary of heritability estimates for phenotypes related to Auditory Processing (AP) of non-speech and speech stimuli and EEG endophenotypes.

Figure 7. Differences in signal extraction in the subcortical auditory pathway depending on the genotype in the promoter region of the *SLC6A4* gene.

Figure 8. Polygenic score (PGS) analysis.

Figure 9. Representation of the Spectro-temporal characteristics of the /oa/ speech stimulus.

Figure 10. PCA results depicting individuals from the GEPLM project (BCN) merged with individuals from the 1000 Genomes Project.

Figure 11. *SLC6A4* gene and its genomic context with the relative location of the polymorphisms analyzed.

Figure 12. Spectro-temporal characteristics of the stimulus and the corresponding elicited response.

Figure 13. Spectro-temporal characteristics of the $\text{FFR}_{(\text{TFS})}$, obtained from the subtracted polarities, are depicted.

Figure 14. Marginal means in all pitch tracking measures estimated for both background conditions and pitch contours.

Figure 15. Principal component analysis, parallel analysis plot. The

Figure 16. Scatterplot between FFR components and Word and Pseudoword reading subtest.

Figure 17. Scatterplot between the performance in the SIN test and Word reading subtest.

Figure 18. Quantile-quantile (Q-Q) plots of the GWAS summary statistics for the four phenotypes analyzed.

Figure 19. GWAS results for the Pitch Magnitude component of the FFR.

Figure 20. GWAS results for the Pitch Tracking component of the FFR.

Figure 21. GWAS results for the F1 /o/ component of the FFR.

Figure 22. GWAS results for the F1 /a/ component of the FFR

Figure 23. Manhattan plots of the gene-based tests of association with the four FFR components.

Figure 24. MAGMA tissue expression analysis results.

Figure 25. Marginal Means of $\text{FFR}_{(ENV)}$ and neural lag estimated for native and non-native bilinguals across both background conditions.

Figure 26. Marginal Means of $\text{FFR}_{(TFS)}$ for native and non-native bilinguals across both background conditions.

Figure 27. Marginal Means of the pitch tracking measures for native and non-native bilinguals across both background conditions.

INDEX OF TABLES

Table 1. Summary statistics and the Wilcoxon paired test are provided for the neural lag, $FFR_{(ENV)}$, and $FFR_{(TFS)}$ between the background conditions.

Table 2. Summary statistics and Repeated Measured ANOVA between the background conditions and the pitch tracking measures.

Table 3. Simple effect between background condition and pitch contour on the cross-correlation results.

Table 4. Spearman correlation between SPIN test and the FFR measures

Table 5. Principal Component Analysis for the FFR parameters in quiet background condition, only on the flat pitch contour.

Table 6. Spearman's correlation between the FFR Components and the different auditory skills.

Table 7. Pearson's correlation between the behavioral measures

Table 8. Descriptive statistics of the polymorphisms analyzed in the *SLC6A4* gene.

Table 9. Association results between markers in the promoter region of the *SLC6A4* gene and FFR parameters.

Table 10. Spearman correlation between the PGS for the traits selected and the FFR components and the behavioral auditory skills.

Table 11: MR ANOVA repeated mean: noise background and bilingual acquisition effects in $\text{FFR}_{(\text{ENV})}$ and neural lag.

Table 12: MR ANOVA repeated mean: noise background and bilingual acquisition effects in $\text{FFR}_{(\text{TFS})}$.

Table 13: ANOVA repeated mean: noise background, pitch contour and bilingual acquisition effects in FFR pitch tracking measures.

Table 14: t test for native and non-native bilinguals across SPIN test, Word/Pseudo word reading, Musical perception, and AM discrimination.

Table of Contents

CHAPTER I: INTRODUCTION	1
1.1 THE AUDITORY SYSTEM	3
1.1.1 Auditory Pathway for Sound Processing	4
1.1.2 Auditory Evoked Potentials	6
1.1.3 Neural oscillations	8
1.2 AUDITORY PERCEPTION ABILITIES	11
1.2.1 Auditory Neural Coding Mechanisms	15
1.2.2 Speech Perception	18
1.2.3 Reading skills and dyslexia	21
1.2.4 Relationship between speech perception and reading skills	24
1.3 FREQUENCY FOLLOWING RESPONSE (FFR)	25
1.3.1 FFR and SIN perception	28
1.3.2 FFR and reading skills	32
1.4 ENVIROMENTAL MODULATORS OF THE AUDITORY PERCEPTION	36
1.4.1 Bilingualism	36
1.4.2 Musical training	39
1.4.3 Other non-modifiable modulators: Sex differences	43
1.5 GENETICS OF SOUND PROCESSING	45
1.5.1 Twin studies	45
1.5.2 Molecular genetics approaches	48
CHAPTER II: OBJECTIVES	57
2.1 GENERAL OBJECTIVE	59
2.2 SPECIFIC OBJECTIVES	60
2.2.1 Study I: FFR and noise	60
2.2.2 Study II: FFR and functional implications	61
2.2.3 Study III: Genetics of the FFR	61
2.2.4 Study IV: Bilingualism as a possible modulator of the neural response	62
CHAPTER III: METHODOLOGY	63
3.1 DESIGN OF THE EXPERIMENTAL SESSION	65
3.2 PARTICIPANTS	65
3.3 STIMULUS AND EXPERIMENTAL DESIGN	66
3.4 NEUROPHYSIOLOGICAL RECORDING	68
3.5 FFR PARAMETERS	69
3.6 BEHAVIORAL MEASURES	72
3.7 DNA COLLECTION AND GENOTYPING	75

3.8 STATISTICAL ANALYSIS.....	79
CHAPTER IV: RESULTS.....	89
4.1 STUDY I: FFR AND NOISE	91
4.2 STUDY II: FFR AND FUNCTIONAL IMPLICATIONS	100
4.3 STUDY III: GENETIC OF THE SOUND PROCESSING	108
4.4 STUDY IV: BILINGUALISM AS A POTENTIAL MODULATOR OF THE NEURAL RESPONSE ..	122
CHAPTER V: DISCUSSION.....	131
5.1 DISCUSSION STUDY I	133
5.2 DISCUSSION STUDY II	139
5.3 DISCUSSION STUDY III	144
5.4 DISCUSSION STUDY IV	155
5.5. GENERAL DISCUSSION	158
5.6 LIMITATIONS AND FUTURE RESEARCH	161
CHAPTER VI: CONCLUSIONS	167
CHAPTER VII: REFERENCES	171
ANNEX	207

CHAPTER I: INTRODUCTION

1.1 THE AUDITORY SYSTEM

Like all mammals, the human being possesses a highly specialized auditory system that plays a significant role in our daily lives. This system enables us to process sounds across a wide range of frequencies and intensities (Robles & Ruggero, 2001). For instance, we are capable of perceiving the faint rustling of a breeze or the resounding screech of a car's brakes during deceleration. Beyond mere auditory perception, these sounds impart vital information, which can prove critical for our survival, from an evolutionary perspective. They facilitate the detection of predatory approaches, or the warning calls emitted by members of a social group, thereby assisting in the avoidance of dangerous situations, and ensuring our safety. Moreover, these auditory cues offer valuable guidance in our daily decision-making processes; for example, the sound of the rain indicates that we should go out dressed/prepared for this weather (Goldstein, 2010).

Moreover, the auditory system is essential in facilitating more intricate processes, such as communication, social interactions, and overall complex cognitive operations (Skipper, 2022; Suresh et al., 2019). Through speech and language, we construct mental representations of the surrounding world, enabling the exchange of ideas, emotions, and knowledge, thus fostering meaningful connections and cooperation within our communities.

On the other hand, the ability to process sounds is not homogeneous, exhibiting a high variability at the population level. It is known that individuals differ from each other in their sound processing capabilities, which can be seen reflected in various contexts. For example, these variations can manifest in skills related to speech, such as speech perception and recognition, the ability to learn and pronounce a new language

accurately, or in musical aptitude, such as learning to play an instrument, detecting different tones, identifying musical notes in a melody and/or being able to sing in tune to a melody, among others. In all these processes an accurate neural representation of the sound signal is needed.

In the last decades, various environmental variables have been studied to explain these individual differences in sound encoding and processing, e.g., socioeconomic status, musical training, and bi/multilingualism have been investigated. However, the role of genetic factors in these variations remains unclear.

For this reason, it is essential to gain a comprehensive understanding of how the auditory system works, starting from the basic mechanism of how an auditory signal is encoded at the different levels of the auditory pathway. Subsequently, a thorough analysis of the factors influencing auditory systems becomes crucial to elucidate and account for the observed differences.

1.1.1 Auditory Pathway for Sound Processing

At the anatomical level, the acoustic features of an auditory stimulus are encoded and processed at different levels of the auditory hierarchy, from the ear through the subcortical structures, to the auditory cortex (Figure 1).

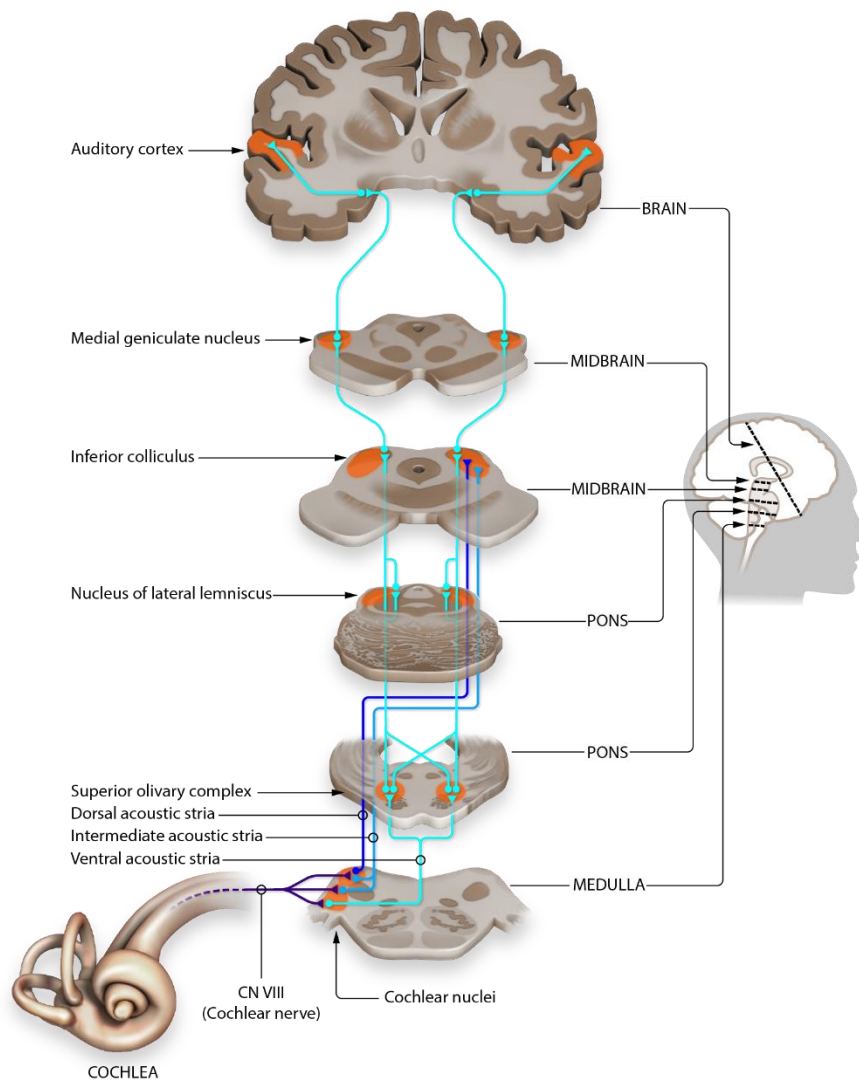


Figure 1. Schematic representation of the human auditory neural pathway. Taken from <https://osf.io/u2gxc/> under a Creative Commons Attribution 4.0 International License.

At the peripheral level, the traveling sound wave, reaches the basilar membrane, which acts like a set of bandpass filters separating complex tones into the fundamental frequency (F_0) and the frequencies of their different harmonics (and formants in the case of speech). Each frequency of a complex tone will stimulate a specific place on the basilar membrane due to its characteristic tonotopy. For example, low frequencies will result in maximum activity at the apex end of the basilar membrane, while high frequencies will stimulate the base of the membrane.

This vibration of the basilar membrane is transformed into electrical signals by the inner hair cell stereocilia in the organ of Corti, situated within the cochlea. These electrical signals encompass the characteristics of the auditory stimulus and journey through the auditory nerve fibers from the cochlea, traversing the subcortical structures in the brainstem, which include the superior olivary nuclei, the inferior colliculus, and the medial geniculate nucleus in the thalamus. Ultimately, these signals undergo processing in the auditory cortex, situated in the temporal lobe of the brain (Goldstein, 2010; Plack, 2005).

Each structure and area within the auditory pathway contributes to sound encoding and participates in specific processing linked to perception. For example, auditory localization, the ability to locate a sound stimulus (whether it is in front, behind, above, below, etc.) is associated with the superior olivary nuclei. Additionally, the inferior colliculus and the medial geniculate nuclei are implicated in responding to distinct attributes of complex stimuli and in perceiving the pitch of a sound. On the other hand, attention and other more intricate processes are tied to the auditory cortex (Goldstein, 2010; Plourde, 2006).

1.1.2 Auditory Evoked Potentials

The electrical activity of the brain might be easily recorded from the scalp through an electroencephalogram (EEG). This technique allows us to explore how the brain encodes and processes an acoustic stimulus in a non-invasive way. Among the different measures that can be extracted from EEG recordings when processing auditory stimuli, the most commonly used are auditory evoked potentials (AEPs). AEPs are electrical changes in the neural activity of the brain in response to sounds, composed by positive and negative

deflections or waves that are time-locked to the stimulus presentation. They provide insights into the various stages of auditory processing, from the initial detection of sound and encoding of the spectro-temporal characteristics of the auditory stimulus to higher-level cognitive functions (Plourde, 2006).

Based on the interval between the stimulus onset and the response occurrence, namely latency, AEPs can be classified as auditory brainstem responses (ABR), middle latency responses (MLR) and late latency responses (LLR) (Figure 2).

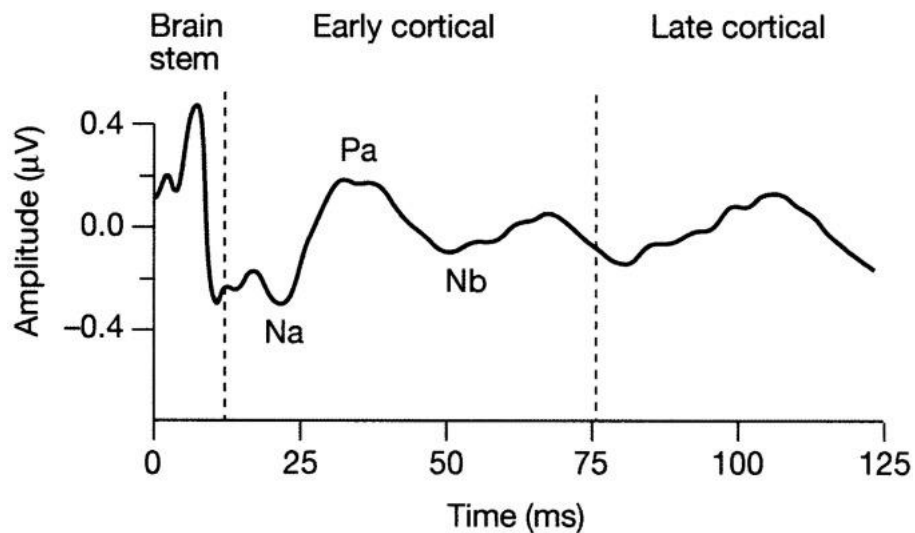


Figure 2. Graphical representation of a typical Auditory Evoked Potential, displaying positive and negative waves deflections across time. The graph is divided into three sections: Brainstem Auditory Evoked Potential (ABR), Early Cortical Auditory Evoked Potential (MLR), and Late Cortical Auditory Evoked Potential (LLR). Taken from Bell et al., 2004.

The earliest series of components of AEPs are the ABR, which emerges within 12 ms after the presentation of an auditory stimulus. This series of components are frequently used for hearing screening and diagnosis, as they reflect the activity of the auditory nerve and

the brainstem. The MLR components spans from approximately 10 to 80 ms after the stimulus presentation. They indicate the engagement of the auditory cortex and its role in processing intricate auditory stimuli, such as speech, providing insights into how the brain interprets and responds to various auditory features. The MLR has been commonly employed to test auditory processing deficits, monitor the effects of anesthesia due to its high sensitivity, localize thalamocortical lesions, and evaluate the effectiveness of cochlear implants through electrical stimulation. The LLR components emerge after the MLR, around 80 ms, and is frequently associated with higher-level cognitive and perceptual processes. This component is often regarded as a biomarker for the maturity of the central auditory system and could be useful in assessing populations with central auditory processing disorders (Ignatious et al., 2023; Macaskill et al., 2022; Paulraj et al., 2015; Plourde, 2006).

All of these components reflect the processing of an auditory stimulus, providing us with real-time information about the evolving activity in each structure that constitutes the auditory pathway. Consequently, they are valuable tools to understand the neural mechanism underlying auditory encoding and processing and, at a higher level, perceptual and cognitive processes (Ignatious et al., 2023).

1.1.3 Neural oscillations

An alternative approach to investigate auditory neural encoding and processing is by studying neural oscillations. The neurons have the capability to fire in a synchronized manner, generating rhythmic oscillations that represent neural activity in the central nervous system. These oscillations are electrical fluctuations occurring at various

frequency ranges, manifesting at various scales, from individual neurons to small clusters of neurons, and even at macroscopic levels within brain regions (Shettigar et al., 2022).

The literature frequently classifies oscillatory activity into two primary categories: evoked and induced oscillation. Evoked oscillations demonstrate phase-locking to external stimuli, such as event-related response. In contrast, induced oscillations are not phase-locked to stimuli. They are labeled as "induced" because they originate from the intricate and self-organizing processes within the brain, rather than being directly triggered by external stimuli. These induced oscillations serve as temporal filters that actively organize the processing of sensory information (David et al., 2006; Gourévitch et al., 2020; Morillon & Schroeder, 2015).

Neural oscillations play a fundamental role in brain functioning, they are linked to a range of essential brain processes including, but not limited to, feature binding, memory retrieval, information transmission between brain regions, and the generation of central patterns governing motor actions. Additionally, neural oscillations are intricately associated with higher-order cognitive states such as sleep, consciousness, and awareness (Gourévitch et al., 2020).

These oscillatory patterns extend their influence on more intricate cognitive processes, including sensory perception, attention, and memory. They are particularly vital within the framework of the perceptual sensory system, where they help organize and modulate the flow of information, allowing the brain to make sense of the sensory input from the external environment (Giraud & Poeppel, 2012; Kay et al., 2009; Morillon & Schroeder, 2015; Wang, 2010; Ward, 2003; Wilsch & Obleser, 2016).

Neural oscillations occur at different frequency ranges, typically classified into several bands, such as delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and gamma (30-100 Hz) (Figure 3). Each frequency band is associated with specific cognitive functions (For a more detailed review, see Gourévitch et al., 2020).

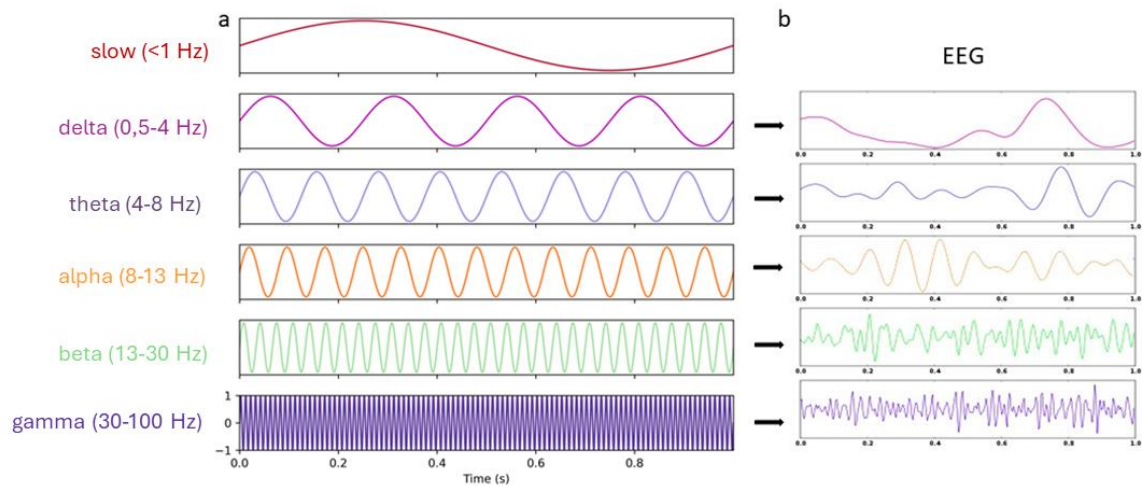


Figure 3. Illustration the oscillatory activity across different frequency bands. Typical sine wave in each frequency band (a). Brain waves visualization, based on different frequency bands, recorded by EEG (b) (Taken from Yuan et al., 2017).

1.2 AUDITORY PERCEPTION ABILITIES

Auditory perception is the process by which the human brain interprets and comprehends sounds detected from the environment. This intricate process involves the recognition, interpretation, and organization of auditory information, enabling individuals to perceive and comprehend various aspects of sound, including pitch, volume, timbre, and localization. Auditory perception plays a pivotal role in both music perception and language acquisition, encompassing both spoken and written forms (Goldstein, 2010).

To comprehend auditory perception, it is necessary to first delve into the spectro-temporal dimensions of auditory stimuli because they provide us with the cues necessary to perceive and interpret auditory information. At the physics level, any sound consists of rapid fluctuations in the air pressure that determine its spectral content. This rapid fluctuation in frequency and amplitude over time is known as the "fine structure," and the pattern or contour of these changes in amplitude, referred to as amplitude modulation, is usually referred as the "envelope" (Joris et al., 2004).

In the spectral domain, the fine structure is observed as the distribution of energy across frequencies within the signal over time. In the time domain, it is reflected as a waveform with amplitude variations across time (Figure 4).

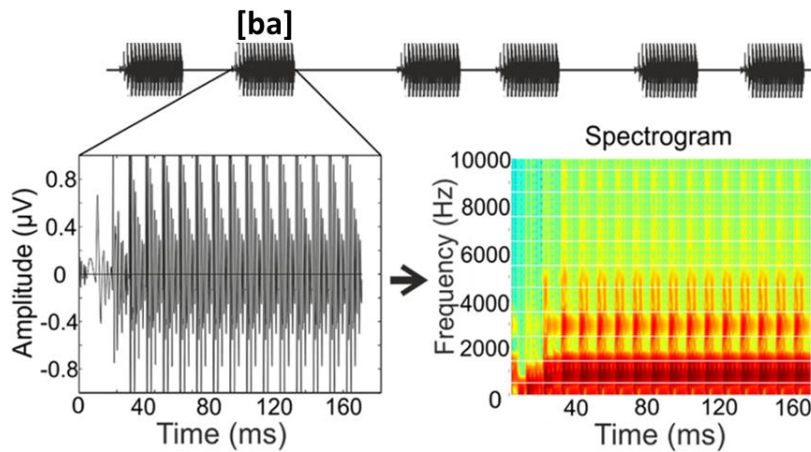


Figure 4. Spectro-temporal characteristics of the /ba/ stimulus. Temporal representation of the stimulus (left). Spectral energy distribution along to the stimulus (right). Taken from Selinger et al. (2016).

Acoustically, the fine structure contains information about the harmonic and, in the case of speech, formant patterns. Harmonics, which are multiples of the fundamental frequency within a complex waveform, play a significant role in music perception. Specifically, the harmonic structure of a musical tone is strongly related to timbre perception, also known as tone colour or tone quality. Timbre is the characteristic that distinguishes between two sounds with identical volume, pitch, and duration, yet they are perceived as distinctly different. This distinction is evident, for example, in how a piano sounds dissimilar to a guitar when both instruments play the same musical note or in the discernible contrast between a human voice and a violin (Goldstein, 2010; Vuust et al., 2022).

In the field of speech perception, sound waves generated by the vocal cords are shaped by the vocal tract, creating resonances or high points in the frequency spectrum identified as formants. To produce the sound of any speech unit or phoneme, whether a

consonant or vowel, the tongue, lips, jaw, teeth, and soft palate alter the shape of the vocal tract, resulting in a specific formant pattern. These formant patterns carry essential information for distinguishing between various speech sounds and, consequently, contribute to the comprehension of speech content (Goldstein, 2010).

The concept of the envelope is more complex and may vary depending on the method used for its extraction. For example, some authors differentiate between the spectral envelope, which involves variations in the relative amplitudes of individual harmonics (spectral domain), and the amplitude envelope, which refers to changes in the loudness of a sound, particularly at the onset and offset of the sound (time domain) (Koelsch, 2011). However, its most common use is in the time domain, often referred to as the temporal envelope, amplitude envelope, amplitude contour, or simply as the envelope, among other terms (Koelsch, 2011; Rosen, 1992).

The envelope, in its typical meaning, plays a primary role in conveying rhythmic variations in the signal, making its timescale and duration essential components of the information it carries, both in speech and music signals (Elhilali, 2019).

In the context of speech signals, it provides prosodic cues. For example, the amplitude onset and offset offer valuable cues for delineating linguistic units such as vowels, syllables, or words. These cues provide significant information about duration, contributing to the understanding of speech rhythm and tempo and assisting in the syllabification process. Similarly, it offers information that helps identify word stress (for a detailed description see, Rosen, 1992). In the field of music perception, the envelope plays a primary role in the detection of periodic pulses or beats, providing the foundational structure of rhythm perception. The pattern of strong and weak beats, also

called accented beats—a perceptual notion—combined with the duration of notes and rests, creates the overall rhythm of a musical piece (Rajendran et al., 2018; Vuust et al., 2022).

The envelope could also contain information about the periodicity of a sound. For example, for any periodic or quasi-periodic sound, be it a phoneme or a musical note, the Fundamental Frequency (F_0) is defined by the repetition rates of the sound waveform. These repetition rates manifest as cycles in the temporal envelope, which captures the changes in amplitude that occur during each cycle. In the case of a speech sound, the fundamental frequency is defined by the rate at which the vocal cords oscillate. In music, the fundamental frequency represents the lowest frequency for a specific note (Aiken & Picton, 2008; Krizman & Kraus, 2019; Moore, 2019).

Pitch, a perceptual quality closely linked to the fundamental frequency, is a fundamental element in music, language, and sound perception, facilitating our ability to differentiate and recognize diverse sounds. Pitch stands as one of the most extensively researched perceptual attributes in auditory perception. In music, melody is defined by sequences of pitch, whereas harmony and tonality are shaped by the simultaneous combinations of various pitches. In speech, variations in pitch contour are closely associated with prosody and speaker identity and recognition, e.g., determining the speaker's gender, age (child or adult), whether a sentence is a statement or a question, and, more importantly, assigning the meaning of words in tonal languages. Furthermore, distinguishing pitch differences among sounds aids in the separation of competing sound sources (Chandrasekaran et al., 2009; Oxenham, 2018; Vuust et al., 2022).

All these spectral and temporal characteristics of an auditory stimulus, whether they are natural sounds, speech sounds, or music sounds, have been theorized to be encoded in the auditory pathway via three coding mechanisms: the place code, the rate code, and the temporal code.

1.2.1 Auditory Neural Coding Mechanisms

The place code

As we saw in the section “Auditory Pathway for Sound Processing”, the spectral patterns of a sound cause a maximum peak vibration at a specific location on the basilar membrane. This tonotopic map is organized and anatomically projected through the central auditory structures, including the auditory cortex. This place representation for frequency is referred to as a place code, conveying information based on the location and anatomical connections of neurons within the representational map (Goldstein, 2010).

Despite the amount of evidence that supports this mechanism, it alone is not sufficient to explain the complexity of auditory perception. This is why two additional theories, the temporal and rate code, have been proposed. These would explain aspects that cannot be elucidated solely by the place code theory.

The rate code

Since synchronizing with the temporal features of the acoustic waveform doesn't encompass the entire spectrum of frequencies that listeners can detect and discriminate, a third coding mechanism is necessary. This mechanism, known as a rate

code, relies on the firing rate of neurons and neuronal assemblies. Specifically, it raises that the frequency information of an auditory signal, e.g. the formants of the speech, is encoding through the firing rate of auditory neurons (Sachs, 1984; Sachs et al., 1986; Su & Delgutte, 2020).

The temporal code

Auditory structures are highly sensitive to the temporal features of acoustic stimuli. In this sense, the temporal patterns in the sound wave will produce a vibrational pattern on the basilar membrane. This vibration will cause the cilia to bend in two directions. When the pressure increases, the cilia bend in one direction, leading to the firing of inner hair cells. Conversely, when the pressure decreases, the cilia bend in the opposite direction, and no firing occurs. Auditory nerve fiber discharges synchronize with the movement of stereocilia on inner hair cells, establishing a faithful representation of the temporal characteristics of the auditory stimulus in the collective firing pattern, commonly referred to as phase-locking properties (Goldstein, 2010).

These three coding mechanisms do not operate in isolation but rather work synergistically, reflecting different features of the acoustic waveform (Nourski & Brugge, 2011).

The study of synchronized neural activity, known as the temporal code, has been widely utilized to explore various processes in psychoacoustics and auditory cognition. In this context, the use of amplitude-modulated (AM) stimuli has been extensive for investigating the temporal processing of auditory signals. For instance, to determine the perceptual boundaries in an acoustic signal based on the rate of AM—e.g., for a given

noise or tones with sinusoidal amplitude modulation—lower rates of AM modulation (around 8-10 Hz) result in each cycle being perceived as a separate event, while higher rates of modulation (10-40 Hz) are perceived as flutter. Different rates within this flutter range are perceived as faster or slower. Sounds with modulation rates higher than 40 Hz are perceived as having a pitch quality (for a detailed review see, Nourski & Brugge, 2011).

AM perception tasks have also been used with more complex stimuli such as speech and music. For example, it has been observed that the perceived quality of a stimulus (very deteriorated/no deterioration), whether it is speech or music, with a modulation amplitude of 41.1 Hz, decreases with an increase in the modulation depth (25, 50, 75, and 100%). This, in turn, is differently affected by the type of stimulus; that is, the music signal is perceived more deteriorated than the speech signal despite having the same deep modulation (100%) (Lamminmäki et al., 2014).

Temporal processing has also been investigated in clinical populations, particularly in cochlear implant users, given the limited spectral resolution of implants compared to acoustically sensitive ears. The restricted spectral resolution of cochlear implants is notably reflected in the poor musical perception typically reported by their users. While implantees demonstrate nearly normal abilities in perceiving rhythm and tone duration, their timbre perception is usually poor (Shannon, 1993 cited in Drennan, 2008; Handel, 1995).

However, the results of cochlear implants in speech perception and recognition are more promising. A longitudinal study assessing several auditory skills at three time points (before the implant or baseline, and two post-implantation follow-ups) revealed

significant improvements in all outcome measures after implantation, including speech in noise, which is considered one of the most challenging conditions in speech perception (Rasmussen et al., 2022). The results can be partially explained by the importance of the temporal envelope in successful speech perception, and these temporal speech cues are effectively captured and/or preserved in cochlear implants.

Extending the findings in the field of cochlear implants, the study on temporal coding has been extensively explored in speech perception and literacy development. Processing and identifying phonemes, the fundamental elements in speech development and subsequently in reading and writing processes, largely depend on accurate neural representations of the temporal features of the stimulus. Deficits in language and literacy development have been associated with challenges in the temporal processing of speech features.

1.2.2 Speech Perception

The perception of speech stands as a critical and extensively researched auditory skill due to its substantial role in our daily lives. Studies on speech intelligibility have observed that disruptions in the amplitude of the speech signal, particularly in its temporal envelope, can compromise the comprehension of speech stimuli. Intriguingly, manipulating the spectral characteristics of speech stimuli, as in auditory chimeras, still allows for speech recognition. This suggests that, relying on temporal cues, individuals can recognize and comprehend speech despite alterations in the spectral aspects of the stimuli (Z. M. Smith et al., 2002) .

Studies on AM discrimination tasks, a perceptual test involving stimuli that vary in amplitude modulation frequency, have revealed that individuals with superior performance also exhibit better adaptation to vocoded speech (Erb et al., 2012).

In this type of task, participants may be required to determine whether two or three stimuli are the same or different, or to identify which one has a higher modulation rate (e.g., which is perceived as faster). The choice of frequency range in these tasks depends on the specific neural generator under evaluation. For instance, AM frequencies above 80-100 Hz primarily originate in subcortical areas, while slower frequencies below 20 Hz arise from cortical areas (Alaerts et al., 2009; Bidelman, 2015; Farahani et al., 2021; Sonck et al., 2023). This allows for the exploration of neural modulation at different levels of the auditory pathway.

Speech exhibits an impressive resistance to distortion, a quality likely attributed to its tendency not to occur in isolation. Instead, it is often influenced by environmental factors such as background noise and reverberations.

Speech in noise (SIN) perception presents particularly challenging circumstances due to the overlap of both the speech and noise signals. At the physical level, this overlap results in the spectro-temporal features of the noise masking the speech signal or target, degrading its quality. This spectral masking effect has a negative impact on both the spectral content and in the temporal envelope of the speech signal, fundamental cues for speech perception (Swaminathan & Heinz, 2012).

Otherwise, the reverberation is a kind of interference resulting from the inherent acoustic properties of an enclosed space. Formally defined, reverberation is marked by

the lingering presence of acoustic energy in the sound field even after its initial production. In the context of a reverberant space, segments of the speech signal reflect off surfaces, leading to a smearing effect in the spectrogram. This smearing distorts the energy within each phoneme, thereby complicating the identification and discrimination of consonants (for more detailed review, see Bidelman, 2017).

Both sources of interferences have different impact in the cues that are important in speech perception. For example, it has been observed that as the noise increases, the phase-locking to the stimulus diminished, while the phase-locking response at the fundamental frequency (voice pitch) remain invariant to increasing reverberation (for an exhaustive revision see Bidelman & Krishnan (2010), and Bidelman (2017) with references therein).

Beyond the degradative effects of noise and reverberations, there are people more or less resistant to these effects, leading to a high level of variability in SIN perception abilities at the population level. For example, it was observed that some life experiences could improve SIN perception, such as bi-multilingualism, music exposure and training (Bidelman & Krishnan, 2010; Coffey, Mogilever, et al., 2017; Rosenthal, 2020). Moreover, some population groups are more susceptible to show problems to process degraded speech, for example, older adults who may experience age-related hearing loss, individuals with specific language impairment whose ability to comprehend and produce speech is affected, or more specifically, individuals with dyslexia, whose language processing is impacted (for more see Boebinger et al., 2015).

1.2.3 Reading skills and dyslexia

In the realm of auditory processing, it has been postulated that the diminished reading abilities observed in individuals with dyslexia may result from abnormal neural encoding of specific speech sounds within the auditory pathway, particularly deficits in the temporal processing of speech features. This irregular encoding potentially leads to inadequate sound representations and subsequently hinders the efficient utilization of these representations (Banai, 2005; Goswami et al., 2002; Goswami, 2011; M. Y. C. Li et al., 2019; Tallal et al., 1993). Studies using AM discrimination tasks have associated reading difficulties with poor performance in AM-beat perception tasks (Goswami, 2019; Lorenzi et al., 2000). This supports the idea that the difficulties observed in dyslexia could be attributed to issues in temporal processing.

Two hypotheses, the “rapid temporal processing” and the “temporal sampling framework”, have been proposed to explain the deficits in temporal processing observed in developmental language disorders including reading disorder. The rapid temporal processing theory (Tallal, 1980) suggests that language and reading difficulties seen in developmental language disorders result from difficulties in rapidly processing ongoing auditory events. This slower rate of information processing, in turn, hinders the ability to accurately segment and identify phonemes. This theory was developed based on years of research, which found that children with language and reading impairments struggled with sequencing and distinguishing brief auditory stimuli or stimuli with short inter-stimulus intervals, but not with longer stimuli or intervals (Rey et al., 2002; Tallal, 2004; Tallal et al., 1993).

Subsequently, an alternative theory was proposed to explain the timing issues observed in phonological processes involved in language and reading skills. The temporal sampling framework theory shifts the focus towards a deficit in slow brain oscillations. It postulates that the problem arises from difficulties in processing the rate of changes in amplitude, referred as “rise time” or commonly recognized as beat perception. Goswami and colleagues (Goswami, 2011, 2015; Goswami et al., 2002) proposed that challenges in perceiving changes in rise time suggest potential difficulties in distinguishing between different modulation frequency ranges in speech. These difficulties might stem from the inefficient phase locking of neuronal oscillations to these frequency ranges. For instance, theta oscillations are important for syllabic perception, delta range frequency is important for prosodic perception and beta-gamma frequency band is important for phoneme processing (Abrams et al., 2009; Ghitza & Greenberg, 2009; Goswami, 2011; Luo & Poeppel, 2007; Poeppel, 2003; Poeppel et al., 2008; Tallal & Gaab, 2006).

Supporting this hypothesis, investigations in population with dyslexia have revealed disruptions in both cortical and subcortical synchronization. Cortical studies have revealed atypical neural oscillations in population with dyslexia, encompassing alpha, beta, delta, theta, and gamma frequencies (Cutini et al., 2016; De Vos et al., 2017; Goswami, 2011; Lizarazu et al., 2015, among others).

An interesting longitudinal study that evaluated the same group of children at 5, 7, and 9 years old (corresponding to pre-reading, 1-year, and 3-year reading instruction, respectively) revealed differences in beta band activity between dyslexics and controls only when participants had already started formal reading instruction (at 7 and 9 years). No differences were observed in their beta synchronization during the pre-reading

stages. These findings suggest that the atypical oscillatory synchronization observed might be a consequence of developing reading difficulties rather than the underlying cause. Regarding theta activity, no relationship was found between oscillatory synchronization and the subsequent development of phonological or reading skills (De Vos et al., 2017a).

Similar results were obtained in an adolescent population with a diagnosis of dyslexia. Differences with controls were only observed in beta and alpha oscillations, but not in theta synchronization (De Vos et al., 2017b). These studies dismiss the hypothesis proposed by Goswami (2011) that postulated theta synchronization as a possible underlying cause of deficient syllable representation in dyslexia.

Despite receiving support from functional neuroimaging studies and research on neuronal oscillations, recent investigations into neuronal oscillations do not align with the temporal sampling theory. Consequently, a reevaluation of previously accepted data supporting the rapid processing theory is required (Lehongre et al., 2011, 2013).

Additionally, both theories lack sufficient predictive power to explain the development of specific language disorders, like dyslexia, and speech perception deficits. It is important to acknowledge that a substantial body of research has consistently shown an association between reading skills and the temporal processing of auditory stimuli. Nevertheless, it remains unclear whether these observed deficits in temporal processing are outcomes or features associated with language disorders rather than a direct cause (for more details, see Protopapas (2014)).

1.2.4 Relationship between speech perception and reading skills

In studies with behavioral measures, reading skills also have been associated with speech in noise perception. Studies in developmental dyslexia have demonstrated that children diagnosed with dyslexia are especially susceptible to the negative effects of background noise, particularly, children with dyslexia fail to successfully separate noise from the signal or target of interest, resulting in a poor SIN perception (Chandrasekaran et al., 2009; Hornickel et al., 2009; White-Schwoch et al., 2015).

Ziegler et al. (2009) found that children with dyslexia lack the concurrent integration of diverse speech signals necessary for reliable speech identification, specifically, to extract regularities, like the speaker's voice characteristics in noisy backgrounds from the incoming sound. The authors argue that SIN perception requires the synchronous integration of different speech cues over diverse time scales and the core of the deficit in dyslexia resides in their poor ability to integrate these speech cues, leaving therefore a poor phonological representation, which lead to the difficulties observed in reading performance.

These findings may suggest that both processes might share some underlying neural mechanisms.

An alternative approach to studying the neural encoding of speech and non-speech sounds is the Frequency-Following Response (FFR), which has been proposed as a potential biomarker for auditory deficits and literacy skills (Hornickel & Kraus, 2013; White-Schwoch et al., 2015).

1.3 FREQUENCY FOLLOWING RESPONSE (FFR)

The frequency following response is a type of AEP, that reflects the synchronous and sustained phase-locking to the spectral and temporal periodic characteristics of an auditory stimulus (Figure 5).

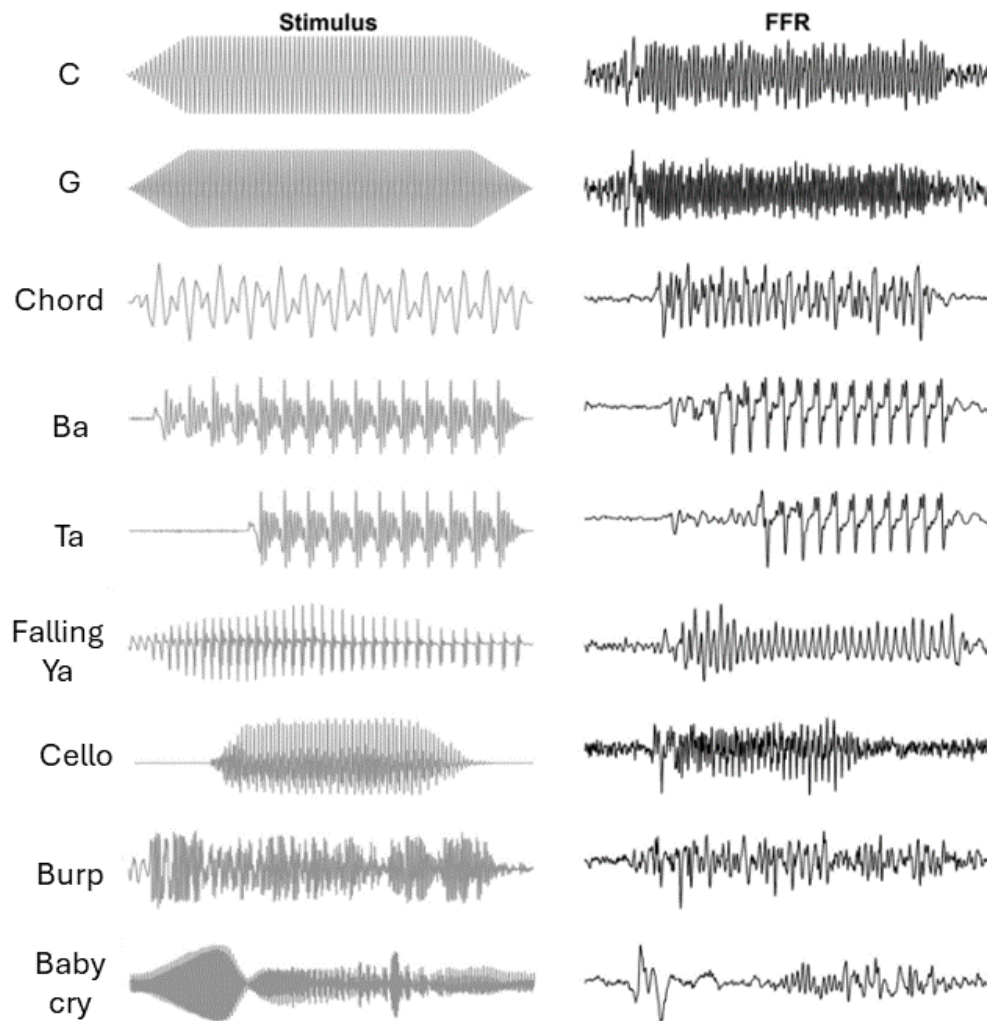


Figure 5. Examples of various types of stimuli and their corresponding FFR are depicted. These stimuli range in nature and complexity, including individual notes (C, G), chords, speech syllables, and baby cries (left). The FFR responses in the time domain for each type of stimulus are illustrated on the right. Taken from Krizman & Kraus (2019).

Regarding the neural origin of the FFR, it was initially thought to be solely of subcortical origin. However, based on neurophysiology studies in animals and non-invasive studies in humans, it has been hypothesized that the FFR arises from multiple generative sources. This response predominantly originates from the inferior colliculus (IC), but contributions from thalamic and cortical regions are also observed. This suggests a more intricate network of neural generators responsible for generating the FFR reflecting different aspects of an auditory stimulus (Abrams et al., 2017; Bidelman, 2018; Coffey et al., 2019; Gorina-Careta et al., 2021; Warrier et al., 2011).

It was hypothesized that the source of the FFR depended on the frequency features of the stimulus, thus, higher frequencies are associated to subcortical structures and lower frequencies are associated with more cortical components. Supporting this hypothesis, Zhang & Gong, (2019), found that the FFR to a sound with a pitch at or above 262 Hz is generated predominantly in subcortical areas, consistent with previous findings where a robust FFR to higher harmonics (~600-700 Hz) was localized mainly in the auditory nerve and the brainstem inferior colliculus. Additionally, weak phase-locking was observed in the bilateral primary auditory cortex with FFR- F_0 above 100 Hz, and this activation were eradicated for FFR > 150 Hz, Meanwhile, FFR to a lower stimulus frequencies (E.G 80 Hz) showed dominant cortical contributions (Bidelman, 2015, 2018; Coffey et al., 2016, 2019).

The FFR can be elicited from simple auditory stimuli like clicks or pure tones to more complex stimuli such as speech sounds. Furthermore, this response accurately preserves the fundamental frequency of speech as well as the related harmonic and formant properties. This characteristic of the FFR holds significance, because allows us to

compare the features of the stimulus with the corresponding response composition, both in the temporal and the spectral domain. Consequently, the FFR provides an objective measure of how precisely the auditory pathway encodes intricate acoustic features (Song et al., 2011).

Studies in newborns have demonstrated that we are born with the ability to track changes in the voice pitch. Specifically, it was observed that newborns, even in their first hours of life, showed similar FFR than adults in response to the F_0 of a speech stimulus, suggesting that pitch encoding could play an important role in our daily lives from the moment that we are born (Arenillas-Alcón et al., 2021; Ribas-Prats et al., 2019).

In the last decades, studies on FFR have gained considerable interest since it can reveal plasticity in the auditory system. The evidence suggests that the FFR can be influenced and modified by short-term and long-term auditory experiences. Indeed, a growing body of evidence shows that musical training, bilingualism, or being a native speaker of a tonal language are associated with a robust FFR (Kraus & Chandrasekaran, 2010; Krishnan et al., 2005; Krizman, Marian, et al., 2012a; Musacchia et al., 2007, 2008; Tallal & Gaab, 2006).

On the other hand, alterations in this response have been found to be closely related to autism spectrum disorder, attentional-deficit/hyperactivity disorder (ADHD), deficits in phonological awareness, dyslexia, and learning and/or reading impairment, among other disorders (Bonacina et al., 2019; Chandrasekaran et al., 2009; Hornickel & Kraus, 2013; King et al., 2002; Krizman & Kraus, 2019; Lam et al., 2017; Otto-Meyer et al., 2018; Talge et al., 2022).

Two primary conclusions can be drawn from these findings. Firstly, the FFR could serve as a potential biomarker for literacy issues or a predictor of future language development. Secondly, the FFR is influenced by environmental factors. This plasticity observed in auditory skills could be enhanced through auditory stimulation and training. Consequently, it can be regarded both as a predictor of long-term plasticity and as an indicator of training-induced plasticity (Chandrasekaran et al., 2009; Musacchia et al., 2007; Slater et al., 2015; Song et al., 2008).

The study of the FFR has yielded profound insights into how the human brain processes sound under normal conditions. Furthermore, evidence suggests that this response could underlie various auditory skills, potentially clarifying why alterations in this response have been associated with different auditory disorders. These disorders encompass reading, spelling, and learning difficulties, as well as issues with speech perception, implying that the FFR might serve as a potential biomarker (Hornickel & Kraus, 2013; Lam et al., 2017; Schochat et al., 2017).

1.3.1 FFR and SIN perception

The ability to segregate the signal or target from the other inputs that compete with the sound or stimuli of interest (noise) is important for successful participation in educational, social, and vocational environments (Anderson et al., 2012). Many studies in the literature have related different parameters of the FFR with performance in SIN perception tasks in different populations, such as in older adult populations (Anderson et al., 2011), in children (Anderson et al., 2010; Anderson & Kraus, 2010), children with specific language impairment (Basu et al., 2010), and learning problems (Cunningham

et al., 2001), and musicians (Parbery-Clark, Skoe, Lam, et al., 2009; Parbery-Clark et al., 2011, 2012), among others.

The voice-tagging hypothesis could explain the relationship between both. According to this hypothesis, a listener inhibits the background noise by focusing on the acoustic features of the speaker's voice. Individuals with a robust neural representation of the speaker's F_0 and their corresponding harmonics and formants should have better performance in tasks that involve SIN processing (Anderson & Kraus, 2010; Rosenthal, 2020).

In a longitudinal study with a sample of fifty-nine preschoolers, significant differences in the spectral amplitude of the F_0 were observed based on performance in a word-in-noise perception task. Specifically, for children who improved their performance on the SIN test over time, an improvement was also observed in their F_0 response. Conversely, for children who experienced a decrease in SIN performance, a lower F_0 was also observed over time. No significant correlations were observed between the performance in SIN and the response to the harmonics (Thompson et al., 2017).

Similar results have been observed in the school population, with some variations. For example, Chandrasekaran et al., (2009) evaluated the correlation between SIN and the spectral amplitude of the F_0 and the second and fourth harmonics (H_2 and H_4) in response to a /da/ stimulus in a sample of twenty-one school children aged between 8-14 years. The authors found significant correlations between the FFR at H_2 and SIN perception only in the consonant-vowel transition portion. No significant correlations were observed in the other comparisons.

Comparable findings were observed in a sample of thirty-eight typically developing children (ages 8-14 years). The group with higher performance in a SIN test (HINT, Biologic Systems Corp., Mundelein, IL) showed a higher spectral amplitude for the F_0 response compared to the group with lower performance. Similarly, a significant correlation was observed between the HINT-FRONT (when the stimulus was displayed in the front position) and the response to the F_0 and H_2 . These results were found only in response to the transition portion of the stimulus, not in the steady-state portion (Anderson et al., 2010).

In children with learning problems (N=18, 9 control, and 9 LP), it has also been observed that reduced speech discrimination are associated with diminished FFR response in the presence of background noise when compared to typically developing children (Cunningham et al., 2001).

In the adult population, similar correlations between SIN perception and different parameters of the FFR have been observed. Song et al., (2011), in a sample of seventeen healthy adults, found a significant and positive correlation between SIN performance and the spectral amplitude of the F_0 when the stimulus was presented in background noise (six-talker babble), but not when it was presented in quiet. These results were specifically identified in the response to the transition portion of the stimulus, but not in steady-state period.

In contrast to these results, in a study of older adults (n=38; ages 60 to 73 years), F_0 amplitude differences were observed only in quiet background conditions, but not in noise. The researcher compared spectral magnitudes of F_0 and harmonics in top and bottom SIN groups, categorized by their SIN test results. The group with top SIN scores

exhibited greater F_0 spectral amplitude than the bottom SIN group only in the quiet condition. No differences in the response to the harmonics were observed between groups, neither among the background conditions (Anderson et al., 2011).

The connection between the encoding of the fundamental frequency and performance on SIN perceptual tasks remains uncertain and, at times, inconsistent across several studies, especially in musician populations, where this relationship is less clear and sometimes inconsistent.

For example, Parbery-Clark, Skoe, & Kraus, (2009), in a sample of sixteen musicians and fifteen non-musicians, found significant correlations between neural transmission delay and stimulus-to-response correlation with SIN perception when it was assessed using the HINT test, but not using QuickSIN, and these results were found only in noise background conditions, but not in quiet backgrounds. Interestingly, SIN performance showed no correlation with the representation of the F_0 or its harmonics.

In a similar study, significant correlations were observed between SIN perception and the response to the F_0 in musicians, but not in non-musicians. No relationship was observed in the spectral magnitude of the harmonic responses (i.e., H_2 - H_5) with SIN perception across all subjects or within each group separately (Parbery-Clark et al., 2011). Furthermore, in a subsequent study, Parbery-Clark et al., (2012) found significant correlation solely between stimulus-to-response correlation and SIN perception. However, neither the spectral amplitude of the F_0 nor the harmonics directly influenced SIN performance.

In contrast, the relationship between the response to the F_0 and SIN perception is more robust when the FFR is measured by magnetoencephalography (MEG). Coffey, Chepesiuk, et al., (2017), evaluated the FFR in response to the /da/ stimulus, measured by EEG and MEG, and then correlated the FFR parameters with performance in a SIN test (HINT). The sample consisted of twenty healthy adults with varying degrees of musical experience. Significant correlations between SIN perception and FFR- F_0 were observed only when the F_0 was obtained by MEG, not when obtained by EEG.

In summary, the variability observed in the results could be due to the diversity of the studies design, the populations analyzed (e.g musicians vs. non-musicians, children with developmental disorders or learning problems vs. controls, older adults, etc.), the tests used to assess SIN perception (HINT vs. QuickSIN), the sample size, or the strategy used to record the FFR (MEG vs. EEG). All of these factors have probably contributed to the inconsistent findings observed across studies and still leaves some questions open about the relationship between the FFR and the performance in SIN perception.

1.3.2 FFR and reading skills

Usually, reading skills, especially dyslexia, have been examined through the neural oscillation paradigm. Being a complex neurodevelopmental disorder, are involved in the development of dyslexia, such as phonological awareness, lexical retrieval, and verbal short-term memory. As a result, it activates distinct neural networks, primarily associated with cortical regions (Finn et al., 2014; M. Y. C. Li et al., 2019). However, recent studies have also identified subcortical structures linked to reading impairments and dyslexia.

A structural connectivity study found that individuals with dyslexia showed reduced responses in the left auditory thalamus, specifically in the medial geniculate body, when processing speech, in comparison to neurotypical readers (Tschentscher et al., 2019). This has led to an alternative hypothesis proposing that subcortical abnormalities in auditory processing could be responsible for the phonological and reading difficulties observed in dyslexia.

FFR studies have also identified different dimensions of subcortical responses associated with reading skills as well as reading impairment.

In a study involving typically developed children with varying levels of reading skills, the auditory brainstem response to a speech syllable stimulus was examined. This study revealed significant correlations between reading performance and two specific parameters of the response. First, there was a correlation with timing in the response, which represents how accurately the brain's activity synchronizes with the stimulus. Better reading performance was associated with more accurate timing. Second, there was a correlation with the encoding of harmonics, reflected in the higher spectral amplitude in response to the stimulus' harmonics. Improved reading skills were linked to better harmonics encoding. However, no relationship was observed between pitch encoding and reading skills (Banai et al., 2009).

Similar findings were reported in another study by Hornickel et al. (2009), where delays in the timing response and reduced spectral magnitude in response to the harmonic components were associated with poorer reading skills.

White-Schwoch & Kraus, (2013) assessed the auditory brainstem response evoked by consonant-vowel stimuli /ba/ and /ga/ in a group of 26 pre-readers children (aged four years) and examined its connection to phonological awareness. The children with higher neural discrimination of the stop consonant /ba/ and /ga/ had better performance in a phonological awareness test. The authors hypothesized that it could serve as a biomarker for subsequent reading development, given the link between phonological awareness and dyslexia. However, due to the study's design, this cannot be confirmed. A follow-up would be required to assess which participants developed dyslexia or reading difficulties and which did not.

In a most recent and interesting follow-up longitudinal study, De Vos and colleagues (2020) recorded the auditory brainstem response, which was elicited by a speech-weighted noise carrier wave with amplitude modulation at 80 Hz. They used this measure to assess how abnormalities in neural synchronization at the subcortical level affect reading development. When comparing the subcortical responses of children at 5 years of age, during the pre-reading stage, they observed a significantly lower response in the children who later developed dyslexia compared to those who did not develop dyslexia (De Vos et al., 2020). This suggests that deficits in subcortical activity synchronization seem to predispose some children to develop dyslexia, even before they started learning to read, making it a risk factor. However, this subcortical effect disappears at later stages of reading acquisition (at 7 and 9 years old), where no differences in subcortical responses were observed between the groups.

These results contradict findings from FFR studies where abnormalities in the subcortical response to speech stimuli were linked to dyslexia. These disparities could be attributed

to methodological distinctions in their study designs, with the most significant distinction being the type of stimulus employed (speech vs. non-speech stimuli) (Banai et al., 2007, 2009; Hornickel et al., 2009; Hornickel & Kraus, 2013).

Dyslexia commonly co-occurs with other conditions, posing challenges for research. For instance, it often correlates with Attention Deficit/Hyperactivity Disorder (ADHD) in 25% to 40% of cases (McGrath & Stoodley, 2019). Likewise, it presents comorbidity rate around 30% to 40% with Specific learning disorders (Ashraf & Najam, 2020; Chung et al., 2020; Willcutt et al., 2019) and 58% with Developmental language disorders or specific language impairments (Snowling et al., 2019; Spanoudis et al., 2019). This substantial comorbidity complicates determining whether observed abnormalities directly contribute to reading deficits or if the relationship is more indirect and mediated by other processes.

1.4 ENVIRONMENTAL MODULATORS OF THE AUDITORY PERCEPTION

The auditory system is a remarkably adaptive and malleable system, highly sensitive to experience (Krishnan et al., 2005; Jeng, 2013). Bilingualism and musicianship have garnered interest in the field of experience-dependent plasticity due to their associations with brain changes in both structure and function. These changes extend beyond auditory areas to include regions linked to other perceptual and high-level cognitive processes.

In the following section, we analyze bilingualism and musical training as environmental modulators of neural systems and their impact on auditory perception. Finally, we include sex differences as a possible non-modifiable modulator of the auditory processing and perception.

1.4.1 Bilingualism

The effects of bilingualism on auditory processing and cognitive functions have been widely documented. Anatomical studies in the auditory cortex have shown that individuals who acquired a second language during early life exhibit significant structural changes in the auditory cortex (P. Li et al., 2014; Ressel et al., 2012). Furthermore, functional studies have revealed increased auditory cortical activation in bilinguals when exposed to pure and speech tones (Vihla et al., 2002).

Likewise, multiple benefits and improvements have been observed at the level of subcortical responses. In a recent study by Jafari et al., (2023), the auditory brainstem response (ABR) to a speech stimulus "/da/" was assessed. Results indicated that multilingual individuals demonstrated superior auditory neural timing processing,

characterized by shorter response latencies, and enhanced neural synchronicity with the stimulus, when compared to monolingual individuals. Additionally, bilinguals exhibit a more robust FFR- F_0 response compared to monolinguals, reflected in higher spectral amplitude in response to the F_0 of the speech stimulus, increased signal-to-noise ratio (SNR), and robust subcortical phase-consistency to the F_0 (Krizman, Marian, et al., 2012b; Krizman et al., 2021; Skoe et al., 2017, 2022).

In the assessment of whether second language acquisition has an impact on the response to the F_0 , it was observed that larger phase-locked responses to the F_0 were displayed by early bilinguals, who had acquired their second language before the age of 9 years, compared to monolinguals. However, no statistically significant differences were observed among early bilinguals based on the age of second language acquisition. Regarding harmonic processing, no discernible differences have been observed between bilingual and monolingual groups (Skoe et al., 2017).

Another advantage observed in the bilingual or multilingual population is their reduced susceptibility to the degrading effects of background noise. Bilingual groups do not exhibit differences in their spectral amplitude response to the F_0 between quiet and noisy background conditions, in contrast to monolinguals who displayed a decreased amplitude response in noisy background conditions compared to quiet conditions (Krizman, Marian, et al., 2012c). However, when assessing neural timing through speech-ABR peak analysis, both multilingual and monolingual groups exhibited similar peak latencies in noisy background (Jafari et al., 2023). These findings suggest that within the bilingual population, noise appears to impact only specific facets of sound

encoding and processing, such as timing, while other aspects remain resilient, as seen in the magnitude of the response to the F_0 .

Despite the numerous advantages of learning a second and/or multiple languages in terms of neural responses, as discussed earlier, it remains unclear how these benefits manifest at a functional level. Studies on executive functions and bilingualism have demonstrated that bilingual populations exhibit enhanced performance in visual and auditory selective attention tasks compared to monolinguals. When assessing the relationship between performance in these tasks and the evaluated electrophysiological response (spectral amplitude of the F_0), a statistically significant correlation was found only within the bilingual group. This correlation was observed in both auditory and visual selective attention tasks but was not present in the monolingual group (Krizman, Marian, et al., 2012c).

Furthermore, it has been observed that bilinguals have enhanced inhibitory control (Bialystok, 2011; Foy & Mann, 2014). However, in a recent study by Krizman et al., (2021), a similar performance was observed in both inhibitory control tasks and selective attention behavioral tasks for both monolingual and bilingual groups. Particularly interesting are the results concerning cortical and subcortical activations within the same study.

Two experimental listening conditions, active and passive attention conditions, were designed to assess potential differences in the neural and cognitive processes required for performing a selective attention task between monolinguals and bilinguals. The researchers found that both groups employed distinct auditory processing strategies in

both conditions. For example, in the active listening condition, both groups exhibited increased cortical phase consistency and decreased subcortical phase-consistency. However, when comparing the two groups, it was observed that monolinguals displayed greater cortical phase-consistency and reduced subcortical phase-consistency, while bilinguals demonstrated the opposite pattern with greater subcortical phase-consistency and lower cortical phase-consistency. Furthermore, a significant relationship between performance in the behavioral test related to selective attention and inhibitory control was only observed in the bilingual group (Krizman et al., 2021).

1.4.2 Musical training

Musical expertise may provide valuable insights into the impact of experience on the adaptive capabilities of auditory functions. A significant body of research on musicians and musical training has been conducted to explore the auditory and cognitive benefits associated with these abilities, shedding light on the reported changes in the brain induced by musical training.

At the anatomical level, observations indicate that musicians who initiate training before the age of seven exhibit a larger corpus callosum compared to non-musicians. This is likely a result of the frequent and intense transfer of sensory-motor information between the motor areas of both hemispheres. Interestingly, this effect was only noted in males, suggesting a potential hormonal contribution (D. J. Lee et al., 2003; Schlaug, Jäncke, Huang, Staiger, et al., 1995). Moreover, musicians with perfect pitch have shown stronger leftward planum temporal asymmetry compared to both non-musicians and musicians without perfect pitch (Schlaug, Jäncke, Huang, & Steinmetz, 1995).

Magnetic Resonance Imaging (MRI) studies have revealed increases in grey matter volumes in various brain areas due to musical practice. Notably, the volume of grey matter in the left hippocampus, posterior cingulate gyrus, and superior temporal cortex, as well as the right middle and superior frontal cortices and supplementary motor area, increases with the number of years of musical practice (Groussard et al., 2014).

Furthermore, this increased density of grey matter extends to areas associated with higher-order cognitive processing, including the right fusiform and mid-orbital gyrus, and left inferior frontal gyrus, among others. Interestingly, a reduction in grey matter density is observed in both perirolandic and striatal areas associated with sensorimotor function as expertise increases, suggesting a potential reflection of highly automated motor skills (James et al., 2014). Additionally, white matter has been found to undergo modifications in musicians compared to non-musicians, revealing more developed white-matter connections in regions linked to fine motor control and sensory processing (for review see Olszewska et al., 2021).

At the functional level, numerous studies have been conducted using various neuroimaging methods to assess functional brain changes in musicians. For instance, fMRI studies have revealed that musicians exhibit enhanced activations in temporal regions associated with improved auditory functions, such as pitch information extraction, processing, and comparison, in comparison to non-musicians. Additionally, musicians show improved activation in brain regions linked to motor areas and language processing, such as the primary and supplementary motor areas and Broca's area. In EEG and MEG studies, musicians display higher amplitude in their event-related potentials (ERPs) in the frontal cortex, as well as a heightened mismatch negativity

response in auditory areas (for a review, see Olszewska et al., 2021), to mention a few examples.

The observed neuroplasticity, evident at both structural and functional levels, has been linked to performance in various cognitive skills and sound processing (for a detailed review see Benz et al., 2016). In the population of school-age children, positive correlations have been reported between musical abilities, reading skills, and phonemic awareness. Instead, a series of longitudinal studies compare the effect of musical vs. painting training programs in several cognitive domains in children, the data showed: i) enhanced speech segmentation observed after two years of musical training in non-musician children aged 8 to 10 years (François et al., 2013); ii) improved performance in reading tasks, specifically in the inconsistent words subset of the European Portuguese Reading Battery, and larger amplitude of specific ERP components elicited in both musical and speech tasks after six months of musical training in children around 8 years old (Moreno et al., 2009); and iii) improvement in intelligence scores was observed after a six-month musical training program, specifically, higher scores in the verbal test of the Wechsler Preschool and Primary Scale of Intelligence Third Edition (WPPSI-III), designed for children between 2 to 7 years old, and more accurate performance in a go/no-go task. Regarding electrophysiological measures, notably greater peak amplitudes in the P2 component were observed during the no-go trials in a go/no-go task. These results were found exclusively in the musical training group, with no similar effects noted in the painting training group (Moreno et al., 2011).

These cognitive benefits of musical training are not limited to the child population. Musical practice has also been explored as a protective factor against cognitive aging.

Extensive meta-analyses have revealed cognitive advantages associated with musical practice, including improvements in early processing of auditory stimuli, as well as more generalized functions such as enhanced episodic memory, executive functions, improved visuospatial ability, and better working memory, among others (for a review see Román-Caballero et al., 2018). These promising findings have motivated the creation of musical training programs as a form of intervention to promote healthy aging (James et al., 2020).

The benefits of musical training have not only been reported at the cognitive level. Recent studies have expanded upon these results, demonstrating that musicianship also affects neural mechanisms at earlier stages in the auditory pathway.

The FFR has highly demonstrated experience-dependent plasticity (Chandrasekaran et al., 2009; Jeng et al., 2013; Krishnan et al., 2005) , and as expected, musical training has been associated with improvements in the ABR and FFR (Bidelman et al., 2011; Kraus et al., 2009). In cross-sectional studies comparing musicians and non-musicians, the musicianship group showed a larger neural pitch response, and this response positively correlated with years of musical training. Additionally, there was a more robust response to the harmonics of the sound stimulus and improvements in the timing response, indicating earlier or faster neural responses (Fujioka et al., 2006; Kraus et al., 2009; K. M. Lee et al., 2009; Musacchia et al., 2007, 2008).

In studies of speech in noise in musicians, it was observed that musicians are less susceptible to the degrading effect of background noise. Parbery-Clark, Skoe, & Kraus, (2009) compared the FFR elicited by the /da/ stimulus presented in both quiet and noisy backgrounds in a sample of sixteen musicians and fifteen non-musicians. The study

revealed that musicians showed a shorter neural delay and greater encoding of the harmonics compared with non-musicians, only in noisy conditions. No significant effects were found in response to the F_0 between groups or across background conditions.

1.4.3 Other non-modifiable modulators: Sex differences

It is well known in auditory neuroscience that the cochlear anatomy differs between males and females, with males having a longer and less steep stiffness gradient (Don et al., 1993; Sato et al., 1991). This difference is hypothesized to lead to functional distinctions in subcortical responses. Specifically, males have been observed to display smaller and later subcortical responses evoked by a click stimulus (Eggermont & Don, 1982).

The use of ABR to clicks or tones has been a very useful tool to assess the integrity of the auditory pathway. It serves as an excellent diagnostic tool for identifying hearing impairments and evaluating the function of the auditory system, among other applications. However, this type of stimuli only captures the onset and phase-locked response and may not fully represent the richness and complexity of the subcortical auditory system when exposed to more naturalistic sounds, such as speech.

A series of studies have been carried out in order to examine the extent to which the observed subcortical differences in the FFR response could be attributed to sex or maturational variables, specifically age. Krizman, Skoe, et al. (2012), examined the FFR elicited by the /da/ stimulus, using a peak-picking analysis approach in a sample of seventy-six participants (38 females) with ages between 22 and 29 years old (for a detailed description of the technique, see Krizman & Kraus, 2019). The results revealed

earlier neural responses (early peak latencies) in females. No sex differences were observed in other FFR parameters, specifically in the consonant-vowel transition, the offset, and the fundamental frequency response. These results were subsequently replicated in an extended sample of five hundred sixteen participants (250 females), that included three age groups, 3-5 years, 14-15 years, and 22-26 years old (Krizman et al., 2019).

Finally, with the purpose of understanding deeply the differences found in the previous studies, the researchers reanalyzed the data reported before (Krizman et al., 2019), complementing it with an analysis of the spectral content of the response for the F_0 and the first formant of the formant transition portion of the stimulus /da/, and the corresponding harmonics presented in this. Additionally, they included consistency measures of the responses by adding polarities to enhance the response for the temporal envelope and subtracting polarities to enhance the response for the fine structure (F_1 and harmonics) (for a detailed description of the technique see Aiken & Picton, 2008; Krizman & Kraus, 2019). Higher amplitude responses were observed in females compared to males, but only for high frequencies (750-1050 Hz). These differences were observed in both the 14-15 years group and the 22-26 years old group. Regarding the response magnitude for the F_0 and F_1 , no statistically significant differences were found between males and females in any of the age ranges, nor were any differences observed in response consistency or any other parameter evaluated (Krizman et al., 2020).

1.5 GENETICS OF SOUND PROCESSING

In the previous chapter, we have explored the contribution of environmental modulators to the individual differences observed in sound processing. The extensive literature available in this field contrast with how little is known about the contribution of genetic factors to these differences. In this chapter we are going to present the contributions from different methodologies to the field, while critically presenting the interest and limitations that each one of these approaches have.

1.5.1 Twin studies

Twin studies have been systematically used to explore the relative contributions of environmental and genetic factors on the individual differences observed in various traits. By quantifying the degree of concordance for a specific trait in a cohort of monozygotic and dizygotic twins, which share a common environment, and 100% and 50% of their genes, respectively, these studies allow the assessment of the heritability (h^2), quantifying the proportion of the variance in the phenotype of interest that can be attributed to additive genetic factors. As a result, twin studies have constituted a milestone in the field of human genetics (for a recent review, see Hagenbeek et al., 2023).

In the auditory processing field, it was observed remarkably high heritability along a wide range of behavioral auditory skills or phenotypes, as it can see in figure 6.

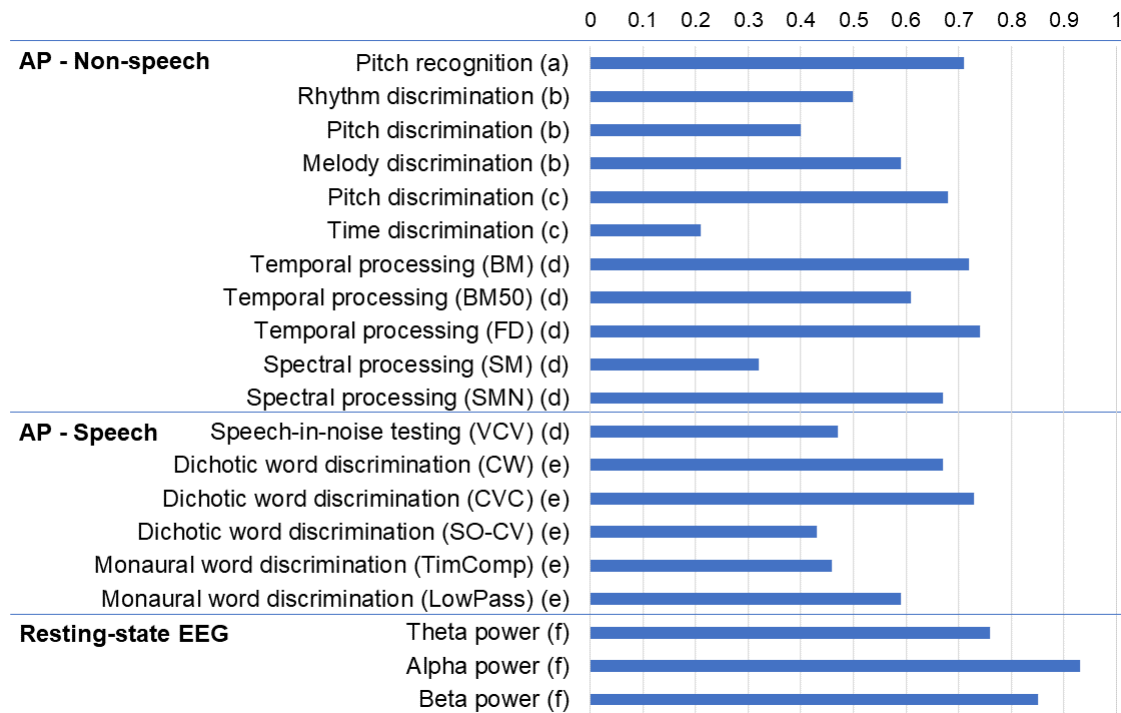


Figure 6. Summary of heritability estimates for phenotypes related to Auditory Processing (AP) of non-speech and speech stimuli and EEG endophenotypes. Values retrieved from: Drayna et al., 2001 (a); Ullén et al., 2014 (b); Oikkonen et al., 2015 (c); Brewer et al., 2016 (d); Morell et al., 2007 (e); Smit et al., 2012 (f). All results were based on twin studies, except (c), that was family-based.

In pitch recognition, Drayna et al., (2001) observed a heritability of 0.71 for this phenotype, measured as the ability to identify notes with incorrect pitch using simple popular melodies. Exploring the auditory discrimination of musical stimuli in the 6,881 twin individuals from the Swedish Twin Registry, Ullén et al., (2014) identified the heritability of pitch, melody, and rhythm discrimination skills ranging from 0.40 to 0.59. Similarly, Brewer et al., (2016) identified strong influences of genetic factors on different non-speech auditory skills, with h^2 values ranging from 0.61 to 0.74 for temporal discrimination and 0.32 to 0.67 for spectral discrimination, depending on the test used.

Relatedly, while not conducted as a traditional twin study, Oikkonen et al., (2015) estimated a h^2 of 0.68 for pitch discrimination skills in a linkage association study involving seventy-six Finnish families, comprising a total of 767 individuals.

In the same manner as the skills related to music perception, speech auditory skills have also demonstrated similarly high heritability values across different phenotypes in twin studies. For example, Morell et al., (2007) reported heritability values of 0.43-0.73 and 0.46-0.59 for dichotic and monaural word discrimination abilities, respectively. Comparable h^2 values (0.47) have also been determined for speech-in-noise skills (Brewer et al., 2016).

In addition to these studies, the heritability of absolute pitch, defined as the ability to identify a musical tone without reference to an external standard has also been explored in twin research (Deutsch, 2013). Theusch & Gitschier, (2011) found a significantly higher concordance rate of absolute pitch among monozygotic twins than among dizygotic twins, indicating substantial genetic influences; however, they did not provide a specific heritability estimate. Although beyond the scope of this thesis, other studies have investigated the heritability of various music-related traits and behaviors, including music achievement, creativity, or practice (For recent review see Wesseldijk et al., 2023).

While a large body of research in twin studies covers various behavioral phenotypes of auditory processing, unfortunately, no twin studies have directly analyzed the neural encoding of sounds or any related EEG phenotype.

Regarding the effects of genetic variants on cognitive assessments obtained through EEG, little is known but their effects are supposed to be stronger than other more

heterogeneous assessments. EEG-based traits constitute so-called intermediate phenotypes, i.e. phenotypes that are closer to the biological modulators of auditory processing than phenotypes based on behavioral or self-reported measurements, and constitute better assessments of innate differences (Anokhin, 2014). As an example, a twin study on resting-state EEG recordings showed remarkably elevated h^2 values, ranging from 0.76 to 0.93 for different power bands (Smit et al., 2012).

Twin studies have long been considered the gold-standard method to quantify the overall contribution of genetic factors to a specific trait. However, statistical methods implemented in twin studies make assumptions on gene-environment correlations and variance explained by shared environment factors that may lead to the overestimation of heritability values (Mayhew & Meyre, 2017). Moreover, heritability estimates need to be interpreted on a population-dependent context. Since they represent the proportion of phenotypic variance due to additive genetic factors, sociocultural differences across populations or developmental stages may vary the variance attributable to environmental influences and, consequently, heritability estimates.

1.5.2 Molecular genetics approaches

Candidate gene studies

As demonstrated in these previous studies, the potential contribution of genetic factors to auditory processing is potentially large. However, if the literature on twin studies is limited, our knowledge of the specific genetic locations on chromosomes or loci involved in the individual differences observed for auditory traits is even less extensive.

The development of the polymerase chain reaction (PCR) and other molecular methods in the decade of the 1980's and 90's allowed the proliferation of candidate gene studies, where an association between pre-specified genes of interest (based on previous knowledge) and a specific phenotype is tested. Although thousands of different traits have been explored, when it comes to the study of genetic variants associated to the individual differences in sound processing at the neural level, the only previous study available was conducted at our research group. Thus, differences in the neural ability to encode pitch were identified among carriers of different alleles in the promoter region of the serotonin uptake gene (*SLC6A4*) (Selinger et al., 2016). Recording the FFR, individuals with higher levels of serotonin available at the synaptic cleft (i.e. carriers of alleles of low serotonin transporter expression) showed a significantly more efficient neural representation of the stimulus features, assessed as higher SNR and pitch strength extraction (Figure 7). This study represents the only evidence so far of a genetic variant modulating brain processing of sounds at subcortical stages. These results were in concordance with findings that linked serotonergic activity in the inferior colliculus (IC) with the brain response to auditory stimuli (Obara et al., 2014).

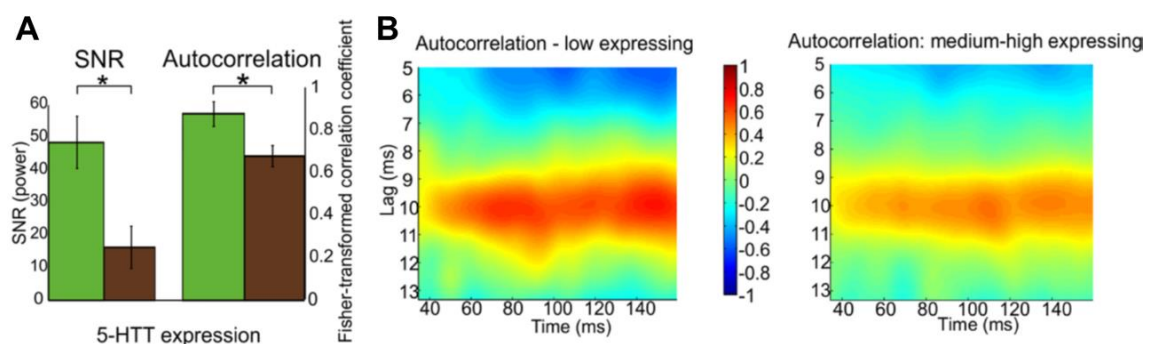


Figure 7. Differences in signal extraction in the subcortical auditory pathway depending on the genotype in the promoter region of the *SLC6A4* gene. Signal-to-noise ratio (SNR)

and pitch strength bar plots, showing significantly higher SNR and pitch strength in subjects with the low 5-HTT-expressing group (green) (A). Auto correlogram of pitch strength, revealing a significantly higher pitch strength in the low-serotonin transporter expression group at the F_0 of the stimulus (B). Results from Selinger et al., 2016.

In addition to this single report on the genetics of the neural encoding of sounds, previous studies have explored genetic variants associated with pitch perception and musical abilities using behavioral measures (See Navarro et al., 2021). Notably, an association was found between genetic variants in the arginine vasopressin receptor 1A (AVPR1A) gene, which mediates the effects of the arginine vasopressin hormone in the brain, and music perception skills in a sample of nineteen Finnish extended families (Ukkola et al., 2009).

However, a limitation of candidate gene studies is that they are hypothesis-driven and, consequently, limited to explore genes already suspected as relevant by our previous knowledge on the biology of the trait. Moreover, most candidate gene studies have limited statistical power and most associations reported have not been replicated in independent cohorts.

Genome-wide studies

With the progressive development of high-through put genotyping technologies, the joint analysis of thousands of markers became feasible and genome-wide scans started to be implemented. Genome-wide studies are hypothesis-free since there is no *a priori* hypothesis on which marker or chromosomal region will be associated with a phenotype of interest.

Among other approaches, family linkage studies provided the first opportunity to conduct genome-wide scans of genetic regions. By genotyping markers distributed across the genome in family pedigrees, linkage studies evaluate the segregation of genetic variants within families compared to the segregation of a trait.

Regarding auditory perception, a linkage study in seven hundred and sixty-seven Finnish families found that markers in the GATA binding protein 2 (*GATA2*) and the protocadherin 7 (*PDCH7*) genes, involved in the development of the cochlea, IC and other brain regions, were significantly associated with music perception skills, especially with pitch and time discrimination (Oikkonen et al., 2015). These results were in line with findings on pitch-production accuracy, that had previously linked chromosomal region 4q23 (where the *GATA2* gene is located) in 1,008 individuals from seventy-three Mongolian families (Park et al., 2012). Moreover, other linkage studies have associated chromosomal regions 6q14.1-6q16.1 and 8q24.21 with absolute pitch (Gregersen et al., 2013; Theusch et al., 2009).

Linkage studies constituted a popular approach to conduct genome-wide hypothesis-free approaches and offer the advantage of avoiding population stratification effects and biases between cases and controls. However, they have very limited statistical power to detect associations that do not have relatively large effect sizes and building large family cohort studies is very challenging.

In the past 25 years, alongside linkage studies, the utilization of single nucleotide polymorphism (SNP) genotyping arrays has facilitated the emergence of genome-wide association studies (GWAS) involving unrelated individuals. Variations at a single position in a DNA sequence, known as SNPs, naturally occur among individuals. These

variations have been linked to differences in characteristics such as traits, susceptibility to diseases, and other genetic attributes.

GWAS provides the opportunity to examine hundreds of thousands of SNPs scattered across the genome for associations with individual differences in specific traits. However, the primary limitation of most GWAS lies in the insufficient statistical power. Given the hundreds of thousands of comparisons conducted in GWAS, researchers must implement stringent multiple comparison procedures. Consequently, large sample sizes are typically required to detect small genetic effects, a common scenario for complex phenotypes. As a result, only a small number of genetic variants are often identified (if any), and there is a frequent lack of replication of results across different studies.

In this sense, Niarchou and colleagues recently published what can be considered the only large scale GWAS on an auditory skill, specifically on musical beat synchronization (Niarchou et al., 2022). In this work, the authors run a GWAS on the ability to clap in time with a musical beat in 606,825 individuals from the “23andMe” platform and identified sixty-nine different loci that reached genome-wide significance ($p < 5 \cdot 10^{-8}$, the usual threshold used in GWAS). Among the regions identified, the top-associated locus was *VRK2*, which encodes for a protein kinase expressed in the brain. However, none of the previous associations from linkage studies (Oikkonen et al., 2015; Park et al., 2012) could be replicated in this newer better-powered genome-wide scan.

post-GWAS analyses

Beyond identifying specific individual SNPs or loci responsible for a small part of the observed variation in the phenotypes of interest, genome-wide data also enable other

types of analyses, so-called genomic architecture of complex traits. These analyses can provide a more comprehensive understanding of how various genetic factors contribute to complex traits.

For instance, GWAS allow the estimation of heritabilities in a set of unrelated individuals from all SNPs in an array instead of estimating them from the concordance between pairs of twins. Thus, the estimated SNP-based heritability for beat synchronization was estimated to be 0.13-0.16 in the study from Niarchou et al., (2022).

SNP-based heritabilities are usually smaller than twin-based heritabilities, since they tend to overestimate the variance attributed to environmental factors (See Yang et al., 2017, for a review). As an illustration of this fact, SNP-based heritabilities for oscillatory power during resting-state EEG have been estimated as 0.11-0.27 in a sample of 8,425 unrelated individuals (Smit et al., 2018), substantially lower than the estimates obtained from twin-based approaches (0.76-0.93, Smit et al., 2012).

In addition to the estimation of the overall contribution of genetic variants to the phenotype of interest, GWAS data also allow the exploration of potential genetic overlap between different traits, i.e. to estimate the degree of shared genetic factors. Moreover, it also allows for analyses at the gene and pathway levels, looking the effect on the phenotype of the genetic variation pooled within genes and across genes in the same physiological process. Gene- and pathway-level analyses have a higher sensitivity to detect biological processes involved in complex traits since the number of comparisons is smaller (less genes or physiological pathways than SNPs across the genome).

Using these approaches, Niarchou et al., (2022) found significant genetic correlations between beat synchronization and traits such as processing speed, motor and executive functions, pointing to shared genetic factors with these phenotypes. Additionally, the authors also revealed an enrichment in genes expressed in brain tissues, especially in cerebellum, cortex, and basal ganglia. Using a conceptually similar approach, Oikonen et al., (2015) found an enrichment for genes involved in inner ear development and in schizophrenia among the associations with pitch and time discrimination skills.

Finally, a post-GWAS approach that has received increasing interest is the elaboration of polygenic scores (PGS). These scores, also called polygenic risk scores (PRS), provide estimations of an individual's genetic liability to a trait based on their genotypes and reference data from a GWAS study (Figure 8). Briefly, a GWAS first identifies alleles associated with a trait of interest in a cohort (Figure 8A). PGS are subsequently calculated for individuals in an independent cohort based on the weighted sum of alleles identified in the discovery GWAS (Figure 8B). Finally, relationships of the PGS with other variables that have been measured can be analyzed in this second cohort (Figure 8C) to highlight potential shared genetic etiology.

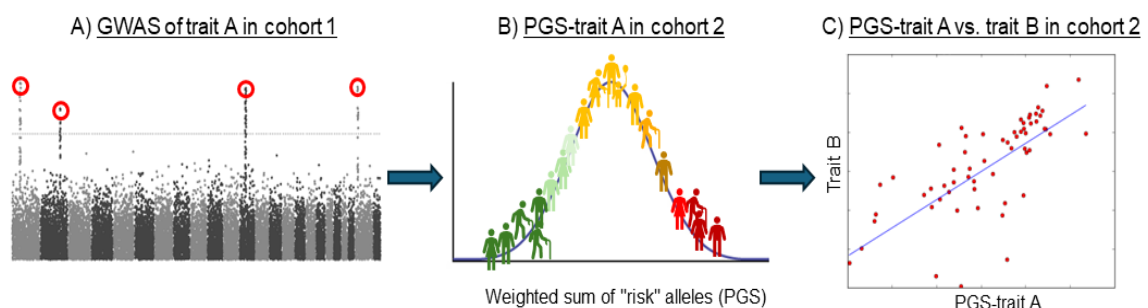


Figure 8. Polygenic score (PGS) analysis. The different stages of PGS analysis are depicted: identifying alleles associated with a trait in a first GWAS (A), calculating a PGS based on these alleles in a second cohort (B), and analyzing the PGS in the second cohort in relation to other phenotypes (C).

Compared to individual markers, PGS captures information from several "risk" alleles, potentially many at once, providing a more comprehensive representation of the genetic factors involved in interindividual differences for a phenotype. Additionally, the implementation of PGS is significant as it informs us about potential relationships between the genetic architecture of different complex traits. However, the development of PGS has been limited until recently due to the need to be designed in large cohorts to achieve adequate power. Most genetic effects for complex traits are small and sample sizes required to detect small effects and identify as many individual SNPs relevant to the phenotype as possible. Nevertheless, PGS designed in large sample sizes can then be applied to small cohorts.

Following this approach, a PGS based on the beat synchronization GWAS has been significantly associated with rhythm, pitch, and melody discrimination in two different studies (Gustavson et al., 2023; Wesseldijk et al., 2022). Additionally, this PGS has also been associated to a more general phenotype of musical engagement in two different cohorts, "consistent with beat synchronization capturing a dimension of musicality" (Niarchou et al., 2022; Wesseldijk et al., 2022).

From previous studies, we have some knowledge on potentially relevant biological pathways involved in the neural encoding of sound. However, the genetic modulators of

the neural correlates of sound processing is still *terra incognita* in the field of cognitive neuroscience.

Results arising from genetic studies would open new insights to understand the biological factors regulating this cognitive process, but also shed some light on the causal factors behind neurodevelopmental problems and be relevant for the study of the complex relationships between our genes and the environment in which we develop. Additionally, these studies have the potential of being especially informative since the effect of genetic variants on physiological endophenotypes with a continuous distribution (such as the FFR and other ERPs) is higher and more homogenous than on a clinical entity (Anokhin, 2014).

CHAPTER II:

OBJECTIVES

2.1 GENERAL OBJECTIVE

It is widely acknowledged that individuals differ in their capacity to process auditory information, which is evident in various contexts such as singing in tune, learning a new language and accurately pronouncing it, understanding speech in challenging environments, and distinguishing different aspects of sound (e.g. notes, tone, rhythm, beats). A precise neural representation is essential for all these processes and can be measured using electrophysiological techniques. However, the exact impact that background noise can have on the neural encoding of sounds and the precise effect that a more accurate neural representation of sounds has on different cognitive abilities are still not clear.

As we have seen in Chapter I, the neural response to an auditory stimulus exhibits high plasticity as a consequence of environmental factors. For example, extensive studies have explored how language exposure and musical training can modify both the anatomical brain structures and their functional responses. On the other side, previous research has identified a relevant role of genetic factors in the individual differences observed for various auditory skills. These skills have shown elevated heritability values and some candidate genes have been proposed, but their replication in independent studies is still very limited. Moreover, (almost) no previous study has targeted the genetic factors modulating the neural encoding of sound stimuli.

The general objective of the present doctoral thesis is to explore the contributions of environmental and genetic factors to the individual differences observed in sound processing during the early stages of auditory processing, specifically in the frequency following response (FFR). Moreover, the relationship between the FFR and functional assessments of auditory and speech perception, such as reading skills, musical abilities, and speech in noise perception, will be assessed while considering the bilingual experience as a potential modulator influencing their respective performances.

2.2 SPECIFIC OBJECTIVES

To achieve the general goal, we designed four studies, each with their specific objectives:

2.2.1 Study I: FFR and noise

The goal of the present study is to assess the effects of background noise on the FFR, and its relationship with perceptual performance in a speech in noise test.

Specific objectives:

- Assess if the background affects the encodings of the temporal envelope of the FFR.
- Assess the effect of the background noise on the fine structure of the FFR.
- Evaluate the effect of the background noise on pitch tracking measures of the FFR.

- Explore the relationship between the FFR and the performance in a behavioral task of speech in noise perception.

2.2.2 Study II: FFR and functional implications

The main purpose of the current study is to evaluate the relationship between the FFR and behavioral measures of reading, musical, and amplitude modulated (AM) discrimination skills. Additionally, we aim to explore if these behavioral measures are related to each other.

Specific Objectives:

- Explore if the FFR correlates with the performance in a reading task.
- Explore if the FFR correlates with the outcomes of a music skill test.
- Evaluate the relationship between the FFR and temporal perception of an acoustic stimulus, measured by an AM-discrimination task.
- Explore if the reading, musical, and AM-discrimination skills are related to each other.

2.2.3 Study III: Genetics of the FFR

The purpose of this study is to explore the effect of genetics on the FFR.

Specific Objectives:

- Replicate the involvement of the promoter of the *SLC6A4* gene in the observed differences for the FFR.

- Identify new genetic loci associated to differences in FFR.
- Identify shared genetic background between the FFR and other related cognitive or behavioral traits.

2.2.4 Study IV: Bilingualism as a possible modulator of the neural response

The goal of the present study is to explore if bilingualism, specifically the age of acquisition of the second language, modulate the FFR and different behavioral measures.

Specific Objectives:

- Evaluate if there are differences in the FFR based on the age of acquisition of the second language.
- Evaluate if there are differences in the FFR elicited in background noise based on the age of acquisition of the second language.
- Explore if there are differences in the performance of the behavioral assessments, taking into account the age of acquisition of the second language.

CHAPTER III: METHODOLOGY

3.1 DESIGN OF THE EXPERIMENTAL SESSION

To achieve the objectives proposed in this doctoral thesis, we established one experimental session per participant composed by an EEG experiment, five tests aimed at evaluating our behavioral skills of interest, and a DNA sample taken through saliva. This experimental session was designed to address all the objectives proposed in each of the four studies constituting this thesis.

Worth highlighting is the fact that this thesis is part of a larger project titled "The Role of Genes and Epigenetic Factors in Modulating Brain Potentials Associated with Speech and Musical Processing (GEPLM)". This broader project encompasses data and analyses that extend beyond the scope of this thesis.

3.2 PARTICIPANTS

One hundred and twenty-nine (N=129) volunteers participated in the present study, of which thirteen were excluded—twelve for having participated in our pilot study and one due to the poor quality of their EEG recording. Finally, the sample comprised one hundred and sixteen (N=116) healthy adults aged between 18 and 35 years [mean age= 20.4 years (SD±3.38); 94 females (80.3%)] with no history of neurological, psychiatric or auditive disorder and psychotropic drug consumption. All the participants reported having European ancestry, this check was made with the purpose to avoid the genetic heterogeneity due to population stratification. All participants were native Spanish speakers. Since the FFR is a response sensitive to auditory experience (Musacchia et al., 2007), consequently participants with more than three years of musical training were excluded.

Before the experiment, we tested hearing thresholds by pure-tone audiometry (frequency range: 250–4000 Hz) using Audiometric ER-3A ABR insert earphones (Etymotic Research, Inc). Participants with thresholds higher than 30 dB sound pressure level (SPL) for all tested frequencies and/or an interaural difference ≥ 10 dB, were excluded from the study.

All participants had an economic compensation of 30€ for their participation and gave their written informed consent at the beginning of the session. This study was approved by the Bioethics Committee of the University of Barcelona (IRB00003099) according to the Ethical Principles of Medical Research Involving Human Subjects (Declaration of Helsinki).

3.3 STIMULUS AND EXPERIMENTAL DESIGN

Stimulus

Auditory brainstem responses were recorded in response to a vowel-vowel syllable /oa/ with a duration of 250 ms (Figure 9), created in *Praat* (Boersma & Weenink, 2022). This stimulus was developed and validated to fully test the encoding of sounds at subcortical level (Arenillas-Alcón et al., 2021).

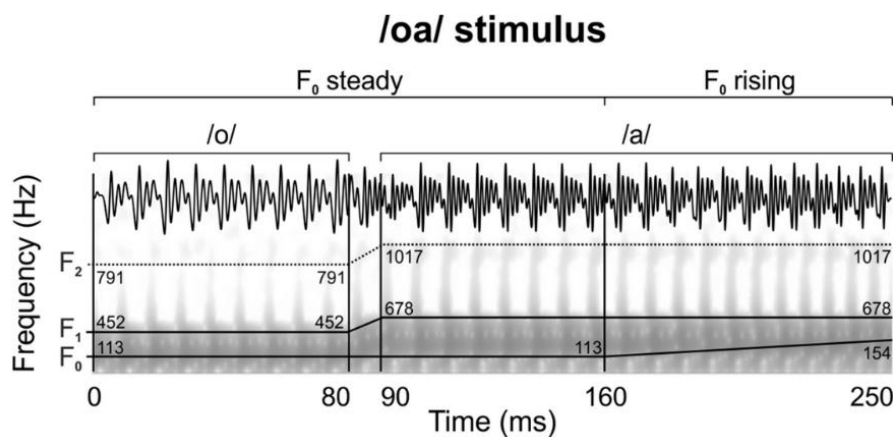


Figure 9. Representation of the Spectro-temporal characteristics of the /oa/ speech stimulus. This includes the time waveform (top) and spectrogram (below). The illustration depicts the pitch contour trajectory (F_0) and the corresponding formant structures (F_1 and F_2).

This stimulus is characterized by starting with a flat pitch contour (from 0 to 160 ms, $F_0 = 113$ Hz) and ending with a rising pitch contour (from 160 to 250 ms, $F_0 = 113$ -154 Hz). The /o/ vowel section lasted 0 to 80 ms ($F_1 = 452$ Hz; $F_2 = 791$ Hz), between the 80 to 90 ms is the formant transition, and the vowel section /a/ begins at 90 ms and ends at 250 ms ($F_1 = 678$ Hz; $F_2 = 1017$ Hz).

Experimental design

Two experimental conditions were used: *In quiet* and *In noise*. The first condition consisted in the presentation of the stimulus in a quiet background and in the latter the stimulus was presented with a looped semantically anomalous speech babble displayed in the background, 5 dB SPL softer than the syllable.

The /oa/ stimulus was presented binaurally, with alternating polarities (in counterbalanced order), at an intensity of 80 dB SLP via Etymotic ER-3A ABR insert earphones (Etymotic Research, Inc., Elk Grove Village, IL-USA). The stimulus onset asynchrony (SOA) was of 325 ms, and only constant timing condition was used since previous studies showed no differences between constant and jittered conditions (Selinger et al., 2016). On the whole, four blocks of 1500 stimuli each were presented, two blocks per each condition, in a counterbalanced and alternate order of

presentation. The participants were instructed to relax and focus on a silent movie while ignoring the auditory stimulation.

3.4 NEUROPHYSIOLOGICAL RECORDING

Data acquisition and pre-processing

The auditory brainstem response was extracted from continuous EEG recording acquired with Curry 7 software and Neuroscan SynAmps RT amplifier (NeuroScan, Compumedics, Charlotte, NC, USA) with a sampling rate of 10 kHz. For the recording, 41 electrodes were used from a scalp Ag/AgCl nylon cap (Quick-Cap; Compumedics, Charlotte, NC, USA) at the standard 10–20 system locations. Four additional electrodes were included at the right and left mastoids and earlobes (M_2 and M_1 , A_2 and A_1 , respectively). The electrooculogram (EOG) was measured with three unipolar electrodes, one placed below the left eye and two electrodes placed on the horizontal outer canthi of the eyes. The ground electrode was located between Fz and FPz, and all the impedances were kept to $<10\text{ k}\Omega$ during the recording session. The data were on-line LowPass filtered at 2000 Hz with a range of DC–2000 and referenced to the nose.

EEG data was pre-processed off-line using EEGLab V. 2021.1 toolbox, running under MATLAB version R2021b (Mathworks). Data were off-line bandpass filtered from 80 to 1500 Hz, re-referenced to the left earlobe (A_1) and epoched from 40.95 ms before the stimulus onset (pre-stimulus period) to 249.975 ms relative to stimulus onset. A total of 6000 epochs were obtained per participant before artifact rejection, divided into 3000 per each condition. Epochs that showed activity that exceeded $\pm 35\text{ }\mu\text{V}$ were rejected automatically as artifacts.

3.5 FFR PARAMETERS

The FFR is commonly used in a general sense, describing different aspects of the speech brainstem responses to formants, harmonics and the fundamental frequency (Aiken & Picton, 2008). For this reason, it is important distinguish between the different kind of the responses referenced as FFR to avoid confusion. For our study, we used the terminology commonly used in FFR studies (e.g. Coffey et al., 2016, 2019; Krizman & Kraus, 2019). We used the FFR temporal envelope ($FFR_{(ENV)}$) as the response that reflects the information related to the fundamental frequency and the FFR temporal fine structure ($FFR_{(TFS)}$), as the response that reflects the spectral information about the harmonic and/or formant structure of the stimulus. In complement, we also included three measures of pitch tracking, that reflect the ability to track pitch stimulus variations in the temporal domain. All these features were obtained for both “in quiet” and “in noise” conditions, separately, using Matlab version R2022a and EEGLAB version 2021.1.

Fundamental frequency ($FFR_{(ENV)}$)

To obtain a measure of pitch encoding in each participant, we estimated the spectral amplitude and the signal-to-noise-ratio (SNR) of the response at the fundamental frequency. To enhance the $FFR_{(ENV)}$ and minimize the cochlear microphonic we averaged the response to stimuli of alternating polarities [(Rarefaction + Condensation)/2] (Aiken & Picton, 2008). Then, a Fast Fourier Transform (FFT with a hanning taper) was computed only on the flat contour of the stimulus (10-160 ms). The spectral amplitude was estimated as a measure of the magnitude of the neural phase-locking at 113 Hz, that corresponds to the fundamental frequency of our stimulus in the section with flat pitch contour.

The SNR was computed as an indicator of the relative spectral magnitude of the response relative to the noise. The SNR was calculated by dividing the mean of the spectral amplitude in a frequency window of ± 5 Hz around the F_0 peak (108 to 118 Hz) by the average of the spectral amplitude of the non-stimulus evoked portion of the response, in our case, frequency windows of ± 28 Hz on either side of the peak (left side 80 to 108 Hz; right side 118 to 146 Hz).

Formant structure ($FFR_{(TFS)}$)

To evaluate the encoding of the $FFR_{(TFS)}$, we estimated the spectral amplitude and the SNR for the first formant (F_1) for both vowels that composed our stimulus, only on the section with a flat pitch contour. To accentuate the $FFR_{(TFS)}$ and minimize the contribution of the $FFR_{(ENV)}$, we subtracted the response to both stimulus polarities [(Rarefaction-Condensation/2)].

The spectral amplitude for the F_1 for the /oa/ flat pitch contour (10 to 160 ms) was calculated in two separate sections of the neural response corresponding to the /o/ section (10-80 ms; $F_1 = 452$ Hz) and the /a/ section (90-160 ms; $F_1 = 678$ Hz). The SNR related to the corresponding spectral amplitudes were calculated in the same way as mentioned above.

Pitch tracking measures

With the objective of exploring how well participants encoded and followed the pitch variations of the stimulus, we estimated three parameters that reflected different aspect of the pitch tracking process, following the recommendations of previous studies (e.g. Ribas-Prats et al., 2019; Russo et al., 2008): The stimulus-to-response cross-correlation,

pitch strength, and pitch error. All these parameters were calculated in the response corresponding to the /a/ stimulus portions, in both pitch contours (a_{flat} = 90 to 160 ms; a_{rising} = 160 to 250 ms) separately.

The stimulus to response cross-correlation was used to assess how the neural response reproduced the morphology of the stimulus waveform. For each participant, the section of the stimulus of interest (flat and rising pitch contour) was cross-correlated with its neural response, taking into account the corresponding neural delay (for more details see Krizman et al., 2019).

Pitch strength is a measure of the response periodicity and reflects the robustness of the neural-phase-locking at the stimulus fundamental frequency (N. Russo et al., 2004). This parameter was obtained through the autocorrelation method with a time-window approach using bins of 40 ms with 1 ms step size (for more details, Jeng et al., 2013). Pitch strength was calculated for the two pitch contour sections separately, by averaging the normalized autocorrelations values across all bins.

In a similar way, we computed the pitch error based on the autocorrelation method. The absolute Euclidian distance between the stimulus F_0 and its response was calculated for each time bin and averaged across all time bins corresponding to the flat and rising pitch contours, separately. This index represents the pitch encoding accuracy along the course of stimulus presentation.

We also included the neural lag, which is not a pitch tracking measure itself but provides us with information about the neural transmission delay. This parameter was calculated

through the cross-correlation and indicated the time point where the maximum Pearson's correlation between stimulus and response occurred.

3.6 BEHAVIORAL MEASURES

Five behavioral tests were applied to the participants to assess the following auditory skills: SIN perception, reading skills, music perception skills, AM discrimination, and language experience and proficiency.

Speech In Noise Perception (SIN)

The Speech in noise test (SPIN) was used to measure the intelligibility of the speech in noise perception, adapted from the original English version (Kalikow et al., 1977) and developed and validated for the Spanish population (Cervera & González-Alvarez, 2010, 2011).

The task consists of presenting fifty sentences, half of them with high predictability (HP) (e.g. "In the castle stands the tower") and the other half with low predictability (LP) (e.g. "Someone has said tower"), while asking the participants to write the last word of the sentence. Each sentence was presented at 80 dB with a background noise 5 dB softer than the stimuli. The background noise was a looped semantically anomalous speech babble. For the analysis of the SPIN results, we used the raw scores obtained from the three different parameters measured in this task: performance in highly predictable sentences, performance in low predictability sentences, and the total score, which is the sum of both subtests. Raw scores were calculated as the sum of correct answers for each category.

Reading Skills

The Word and Pseudoword Reading subtests of the *“Batería de Evaluación de los Procesos Lectores en Secundaria y Bachillerato, revisada”* (PROLEC-SE-R) were used to assess the reading performance across Word and Pseudoword list subtest. This test is commonly used to assess reading problems in Spanish students and to evaluate the level of reading development (Cuetos Vega et al., 2016).

The task involved presenting four lists of word for the Word subtest and two lists of pseudo-words for the Pseudo-word subtest. Participants were instructed to read each list aloud as quickly as possible. The time taken and errors made during the performance of each list were recorded, and overall scores for each subtest were determined based on both accuracy and reading time. Separate scores for accuracy and for time were also calculated for each subtest.

Musical skills

The *Profile of Music Perception Skills-short version* (PROMS-S) battery assesses musical perception skills across eight modalities: melody, pitch, timbre, tuning, rhythm, rhythm to melody, accent, and tempo. For each modality subtest two stimuli were presented (e.g. standard and comparison melodies) and the participant had to discriminate if the stimuli were the same or different. These eight subtests were grouped into three main domains: Tonal (composed by the melody and pitch subtests), Temporal (includes the rhythm, embedded rhythm, metric accent, and tempo subtests), and Qualitative (composed by the timbre and tuning subtests).

The PROMS battery reported good psychometric properties, with a test-retest reliability of $r > 0.85$ for the overall score, and coefficients between 0.56 to 0.85 for the individual subtest (Law & Zentner, 2012). Regarding the evidence of validity of the instrument, different studies reported adequate convergent validity between PROMS with Tonal and Rhythm subtest of the Advanced measures of Musical Aptitude Profile ($r = 0.68$ and 0.42 , $p < 0.01$); Rhythm subset of the musical Ear Test ($r = 0.04$, $p < 0.01$); Tempo subset of Musical Aptitude Profile ($r = 0.33$, $p < 0.05$); Timbre (monophonic) ($r = 0.40$ to 0.53 , $p < 0.01$), and harmonic closure judgment accuracy ($r = 0.36$, $p < 0.01$). for more detail see Kunert et al., 2016; Law & Zentner, 2012; Zentner & Strauss, 2017).

The Amplitude Modulation discrimination

The AM- Discrimination task consisted in the modulation of the acoustic signal at a specific frequency, and it allows to study temporal perception of sounds. Due the relevance of the temporal perception in the field of the auditory perception group (for a detailed revision see chapter I, section 1.2), We included an amplitude modulation (AM) discrimination task, developed by our research group, utilizing the beta-gamma frequency band due to its significant role in phonemic processing and perception (Goswami, 2011).

The test consisted in the presentation of three sounds, two with the same AM rate (standard) and one with a higher modulation rate (deviant), and the participant should indicate which one is different (or perceptually faster). The stimuli used were white noise, bandpass filtered from 200 to 20.000 Hz, with an amplitude modulation rate of 20- 40 Hz. MATLAB version R2021b was employed both for generating the stimuli and implementing the task. To obtain a measure of the performance in temporal

discrimination of each participant, we used the Weber Fraction (WF). This index estimates the minimum detectable change in the modulation rate by calculating the geometric mean of the last 12 reversals.

Language experience and proficiency

The *Language Experience and Proficiency Questionnaire* (LEAP-Q) is a questionnaire to assess the language experience and proficiency in more than one language (bilingualism and multilingualism). In this test, participants have to list all of the languages that they know in order of 1) dominance and 2) acquisition. Through a quantitative scale, it measures the level of proficiency in understanding, speaking, reading, and writing domains, for each language, taking into account the age of acquisition of each language, the educational degree, context, and exposition for each language (e.g. Country, family, school, interacting with friends, watching tv, listening radio, etc.). Based on the score obtained, it could calculate different measures of Bilingualism/multilingualism, and different category of them.

The survey was validated in two studies, proving to be a reliable and effective instrument for evaluating the language characteristics of multilingual, neurologically intact adults in research environments (Marian et al., 2007). For the present study, we used the Peninsular Spanish Version of the LEAP-Q, available in *LEAP-Questionnaire « Bilingualism and Psycholinguistics Research Laboratory*, n.d.

3.7 DNA COLLECTION AND GENOTYPING

Two milliliters of saliva were collected using Oragene OG-500 collection kits (DNA Genotek) and following manufacturer's instructions. Samples were stored at 4°C at the

Neurogenetics laboratory at the Department of Clinical Psychology and Psychobiology, Barcelona University (UB), until all participants finished the experiments.

Collection tubes were then sent to the USC node of the National Genotyping Center in Santiago de Compostela (CeGen-FPGMX) for DNA extraction following manufacturer's specifications. DNA quality and concentration were subsequently assessed by means of fluorometric methods (Quant-iT™ assays, Invitrogen) and concentrations were standardized to 50ng/μl. DNA samples were shipped to two different genotyping facilities: to the Fundació Mútua de Terrassa (FMT) to type the polymorphisms in the promoter region of the *SLC6A4* gene, and to the National Genotyping Center node in Madrid (CeGen-CNIO) for SNP array genotyping.

Polymorphisms in the promoter region of the *SLC6A4* gene genotyped at the FMT included the VNTR 5-HTTLPR and the A/G SNP rs25531. The 5-HTTLPR (5-HTT linked polymorphic region) is a 20-23 bp variable number tandem repeat with two alleles: a short (S) variant comprising 14 copies of the repeat unit, and a long (L) variant comprising 16 copies. The rs25531 SNP is only present in the L allele. Both polymorphisms were determined through PCR amplification, followed by digestion of the PCR products using the MspI restriction enzyme and agarose gel electrophoresis.

Genome-wide SNP genotyping was conducted at the CeGen-CNIO using the Infinium™ Global Screening Array-24 v2.0 (GSA) BeadChip (Illumina). The GSA array includes ~650,000 genetic markers (mostly SNPs, but also short insertion/deletions) distributed across all chromosomes, including sexual chromosomes and mitochondrial genome. From the intensity files (.idt) generated after the determination, genotypes were called using GenomeStudio 2.0.3 software, PLINK Input Report Plug-in v2.1.4, GSA v.3.0 +MD

v.1.0 cluster file (manually-curated by the CeGen-ISCI Unit), and the GSAMD-24v3-0-EA_20034606_A1 manifest file. An overall call rate of 0.901 was obtained at this stage.

SNP array Quality Control and Imputation

Genotype quality control (QC) was performed with PLINK v1.9 (Chang et al., 2015). Specifically, markers that exceeded a 5% missingness rate, did not fit Hardy-Weinberg proportions ($p < 10^{-5}$), or had a minor allele frequency (MAF) lower than 1% were removed from the dataset. Additionally, samples with more than 1% missing data were also removed.

Additional QC steps were conducted to avoid potential biases in the results due to the inclusion of related individuals or population outliers. An Identity by Descent (IBD) analysis detected one duplicated sample and a pair of first-degree relatives, that were removed from the dataset. A Principal Component Analysis (PCA) of our dataset merged with continental populations from the 1000 Genomes Project (1000 Genomes Project Consortium et al., 2015) was performed and 6 individuals were removed from the dataset due to partial non-European ancestry (Figure 10). Both IBD and PCA analyses were run in 238,906 unlinked variants obtained after linkage disequilibrium (LD) pruning, using an r^2 threshold of 0.3.

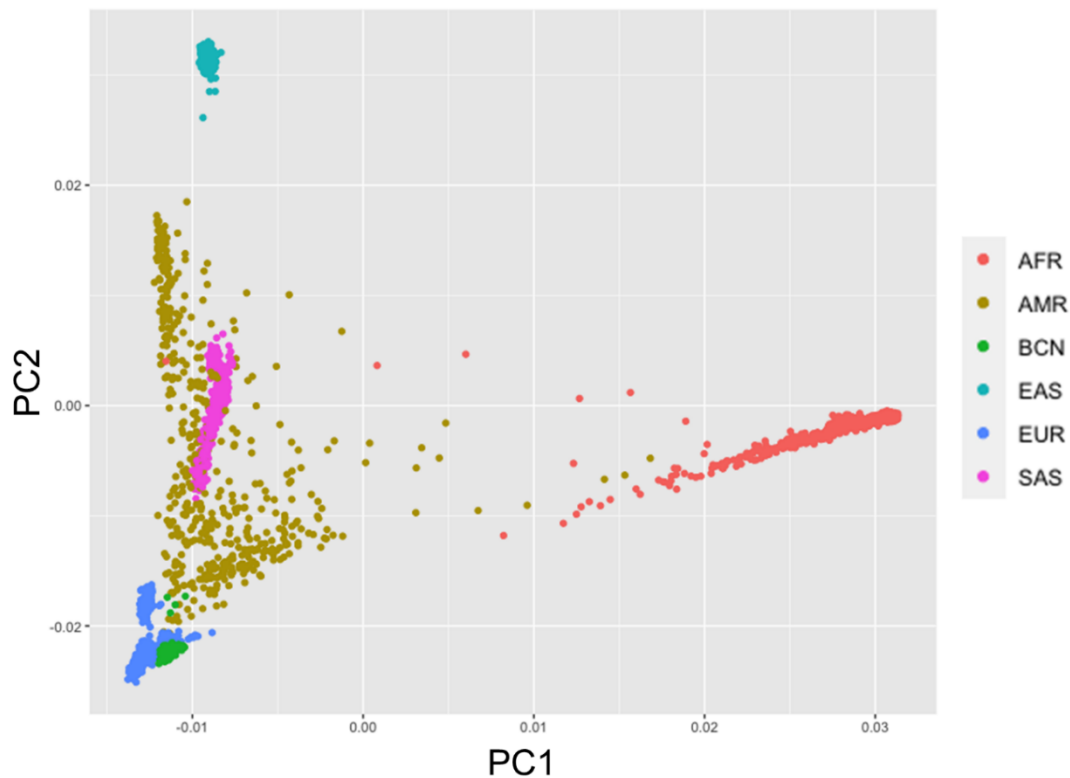


Figure 10. PCA results depicting individuals from the GEPLEM project merged with individuals from the 1000 Genomes Project. Population acronyms correspond to their continental affiliation (AFR: African; AMR: American; EAS: East Asian; EUR: European; SAS: South Asian). Results depict first and second PCs.

After quality control, information on all non-autosomal (i.e. sex chromosomes and mitochondrial genome) chromosomes was removed from further analyses. At this stage, a total of 123 samples and 493,712 genetic markers across all autosomes were available for downstream analyses. Call rate was 0.9991.

Next, genotype imputation was conducted on the genomic data. Based on the genotypes observed in our dataset, genetic imputation allows for the prediction of genotypes at non-typed positions taking advantage of the linkage disequilibrium (LD)

patterns observed across the human genome. In our case, imputation was conducted using the TOPMed Imputation Server (<https://imputation.biodatacatalyst.nhlbi.nih.gov/#!>) using their TOPMed reference panel, which includes reference data from ~180,000 individual genomes (<https://topmed.nhlbi.nih.gov/>). Imputation algorithms used were Eagle v2.4 for phasing and Minimac4 for imputation. After imputation, 292M imputed SNPs were obtained, and ~204M were kept for analyses for their good imputation quality ($R^2 > 0.8$).

3.8 STATISTICAL ANALYSIS

Based on the objectives of the present thesis, the corresponding statistical analysis was selected for each study, aligning with the specific goals of each study.

All statistical analyses were performed using Jamovi project (2022) version 2.3, and plots were made in MATLAB version R2022a and RStudio Version 4.1.2. For the genetic analyses, the specific software, toolboxes used, and their corresponding versions are detailed in their respective sections within the statistical analysis.

3.8.1 Study I: FFR and Noise

The purpose for the first study was to evaluate the effect of the background noise on the neural encoding of the syllable /oa/, reflected in the parameters of the FFR. Additionally, the study aimed at evaluating the relationship between $FFR_{(ENV)}$, $FFR_{(TFS)}$, pitch tracking measures, and performance in a behavioral test of speech in noise perception (SPIN test).

To achieve these goals, first, descriptive statistics and the Kolmogorov-Smirnov normality test with Lilliefors' correction were performed to gain an overview understanding of the sample's characteristics and data behavior. Based on the results of the normality test, the Wilcoxon signed-rank test was used to assess whether the background condition affected the neural encoding of the $FFR_{(ENV)}$, $FFR_{(TFS)}$, and the neural lag. Repeated measures ANOVA were conducted to test the main effects of the background condition (in quiet vs. in noise) and the pitch contour (flat vs. rising), as well the interaction effect between both, across all estimated pitch tracking parameters (cross-correlation, pitch strength, and pitch error).

Finally, Spearman correlations were estimated between the FFR parameters and the SPIN test, for both background conditions.

3.8.2 Study II: FFR and Functionals Implications

To explore whether there were functional-level relationships between our electrophysiological measurements and the evaluated auditory skills, in addition to assessing if these auditory skills are related, a Principal Component Analysis (PCA) was performed to create components that summarized and reflected different aspects of the FFR.

First, we identified and removed variables redundant for our analysis with Spearman's correlation coefficients higher than 0.9, and then we carried out an exploratory PCA with the following variables: F_0 spectral amplitude, F_0 SNR, F_1 /o/ spectral amplitude, F_1 /o/ SNR, F_1 /a/ spectral amplitude, F_1 /a/ SNR, and Neural Lag. Additionally, cross-correlation,

pitch strength, and pitch error in the flat pitch contour segment were included in the analysis.

Each component was created by computing the weighted sum of each variable in the component with its corresponding loading. Second, Spearman's correlations were performed between all the created components and the scores in each behavioral test used, in addition with Pearson's correlations between the scores in all of the behavioral measures utilized.

3.8.3 Study III: Genetics of the Sound Processing

SLC6A4 association analyses

In order to confirm previous links between genetic variation in the promoter of the *SLC6A4* gene and the neural encoding of sounds, a genetic association study between polymorphisms in this genomic region and different FFR parameters was conducted.

Since, the rs25531 SNP is only present in the long (L) allele of the 5-HTTLPR, alleles in these two polymorphisms were grouped according to their level of 5-HTT expression as low (S and L-G haplotypes) and high (LA/LG, LA/S, LA/LA) expression haplotypes, similarly to previous studies (Hu et al., 2005; Selinger et al., 2016). In addition to the 5-HTTLPR and rs25531 variants, we also included in these analyses other markers in the promoter region of the *SLC6A4* gene that were included in the genome-wide SNP array. We extracted from the GWAS dataset all markers that were located between positions 30,235,481..30,238,507 on chromosome 17. These are the genomic coordinates of the *SLC6A4* gene promoter, as stated in assembly GRCh38.p14, the latest version of the human genome reference (<https://www.ncbi.nlm.nih.gov/gene/6532>, accessed on Sept

19th 2023). Thus, 3 additional SNPs were added to the genetic dataset for this goal: rs2020933, rs25532, and rs73263896. A schematic representation of SLC6A4 gene and the location of the polymorphisms analyzed is depicted in Figure 11.

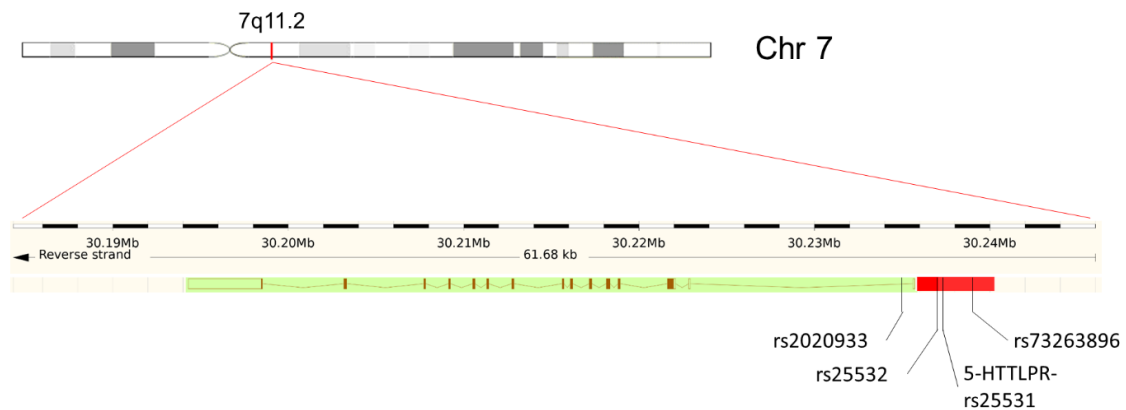


Figure 11. *SLC6A4* gene and its genomic context with the relative location of the polymorphisms analyzed. Information extracted from ensembl.org and genecards.org, accessed on Dec 19th 2023.

Regarding phenotypes, we selected the SNR at the fundamental frequency as the main dependent variable to be analyzed since it was the FFR parameter that was associated with variants in the *SLC6A4* promoter in the previous report. Additionally, we also analyzed the involvement of these polymorphisms in the individual differences observed for the FFR Principal Components constructed in Study II of this thesis. These analyses explored the role of the serotonergic system in the different domains of the FFR: pitch magnitude, pitch tracking and formant encoding.

Association analyses were conducted by means of linear regression models with FFR phenotypes as dependent variables and alleles as independent variables under an

additive genetic model. Regression analyses also included sex, age and the first three PCs from the PCA to avoid population stratification effects (see previous section) as covariates. To assess the stability and robustness of the results, a false-discovery rate (FDR) correction and a permutation test with 10,000 iterations was conducted on the nominally significant (i.e. $p < 0.05$) results. All analyses were run using PLINK v1.9.

Genome-wide association analyses

The phenotypes pitch magnitude, pitch tracking, and formant encoding (F_1 /o/ and F_1 /a/), calculated in Study II, were analyzed as the main traits for genome-wide analysis. A rank-based inverse normal transformation was applied to remove effects of outliers, normalize the distributions, and make effect sizes of the 4 different phenotypes comparable with each other.

Association analyses were conducted with PLINK v1.9 using linear regression of the phenotypes on imputed genotypes assuming an additive model between phenotypes and genotype dosages. We adjusted all models for age, sex, and the first three genotype principal components.

We used the Functional Mapping and Annotation of Genome-Wide Association Studies (FUMA GWAS) webtool (<https://fuma.ctglab.nl/>) to create genome-wide plots, such as Manhattan plots and Q-Q plots and to identify independent (not in LD) significant SNPs at the conventional genome-wide significance threshold ($p < 5 \cdot 10^{-8}$), and the nearest genes based on position.

To provide a biological context of the loci identified in the GWAS analyses, we checked the function and tissue expression profile of the overlapping protein-coding genes using

the data available in the NCBI Gene (<https://www.ncbi.nlm.nih.gov/gene/>) and GTEx Analysis Release V8 (<https://www.gtexportal.org/home/>) databases. Moreover, we also checked previous associations of SNPs in the region with potentially related phenotypes as reported in the GWAS Catalog (<https://www.ebi.ac.uk/gwas/home>).

Post-GWAS analyses: Gene, pathway, and tissue enrichment.

Gene, pathway, and tissue enrichment

To annotate, visualize and prioritize genes, pathways and tissues involved in the genetic architecture of our main phenotypes, we conducted different downstream analyses which involved several analytical approaches implemented in the FUMA webtool (Watanabe et al., 2017, 2019). Specifically, we performed the Multi-marker Analysis of GenoMic Annotation (MAGMA) v1.09 gene (Leeuw et al., 2015), gene-set, and tissue expression analysis. MAGMA is a flexible tool for gene and gene-set analysis of GWAS genotype data. MAGMA's gene analysis uses a multiple regression approach to properly incorporate LD between markers and to detect multi-marker effects. Gene and gene-set analysis are statistical methods for analyzing multiple genetic markers simultaneously to determine their joint effect. These methods can be used when the effects of individual markers are too weak to detect, which is a common problem when studying polygenic traits. SNPs in our dataset were mapped to 18,620 protein coding genes and gene-wide significance was set at $p = 0.05/18620 = 2.68 \cdot 10^{-6}$.

In addition, we used the MAGMA gene-property analyses (implemented in FUMA) to identify tissue specificity of the phenotype analyzed. This analysis aims to test relationships between tissue specific gene expression profiles and phenotype-gene associations using the full distribution of SNP p-values. This analysis has been performed

using the 53 specific tissue types available in GTEx v8 and setting a tissue-wide significance level of $p = 0.05/53 = 9.43 \cdot 10^{-4}$.

Polygenic scores

A polygenic score (PGS) aggregates the effects of many genetic variants into a single number which predicts genetic predisposition for a phenotype. PGS are typically composed of hundreds-to-millions of genetic variants which are combined using a weighted sum of allele dosages multiplied by their corresponding effect sizes, as estimated from a relevant genome-wide association study.

We used the `pgsc_calc` tool in our imputed genotypes (Lambert et al., 2021). It is a bioinformatics best-practice analysis pipeline for calculating PGS on samples with imputed genotypes using existing scoring files from the PGS Catalog, an open database of published PGS (<https://www.pgscatalog.org/>). Each PGS in the Catalog is consistently annotated with relevant metadata; including scoring files (variants, effect alleles/weights), annotations of how the PGS was developed and applied, and evaluations of their predictive performance. We included in our analyses the following defined PGS from the PGS Catalog: intelligence (Hatoum et al., 2023; Privé et al., 2022); attention deficit hyperactivity disorder (Lahey et al., 2022); autism spectrum disorder (Grove et al., 2019; Gui et al., 2022); executive function measurement (Hatoum et al., 2023); and white matter volume measurement (Tanigawa et al., 2022). The corresponding PGS were calculated for the nine phenotypes listed above for all samples included in our dataset (linear sum of weights and dosages). Then, we computed a phenotype correlation matrix of each one of these PGS with the FFR components

studied (pitch magnitude, pitch tracking and formant encoding (F_1 /o/ and F_1 /a/) and with the behavioral assessments.

3.8.4 Study IV: Bilingualism as a potential modulator of the neural response

The main purpose of the present study is to explore the influence of the linguistic experience in the neural response, reflected in the parameters of the FFR and the behavioral measures used.

To achieve these goals, first, we created a multicategory variable based on the performance and the age at which participants became fluent in their second language. These age cutoffs were selected to reflect typical milestones in language development proposed in previous studies (Krizman et al., 2014; Krizman, Marian, et al., 2012a; Pelham & Abrams, 2014; Shi, 2010) and defined the following 4 groups: Native monolinguals (average score <8 in the second language in speaking, listening, reading, and writing domains), Native bilinguals (the second language (in dominance) was acquired at age 2 or younger), Early bilinguals (the age of acquisition of the second language was 2-7 years old), and Late bilinguals (age of acquisition of the second language > 7).

Given the hypothesis that the acquisition of the first language may influence the early development of the perceptual system, impacting the acquisition of a second language (Sebastián-Gallés & Soto-Faraco, 1999), and considering findings suggesting that the age of second language acquisition enhances brain connectivity, reflected in the white matter, facilitating the quicker transmission of semantic information (Mohades et al., 2012), we reclassified our sample into two groups based on the multicategory variable

derived from the LEAP-Q (see above): native bilinguals and non-native bilinguals (including early and late bilinguals). Monolingual participants were excluded in this study. This recategorization allows for a more effective cross-group comparison, considering that our sample was recruited in the city of Barcelona, where the population is predominantly bilingual in Catalan/Spanish (approximately 94% in our sample), especially native and early bilinguals.

MR ANOVA was carried out for language experience (native/non-native bilinguals) in the $FFR_{(ENV)}$, $FFR_{(TFS)}$ and the measures of pitch tracking. Despite the rejection of the normality assumption for some of the FFR parameters (see results in Study I, Table 1), we chose this approach because it allowed us to include a between-subjects factor in a within-subjects model as well as to test the main effects and interaction effects. Additionally, recent simulation studies have observed the robustness of MR ANOVA, even in the presence of non-normality (Blanca et al., 2017, 2023; Islam & Abbas, 2022).

Finally, with the aim to evaluate if the native bilinguals showed different performance than non-native bilinguals in SIN perception, reading skills, musical skills and AM discrimination, independent samples t-test were estimated between both bilingual acquisition group, across all of the auditory skills evaluated. Notewise, in those cases in which the homogeneity of the variance was rejected, the *Welch's t* was reported (Poncet et al., 2016).

CHAPTER IV:

RESULTS

4.1 STUDY I: FFR AND NOISE

Effects of background noise on the FFR

To investigate the impact of background noise on the FFR evoked by our novel verbal stimulus, we compared the different measures of the FFR: $FFR_{(ENV)}$, $FFR_{(TFS)}$, Neural Lag, and the pitch tracking measures, "in quiet" and "in noise" background conditions.

Based on the results, it can be seen that the background noise affects the transmission delay, reflected in a longer neural lag 'in noise' compared to the 'in quiet' background condition, with a medium effect size ($W=4704$, $p<0.001$, *Cohen's d* =0.44) (Table 1).

The Grand average waveforms of $FFR_{(ENV)}$ were obtained in quiet and in noise background conditions (Figure 12). In the noise condition, the amplitude response at the F_0 (113 Hz) was higher compared to the quiet condition. However, in the noise condition, the spectral amplitude of the harmonics was lost, unlike the quiet condition where it exhibited peaks around 226, 339, and 452 Hz (Figure 12c). For the F_0 spectral amplitude, statistical comparison revealed a significantly higher amplitude, ($W=5138$, $p<0.001$, *Cohen's d* =0.51) and SNR ($W=5799$, $p<0.001$, *Cohen's d* =0.71), in the background noise condition compared to the quiet condition, with a medium to high effect size (Table 1).

Table 1. Summary statistics and the Wilcoxon paired test are provided for the neural lag, $FFR_{(ENV)}$, and $FFR_{(TFS)}$ between the background conditions.

Measure	Median	Mean	95% interval	Confidence	Normalit y test p - value	W- Wilcoxon (p -value)	Effect size $Cohen's d$
Neural Lag (ms; from 10 to 250 ms)							
In quiet	8.437	8.815	8.444	9.158	<0.001	4704 (<0.001)	0.44
In noise	9.862	9.737	9.476	9.999	<0.001		
FFR _(ENV)							
F ₀ spectral amplitude (μV; from10 to 160 ms)							
In quiet	0.026	0.029	0.026	0.032	<0.001	5138 (<0.001)	0.51
In noise	0.036	0.040	0.036	0.044	<0.001		
F ₀ SNR (from 10 to 160 ms)							
In quiet	3.411	3.245	2.572	3.917	0.111	5799 (<0.001)	0.71
In noise	6.830	6.452	5.784	7.119	<0.001		
FFR _(TFS)							
F ₁ /o/ spectral amplitude (μV; from10 to 80 ms)							
In quiet	0.014	0.016	0.013	0.019	<0.001	5843 (<0.001)	0.72
In noise	0.007	0.010	0.007	0.013	<0.001		
F ₁ /o/ SNR (from10 to 80 ms)							
In quiet	528.877	420.369	366.828	473.910	<0.001	5575 (<0.001)	0.64
In noise	327.968	190.146	107.091	273.201	<0.001		
F ₁ /a/ spectral amplitude (μV; from90 to 160 ms)							
In quiet	0.007	0.008	0.007	0.009	<0.001	3451 (0.874)	-
In noise	0.007	0.008	0.007	0.010	<0.001		
F ₁ /a/ SNR amplitude (from90 to 160 ms)							
In quiet	456.260	329.746	263.424	396.068	<0.001	3115 (0.445)	-
In noise	450.038	351.278	290.012	412.545	<0.001		

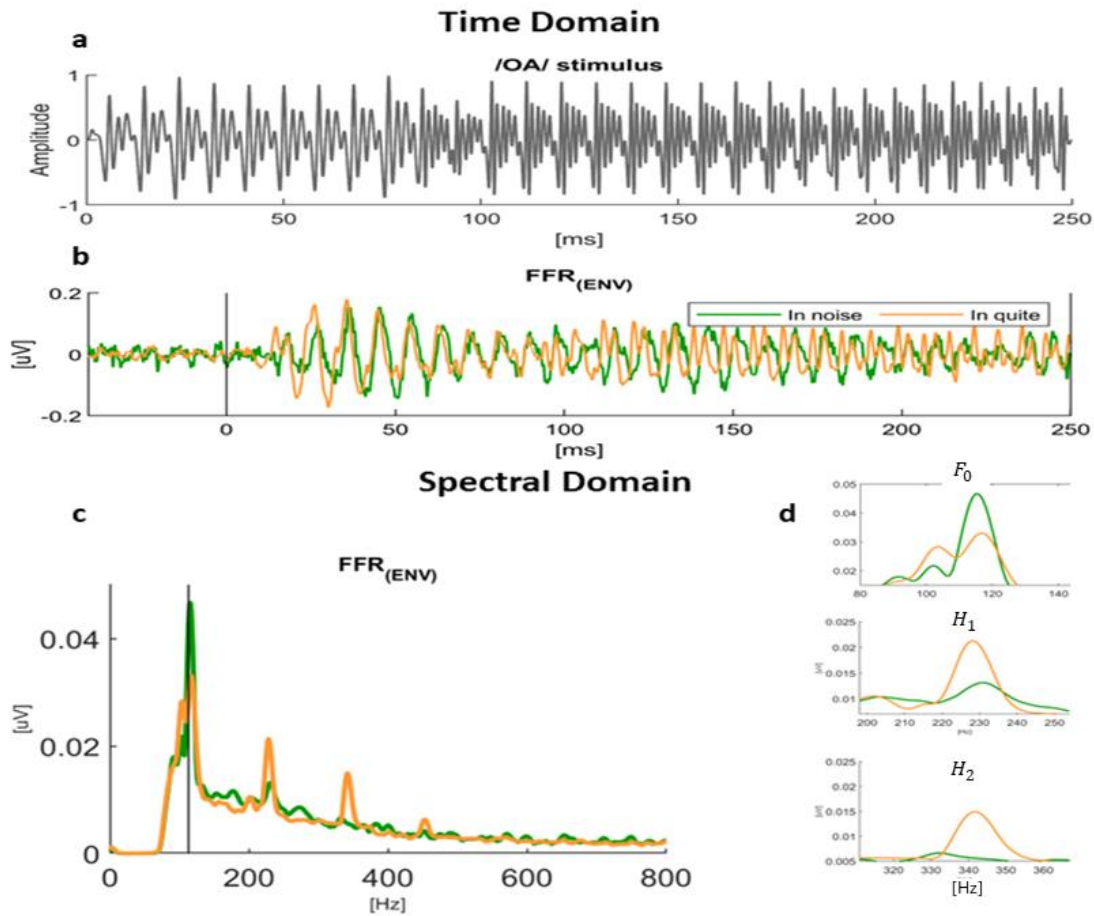


Figure 12. Spectro-temporal characteristics of the stimulus and the corresponding elicited response. Temporal representation of the stimulus (a); Grand averages time-domain waveform of the FFR_{ENV} obtained in the flat pitch contour (b); Spectral amplitude FFR_{ENV} in response to the $F_0=113$ Hz, extracted from the flat pitch contour (c); Spectral amplitude for the $F_0=113$ Hz, $H_1=226$ Hz, and $H_2=339$ Hz respectively (d). All of the estimations were made for both, in quiet (orange) and in noise (green) condition.

The Grand average waveforms of the FFR_{TFs} were calculated by subtracting polarities to accentuate higher frequencies corresponding to harmonic and formant components in the response. This was done separately for both quiet and noise background conditions (Figure 13).

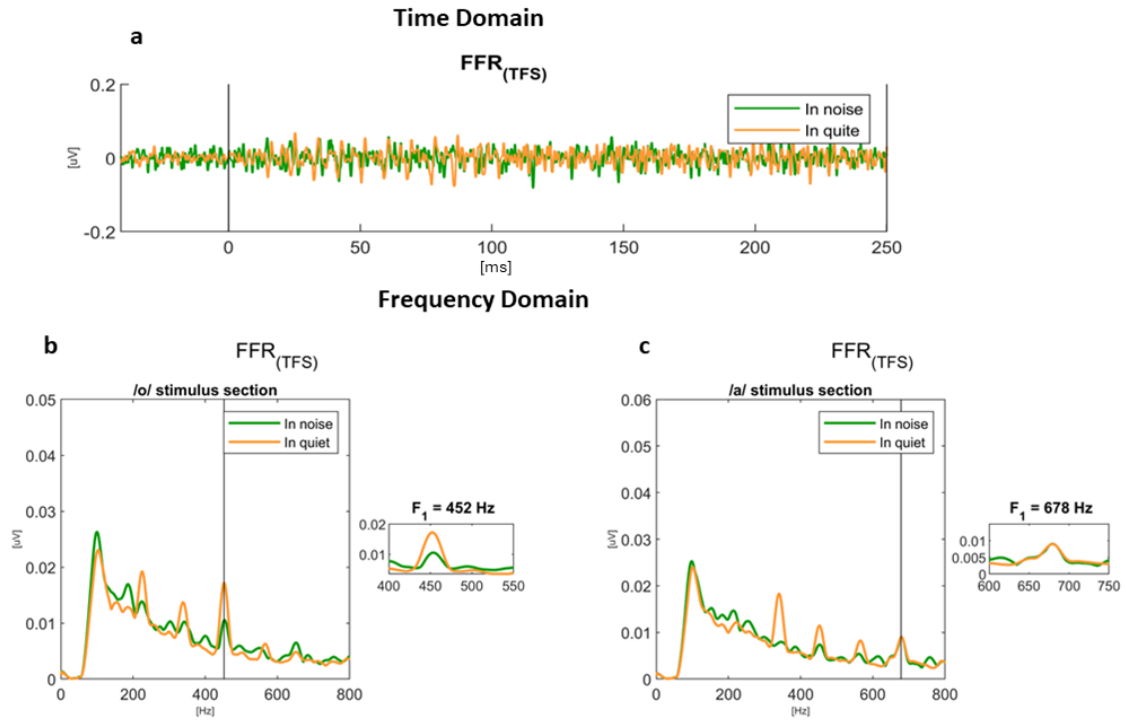


Figure 13. Spectro-temporal characteristics of the $FFR_{(TFS)}$, obtained from the subtracted polarities, are depicted. Grand average time-domain waveforms of the $FFR_{(TFS)}$ for 'in quiet' and 'in noise' background conditions (a); FFR spectral amplitude obtained in response to the /o/ stimulus segment (b left), The response at the F_1 of the /o/ corresponding to 452 Hz (b right); FFR spectral amplitude in response to the /a/ stimulus segment (c left), The response at the F_1 of the /a/ corresponding to 678 Hz (c right). All estimations were made for both, in quiet (orange) and in noise (green) condition.

Spectral analysis revealed the impact of noise on the $FFR_{(TFS)}$, manifested as reduced peak amplitudes in the noise condition compared to the quiet background condition, evident for both the F_1 of /o/ and /a/ (Figure 13 b-c). Specifically, the spectral amplitude values and SNR in the /o/ vowel section were higher in the in quiet condition with medium-high effect sizes ($W=5843$, $P<0.001$, *Cohen's* $d=0.72$; $W=5575$, $P<0.001$, *Cohen's*

$d=0.64$, respectively), but no significant differences were observed for the /a/ vowel section between background conditions (Table 1).

To explore how the background noise affected the pitch tracking measures estimated, MR ANOVA was used to test the main effects of the background condition (in quiet vs. in noise) and pitch contour (flat vs. rising), as well as their interaction effect.

The ability to track pitch changes, as reflected in cross-correlation, is influenced by both background noise and pitch contour. Significant main effects for background noise and pitch contour were observed ($F=4.06$, $p=0.046$, $F=659.63$, $p<0.001$; respectively). Higher cross-correlation was observed in the noise condition compared to the quiet condition, but caution is advised in interpretation because the *p-value* is too close to the nominal threshold of statistical significance. Likewise, the cross-correlation was significant in the flat pitch contour in comparison with rising pitch contour, with high effect size (Table 2, Figure 14 a).

The interaction effect between the background condition and pitch contour was significant ($F=5.16$, $p=0.025$). Analyses of the simple effects showed significant differences between the flat and rising pitch contour, for both background conditions. It is interesting to highlight that the mean of cross-correlation values for the rising pitch contour remained stable in quiet and in noise conditions, but not in the /a/ flat pitch contour (Table 3).

Table 2. Summary statistics and Repeated Measured ANOVA between the background conditions and the pitch tracking measures.

Measure	Median	Mean	95% interval	Confidence	Normality test p - value	MR-ANOVA $F_{(1)}$ (p -value)	Effect size η^2_G
Pitch Tracking Measures							
Cross-correlation (<i>Pearson's r</i>)							
Main effect background condition							
In quiet	0.123	0.125	0.120	0.129	0.774	4.055 (0.046)	0.007
In noise	0.130	0.130	0.126	0.134	0.157		
Main effect pitch contour condition							
/a/ flat	0.175	0.172	0.167	0.178	0.652	659.6 (<0.001)	0.661
/a/ rising	0.083	0.083	0.079	0.086	0.988		
Interaction effect background condition*pitch contour						5.157 (0.025)	0.008
Pitch strength (r)							
Main effect background condition							
In quiet	0.617	0.604	0.582	0.626	0.001*	12.901 (<0.001)	0.027
In noise	0.549	0.563	0.539	0.586	0.012*		
Main effect pitch contour condition							
/a/ flat	0.581	0.583	0.563	0.603	0.539	1.472	-
/a/ rising	0.579	0.583	0.564	0.603	0.401	(0.227)	
Interaction effect background condition*pitch contour						0.493 (0.484)	-
Pitch error (Hz)							
Main effect background condition							
In quiet	5.2	6.22	5.54	6.90	0.002*	75.889 (<0.001)	0.186
In noise	9.24	9.89	9.17	10.6	0.025*		
Main effect pitch contour condition							
/a/ flat	7.62	8.06	7.48	8.63	0.621	0.022	-
/a/ rising	7.60	8.06	7.50	8.62	0.848	(0.884)	
Interaction effect background condition*pitch contour						1.170 (0.282)	-

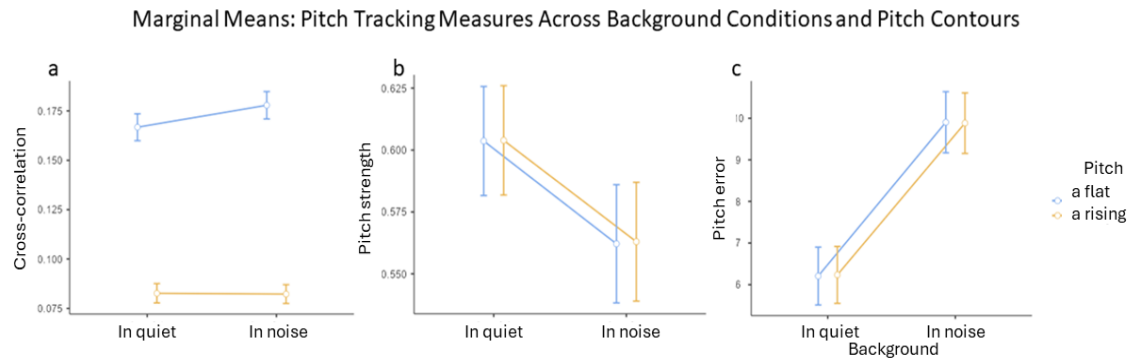


Figure 14. Marginal means in all pitch tracking measures estimated for both background conditions and pitch contours. The mean cross-correlation estimated in quiet and in noise for /a/ flat and /a/ rising pitch contour (a). The mean estimated in both background conditions for /a/ flat and /a/ rising pitch contour (b). Marginal mean pitch strength estimated in quiet and in noise for both pitch contours (c).

Regarding pitch strength and pitch error, main effects were observed only in the background condition ($F=12.90$, $p<0.001$; $\eta^2_G=0.027$; $F=75.89$, $p<0.001$, $\eta^2_G=0.186$, respectively) but not for the pitch contour. Post-hoc tests (with Bonferroni's correction) revealed higher pitch strength and lower pitch error in the quiet condition than in the noise condition (Figure 14 b-c), with small to moderate effect size, respectively. There were no interaction effects between the background condition and pitch contour in terms of both pitch strength and pitch error (Table 2).

Table 3. Simple effect between background condition and pitch contour on the cross-correlation results.

Simple Effect (Interaction)		Mean	S.e	Mean difference	$t_{(115)}$ <i>p-value</i>
In quiet	/a/ flat	0.167	0.003	0.084	20.440
	/a/ rising	0.083	0.002		<0.001^{*1}
In noise	/a/ flat	0.178	0.003	0.096	21.138
	/a/ rising	0.082	0.002		<0.001^{*1}

¹*p-value* with Bonferroni correction

FFR and Speech in noise perception

With the objective of evaluating the relationship between perceptual performance in a speech in noise test and the FFR parameter, Spearman correlations were estimated between the overall score of the SPIN test, as well as the High Predictability (HP) and Low Predictability (LP) subtests, across all of the FFR measures calculated. The results are presented in Table 4.

Significant correlations were observed between the HP subtest of the SPIN test and the F_1 spectral amplitude in the /a/ section of the stimulus, both in quiet and in noise background conditions ($\rho=0.213$, $p\text{-value}=0.022$; $\rho=0.189$, $p\text{-value}=0.042$, respectively). However, after applying the Bonferroni correction, these correlations did not reach statistical significance. For all other comparisons, no significant correlations were observed between our FFR measures and performance in the SPIN test, either in the overall score or in the subtests (Table 4).

Table 4: Spearman correlation between SPIN test and the FFR measures

FFR parameter		Speech in noise test (SPIN)		
		HP sub-test	LP sub-test	Overall
<i>Rho correlation_(df=114) (p-value)</i>				
Neural lag	In quiet	-0.035 (0.713)	0.017 (0.854)	-0.018 (0.844)
	In noise	0.034 (0.719)	-0.097 (0.301)	-0.100 (0.285)
FFR_(ENV)				
F ₀	In quiet	0.046 (0.623)	0.070 (0.455)	0.091 (0.332)
	In noise	0.097 (0.302)	0.017 (0.855)	0.066(0.484)
F ₀ SNR	In quiet	0.092 (0.325)	0.127 (0.176)	0.155 (0.096)
	In noise	0.142 (0.128)	-0.003 (0.975)	0.069 (0.461)
FFR_(TFS)				
F1/o/	In quiet	-0.059 (0.531)	-0.086 (0.360)	-0.098 (0.295)
	In noise	-0.110 (0.242)	-0.029 (0.758)	-0.089 (0.340)
F1/o/ SNR	In quiet	-0.027 (0.772)	0.127 (0.174)	0.088 (0.350)
	in noise	-0.074 (0.432)	0.002 (0.983)	-0.043 (0.622)
F1/a/	In quiet	0.213 (0.022) *	0.039 (0.680)	0.166 (0.076)
	In noise	0.189 (0.042) *	-0.033 (0.724)	0.110 (0.239)
F1/a/ SNR	In quiet	0.107 (0.251)	0.012 (0.895)	0.094 (0.315)
	in noise	0.124 (0.185)	0.012 (0.897)	0.109 (0.243)
Pitch Tracking measures				
Cross-correlation /a/ flat	In quiet	0.015 (0.870)	0.156 (0.095)	0.168 (0.072)
	in noise	0.105 (0.264)	-0.040 (0.673)	0.029 (0.760)
Cross-correlation /a/ rising	In quiet	-0.070 (0.454)	0.053 (0.576)	-0.027 (0.777)
	in noise	0.048 (0.612)	0.009 (0.927)	0.021 (0.822)
Pitch Strength /a/ flat	In quiet	0.057 (0.546)	0.103 (0.272)	0.134 (0.152)
	in noise	0.108 (0.251)	-0.005 (0.955)	0.054 (0.564)
Pitch Strength /a/ rising	In quiet	0.049 (0.602)	0.105 (0.263)	0.131 (0.161)
	in noise	0.117 (0.210)	-0.009 (0.927)	0.059 (0.527)
Pitch error /a/ flat	In quiet	0.079 (0.398)	-0.100 (0.285)	-0.076 (0.417)
	in noise	-0.011 (0.910)	0.026 (0.779)	0.004 (0.965)
Pitch error /a/ rising	In quiet	0.085 (0.367)	-0.086 (0.357)	-0.063 (0.502)
	in noise	0.002 (0.983)	0.045 (0.630)	0.027 (0.770)

Finally, for exploratory purposes, the differences between each parameter in quiet and in noise were estimated and correlated with the SPIN scores. Non-significant correlations were observed in all of the comparisons.

4.2 STUDY II: FFR AND FUNCTIONAL IMPLICATIONS

To investigate the relationship between the FFR and reading skills, musical skills, and acoustical temporal perception, a two-step analysis was conducted. First, a Principal Component Analysis (PCA) was performed to reduce the dimensionality of the estimated FFR parameters while retaining as much of the original variability, capturing their essential information. Subsequently, Spearman correlations were computed to examine the association between the derived components and the selected behavioral measures.

Finally, Pearson's correlations were conducted to explore relationships among the behavioral measures themselves.

Principal component analysis (PCA)

Spearman's correlation analysis was conducted among all FFR parameters to identify and eliminate redundant variables. Cross-correlation, pitch strength, and pitch error in the /a/ rising pitch contour were removed due to redundancy ($\rho \geq 0.9$) (Annex 1). A Principal Component Analysis (PCA) was then performed on our electrophysiological measurements to reduce the number of FFR-based parameters analyzed while retaining most of the relevant information. The results are presented in Figure 15 and Table 5.

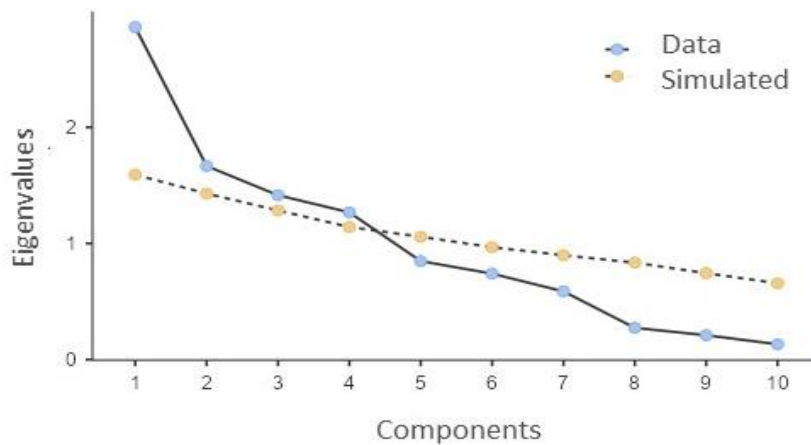


Figure 15. Principal component analysis, parallel analysis plot. The parallel analysis suggests the extraction of four components, which is the point where the eigenvalues from the simulated data (in orange) intersect with the eigenvalues obtained from the empirical data (in blue).

According with the parallel analysis, four eigenvalues exceeded the random data eigenvalues, indicating that we should retain four components (Figure 15).

Based on the component loading information (Table 5) and considering what each FFR parameter represents and measures, the components were constructed in the following manner:

1. The Pitch Magnitude component, which includes the spectral amplitude of the F_0 , SNR of this spectral amplitude, and the cross-correlation.
2. The Pitch Tracking component, which includes the neural lag, pitch strength, and pitch error.
3. First formant /o/ component, which includes the spectral amplitude and SNR of the first formant elicited by the /o/ part of the stimulus.

4. First formant /a/ component, which includes the spectral amplitude and SNR of the first formant for the /a/ stimulus section.

Table 5: Principal Component Analysis for the FFR parameters in quiet background condition, only on the flat pitch contour.

FFR parameter	Component loadings*				uniqueness
	1 25.8% ¹	2 17.4% ¹	3 15% ¹	4 14% ¹	
F_0 spectral amplitude	0.886				0.241
F_0 SNR	0.823		-0.331		0.196
F_1 /o/ spectral amplitude			0.932		0.117
F_1 /o/ SNR		0.492	0.506		0.492
F_1 /a/ spectral amplitude			0.328	0.852	0.131
F_1 /a/ SNR amplitude		0.302	-0.323	0.792	0.178
Neural Lag		0.761			0.347
Cross-correlation	0.619				0.602
Pitch strength	0.692	0.581			0.164
Pitch error	-0.479	-0.650			0.321

*Varimax rotation. ¹ % of the variance

FFR components and their functional relationship

Correlation analyses between the FFR Components and the different auditory skills revealed only significant results with reading abilities (Table 6). Specifically, the pitch magnitude and F_1 /o/ components were associated with the performance in the Word and Pseudo Word reading subtests of the PROLEC-SE-R battery, whereas the pitch tracking and F_1 /a/ components were not.

Table 6: Spearman's correlation between the FFR Components and the different auditory skills.

Behavioral measure	Components			
	Pitch magnitude	Pitch tracking	First formant /o/	First formant /a/
<i>Rho correlation_(df=114) (p-value)</i>				
Words (W) and Pseudowords (PW) Reading sub-test (Prolec-SE-R)				
W reading overall	-0.105 (0.260)	0.019 (0.841)	0.203 (0.029)	0.022 (0.816)
W reading accuracy	0.015 (0.872)	0.017 (0.853)	0.023 (0.806)	0.055 (0.560)
W reading time	0.102 (0.276)	-0.022 (0.814)	-0.205 (0.028)	-0.022 (0.811)
PW reading overall	-0.186 (0.046)	0.120 (0.201)	0.221 (0.017)	-0.035 (0.707)
PW reading accuracy	-0.164 (0.078)	-0.070 (0.454)	0.156 (0.095)	-0.058 (0.540)
PW reading time	0.168 (0.072)	-0.142 (0.129)	-0.214 (0.021)	0.039 (0.678)
Music Perception Skills (PROMS)				
Tonal domain	0.111 (0.237)	-0.116 (0.215)	-0.179 (0.055)	-0.065 (0.485)
Qualitative domain	0.178 (0.056)	-0.071 (0.446)	-0.172 (0.065)	-0.107 (0.255)
Temporal domain	0.048 (0.612)	-0.046 (0.627)	-0.093 (0.319)	-0.122 (0.191)
Total PROMS	0.115 (0.219)	-0.073 (.439)	-0.178 (0.056)	-0.123 (0.188)
AM discrimination task				
Total AM	0.030 (0.752)	0.100 (0.284)	0.005 (0.955)	0.114 (0.223)

Significant correlations were observed between the pitch magnitude component and the Pseudowords reading subtest. Specifically, higher spectral amplitude and SNR of the F_0 were associated with lower performance in the Pseudoword reading overall score ($\rho = -0.186$, $p\text{-value} = 0.046$). In contrast, higher scores on the F_1 /o/ vowel component were associated with better performance in both the Words and Pseudowords reading overall scores ($\rho = 0.203$, $p\text{-value} = 0.029$ and $\rho = 0.221$, $p\text{-value} = 0.017$, respectively). This overall better performance is explained by a faster execution of the tasks and not by an increased accuracy, since a higher score in this component was also associated with less time spent on reading tasks in both subtests ($\rho = -0.205$, $p\text{-value} = 0.028$ and $\rho = -0.214$, $p\text{-value} = 0.021$, respectively for the Words and Pseudowords reading time), indicating a potential relationship between component performance and reading efficiency (Figure 16). Non-significant correlations were observed for all other performed contrasts.

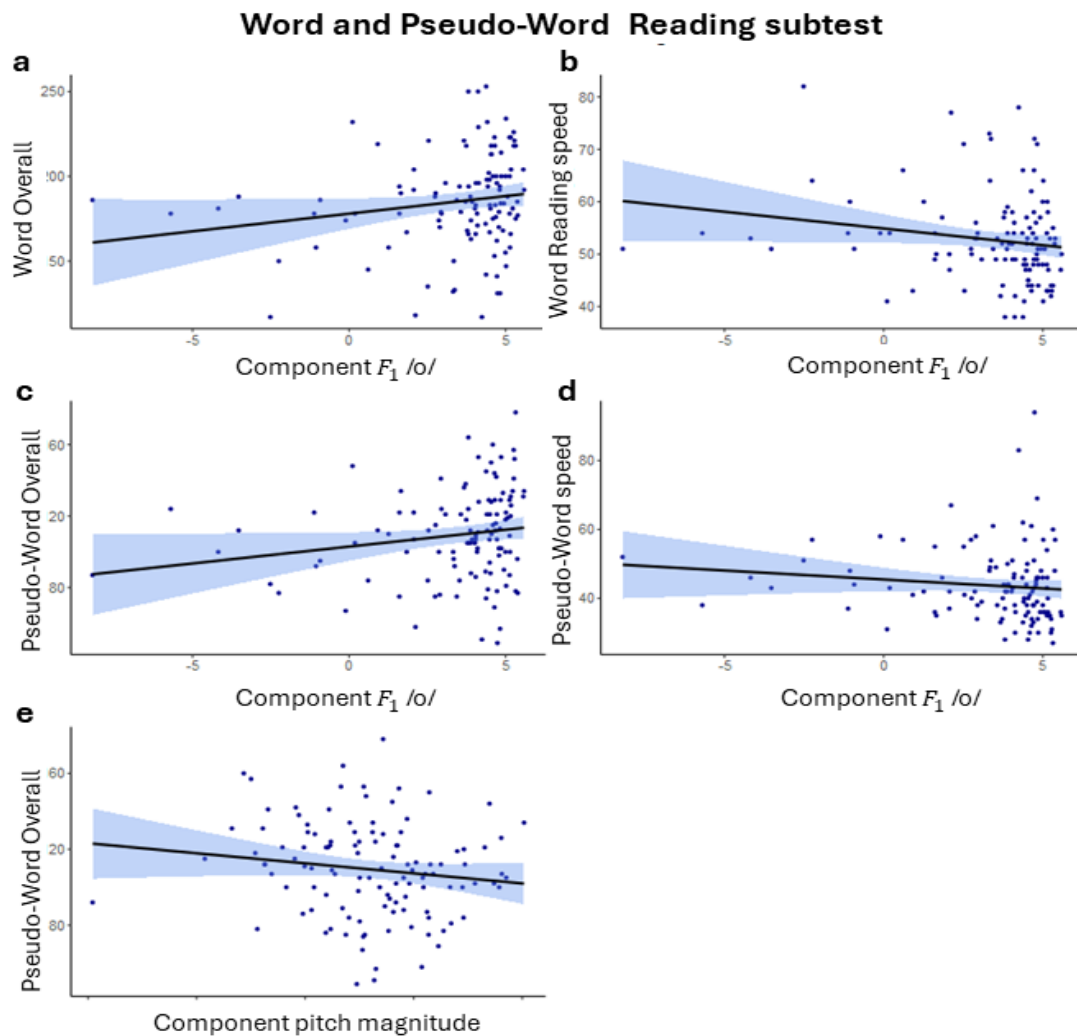


Figure 16. Scatterplot between FFR components and Word and Pseudoword reading subtest. Relationship between Component F_1 /o/ and the overall score of the Word reading subtest (a), Word reading speed (b), Pseudo-Word reading overall score (c) and Pseudo-Word speed (d). Relationship between the component of pitch magnitude and Pseudo-Word overall score (e).

Relationship between the behavioral measures

As discussed in the "Auditory Perception Abilities" section (Chapter I, section 1.2), a substantial body of evidence supports the idea that temporal processing and perception are crucial for musical perception, reading skills, and speech in noise perception.

Similarly, there is evidence suggesting an association between SIN perception and reading skills, hypothesizing a potential shared neural basis. To explore these relationships within our sample, Pearson correlations were calculated between our behavioral measures.

Significant correlations were observed between the Word Reading subtest and Speech in Noise (SIN) perception ($r=0.198$, $p < 0.05$) (Table 7). Specifically, the SPIN test was correlated with the time of word reading performance, indicating that higher speech in noise performance is associated with less reading time ($r=-0.196$, $p = 0.035$). When exploring whether the type of predictability was related to the word reading process, only the low predictability subtest correlated with both overall Word reading and reading time, indicating that higher performance in speech in noise perception for sentences with low predictability is associated with lower (i.e. faster) word reading time ($r=-0.192$, $p = 0.038$) (Figure 17).

No statistically significant correlations were identified between performance on the Amplitude Modulated task and the remaining behavioral measures, nor were any significant correlations observed among the various assessed listening skills (Table 7).

Table 7: Pearson's correlation between behavioral measures

	AM	W-Reading	PW-Reading	SPIN
	<i>Pearson's correlation_(df=115) (p-value)</i>			
AM	-			
W-Reading	-0.046 (0.626)	-		
PW-Reading	-0.058 (0.537)	0.688 (<0.001)	-	
SPIN	0.154 (0.097)	0.198 (0.032)	-0.008 (0.928)	-
PROMS Overall	-0.060 (0.523)	-0.086 (0.355)	-0.113 (0.226)	0.134 (0.150)
Tonal Domain/PROMS	-0.053 (0.570)	-0.088 (0.348)	-0.153 (0.100)	0.064 (0.492)
Qualitative Domain/PROMS	0.006 (0.947)	-0.064 (0.492)	-0.114 (0.222)	0.084 (0.369)
Temporal Domain/PROMS	-0.083 (0.376)	-0.078 (0.403)	-0.074 (0.429)	0.159 (0.087)

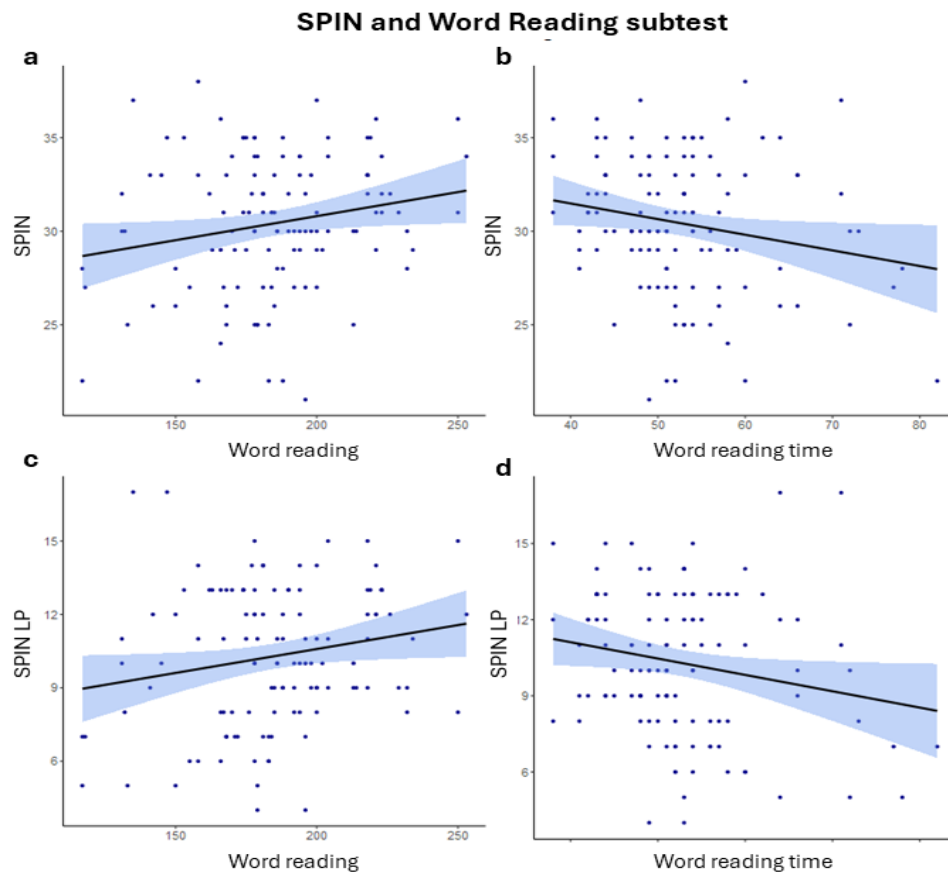


Figure 17. Scatterplot between the performance in the SIN test and Word reading subtest. Relationship between the overall score in the SPIN test and the overall performance in Word reading subtest (a), and word reading time (b). Relationship between the Low Predictability sentences of the SPIN test and the overall performance in Word reading subtest (c), and word reading time (d).

4.3 STUDY III: GENETIC OF THE SOUND PROCESSING

SLC6A4 association analyses

A total sample of one-hundred and twelve individuals from the GEPLM study fulfilled all QC criteria and had valid genotype and phenotype data to be included in these analyses. Characteristics of the polymorphisms analyzed in the *SLC6A4* promoter are summarized in Table 8. The three markers extracted from the SNP array showed relatively low frequencies (MAF<0.10) and will likely retain limited power to detect genetic effects. All markers fitted HWE assumptions ($p>0.7$).

Although these markers are all in a relatively small (<4 Kb) chromosomal area, they did not show substantial amounts of linkage disequilibrium (LD). Most pairs of markers showed no relevant signs of LD ($R^2 < 0.1$), with the only exception of markers rs2020933 and rs73263896, that showed a substantial LD between them ($R^2=0.60$), and markers rs25532 and 5-HTTLPR with a limited amount of LD ($R^2=0.11$).

Table 8. Descriptive statistics of the polymorphisms analyzed in the *SLC6A4* gene.

Marker	ChrPos	N	A1	A2	MAF	Genots	HWE-p
rs2020933	30234737	105	A	T	0.062	0/13/92	1
rs25532	30237152	101	T	C	0.094	1/17/83	1
5-HTTLPR/rs25531	30237328	111	High	Low	0.478	24/58/29	0.705
5-HTTLPR	30237328	111	S (Short)	L (Long)	0.473	26/53/32	0.704
rs73263896	30238408	105	C	T	0.038	0/8/97	1

The ChrPos indicates the Chromosomal position in chromosome 17 according to assembly GRCh38.p14 (since 5-HTTLPR varies in size depending on the number of tandem repeats, relative position for the 5-HTTLPR marker has been established based on rs25531). N is the number of individuals with valid genotypes. A1/A2 are the alleles sorted by frequency (A1 minor / A2 major). MAF, indicates the minor allele (A1)

frequency. Genots is the genotype absolute frequencies (A1A1/A1A2/A2A2). HWE-p is the probability of the Hardy-Weinberg equilibrium exact test.

We first tried to replicate the previous association between variants in the SLC6A4 promoter and SNR at the fundamental frequency. Thus, we observed two markers, SNPs rs2020933 and rs73263896, that were nominally associated with differences in their SNR values (Table 9). Under an additive genetic model, minor alleles at both SNPs were significantly associated with a decreased SNR, with a mean difference of 3 units (almost a SD) for each copy of the minor allele that an individual carries.

Table 9. Association results between markers in the promoter region of the SLC6A4 gene and FFR parameters.

Phenotype	Marker	A1	Beta	95% Confidence interval		<i>P-value</i>
				lower	upper	
SNR	rs2020933	A	-2.794	-5.098	-0.4905	0.01947
	rs25532	T	0.6542	-1.171	2.479	0.4841
	5-HTTLPR/rs25531	High	-0.2058	-1.319	0.9072	0.7179
	rs73263896	C	-3.256	-5.985	-0.5266	0.0215
Pitch magnitude	rs2020933	A	-2.493	-4.552	-0.4339	0.01968
	rs25532	T	0.5826	-1.048	2.213	0.4856
	5-HTTLPR/rs25531	High	-0.1777	-1.173	0.8172	0.7271
	rs73263896	C	-2.894	-5.334	-0.4533	0.02228
Pitch tracking	rs2020933	A	-1.338	-3.67	0.9932	0.2635
	rs25532	T	0.227	-1.556	2.01	0.8035
	5-HTTLPR/rs25531	High	0.1878	-0.9206	1.296	0.7406
	rs73263896	C	-1.712	-4.469	1.044	0.2265
F1 /o/	rs2020933	A	-0.02143	-1.296	1.254	0.9738
	rs25532	T	-0.04289	-1.038	0.952	0.9328
	5-HTTLPR/rs25531	High	0.1698	-0.4358	0.7754	0.5839
	rs73263896	C	0.01394	-1.495	1.523	0.9856
F1 /a/	rs2020933	A	-0.9416	-2.85	0.9671	0.3361
	rs25532	T	0.3143	-1.18	1.809	0.6812
	5-HTTLPR/rs25531	High	-0.536	-1.442	0.3702	0.2493
	rs73263896	C	-2.083	-4.314	0.1482	0.07045

Linear regression results included age, sex, and the three first PCs to avoid population stratification as covariates (see Chapter III: Methodology, section 3.8). A1 is the minor allele. Analyses were conducted under an additive model.

Additionally, we also run association tests between variants in the *SLC6A4* promoter and the four FFR component calculated and defined in Study II of this thesis (Table 5). Based on our results, associations were only observed between SNPs rs2020933 and rs73263896 and the component pitch magnitude, which encompasses the spectral amplitude and their corresponding SNR in response to the F_0 . These two SNPs exhibited associations similar to those observed for SNR, demonstrating a negative effect of their minor alleles on the outcome, with comparable p-values.

Although these associations were consistent across related FFR phenotypes, they showed limited robustness. To limit the impact of multiple comparisons on Type I error, an FDR adjustment was run, and all the associations detected still retrieved nominally significant *p-values* ($p=0.043-0.045$). However, none of them withheld a permutation test after 10,000 iterations ($p=0.070-0.079$).

Genome-wide association analyses

For genome-wide analyses, one hundred and one participants were retained with complete genomic and FFR information after stringent QC criteria were applied. Before delving into the association results, we initially visually examined the quantile-quantile (Q-Q) plots of the genome-wide association tests. No significant deviations from the expected p-values were observed (Figure 18).

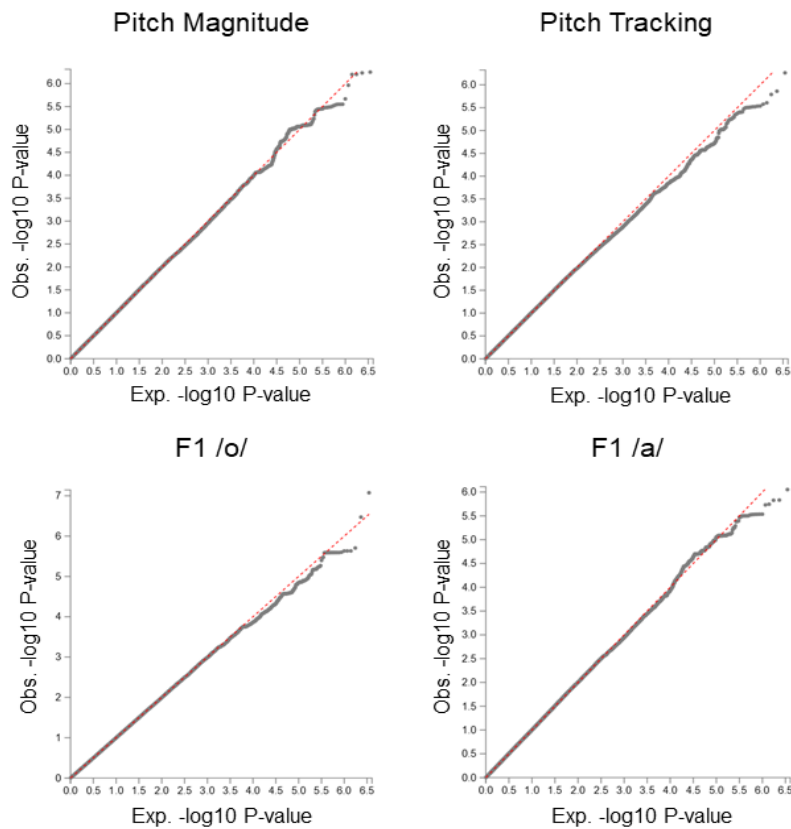


Figure 18. Quantile-quantile (Q-Q) plots of the GWAS summary statistics for the four phenotypes analyzed. Plots extracted from FUMA.

Regarding GWAS results, we did not find genome-wide significant results at the conventional significance threshold used in this kind of studies ($5 \cdot 10^{-8}$). However, several interesting SNPs were associated with the different FFR components at suggestive probability values.

For the pitch magnitude component, the top associated locus SNP was a region on chromosome 8, with the lead SNP rs4871957 ($p=5.61 \cdot 10^{-7}$; MAF=0.164) (Figure 18). This genetic variant, and several others in tight LD, map to the *EBF2* gene, which encodes for a transcription factor that shows a maximal expression in nerval tissue. The following

two top associated loci correspond to regions on chromosomes 2 and 16 ($p=2.16 \cdot 10^{-6}$ and $p=2.82 \cdot 10^{-6}$, for the respective lead SNPs), in regions that harbor the *RN7SL423P* and the *RBFOX1* genes. Many genetic variants in both genes have been associated with different cognitive, behavioral and neuroanatomic assessments, ranging from intelligence and educational attainment to substance use disorder, neuroticism and hippocampal volume in many published GWAS (for a thorough list, check GWAS Catalog entries <https://www.ebi.ac.uk/gwas/genes/RN7SL423P> and <https://www.ebi.ac.uk/gwas/genes/RBFOX1>). Moreover, the *RBFOX1* gene encodes for an RNA-binding protein that regulates tissue-specific alternative splicing and is highly expressed in different brain tissues.

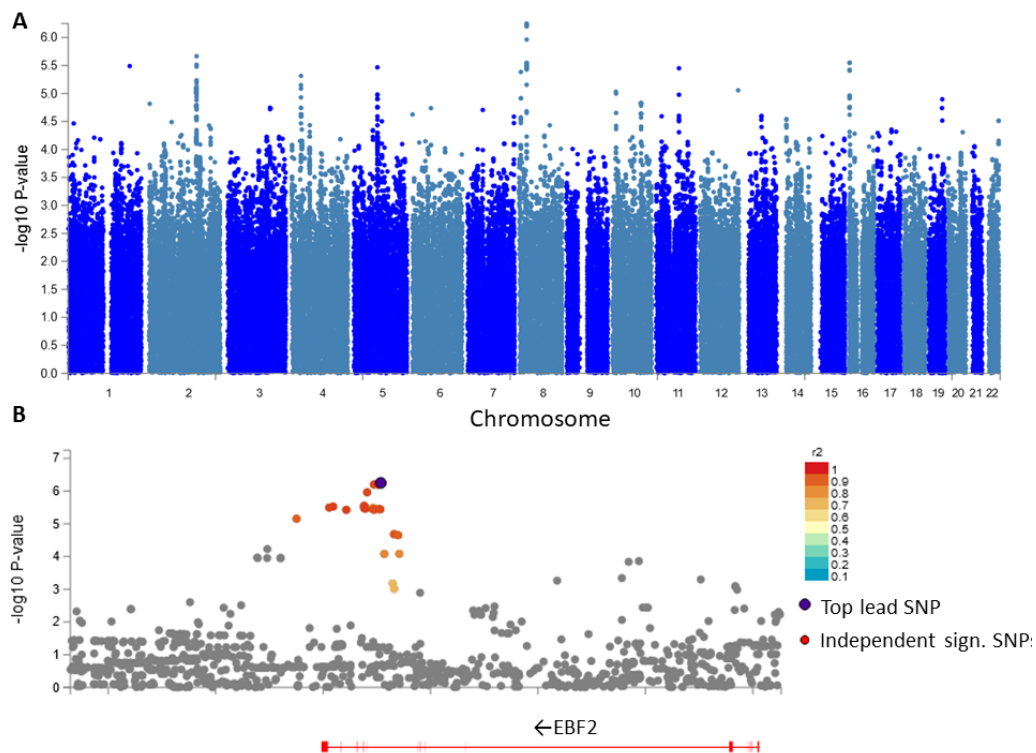


Figure 19. GWAS results for the Pitch Magnitude component of the FFR. Manhattan plot of the genome-wide analyses (A). Regional plot of the top associated region on chromosome 8 with the relative location of mapped genes at the bottom (B).

The top associated locus for the pitch tracking component was a region on chromosome 6, with the lead SNP rs6922782 ($p=5.45 \cdot 10^{-7}$; MAF=0.332) (Figure 20). This region maps to the *PHACTR1* gene, which is involved in tubule formation and is highly expressed in different brain tissues. The following two top associated loci correspond to regions on chromosomes 17 and 7 ($p=1.38 \cdot 10^{-6}$ and $p=1.62 \cdot 10^{-6}$, for the respective lead SNPs), in regions that harbor the *TBX4* and the *PMS2* genes.

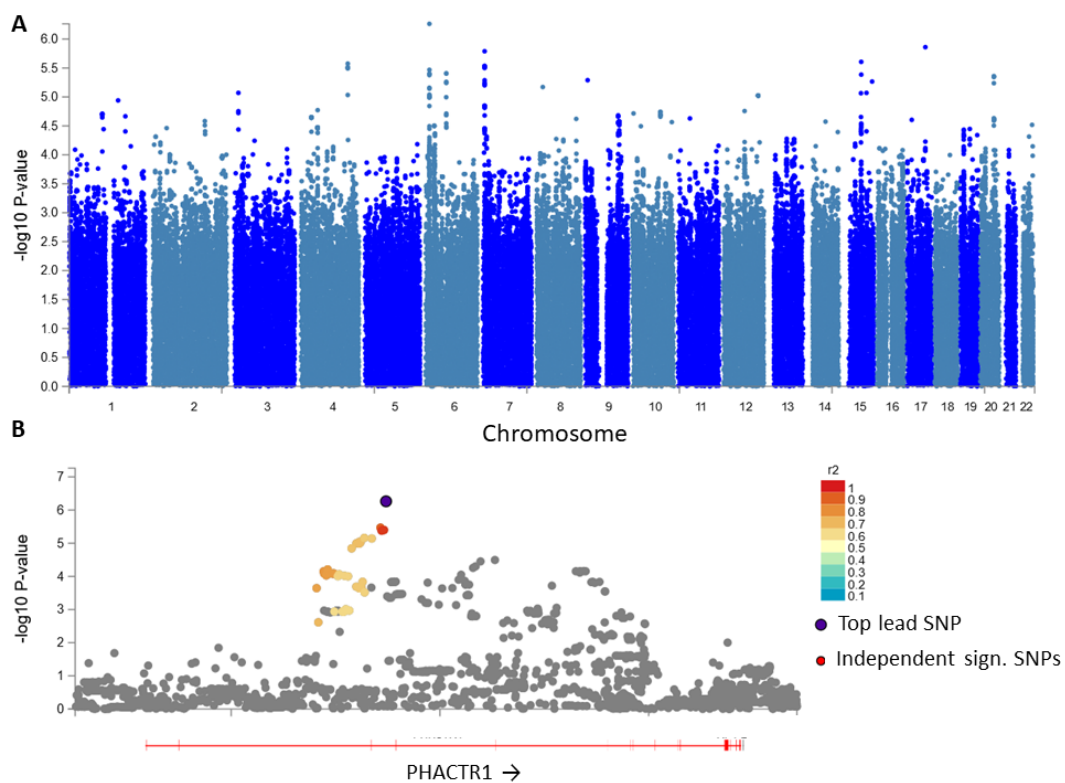


Figure 20. GWAS results for the Pitch Tracking component of the FFR. Manhattan plot of the genome-wide analyses (A). Regional plot of the top associated region on chromosome 6 with the relative location of mapped genes at the bottom (B).

The F1 /o/ component showed the lowest p-value of all association tests conducted for the FFR phenotypes. Specifically, the lead SNP rs12611119 on chromosome 19 showed a p-value close to the conventional threshold for genome-wide significance ($p=8.40 \cdot 10^{-7}$).

⁸; MAF=0.374) (Figure 21). This genetic variant maps to the *SULT2B1* gene, which encodes for a sulfotransferase that catalyzes the sulfate conjugation of many hormones, neurotransmitters, and drugs. The following top associated loci correspond to regions on chromosomes 15, which overlaps with the coding region of the *LIPC* gene, and 12, without any protein-coding gene nearby ($p=1.97 \cdot 10^{-6}$ and $p=2.33 \cdot 10^{-6}$, for the respective lead SNPs).

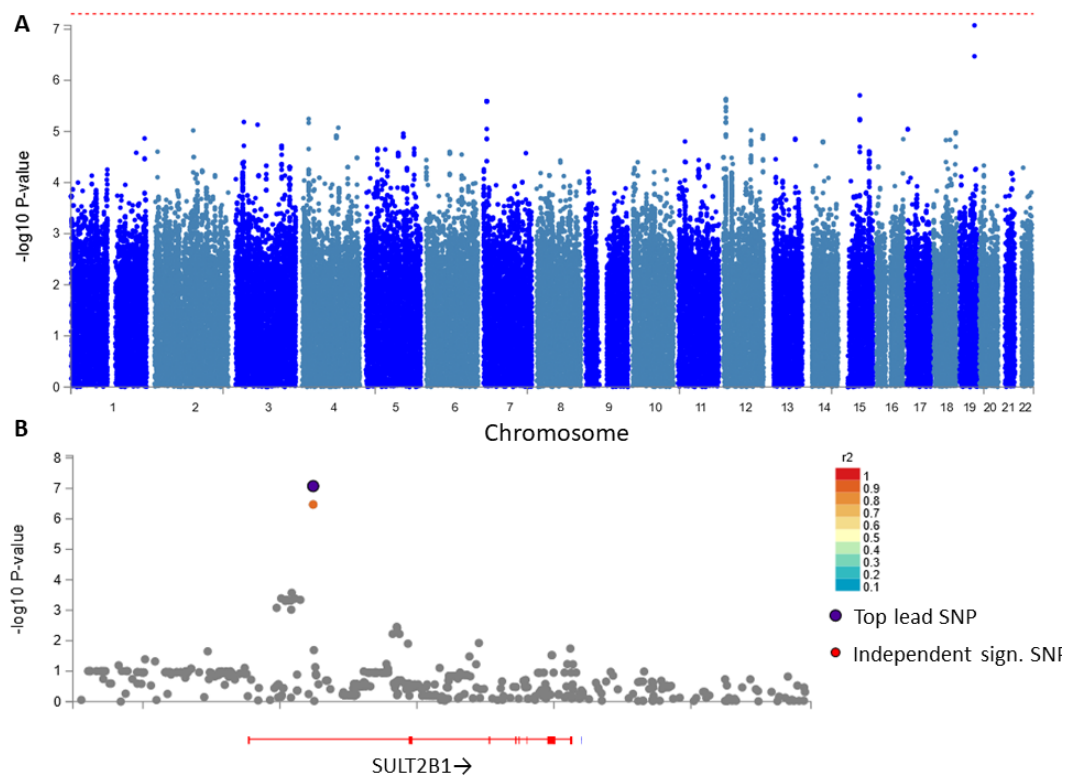


Figure 21. GWAS results for the F1 /o/ component of the FFR. Manhattan plot of the genome-wide analyses. Red dotted line marks the genome-wide significance threshold ($5 \cdot 10^{-8}$) (A). Regional plot of the top associated region on chromosome 19 with the relative location of mapped genes at the bottom (B).

Regarding the F1 /a/ component, the lead SNP of the top associated locus was rs11846871 on chromosome 14 ($p=8.85 \cdot 10^{-7}$; MAF=0.477) (Figure 22). This marker, and

others in tight LD, lie in the vicinity of the non-coding RNAs *RP11-789A21.1* and *RNU3P3*. Additionally, SNPs in regions on chromosomes 14 and 7 were the following two top associated loci ($p=1.80 \cdot 10^{-6}$ and $p=1.86 \cdot 10^{-6}$, for the respective lead SNPs). The former is located in a region that harbor the *PSMC1* and *NRDE2* genes, while the latter has no known gene in its vicinity.

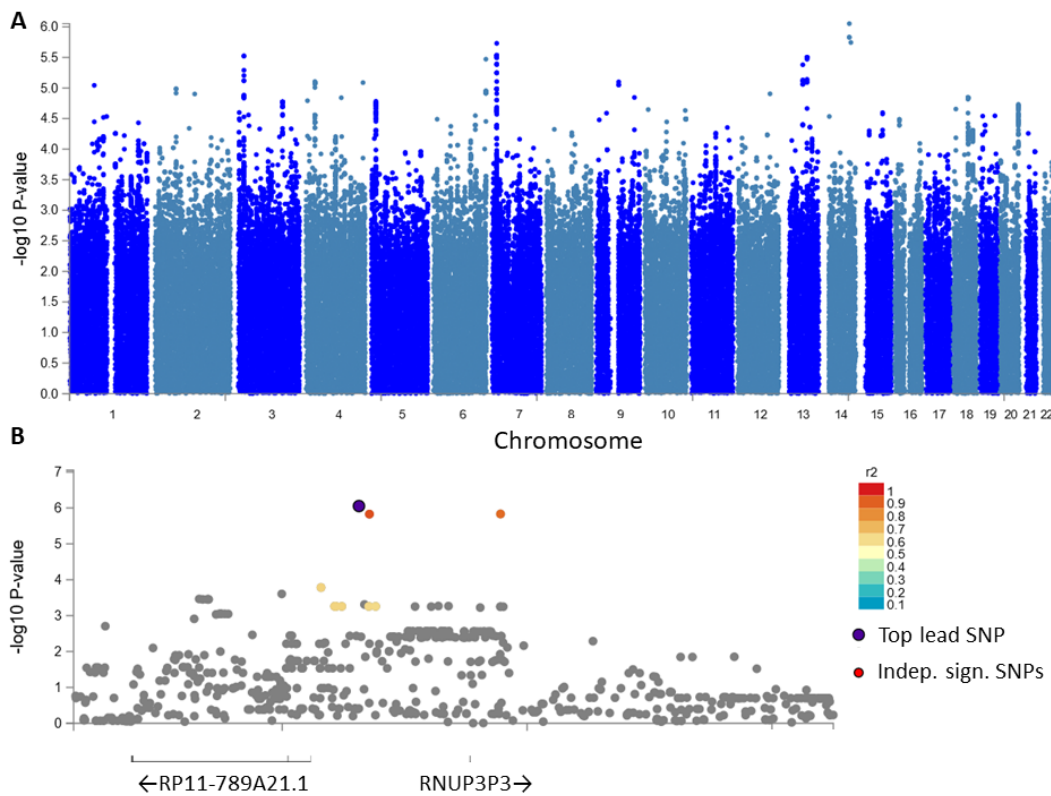


Figure 22. GWAS results for the F1 /a/ component of the FFR. Manhattan plot of the genome-wide analyses (A). Regional plot of the top associated region on chromosome 14 with the relative location of mapped genes at the bottom (B).

Post-GWAS analyses

While the previous group of genome-wide analyses tried link variation at individual markers and differences in FFR responses, the sets of analyses in this section attempt at

identifying genetic variation at the gene level involved in the observed differences, relate it to gene expression levels in different tissues, or reveal genetic shared etiology with other phenotypes.

We first conducted gene-based tests to identify genes potentially associated with our traits of interest. These analyses test the association of the overall genetic variation across the gene region, instead of analyzing SNPs at the individual level. Similarly to GWAS analyses, these tests did not reveal associations at the gene-wide significance level ($p < 2.68 \cdot 10^{-6}$) but identified several genes that are potential candidates to be involved in the observed differences (top five genes associated for every FFR component are labeled in Figure 23).

Among the different genes identified, it is worth highlighting some potentially interesting candidates. The *OPALIN* gene on chromosome 10 is the top associated gene with the pitch magnitude component of the FFR ($p = 2.74 \cdot 10^{-5}$) and it is also the gene with the lowest probability for all gene-based tests and all FFR components. The protein encoded by this gene, the oligodendrocytic myelin paranodal and inner loop protein, is involved in the differentiation of oligodendrocytes and its expression is restricted to the CNS, where it is broadly expressed in the spinal cord and all cortical and subcortical tissues (see, for example, GTEx expression data <https://www.gtexportal.org/home/gene/OPALIN>).

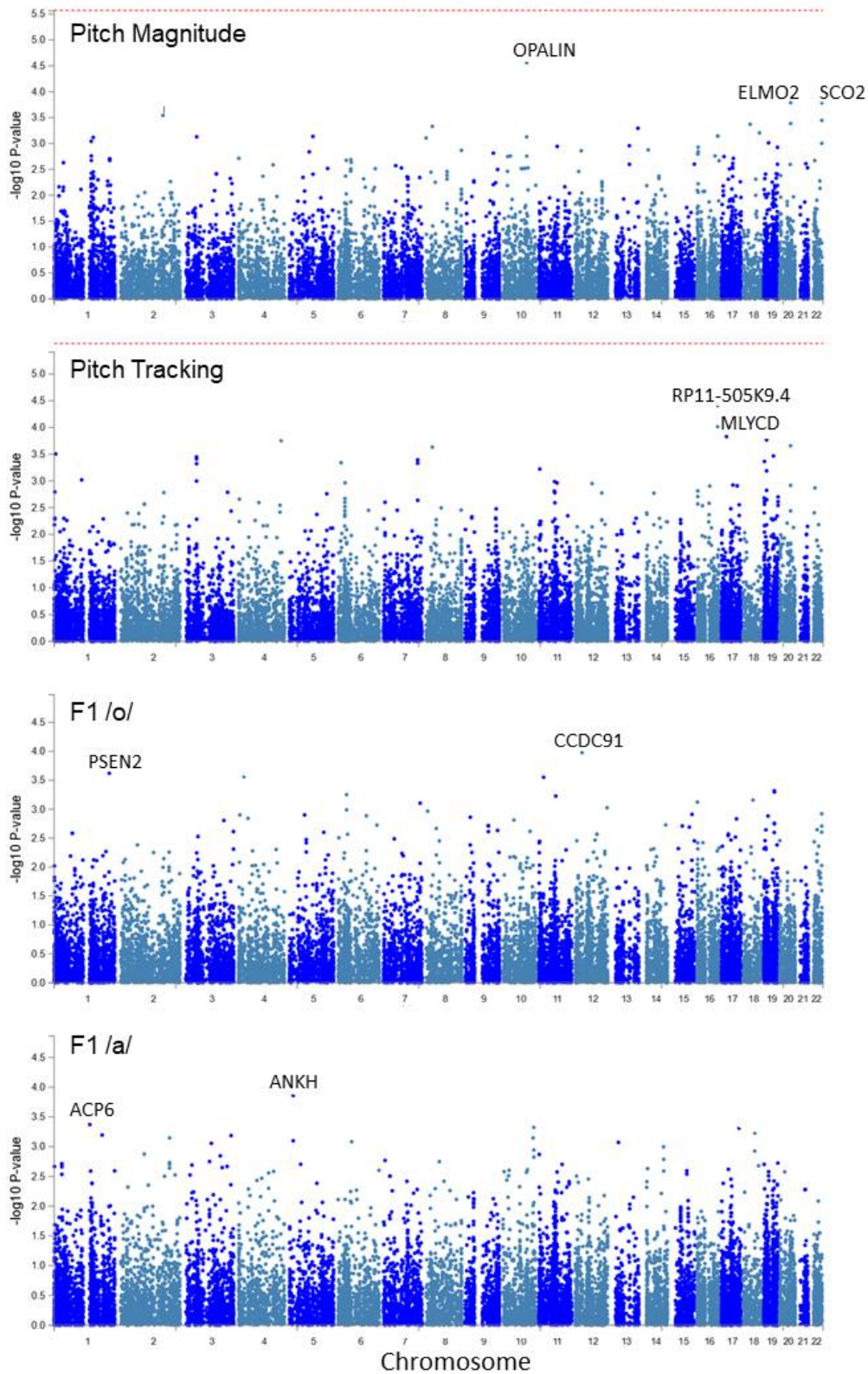


Figure 23. Manhattan plots of the gene-based tests of association with the four FFR components. Red dotted lines mark the gene-wide significance threshold ($2.68 \cdot 10^{-6}$).

It is also worth highlighting that the pitch magnitude and the pitch tracking components share some of the top associated genes. Specifically, the two top associated genes with the pitch tracking component, *RP11-505K9.4* ($p=3.90 \cdot 10^{-5}$) and *MLYCD* ($p=9.64 \cdot 10^{-5}$), were also among the top genes associated with the pitch magnitude component ($p=7.14 \cdot 10^{-4}$ and $p=7.23 \cdot 10^{-4}$, respectively). Moreover, the *EBF2* gene was also among the genes with the lowest probability values for both components ($p=4.68 \cdot 10^{-4}$ and $p=2.34 \cdot 10^{-4}$, for the pitch magnitude and the pitch tracking components) in addition to the findings from the GWAS analyses in the previous section.

Tissue Expression analyses were conducted to identify tissues whose expression patterns are enriched for genes at the associated loci. These analyses were conducted in MAGMA using the GTEx v8 reference data on gene expression patterns across 53 specific tissue types.

Results did not reveal tissues significantly enriched for genes identified in our genetic dataset at the corrected p-value of 9.43×10^{-4} (Figure 24). Nevertheless, both the pitch magnitude and the pitch tracking components of the FFR showed nominally significant enrichments of expression of genes expressed in brain tissues. Specifically, the pitch magnitude component showed associations with expression patterns in the two samples from the cerebellum in the GTEx dataset ($p=0.0076$ and $p=0.020$). In turn, genetic variants associated to the pitch tracking component are enriched in the expression patterns of the spinal cord samples ($p=0.0067$) and show some biased expression in the mesencephalic region of the substantia nigra, although it did not reach nominal significance ($p=0.077$).

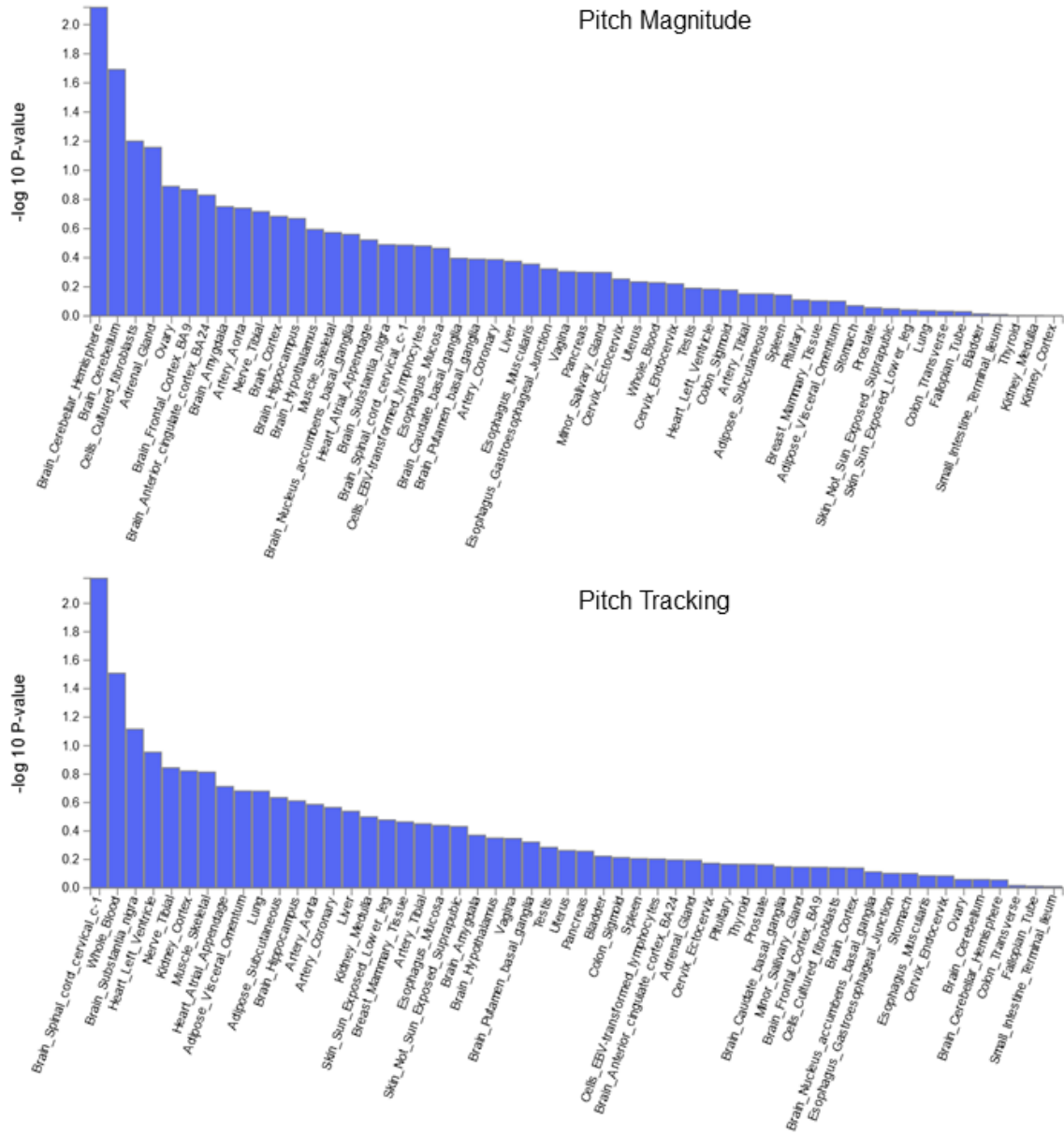


Figure 24. MAGMA tissue expression analysis results. Gene-property analyses were performed on the full distribution of SNP p-values for gene expression based on GTEx v8 53 tissue types.

Polygenic Scores (PGS)

For each individual, a polygenic risk score (PGS) was calculated based on the cumulative effect of multiple genetic variants associated with the traits of interest. In our study, we selected traits/diseases including intelligence, ADHD, autism spectrum disorder, executive function, and white matter volume from the PGS catalog. Subsequently, we correlated each PGS with the components calculated in Study II and with the behavioral measures used in the present thesis. The results are shown in Table 10.

For all contrasts performed, non-significant correlations were observed between the PGS calculated with our FFR components and reading skills, musical perception skills, AM discrimination, and speech-in-noise perception, with the exception of the speech-in-noise perception low predictability subtest and PGS002746, corresponding to the trait attention deficit hyperactivity disorder (ADHD) ($\rho = -0.194$, $p\text{-value} = 0.044$). This negative correlation indicates that a higher polygenic risk score is associated with lower speech-in-noise perception for low predictability sentences.

Table 10. Spearman correlation between the PGS for the traits selected and the FFR components and the behavioral auditory skills.

Phenotypes	PGS							
	PGS 000327 ¹	PGS 002790 ¹	PGS 003724 ²	PGS 002746 ³	PGS 003723 ⁴	PGS 001541 ⁵	PGS 001640 ⁵	
<i>Rho's Spearman</i> (df=105) (<i>p-value</i>)								
FFR components								
Pitch Magnitude	-0.126 (0.195)	-0.126 (0.194)	-0.014 (0.888)	0.007 (0.940)	0.071 (0.465)	-0.073 (0.455)	0.041 (0.675)	
Pitch tracking	0.098 (0.313)	0.011 (0.912)	0.082 (0.403)	0.071 (0.469)	-0.083 (0.397)	-0.085 (0.383)	-0.100 (0.304)	
F1 /o/	0.005 (0.963)	-0.056 (0.566)	0.003 (0.976)	-0.165 (0.090)	-0.067 (0.493)	-0.084 (0.388)	-0.055 (0.575)	
F1 /a/	0.021 (0.831)	0.108 (0.267)	-0.029 (0.762)	0.100 (0.305)	0.0004 (0.997)	-0.092 (0.348)	-0.030 (0.756)	
<i>Rho's Spearman</i> (df=106) (<i>p-value</i>)								
Behavioral Measures								
SIN test	Low	-0.099 (0.307)	-0.016 (0.870)	-0.071 (0.463)	-0.194 (0.044)	-0.026 (0.786)	-0.085 (0.381)	-0.034 (0.726)
	High	0.076 (0.436)	0.052 (0.596)	-0.056 (0.561)	0.001 (0.989)	0.049 (0.618)	-0.098 (0.311)	-0.075 (0.439)
	Overall	-0.019 (0.839)	0.023 (0.813)	-0.137 (0.157)	-0.126 (0.194)	-0.012 (0.898)	-0.130 (0.178)	-0.054 (0.576)
W-reading	Accuracy	-0.012 (0.895)	-0.003 (0.977)	0.071 (0.466)	0.028 (0.775)	0.052 (0.595)	0.076 (0.431)	0.018 (0.850)
	Time	-0.088 (0.364)	-0.132 (0.174)	-0.104 (0.285)	0.126 (0.195)	0.047 (0.626)	-0.019 (0.838)	-0.071 (0.465)
	Overall	0.093 (0.352)	0.138 (0.156)	-0.107 (0.270)	-0.119 (0.222)	-0.051 (0.602)	0.019 (0.839)	0.046 (0.435)
PW-reading	Accuracy	-0.024 (0.809)	-0.048 (0.621)	-0.045 (0.644)	0.043 (0.652)	0.037 (0.705)	-0.086 (0.376)	0.039 (0.688)
	Time	-0.097 (0.314)	-0.008 (0.933)	0.012 (0.902)	0.112 (0.248)	-0.017 (0.862)	-0.062 (0.527)	-0.071 (0.467)
	Overall	0.093 (0.337)	0.006 (0.949)	-0.007 (0.941)	-0.095 (0.328)	0.027 (0.780)	0.040 (0.678)	0.085 (0.382)
AM	0.029 (0.762)	-0.1381 (0.154)	0.136 (0.161)	-0.138 (0.154)	0.166 (0.085)	0.108 (0.266)	0.107 (0.270)	
Proms	Qualitative	0.034 (0.729)	0.139 (0.150)	-0.014 (0.888)	0.067 (0.489)	0.011 (0.914)	0.0006 (0.995)	-0.005 (0.953)
	Tonal	-0.182 (0.059)	-0.002 (0.986)	0.008 (0.936)	0.113 (0.243)	0.096 (0.324)	0.022 (0.822)	0.040 (0.679)
	Temporal	-0.073 (0.451)	0.083 (0.395)	-0.049 (0.612)	-0.025 (0.797)	-0.064 (0.509)	0.021 (0.832)	-0.138 (0.1546)
	Overall	-0.059 (0.543)	0.106 (0.273)	-0.027 (0.782)	0.044 (0.645)	-0.006 (0.949)	0.013 (0.891)	-0.083 (0.393)

¹ Autism Spectrum Disorder; ² Intelligence; ³ Attention Deficit Hyperactivity Disorder; ⁴ Executive function; ⁵ White Matter Volume measurement.

4.4 STUDY IV: BILINGUALISM AS A POTENTIAL MODULATOR OF THE NEURAL RESPONSE

A substantial body of research has extensively investigated the impact of bilingualism on auditory processing and cognitive functions (see Chapter I, section 1.4.1). Moreover, the FFR has demonstrated sensitivity to the acquisition of second and/or third languages. Therefore, the current study aims to explore the influence of bilingualism, particularly focusing on the age of acquisition of the second language, in the FFR as well as in the performance of the auditory skills evaluated in this thesis.

Descriptive statistics showed that 7 participants were monolingual (6.1%), 56 were native bilingual (48.7%), 48 were early bilingual (41.7%), and 4 were late bilingual (3.7%).

After recategorized the current sample into two groups (see chapter III, section 3.8), the sample was composed by fifty-six native bilinguals (51.9%) and fifty-two non-native bilinguals (48.1%).

FFR and age of acquisition of the second language

With the aim of evaluating whether differences existed in $FFR_{(ENV)}$ and neural lag between native and non-native bilinguals, a MR ANOVA was conducted.

Main effects of the background conditions were observed in the spectral amplitude response of the F_0 , SNR of the F_0 response, and the neural lag, with a moderate effect size in the F_0 and neural lag, and large effect size for F_0 SNR ($F=23.94$, $p<0.001$, $\eta^2_G=0.081$; $F=58.92$, $p<0.001$, $\eta^2_G=0.161$; $F=20.09$, $p<0.001$, $\eta^2_G=0.078$). These findings are consistent with those reported in Study I of this thesis. No main effect of the

bilingual acquisition nor interaction effect were found in the $FFR_{(ENV)}$ and neural lag (Table 11 and Figure 25).

Table 11: MR ANOVA repeated mean: noise background and bilingual acquisition effects in $FFR_{(ENV)}$ and neural lag.

Parameter	Contrast	Condition	$F_{(1)} (p\text{-value})$	η^2_G
F_0 spectral amplitude	Main effect background condition	<u>In quiet</u> <u>In noise</u>	23.940(<.001)	0.081
	Main effect Bilingual-acquisition	<u>native</u> <u>non-native</u>	0.310(0.579)	-
	Interaction effect		1.398 (0.240)	-
F_0 SNR	Main effect background condition	<u>In quiet</u> <u>In noise</u>	58.925(<.001)	0.161
	Main effect Bilingual-acquisition	<u>native</u> <u>non-native</u>	13232 (0.269)	-
	Interaction effect		1.513 (0.221)	-
Neural Lag	Main effect background condition	<u>In quiet</u> <u>In noise</u>	20.099(<.001)	0.078
	Main effect Bilingual-acquisition	<u>native</u> <u>non-native</u>	0.271(0.603)	-
	Interaction effect		2.107(0.150)	-

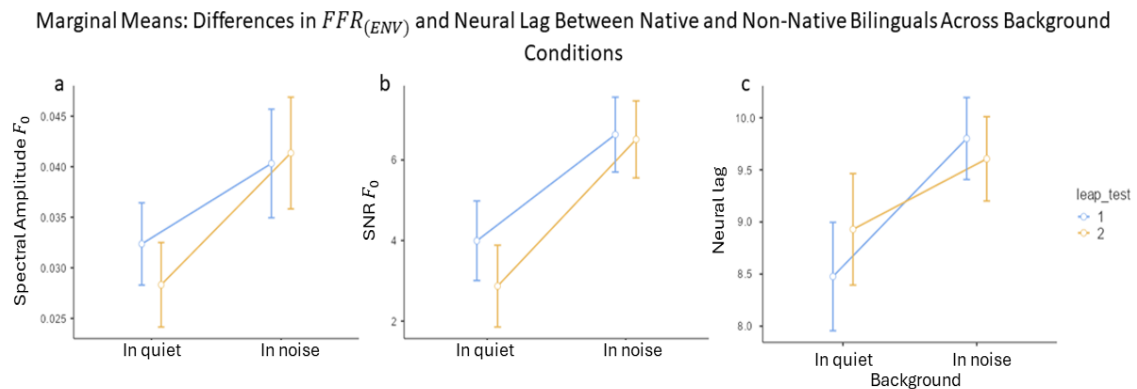


Figure 25. Marginal Means of $FFR_{(ENV)}$ and neural lag estimated for native and non-native bilinguals across both background conditions. The mean spectral amplitude of the response at the F_0 for both bilingual groups, estimated in quiet and in noise background

conditions (a). The mean SNR of the F_0 response between native and non-native bilinguals, estimated in both background conditions (b). The mean neural lag, estimated in both bilingual groups across background condition (c). The blue line represents the native bilingual group and in orange the non-native bilingual group (a, b and c).

To assess whether the $FFR_{(TFS)}$ differs according to the age of bilingualism acquisition a MR ANOVA was estimated between the $FFR_{(TFS)}$ in quiet and in noise, including the age of acquisition as a between-subject effect.

Main effects were found for the background condition in the Spectral amplitude F_1 and F_1 SNR in response to the /o/ part of the stimulus, with effect size between small to moderate ($F=42.33$, $p < 0.001$, $\eta^2_G=0.028$; $F=23.74$, $p < 0.001$, $\eta^2_G=0.088$). The main effect of the age of bilingual acquisition and the interaction effect were not significant. Regarding the response for the /a/ stimulus section, the main effect of the background condition and age of second language acquisition were not significant for the spectral amplitude F_1 nor F_1 SNR. The interaction effect (Background condition * Bilingual-acquisition) was significant in the F_1 SNR /a/, with a small effect size ($\eta^2_G= 0.011$) (Table 12, Figure 25). All of the results reported are consistent with the findings reported in Study I.

Table 12: MR ANOVA repeated mean: noise background and bilingual acquisition effects

in $FFR_{(TFS)}$.

Parameter	Contrast	Condition	$F_{(1)} (p\text{-value})$	η^2_G
F_1 spectral amplitude /o/	Main effect background condition	In quiet In noise	42.335(<0.001)	0.028
	Main effect Bilingual-acquisition	native non-native	0.412(0.522)	-
	Interaction effect		0.731(0.394)	-
F_1 SNR /o/	Main effect background condition	In quiet In noise	23.741(<0.001)	0.088
	Main effect Bilingual-acquisition	native non-native	0.136(0.713)	-
	Interaction effect		0.009(0.921)	-
F_1 spectral amplitude /a/	Main effect background condition	In quiet In noise	0.240(0.626)	-
	Main effect Bilingual-acquisition	native non-native	0.0079(0.778)	-
	Interaction effect		0.273(0.602)	-
F_1 SNR /a/	Main effect background condition	In quiet In noise	0.746(0.390)	
	Main effect Bilingual-acquisition	native non-native	0.148(0.702)	
	Interaction effect		4.276(0.041)	0.011

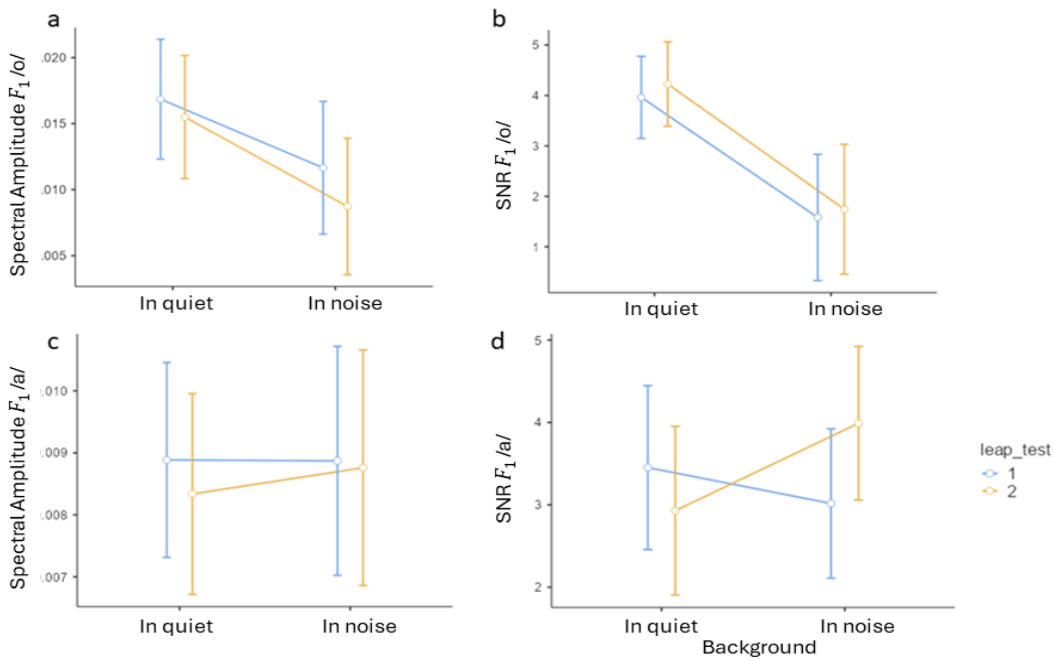
Marginal Means: Differences in the $FFR_{(TFS)}$ Between Native and Non-Native Bilinguals Across Background Conditions

Figure 26. Marginal Means of $FFR_{(TFS)}$ for native and non-native bilinguals across both background conditions. The mean spectral amplitude of the response at the F_1 of the /o/ for both bilingual groups, estimated in quiet and in noise background conditions (a). The mean SNR of the F_1 /o/ response between native and non-native bilinguals, estimated in both background conditions (b). The mean spectral amplitude of the response at the F_1 of the /a/ for both bilingual groups, across both background conditions (c). The mean SNR of the F_1 /a/ response between native and non-native bilinguals, estimated in both background conditions (d). The blue line represents the native bilingual group and in orange the non-native bilingual group (a b c and d).

For cross-correlation, significant main effect in the background noise, with small effect size, and main effect of pitch contour, with large effect size, were observed, indicating higher cross-correlation in the noise condition compared with in quiet, and in the /a/ flat pitch contour compared with rising pitch contour, separately ($F=4.671$, $p=0.033$, $\eta^2_G=0.043$; $F=650.09$, $p < 0.001$, $\eta^2_G=0.861$, respectively). No significant interaction effects were found between them. The main effect of bilingual acquisition was not significant, but the interaction effect with background conditions was significant but with a small effect size ($F = 6.21$, $p = 0.014$, $\eta^2_G=0.004$) (Table 13).

Analyzing the simple effect, significant differences were observed only in the non-native bilingual group in both quiet and in noise background conditions. This indicates that cross-correlation was higher in the noise condition compared to the quiet condition in the non-native bilingual group but not in the native bilingual group (Figure 27a). All other interaction effects (background*pitch contour, pitch contour*bilingual acquisition,

background**pitch contour**bilingual acquisition) were non-significant (see Figure 27b-c).

For pitch strength, the main effect of the background condition was significant, revealing that pitch strength is higher in quiet conditions compared to in noise background conditions with a small effect size ($F=12.80$, $p < 0.001$, $\eta^2_G=0.029$). The main effects of pitch contour and bilingual acquisition were not significant, nor were the interaction effects significant in any of the contrasts carried out (background* bilingual acquisition, pitch contour*bilingual acquisition, background*pitch contour, and background*pitch contour* bilingual acquisition) (Table 13).

Regarding Pitch error, significant main effect was observed in the background condition, with higher pitch error in noise background compared to in quiet with large effect size ($F=67.72$, $p < 0.001$, $\eta^2_G=0.193$). Non-significant main effects were observed in the pitch contour and acquisition of the second language, nor were interaction effects observed in any of the contrasts made (Table 13).

Table 13: ANOVA repeated mean: noise background, pitch contour and bilingual acquisition effects in FFR pitch tracking measures.

Parameter	Contrast	Condition	$F_{(1)} (p\text{-value})$	η^2_G
Cross-correlation	Main effect background condition	In quiet In noise	4.671(0.033)	0.003
	Main effect pitch contour	/a/ flat /a/ rising	650.09(<0.001)	0.667
	Main effect Bilingual-acquisition	native non-native	0.150(0.699)	-
	Interaction effect: background*bilingual acquisition		6.211(0.014)	0.004
	Interaction effect: pitch contour* bilingual acquisition		0.156(0.693)	-
	Interaction effect: background*pitch contour		3.639(0.059)	-
	Interaction effect: background*pitch contour* bilingual acquisition		0.591(0.444)	-
Pitch strength	Main effect background condition	In quiet In noise	12.803(<0.001)	0.029
	Main effect pitch contour	/a/ flat /a/ rising	2.623(0.108)	-
	Main effect Bilingual-acquisition	native non-native	0.740(0.391)	-
	Interaction effect: background*bilingual acquisition		0.017(0.895)	-
	Interaction effect: pitch contour* bilingual acquisition		1.343(0.249)	-
	Interaction effect: background*pitch contour		1.237(0.268)	-
	Interaction effect: background*pitch contour* bilingual acquisition		0.287(0.593)	-
Pitch error	Main effect background condition	In quiet In noise	67.724(<0.001)	0.193
	Main effect pitch contour	/a/ flat /a/ rising	1.71e-5(0.997)	-
	Main effect Bilingual-acquisition	native non-native	1.31(0.255)	-
	Interaction effect: background*bilingual acquisition		1.098(0.297)	-
	Interaction effect: pitch contour* bilingual acquisition		0.917(0.341)	-
	Interaction effect: background*pitch contour		0.971(0.327)	-
	Interaction effect: background*pitch contour* bilingual acquisition		0.143(0.706)	-

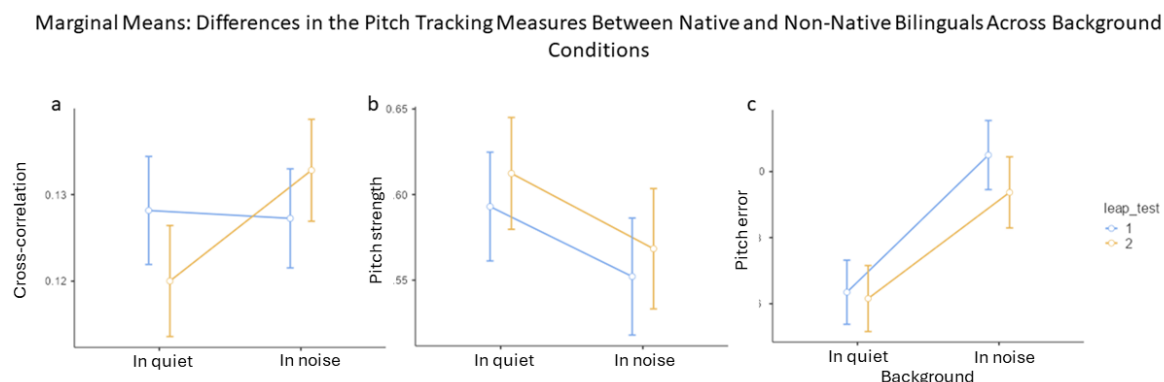


Figure 27. Marginal Means of the pitch tracking measures for native and non-native bilinguals across both background conditions. The mean cross correlation estimated in quiet and in noise background condition for native and non-native bilinguals (a). The mean Pitch strength estimated in both background conditions, for both bilingual groups (b). The mean pitch error estimated in both background conditions, for both bilingual groups (c). The blue line represents the native bilingual group and in orange the non-native bilingual group (a b, and c).

Behavioral measures and age of acquisition of the second language

In order to assess whether native bilinguals perform differently from non-native bilinguals in the SIN perception, Word/PseudoWord reading, Musical skills and AM discrimination, an independent samples t-test was conducted between the two bilingual acquisition groups for each of the auditory skills that were assessed. The results are presented in Table 14.

Non-significant differences were observed between natives/non-native bilingual groups across all of the auditory skills assessed in the present thesis.

Table 14: t test for native and non-native bilinguals across SPIN test, Word/Pseudo word reading, Musical perception, and AM discrimination.

Measure	<i>Levene's</i> _(1,106) <i>p-value</i>	Mean difference	<i>t</i> ₍₁₀₆₎ (<i>p-value</i>)	Effect size <i>Cohen's d</i>
SPIN	0.159	-0.478	-1.101(0.273)	-
Word Reading	0.038*	1.875	0.340(0.735)	-
PseudoWord Reading	0.594	-1.619	-0.328(0.747)	-
PROMS overall	0.336	-0.390	-0.249(0.804)	-
Tonal domain/PROMS	0.604	0.207	1.063(0.290)	-
Qualitative domain/PROMS	0.295	-0.127	-0.664(0.508)	-
Temporal domain/PROMS	0.078	-0.125	-0.647(0.519)	-
AM discrimination	0.362	2.047	1.171(0.244)	-

CHAPTER V: DISCUSSION

5.1 DISCUSSION STUDY I

In the current study, a speech-evoked FFRs paradigm was employed to investigate the impact of background noise on early neural responses in the auditory pathway, as well as its functional consequences on the perceptual performance in a speech-in-noise task. To achieve these goals, participants first underwent a Spanish-adapted and validated Speech-in-Noise test (Cervera & González-Alvarez, 2010, 2011). Subsequently, the subjects were instructed to focus on a silent movie while the auditory stimulus /oa/ was presented in both quiet and noisy background conditions, whilst their brain electrical activity was recorded using an EEG.

Effects of background noise on the FFR

Intuitively, one might expect background noise to degrade the response to the F_0 and its corresponding SNR; however, this effect was not observed in the present study. The findings showed that the response to the fundamental frequency and its magnitude, reflected in the SNR, are significantly enhanced in background noise conditions compared to in quiet.

These outcomes are not entirely surprising, as the response to the fundamental frequency remained robust in background noise conditions. This observation has been documented in various populations, including children without a history of medical or learning problems (N. Russo et al., 2004), healthy adults (X. Li & Jeng, 2011; Smalt et al., 2012; Song et al., 2011), bilingual individuals (Krizman, Marian, et al., 2012), and musicians (Parbery-Clark, Skoe, & Kraus, 2009). Moreover, studies in animals, particularly

albino guinea pigs, have indicated that the FFR- F_0 response to the vowel exhibits greater resistance to noise background (Cunningham et al., 2002).

In the cited studies, no differences were observed in the response amplitudes of the F_0 in the spectral domain between quiet and noise background conditions. However, some variations were noted among studies, revealing differences in responses in the time domain through peak analysis.

In the particular case of our study, not only there were no differences observed between the two background conditions, but the F_0 was significantly enhanced in the noise background. Similar results were found by Prévost et al., (2013) with healthy adults population. In their study, the spectral amplitude of the F_0 was tested in quiet and in different degrees of signal-to-noise ratio (SNR) of Gaussian white noise background. The results revealed that the F_0 evoked in noise was significantly higher than the response elicited in quiet, in addition with suppressed formant responses.

Different hypothesis and mechanisms have been proposed to explain these results. Firstly, it is suggested that stochastic resonance occurs in the auditory system, wherein neural responses synchronize with the target signal in the presence of noise. This synchronization compensates for the signal degradation caused by noise and enhances the representation of the signal in the auditory system (Cunningham et al., 2002; Prévost et al., 2013). The precise neural mechanisms underlying this phenomenon remain unclear. Secondly, in typical brainstem experiments, higher stimulus levels (>70-80 dB SPL) are often utilized, resulting in broad cochlear excitation as the intensity increases. Consequently, the resilience of the F_0 within the FFR may not solely indicate stochastic resonance, but also reflect the synchronization of low-frequency "tails" from basal, high-

frequency neurons with the common, high-intensity F_0 across cochlear channels. Multiple instances of phase locking at F_0 across the cochlear structure reinforce each other, providing robustness and redundancy in perceiving pitch amidst noise interference (Bidelman, 2017).

Another potential explanation involves top-down modulation of the brainstem response. Studies in bats and mice, using single-unit recording, propose that corticofugal projections from the auditory cortex to the central nucleus in the inferior colliculus may aid in segregating target and noise stimuli. In humans, numerous studies have identified descending projections from the auditory cortex to the inferior colliculus, from the thalamus to the inferior colliculus, from the inferior colliculus to the superior olive and cochlear nucleus, and from the superior olive to the hair cells in the cochlea (for an extensive review, see Asilador & Llano, (2021), and references therein). This extensive pathway of descending projections influences processing at early stages of auditory processing, such as the cochlea. An illustrative example of this top-down modulation is reported by Forte et al., (2017), where attentional processes modulated the brainstem response to the pitch of continuous speech in challenging conditions.

Given the notion that prior experiences modulate the way stimuli are perceived via top-down processes, the task of segregating a stimulus from a multi-talker background is a common experience in daily life. Therefore, it could be anticipated that afferent projections from the auditory cortex to brainstem structures might be enhanced to the extent of being activated almost automatically in the presence of noise, even pre-attentively, as in our study.

Regardless of which hypothesis explains our results, it is evident from our data that noise affects the elicited FFR in distinct ways. On the one hand, it enhances the signaling in response to the fundamental frequency and increases the magnitude of this response (Signal-to-Noise Ratio-SNR). On the other hand, it has a detrimental effect on the coding of the formants, resulting in significantly lower amplitudes in background noise conditions. Furthermore, this adverse impact of noise on the response is also reflected in reduced pitch strength, increased pitch error, and a higher neural transmission delay. It should be noted that a higher cross-correlation was observed with background noise. However, it is necessary to interpret these results with caution due to the *p-value* being very close to the conventional threshold for statistical significance and could reflect a spurious correlation.

FFR and Speech in noise perception

According to the second objective of this study, which aims to investigate the relationship between FFR parameters and perceptual performance in a SIN task, non-significant correlations were observed between the amplitude of the F_0 elicited and their SNR, the pitch tracking measures, and the performance in a SIN test. Only the amplitude of the F_1 in the /a/ vowel was statistically correlated with the performance of the SIN task with high predictability sentences; however, after applying the Bonferroni correction, these correlations did not reach statistical significance.

The relationship between FFR parameters and SIN perception is not entirely clear and exhibits variability across different studies, as discussed in Chapter I, Section 1.3.1. One

possible explanation for this variability could be attributed to the characteristics of the sample used, the sample size, and the nature of the comparisons made.

Typically, the associations between SIN perception and specific FFR parameters have been observed in studies comparing dichotomous groups or extreme values. Examples include comparisons between musicians and non-musicians, bilingual and non-bilingual individuals, high and poor performers in SIN tasks, populations with reading/listening or learning impairments versus normally developed children, and clinical versus non-clinical populations (Anderson et al., 2010, 2011; Anderson & Kraus, 2010; Basu et al., 2010; Chandrasekaran et al., 2009; Coffey, Mogilever, et al., 2017; Musacchia et al., 2007). However, in correlational studies, this association is not as straightforward or consistent (Parbery-Clark, Skoe, & Kraus, 2009; Rosenthal, 2020).

Dichotomizing the sample may magnify smaller differences that might not be apparent when treating variables as continuous, as seen in correlational studies, even with remarkable sample sizes such as in our study. Moreover, various factors should be considered when comparing our study results with previous findings. The cortical nature of SIN perception has been proposed, possibly explaining divergent results, such as in Coffey, Chepesiuk, et al. (2017) study, where a significant relationship between F_0 amplitude and SIN task performance was found only when F_0 was estimated from MEG recordings but not from EEG recordings.

The choice of SIN assessment tests, such as HINT or Quick, can also influence outcomes, as indicated in the systematic review by Rosenthal (2020). Additionally, the type of stimulus used is crucial, with prior correlational studies consistently showing a significant relationship between the response to the consonant-vowel transition portion of the /da/

stimulus and SIN performance tasks. However, due to the characteristics of our vowel-vowel stimulus, this specific comparison could not be made in our study.

All these factors may contribute to the higher variability found in the reported outcomes, even leading to the lack of replicability of these findings in some cases (Parbery-Clark et al., 2011, 2012; Parbery-Clark, Skoe, & Kraus, 2009; Thompson et al., 2017).

5.2 DISCUSSION STUDY II

In the current study, three behavioral tests were administered to measure various auditory skills, including reading skills, musical abilities, and sound temporal discrimination. The primary objective was to explore potential functional relationships between these behavioral measures and the frequency following response (FFR). Additionally, the study aimed to clarify the relationships among the behavioral measures themselves.

As previously mentioned in Chapter III, section 3.8, the first step involved executing a Principal Component Analysis with the FFR measures estimated. This analysis resulted in four components encompassing all our electrophysiological measures. The subsequent steps involved estimating Spearman and Pearson's correlations to cover the objectives proposed in this study.

FFR components and their functional relationship

Based on our results, the F_1 /o/ component showed a significant correlation with reading performance. Specifically, increased spectral amplitude and SNR in response to the F_1 for the /o/ vowel consistently correlated with higher scores in both word and pseudoword subtests of the PROLEC-SE-R battery. Conceptually, this relationship is expected as the encoding of formants is a fundamental characteristic distinguishing one letter/vowel from another. These formants act as basic units of speech sounds, enabling accurate recognition and the acquisition of explicit knowledge about phoneme construction. Such processes play a crucial role in the development and acquisition of reading skills.

Exploring this relationship in more detail, it was observed that higher scores in the F_1 /o/ component were associated with shorter reading time, but not with higher accuracy in the reading process. These findings were consistent in both word and pseudoword reading subtests.

One potential interpretation of these findings may be associated with the intricate processes of automatization. Reading proficiency is a sophisticated cognitive task at a high level, encompassing the swift and efficient identification of sentences while simultaneously retrieving and integrating phonological, orthographic, and semantic information. An adequate neural representation of the fine structure of speech stimuli forms the fundamental basis for phoneme identification and recognition. In line with this notion, a precise neural representation of phonemes could optimize the experience of automatic sound processing, liberating cognitive resources that can be reallocated to critical cognitive processes such as phonological processing, working memory, attention, among others. These processes are essential for reading performance, and the automatization could contribute to the enhancement of reading fluency.

Lam et al., (2017) investigated the performance of eighty-seven children with reading impairment in tasks related to reading fluency, automaticity in reading context (Rapid Automated Naming-RAN), and processing speed. They assessed the relationship between these factors and neural stability calculated from the FFR, reflecting the precision or stability of the FFR across multiple events or trials. The study revealed that a more stable neural response was associated with better RAN scores, higher processing speed, and enhanced reading fluency. Subsequently, the researchers conducted a mediation analysis to explore the nature and direction of these relationships. The results

suggested that neural stability predicts RAN performance, and this, in turn, predicts reading fluency, indicating a the role of RAN as a mediator connecting auditory neural stability and reading fluency.

Linking with our results, we identified a significant but small correlation between the F_1 /o/ component and reading performance (approximately 0.2). We hypothesized that the neural representation of formants, captured by the F_1 /o/ component, could potentially influence reading speed through Rapid Automatized Naming (RAN), akin to the observed associations in the study by Lam et al. (2017). Just as neural stability was found to predict RAN performance in their research, we posit that the F_1 component's neural representation of formants might similarly contribute to reading speed through RAN. Further investigations and analyses are warranted to delve deeper into the intricate relationships among these variables.

On the other hand, the pitch magnitude and pitch tracking components do not appear to be related to reading skills, musical perception, and AM discrimination tasks. There is an exception, which is a negative correlation between the pitch magnitude component and pseudoword reading subtest. However, considering that the p-value is close to the critical threshold (0.046) and that a negative correlation does not have a straightforward theoretical ground, we consider this to be a spurious correlation.

No relationship was observed between our electrophysiological component and musical perception. While one might intuitively anticipate that our electrophysiological components should, to some extent, be associated with perceptual musical processes, such a connection was not evident in our data. One possible explanation is offered in a review conducted by Koelsch, (2011), where it is argued that auditory information

originating from the cochlea, basilar membrane, and subcortical structures undergoes transformation from acoustic features (e.g., frequencies, periodicity, etc.) into auditory percepts, primarily within the primary auditory cortex (PAC). In the context of our data, this interpretation suggests that the processing and encoding of acoustic features, as reflected in the FFR, may not be comprehensive enough to elucidate or comprehend more intricate mechanisms at the perceptual level, such as music perception, at least as measured in the present study.

Relationship between the behavioral measures

Regarding the relationship between the auditory skills measures in our study, only speech-in-noise perception was significantly correlated with word reading performance, supporting the hypothesis that both skills are related, as has been suggested in previous studies (Vanden Bempt et al., 2022; Ziegler et al., 2009).

Upon exploring this correlation in detail, we found that higher performance in SPIN for low predictability sentences was associated with shorter times in word reading performance, supporting the idea that both processes, might share some underlying neural mechanisms. However, based on our results, the relationship between reading and SIN performance is not explained by the FFR parameter estimated in the present thesis.

Finally, it was expected that the amplitude modulation (AM) discrimination task would show correlations with SIN perception, reading skills, and musical perception, given the understanding that temporal perception is crucial for successful performance in these domains. However, these relationships were not observed in our data, indicating that, at

least in the beta-gamma frequency band (20-40 Hz), there is no correlation with these auditory skills in a healthy, normal young-adult population, such as our sample.

5.3 DISCUSSION STUDY III

The purpose of the current study was to identify genetic factors that modulate the neural encoding of sounds. We addressed this goal at two levels: one hypothesis-driven approach to replicate the involvement of the promoter region of the *SLC6A4* gene (the only genetic region associated so far in the previous literature) and a hypothesis-free approach to detect novel genomic regions relevant for this neurophysiological process. For the purpose of this study, we select the FFR component calculated in study II, as phenotypes of interest.

In the candidate gene approach, we replicated the involvement of genetic variants in the promoter region of the serotonin transporter gene (*SLC6A4*) in our dataset. Two different markers, SNPs rs2020933 and rs73263896, were associated with differences in the SNR and pitch magnitude component values even after FDR adjustment. Despite this, the specific polymorphisms that were associated do not match those that were previously associated by Selinger et al. (2016). In that previous study, the combination of the 5-HTTLPR minisatellite and SNP rs25531 was associated with FFR differences in SNR and pitch strength, but we did not find equivalent results in this thesis.

Our results support the importance of the activity of the serotonin transporter, and consequently of the serotonergic system, in the neural encoding of sounds. These results are in concordance with findings that linked serotonergic activity in the inferior colliculus with the brain response to auditory stimuli (Hurley & Hall, 2011; Obara et al., 2014). Postmortem studies have described a high density of serotonergic fibers that innervate the IC, especially in its central nucleus (Lavezzi et al., 2015). According to our results, differences in the availability of serotonin at the synaptic cleft in the

serotonergic afferences to the IC, mediated through differences in serotonin uptake, might regulate the neural encoding of sounds, specifically to speech stimuli. Thus, genes encoding for other modulators of the serotonergic activity in the IC, such as the HTR1A and HTR1B inhibitory autoreceptors, may also be additional candidates to modulate the mechanism of encoding the physical properties of incoming sounds (A. R. Smith et al., 2014). This role of serotonin is thought to be executed through the modulation of the activity of GABAergic inhibitory and glutamatergic excitatory synaptic inputs that may be activated simultaneously in the IC by sound stimulation (Ayala & Malmierca, 2018; Obara et al., 2014), providing additional biological actors involved in this process.

Although we replicated the involvement of variants in the *SLC6A4* gene promoter, this replication is only partial since the specific variants associated are not the same across studies. These discrepancies cannot be attributed to differences in allele frequency or in LD patterns between studies, as could be the case if there were differences in the population background of participants, because both studies recruited participants from the same population (mostly college students at the Mundet Campus of the University of Barcelona). Comparing the methodology implemented in Selinger et al. (2016) with ours did not reveal remarkable differences that could explain the observed differences. Both studies used a similar experimental design, with just some minor differences in the stimulus used: a /ba/ syllable instead of an /oa/ and a fundamental frequency of 100Hz instead of 113Hz in our study. Regarding sample sizes (and, thus, statistical power), in this thesis we almost doubled the number of participants in the study, from 58 in Selinger et al. (2016) to 101 in our study.

The consistent association of variants in the promoter of the *SLC6A4* gene but with discrepancies in the specific associated marker/s may point to a more complex regulation of the expression of this gene. In this sense, other genetic variants in different regulatory elements across the gene region (including alternative non-coding exons) have been shown to influence transcriptional and post-transcriptional processing of the transporter (for a review, Iurescia et al., 2017). Moreover, environmental exposures along the lifespan that have an impact on brain function, such as language exposure or musical training in the case of the subcortical encoding of sounds, often involve epigenetic modifications that provide an additional level of regulation of gene expression. In this sense, the methylation profile in a CpG island region in the promoter region of the *SLC6A4* gene regulates mRNA levels and has also shown interactions with the 5-HTTLPR polymorphism in relation to both the effect of environmental stressors and the regulation of the *SLC6A4* gene expression (Iurescia et al., 2017; Wankerl et al., 2014). Moreover, this complex interaction between genetic and epigenetic mechanisms in the same gene of the serotonergic system has also been demonstrated for the *HTR1A* and *HTR1B* genes (Xu et al., 2022).

Future approaches to characterize the regulatory effect of the serotonin transporter in the neural encoding of sounds should disentangle the complex regulation of the expression of this gene. On one side, genetic analyses should not be limited to one or few variants in the *SLC6A4* gene (in the literature most studies focus only on the 5-HTTLPR), but include the different genetic variants across the gene that have been described as having an effect on gene expression. Full sequencing of the gene might even be considered as a useful approach. On the other side, the methylation profile

across the CpG island in the promoter region should also be assessed and be included in a joint analysis with the genetic layer of information to provide a comprehensive picture of the epigenetic/genetic regulators of the serotonin transporter that influence the FFR. To get an even more comprehensive understanding of this mechanism and expand it to the overall serotonergic system (and not just to the transporter), future analyses could also incorporate genetic and epigenetic information on other potentially relevant genes such as *HTR1A* and *HTR1B*.

In the hypothesis-free genome-wide approach, we attempted at identifying individual SNPs and genes that were associated with differences in the FFR components that were designed in Study II. Moreover, we also wanted to detect overlap in the genetic determinants between our phenotypes of interest and other potentially related phenotypes, to shed some light on the genetic architecture of sound processing.

Overall, we did not detect association results that reached the stringent significance thresholds that genome-wide analyses usually implement given the extremely high number of comparisons that most of these analyses entail. However, several chromosomal regions (and their associated genes) showed suggestive associations with different FFR parameters and at different levels, emerging as suitable candidates to be involved in the neural encoding of sounds. The discussion hereon will focus on the subset of the top associated genes in both the SNP-based and gene-based analyses that also show some supportive evidence from function, expression patterns, or previous reports in the literature.

The EBF transcription factor 2 (*EBF2*) gene arose as one of these candidates to be involved. The top lead SNP in the GWAS analyses on the pitch magnitude component

(rs4871957) and many others in the same LD block lie within the coding region of this gene. Moreover, this gene was also one of the top associated genes both pitch magnitude and pitch tracking components in the gene-based analyses. *EBF2* belongs to a family of genes that encode for transcription factors important in different developmental processes (*EBF2 EBF transcription factor 2 [Homo sapiens (human)] - Gene - NCBI, s. f.*) and show maximal expression levels in the nerval tissue (*GTEX Portal, s. f.-a*) there are no previous reports on GWAS involving genetic variants in this gene with cognitive, behavioral or neuroanatomic phenotypes.

Markers in the *RBFOX1* gene were also among the top-associated loci with the pitch magnitude component and mark this gene as a good candidate given also its expression profile and its previous associations. The *RBFOX1* (RNA binding fox-1 homolog 1) gene encodes for an RNA-binding protein that regulates alternative splicing (*RBFOX1 RNA binding fox-1 homolog 1 [Homo sapiens (human)] - Gene - NCBI, s. f.*) and is selectively expressed in the brain. It is expressed in all 13 different CNS tissues in the GTEx v8 database, especially in telencephalic and cerebellar areas (*GTEx Portal, s. f.-b*). Moreover, variants in this gene have been repeatedly associated with different cognitive and behavioral as summarized in the GWAS Catalog. Reports on this gene include associations on educational attainment (N=30 counts), neuroticism (N=21), smoking behavior (N=20), depressive symptomatology (N=28), intelligence (N=12), or math ability (N=8), among many others (*GWAS Catalog, s. f.*).

Another suitable candidate identified in this Study is the *OPALIN* gene on chromosome 10. It constituted the top association in all gene-based analyses and showed its potential effects on the pitch magnitude component of the FFR. This gene encodes for a myelin

protein specific of the CNS that regulates gene expression during the differentiation of oligodendrocytes and promotes oligodendrocyte terminal differentiation. Its expression is limited to CNS, where it is expressed in all tissues (*GTEx Portal*, s. f.-c).

Finally, the genomic region on chromosome 16 that contains the *RP11-505K9.4* and *MLYCD* genes also deserves further studies. These two genes constituted the two top associated genes with the pitch tracking component in the gene-based tests and also among the top genes associated with the pitch magnitude component. *RP11-505K9.4* has no known function, is mostly expressed in the testis and no previous study has associated it with any trait. *MLYCD* participates in fatty acid oxidation, is not particularly expressed in the brain, and all previous associations of this gene are related to metabolic outcomes, and none to cognitive or behavioral phenotypes. Although the biological relationship of these genes with the neural encoding of sounds is not clear, it is remarkable that they show associations with both FFR components. Since the pitch magnitude and the pitch tracking components are not substantially correlated (it is an assumption of the PCA analysis), it would be unlikely that the association levels observed ($p \sim 10^{-4}$ - 10^{-5}) show up spuriously for two different independent variables for both genes. FUMA results indicate that the associations in these yes is true, are independent, I check it for curiosity genes are also independent since there is no substantial LD among the associated regions, but it cannot be discarded that they are capturing the same association signal since these two genes are less than 5Kb away from each other.

In addition to SNP-based and gene-based analyses, we employed an alternative approach. Initially, we computed the polygenic scores (PGS) for five traits included in the PGS Catalog: Intelligence, ADHD, Autism Spectrum Disorder (ASD), Executive

Function, and White Matter Volume. Subsequently, we correlated these scores with the four FFR components calculated in Study Two, alongside four behavioral measures of auditory skills.

We only observed a significant correlation between speech-in-noise perception, particularly with low predictability sentences, and the polygenic risk of ADHD. It would indicate a shared degree of etiology between genetic variants that are associated with the risk for ADHD and the genetic variants that modulate SIN perception, especially when there is no contextual clue and attentional processes are essential. This finding aligns with prior research indicating that individuals with ADHD often exhibit diminished performance in tasks involving speech-in-noise perception (Blomberg et al., 2019; Jayawardena et al., 2020; Laffere et al., 2021).

As presented in the introduction, the influence of genetics in auditory processing of both speech and non-speech sounds is potentially very large, as reflected in the estimated heritabilities from twin studies ranging from 40% to 70% in most cases. However, this remarkable proportion of the variance in these traits that is determined by the individual genetic makeup of an individual is probably highly polygenic. Most complex traits, as is the case of the phenotypes studied in this thesis, are due to the effect of multiple genes each one of them with a relatively small effect size. In this context, the trend in the GWAS field has been to increase sample sizes to overcome the limitations imposed by the small effect sizes and the high number of variants analyzed (see Abdellaoui et al., 2023 for a recent review). Using height as a model to understand the role of genetics in the architecture of polygenic traits (it is easily measured and has one of the highest heritabilities for a complex trait), efforts had been successful at identifying variants

involved in height but were not able to explain a substantial amount of the heritability until a recent GWAS in 5.4 million individuals identified 12,000 SNPs that jointly explained most of the heritability in European populations (Yengo et al., 2022).

In this context, is it feasible to conduct genomic analyses in small samples? It may be the point for some clinical traits (e.g. height, blood pressure, BMI, ...) that have already been extensively studied in massive cohorts, which often also offer their data in public repositories. However, many phenotypes are difficult to obtain at large scales for various reasons: economic cost, long experimental sessions to conduct the assessment, limited access to technique or equipment, specificity of the biological process studied, ... Anyway, some of this ‘smaller-scale’ phenotypes (such as the event-related imaging techniques used in this thesis) provide valuable information in many neurophysiological processes that are difficult to characterize otherwise.

Another consideration that emanates from our study is the FFR parameter that seems to be more closely influenced by genetic factors and, consequently, constitutes a more informative biomarker of neural encoding of sounds. The identification of biomarkers that are in a predictable biological path from gene to behavior constitutes one of the challenges in genetic psychophysiology research (Anokhin, 2014; Lenzenweger, 2013).

Based on our and previous research on the genetic variation in the *SLC6A4* promoter region, assessments of the spectral encoding of the fundamental frequency appear to be more informative about these complex genotype-phenotype relationships. Selinger et al. (2016) found associations with SNR and pitch strength, and we found them with SNR and the pitch magnitude component, which includes the spectral amplitude of the F_0 , SNR of this spectral amplitude, and the cross-correlation.

Our findings in the GWAS and post-GWAS analyses confirm pitch magnitude assessments as suitable biomarkers and extend these observations to pitch tracking parameters as they both seemed to be more sensitive to genetic effects. These two components were systematically showing not only the more ‘statistically’ relevant genes (i.e. the ones with lowest p-values in the analyses) but also the ones that provided candidates that were more plausible from the functional and expression points of view. Moreover, the tissue expression analyses conducted in FUMA indicated that the genetic variants associated to both the pitch magnitude and the pitch tracking components (i.e. the full distribution of SNP-associated p-values across the genome for each phenotype) were nominally associated with an enrichment of genes expressed in brain tissues, particularly relevant for cerebellar expression in the case of pitch magnitude.

The fact that pitch magnitude and pitch tracking might serve as more informative biomarkers for studying the genetics of sound neural encoding could be explained by several factors. On one hand, both components exhibit high variability. The variances for pitch magnitude and pitch tracking components range from 10 to 15, whereas for both components summarizing the F_1 response, variances fall between 5 and 9. The lower variability in the response to the formants could potentially indicate less informativeness.

Another explanation for this difference in informativeness across FFR components in the genetic analyses could be that these components capture aspects of the FFR that might have different heritabilities. Thus, if the proportion of the overall variance in one FFR component that could be attributed to genetic factors is substantially smaller than for other components, the statistical power for that component will also be substantially

smaller (i.e. there is less effect size to be detected). Unfortunately, we could not calculate SNP-based heritability for our traits of interest in the GEPLEM cohort. SNP-based heritabilities are estimated in a set of unrelated individuals from all SNPs in an array and the calculations are based on the correlation between the matrices of genetic and phenotypic similarity (Yang et al., 2017). In our sample of unrelated individuals from a cosmopolitan population, the values of the matrix of genetic similarity (i.e. of genetic relatedness) were small and SNP-based h^2 estimates provided unreliable values with large SE (data not shown). However, just slightly larger sample sizes (in the order of 200-300) would start providing reliable estimates for most phenotypes, since the number of values in the genetic matrix of similarity increases exponentially with the increase in the number of individuals (Elhezzani, 2018).

As previously pointed out for the *SLC6A4* gene, not all biological differences may be genetically regulated in origin. Epigenetic mechanisms of gene regulation may be relevant in these cognitive processes since they can be shaped by experience and can alter gene transcription in the absence of direct alterations to the DNA sequence. The field of epigenomics has bloomed in recent years and has allowed the identification of epigenetic regulators of many diseases (specially in cancer) and the development of potential novel therapies. Some good reviews on the field of neuroepigenomics include Nayak et al., (2022) and Satterlee et al., (2015). In psychology, it has arisen high expectations since certain environmental exposures can lead to intergenerational inheritance of behavioral phenotypes, including examples induced by sensory stimulation (Dias & Ressler, 2014; Vassoler et al., 2013). However, there are no studies about potential epigenetic regulators of acoustic processing. The only study on a related

topic is a report on the DNA methylation changes in response to chronic low intensity environmental noise stress in rats (Guo et al., 2017). Future studies on the epigenetic modulators of the electrophysiological assessments of neural encoding of sounds could supplement the information retrieved by genetic studies to jointly build a more comprehensive understanding of the biological determinants of sound processing.

5.4 DISCUSSION STUDY IV

The impact of bilingualism on auditory processing has been extensively documented, serving as a representative example of brain plasticity (for more details, see Chapter I, Section 1.4.1). The superior auditory neural responses of bilingual individuals, such as enhanced neural synchronicity, shorter response latency, and robust FFR- F_0 , have been well-documented even in the early stages of the auditory pathway (Jafari et al., 2023; Krizman et al., 2021; Skoe et al., 2022). However, the benefits of acquiring a second language do not affect all bilinguals equally, with the age of second language acquisition emerging as one of the most significant environmental factors (Pirani, 1998; Sebastián-Gallés & Soto-Faraco, 1999; Shi, 2010; Skoe et al., 2017).

In the present study we aimed to explore the age of acquisition of the second language as a potential modulator for the FFR response, elicited in both quiet and noise background conditions. We specifically focused on comparing native and non-native bilinguals.

As a secondary objective, we examined the impact of native and non-native bilingualism on performance in speech-in-noise perception, reading skills, and musical perception skills.

Our findings indicated the absence of a significant main effect attributed to the age of bilingualism across diverse FFR measures, encompassing $FFR_{(ENV)}$, $FFR_{(TFS)}$, Pitch Tracking measures, and Neural Transmission Delay. However, interaction effects surfaced between background condition and the age of second language acquisition. Notably, these interactions manifested in specific metrics, such as the signal-to-noise ratio (SNR)

of the response elicited by the F1 corresponding to the /a/ segment stimulus and in the cross-correlation.

The interaction effects identified in this study suggest distinct impacts of noise on native and non-native bilinguals. In the case of native bilinguals, a robust neural response to the F1 in the /a/ vowel, as indicated by the signal-to-noise ratio (SNR), was observed in the quiet background condition. However, this response diminished in the presence of noise. Conversely, non-native bilinguals exhibited lower SNR in the quiet background condition, with a stronger neural response to the F1 /a/ observed in noisy conditions. This interaction effect interaction between background condition and the age of acquisition of the second language is small, accounting for only 1.1 % of the observed variability in this FFR parameter.

In the case of cross-correlation, the observed interaction effect suggests that, for native bilinguals, the cross-correlation remains consistent across both background conditions. However, for the non-native group, the cross-correlation increased in noise compared to the quiet background condition. It is important to note that this interaction effect explains only 0.4 % of the variance observed in the cross-correlation.

Although the observed effects were small, they reached statistical significance. When interpreted with caution and in conjunction with additional information, they offer insights into the relationships investigated in this thesis. For instance, in Study I of this thesis, it was noted that the spectral amplitude response to the fundamental frequency (F_0) and the cross-correlation were increased in background conditions. Both measures are interrelated as they reflect different aspects of pitch processing. Complementing these findings with the current study, which investigates how noise affects these

responses differently based on bilingualism nature (native/non-native), we can cautiously hypothesize that native bilinguals exhibit a more stable neural representation of changes in the auditory signal, as reflected in the cross-correlation, with relative independence from background noise conditions. In contrast, non-native bilinguals show a more variable neural representation influenced by the background condition.

Considering the relative magnitude of the response to the first formant, specifically in the /a/ segment, non-native bilinguals exhibit an increased response in noise conditions. While it may seem that noise conditions enhance the response of non-native bilinguals, at least in the indicated frequency-following response (FFR) measures, it would be intriguing to explore the potential functional impact of this phenomenon. Perhaps a better response translates into greater neural resources dedicated to processing and/or encoding the auditory signal in challenging conditions, potentially leaving fewer resources for high-level cognitive functions. This is merely a hypothesis, and further exploration of the functional consequences of this unlikely signaling on unmeasured high-level cognitive processes is warranted.

Regarding the secondary objective of the current study, which explores the impact of native/non-native bilingualism on performance in speech-in-noise perception, reading skills, and musical perception skills, no significant differences were observed between the two bilingual groups. Thus, according to our results, there is no obvious functional impact of the age of acquisition of the second language among bilinguals on these different auditory domains, at least with the limitations imposed by the specific tasks and tests that we have implemented.

5.5. GENERAL DISCUSSION

The general objective of the current doctoral thesis was to investigate how environmental and genetic factors contribute to the individual variations seen in sound processing during the initial phases of auditory processing, specifically the FFR, and how this response could be related to more functional assessment of auditory and speech perception, such as reading skills, musical abilities, and speech in noise perception, considering the bilingual experience as a potential influencing factor in their respective performance.

To explore the genetic factors contributing to our endophenotypes of interest, we employed various approaches aimed at covering different aspects where this influence could be observed, as discussed in the previous section (Section 5.3). For example, from a hypothesis-driven approach as like candidate genes study, we found that the SLC6A4 gene that encode the serotonin transporter is related with the pitch magnitude response. This finding reinforces the hypothesis that the amount of serotonin available in the synaptic cleft plays a significant role in the neural encoding of the sound, at least at the subcortical level, partially replicating the findings reported by Selinger in 2016.

In genome-wide approaches, including SNP-based and gene-based analysis, the relationship with our endophenotypes is less clear. This ambiguity may stem from limitations inherent in our study, with the primary constraint being the sample size necessary to reach statistical power in this type of studies. Nevertheless, in our study we identify novel candidate genes as EBF2, RBFOX1, OPALIN and RP11-505K9.4/MLYCD, that could serve as a starting point for future studies in the field of the neural encoding of the sound (for further details, refer to section 5.3). Of particular interest were the results

obtained with the polygenic score, where a significant correlation was observed between the polygenic risk of ADHD and performance in our speech-in-noise perception task. This finding supports previous evidence indicating that individuals with ADHD often struggle with speech-in-noise perception tasks. These results could suggest a shared underlying genetic basis between both.

Beyond the potential limitations stemming from the sample size of this study, there are some noteworthy points to highlight. For instance, both the pitch response and its magnitude consistently demonstrate associations with various genetic variants across all genetic approaches employed. This includes associations with the *SLC6A4* gene identified in our candidate gene study, as well as an increased presence of genes expressed in brain tissues, with a specific emphasis on cerebellar expression, as observed in our tissue expression analysis. This suggests that these endophenotypes could be more informative biomarkers for investigating the genetics of neural sound encoding compared to the encoding of formants, which do not appear to have a significant genetic influence.

In summary, based on our studies, we observed that the pitch response, considering both pitch magnitude and pitch tracking measures, are the most informative endophenotypes at the genetic level. However, what does this translate to at a more functional level? To address this question, we investigated the relationship between all of our endophenotypes and performance in four behavioral tasks: speech-in-noise perception, reading skills, musical abilities, and temporal discrimination tasks.

In contrast to the outcomes of our genetic analysis, our investigation revealed no significant correlations between pitch magnitude and pitch tracking measures with

speech-in-noise perception, reading skills, or musical abilities. This incongruity may suggest that the subcortical response measured by these parameters may not fully capture the intricacies of the processes examined in our study. However, our findings indicate a noteworthy association between formant encoding and speech-in-noise perception as well as reading skills, although such associations were not observed with musical abilities or temporal discrimination. These results underscore the significance of formant encoding in language-related capacities like speech-in-noise perception and reading skills, yet indicate its lack of influence on musical abilities. Although the observed relationships are modest in magnitude, they may still be relevant given our attempt to link subcortical measures with complex cognitive processes.

Finally, we investigated the potential influence of bilingual experience on both our electrophysiological and behavioral measures examined in this thesis. However, our findings did not support the hypothesis that the age of acquisition of a second language directly impacts our electrophysiological measures. Furthermore, we observed no significant differences in the behavioral performance of the assessed skills between native and non-native bilingual groups.

5.6 LIMITATIONS AND FUTURE RESEARCH

The four studies discussed in this doctoral thesis shed light on several important aspects of sound processing in the early stages of the auditory pathway. These include genetic influences, the effects of background noise on this response, their correlation with more functional measures of auditory and speech perception, and how the age of acquisition of a second language could mediate both electrophysiological responses and behavioral measures. However, certain limitations and potential areas for future research were encountered.

In the experimental design of our EEG experiment, we selected the /oa/ stimulus for several reasons. Firstly, it was chosen to facilitate various comparisons, such as comparing different pitch contours while controlling other parameters like durations and formants. Secondly, we ensured that no higher frequencies comprised the first formant of our stimulus, considering that the FFR tends to decrease as frequency increases. Additionally, the harmonics and formants of our stimulus matched, aiding in controlling the influences of other frequencies, such as harmonics, when evaluating the formant response. However, this stimulus does not allow for the assessment of the response to consonant-vowel transitions, a part where FFR studies often demonstrate sensitivity to experience and its relation to speech-in-noise perception. A potential solution to this limitation could involve creating and testing a stimulus that incorporates this consonant-vowel parameter.

Another potential limitation of the present study is the predominance of females in our sample (80.3%). This sex imbalance could potentially introduce bias into our results,

given that sex differences have been noted in subcortical responses such as ABR. Studies have shown that this sexual dimorphism in the FFR is observed at higher frequencies (≥ 750 Hz) and in peak latencies using peak-picking analysis. However, in our study, all evaluated frequencies are below 750 Hz, with our highest frequency being 678 Hz. Moreover, we primarily analyzed the spectral content of the response. Therefore, the likelihood of our data being affected by this bias is minimal. Nevertheless, we recommend that future studies strive for a more balanced distribution of sexes in the sample to ensure control over potential sex-related biases, or perhaps to replicate the lack of differences reported in previous studies.

One of the primary objectives of this thesis was to investigate the neural encoding of auditory information in the early stages of the auditory pathway and its correlation with more functional assessments. To achieve this goal, we selected the FFR as our focal point of study due to its rich and versatile nature, offering insights into various auditory processing mechanisms. For instance, it demonstrates precise synchronization with stimuli, faithfully replicates spectro-temporal characteristics, and exhibits high sensitivity to auditory experiences such as bilingualism and musical training. These attributes underscore the importance of accurate neural representation of auditory stimuli, forming the basis for subsequent high-order cognitive processes. However, our results suggest that the FFR alone may not suffice to comprehend more complex processes like speech-in-noise perception, reading performance, or musical abilities, perhaps due to its subcortical nature. In this regard, analyzing responses to more complex and realistic stimuli, such as ongoing speech, could be interesting. Advanced analysis techniques have been developed to study auditory neural responses to more

complex stimuli, both at the subcortical level, as evidenced in the work by Forte et al. (2017), and at the cortical level, as demonstrated by studies like those by Crosse et al. (2021), Fiedler et al. (2019), and Jessen et al. (2021).

Moreover, incorporating potential modulator variables between neural responses and behavioral measures could prove beneficial. For instance, Rapid Automatization Naming (RAN) has been suggested as a modulator variable influencing neural stability and reading fluency, as proposed by Lam et al. (2017). By including such variables, we can explore the mediating factors between electrophysiological responses and the more complex cognitive processes examined in this thesis. This approach opens avenues for testing more sophisticated processing models, which, in our view, would offer a more appropriate and realistic approximation for the types of auditory skills under evaluation.

Another limitation of the current study is the small sample size for the genetic analyses, as we mentioned in section 5.3, as a consequence our GWAS analysis is underpowered. Nevertheless, we consider that implementing genome-wide approaches still provide novel and relevant information. We have identified novel candidate genes to be involved in the neural encoding of sounds: EBF2, RBFOX1, OPALIN, and the RP11-505K9.4/MLYCD region. Moreover, we have shown that variants associated to some FFR components are enriched in genes expressed in brain tissues and evidence of some shared genetic mechanisms between SIN perception and ADHD. Although these results need to be considered with caution because of the limited statistical power, they still have provided hints into molecular mechanisms that may regulate the subcortical encoding of speech sounds. Interestingly, most of these novel genes exert their biological role through different mechanisms of gene expression regulation (e.g. transcription factors,

regulating alternative splicing, ...). These molecular mechanisms have been identified because we implemented a hypothesis-free genome-wide approach and would have remained unknown if we had stuck to the traditional candidate gene approach because of statistical power issues. Candidate gene approaches in genetic psychophysiology have provided interesting clues (Anokhin, 2014) but have mostly been limited to a 'usual suspects' shortlist of genes involved in neurotransmission or as growth factors (e.g. COMT, SLC6A4, MAOA, BDNF, ...). The genome-wide approach implemented here comes at the cost of being underpowered and the first and foremost future efforts should focus on increasing the sample size (and, consequently, the power) and replicating the associations described here in independent samples.

Finally, the impact of bilingualism on auditory processing and cognitive functions has been extensively documented. Therefore, we included the LEAP-Q questionnaire to obtain a more comprehensive measure of language proficiency than a simple binary question like "Are you bilingual? Yes or no." However, the bilingual nature of our sample limited our ability to compare individuals beyond native and non-native bilingual categories. Most participants were native and early bilinguals, potentially overlooking the true impact of bilingualism on the FFR and the behavioral measures assessed in this thesis. It is likely that the majority of our sample acquired their second language at an early age, possibly before the age of 7, resulting in minimal differences during these developmental stages. The impact of bilingualism may manifest later in development. Hence, it is crucial to include participants with varying degrees of language proficiency, such as monolinguals, late bilinguals, and multilinguals, to enhance the variability of language backgrounds. This approach would facilitate more comprehensive

comparisons that accurately capture the influence of language experience on the measures utilized in this study.

CHAPTER VI: CONCLUSIONS

Based on the findings derived from the comprehensive investigation encompassing the four studies conducted, strategically aligned with the objectives delineated in this thesis, the following conclusions can be drawn:

1. The impact of noise on the Frequency Following Response (FFR) varies significantly. While it enhances responses to the fundamental frequency and improves pitch contour tracking, it also exerts detrimental effects on fine structure, degrading responses to formants and/or harmonics, and increasing neural transmission delay.
2. Regarding the association between FFR and Speech-in-Noise (SIN) perception, only the response to the first formant of /a/ showed a correlation with performance in SIN perception for sentences with high predictability. Other FFR measures did not predict performance in the SIN test adequately.
3. The component encompassing the response to the first formant of /o/ correlates with higher scores in word and pseudoword reading subtests, specifically with reduced reading time. No associations were observed between FFR components and amplitude modulation discrimination or musical skills.
4. Significant correlations were found between SIN perception and reading skills, suggesting a shared neural basis between these two abilities, as posited in previous research.
5. The association of genetic variants in the promoter region of the serotonin transporter gene (SLC6A4) with differences in neural encoding of sounds, particularly FFR responses to the fundamental frequency, was confirmed. Additionally, our results suggest complex regulation of gene expression.

6. Several novel candidate genes were identified to potentially contribute to individual differences in sound encoding, including EBF2, RBFOX1, OPALIN, and the RP11-505K9.4/MLYCD region. Although not significant at the genome-wide level, these genes showed suggestive associations supported by functional, expression, and association data.
7. Polygenic score analyses indicated shared genetic etiology between SIN perception skills and ADHD, with higher polygenic scores for ADHD associated with lower SIN perception in low predictability contexts.
8. Parameters related to pitch magnitude and pitch tracking components extracted from the FFR appear to be more informative intermediate phenotypes for genetic studies.
9. Despite being underpowered for most analyses, genome-wide analyses still yielded valuable insights in small samples with detailed phenotyping.
10. The nature of bilingualism (native/non-native) alone does not significantly impact the FFR response. However, an interaction effect was observed with background noise, indicating differential effects on native and non-native bilinguals.
11. Native/non-native bilingualism does not significantly affect performance in speech-in-noise perception, reading skills, or musical perception skills, at least in our sample of young adults without psychiatric or neurological disorders, with a high educational level, and homogeneous socioeconomic status.

CHAPTER VII:

REFERENCES

- 1000 Genomes Project Consortium, Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., Korbel, J. O., Marchini, J. L., McCarthy, S., McVean, G. A., & Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature*, 526(7571), 68-74.
<https://doi.org/10.1038/nature15393>
- Abdellaoui, A., Yengo, L., Verweij, K. J. H., & Visscher, P. M. (2023). 15 years of GWAS discovery: Realizing the promise. *American Journal of Human Genetics*, 110(2), 179-194.
<https://doi.org/10.1016/j.ajhg.2022.12.011>
- Abrams, D. A., Nicol, T., White-Schwoch, T., Zecker, S., & Kraus, N. (2017). Population responses in primary auditory cortex simultaneously represent the temporal envelope and periodicity features in natural speech. *Hearing Research*, 348, 31-43.
<https://doi.org/10.1016/j.heares.2017.02.010>
- Abrams, D. A., Nicol, T., Zecker, S., & Kraus, N. (2009). Abnormal Cortical Processing of the Syllable Rate of Speech in Poor Readers. *The Journal of Neuroscience*, 29(24), 7686-7693. <https://doi.org/10.1523/JNEUROSCI.5242-08.2009>
- Aiken, S. J., & Picton, T. W. (2008). Envelope and spectral frequency-following responses to vowel sounds. *Hearing Research*, 245(1-2), 35-47.
<https://doi.org/10.1016/j.heares.2008.08.004>
- Alaerts, J., Luts, H., Hofmann, M., & Wouters, J. (2009). Cortical auditory steady-state responses to low modulation rates. *International Journal of Audiology*, 48(8), 582-593.
<https://doi.org/10.1080/14992020902894558>
- Anderson, S., & Kraus, N. (2010). Objective Neural Indices of Speech-in-Noise Perception. *Trends in Amplification*, 14(2), 73-83. <https://doi.org/10.1177/1084713810380227>
- Anderson, S., Parbery-Clark, A., White-Schwoch, T., & Kraus, N. (2012). Aging Affects Neural Precision of Speech Encoding. *Journal of Neuroscience*, 32(41), 14156-14164.
<https://doi.org/10.1523/JNEUROSCI.2176-12.2012>

- Anderson, S., Parbery-Clark, A., Yi, H.-G., & Kraus, N. (2011). A Neural Basis of Speech-in-Noise Perception in Older Adults. *Ear & Hearing, 32*(6), 750-757.
<https://doi.org/10.1097/AUD.0b013e31822229d3>
- Anderson, S., Skoe, E., Chandrasekaran, B., Zecker, S., & Kraus, N. (2010). Brainstem correlates of speech-in-noise perception in children. *Hearing Research, 270*(1-2), 151-157.
<https://doi.org/10.1016/j.heares.2010.08.001>
- Anokhin, A. P. (2014). Genetic psychophysiology: Advances, problems, and future directions. *International Journal of Psychophysiology, 93*(2), 173-197.
<https://doi.org/10.1016/j.ijpsycho.2014.04.003>
- Arenillas-Alcón, S., Costa-Faidella, J., Ribas-Prats, T., Gómez-Roig, M. D., & Escera, C. (2021). Neural encoding of voice pitch and formant structure at birth as revealed by frequency-following responses. *Scientific Reports, 11*(1), 6660. <https://doi.org/10.1038/s41598-021-85799-x>
- Ashraf, F., & Najam, N. (2020). An epidemiological study of prevalence and comorbidity of non-clinical Dyslexia, Dysgraphia and Dyscalculia symptoms in Public and Private Schools of Pakistan. *Pakistan Journal of Medical Sciences, 36*(7), 1659-1663.
<https://doi.org/10.12669/pjms.36.7.2486>
- Asilador, A., & Llano, D. A. (2021). Top-Down Inference in the Auditory System: Potential Roles for Corticofugal Projections. *Frontiers in Neural Circuits, 14*.
<https://www.frontiersin.org/articles/10.3389/fncir.2020.615259>
- Ayala, Y. A., & Malmierca, M. S. (2018). The effect of inhibition on stimulus-specific adaptation in the inferior colliculus. *Brain Structure & Function, 223*(3), 1391-1407.
<https://doi.org/10.1007/s00429-017-1546-4>
- Banai, K. (2005). Brainstem Timing: Implications for Cortical Processing and Literacy. *Journal of Neuroscience, 25*(43), 9850-9857. <https://doi.org/10.1523/JNEUROSCI.2373-05.2005>

- Banai, K., Abrams, D., & Kraus, N. (2007). Sensory-based learning disability: Insights from brainstem processing of speech sounds. *International Journal of Audiology*, 46(9), 524-532. <https://doi.org/10.1080/14992020701383035>
- Banai, K., Hornickel, J., Skoe, E., Nicol, T., Zecker, S., & Kraus, N. (2009). Reading and Subcortical Auditory Function. *Cerebral Cortex*, 19(11), 2699-2707. <https://doi.org/10.1093/cercor/bhp024>
- Basu, M., Krishnan, A., & Weber-Fox, C. (2010). Brainstem correlates of temporal auditory processing in children with specific language impairment: Brainstem correlates of temporal processing. *Developmental Science*, 13(1), 77-91. <https://doi.org/10.1111/j.1467-7687.2009.00849.x>
- Bell, S. L., Smith, D. C., Allen, R., & Lutman, M. E. (2004). Recording the middle latency response of the auditory evoked potential as a measure of depth of anaesthesia. A technical note. *British Journal of Anaesthesia*, 92(3), 442-445. <https://doi.org/10.1093/bja/ae074>
- Benz, S., Sellaro, R., Hommel, B., & Colzato, L. S. (2016). Music Makes the World Go Round: The Impact of Musical Training on Non-musical Cognitive Functions—A Review. *Frontiers in Psychology*, 6. <https://www.frontiersin.org/articles/10.3389/fpsyg.2015.02023>
- Bialystok, E. (2011). Reshaping the Mind: The Benefits of Bilingualism. *Canadian journal of experimental psychology = Revue canadienne de psychologie experimentale*, 65(4), 229-235. <https://doi.org/10.1037/a0025406>
- Bidelman, G. M. (2015). Multichannel recordings of the human brainstem frequency-following response: Scalp topography, source generators, and distinctions from the transient ABR. *Hearing Research*, 323, 68-80. <https://doi.org/10.1016/j.heares.2015.01.011>
- Bidelman, G. M. (2017). Communicating in Challenging Environments: Noise and Reverberation. En N. Kraus, S. Anderson, T. White-Schwoch, R. R. Fay, & A. N. Popper

- (Eds.), *The Frequency-Following Response* (Vol. 61, pp. 193-224). Springer International Publishing. https://doi.org/10.1007/978-3-319-47944-6_8
- Bidelman, G. M. (2018). Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. *NeuroImage*, 175, 56-69.
<https://doi.org/10.1016/j.neuroimage.2018.03.060>
- Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Musicians demonstrate experience-dependent brainstem enhancement of musical scale features within continuously gliding pitch. *Neuroscience Letters*, 503(3), 203-207.
<https://doi.org/10.1016/j.neulet.2011.08.036>
- Bidelman, G. M., & Krishnan, A. (2010). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Research*, 1355, 112-125.
<https://doi.org/10.1016/j.brainres.2010.07.100>
- Blanca, M. J., Alarcón, R., & Arnau, J. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, 29.4, 552-557. <https://doi.org/10.7334/psicothema2016.383>
- Blanca, M. J., Arnau, J., García-Castro, F. J., Alarcón, R., & Bono, R. (2023). Non-normal Data in Repeated Measures ANOVA: Impact on Type I Error and Power. *Psicothema*, 35.1, 21-29. <https://doi.org/10.7334/psicothema2022.292>
- Boebinger, D., Evans, S., Rosen, S., Lima, C. F., Manly, T., & Scott, S. K. (2015). Musicians and non-musicians are equally adept at perceiving masked speech. *The Journal of the Acoustical Society of America*, 137(1), 378-387. <https://doi.org/10.1121/1.4904537>
- Boersma, P., & Weenink, D. (2022). *Praat: Doing phonetics by computer* (6.2.09) [Windows]. University of Amsterdam. <http://www.praat.org>
- Bonacina, S., Otto-Meyer, S., Krizman, J., White-Schwoch, T., Nicol, T., & Kraus, N. (2019). Stable auditory processing underlies phonological awareness in typically developing preschoolers. *Brain and Language*, 197, 104664.
<https://doi.org/10.1016/j.bandl.2019.104664>

- Brewer, C. C., Zalewski, C. K., King, K. A., Zobay, O., Riley, A., Ferguson, M. A., Bird, J. E., McCabe, M. M., Hood, L. J., Drayna, D., Griffith, A. J., Morell, R. J., Friedman, T. B., & Moore, D. R. (2016). Heritability of non-speech auditory processing skills. *European Journal of Human Genetics*, 24(8), 1137-1144. <https://doi.org/10.1038/ejhg.2015.277>
- Cervera, T., & González-Alvarez, J. (2010). Lists of Spanish Sentences with Equivalent Predictability, Phonetic Content, Length, and Frequency of the Last Word. *Perceptual and Motor Skills*, 111(2), 517-529. <https://doi.org/10.2466/28.PMS.111.5.517-529>
- Cervera, T., & González-Alvarez, J. (2011). Test of Spanish sentences to measure speech intelligibility in noise conditions. *Behavior Research Methods*, 43(2), 459-467. <https://doi.org/10.3758/s13428-011-0063-2>
- 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(3), 311-319. <https://doi.org/10.1016/j.neuron.2009.10.006>
- Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, 4, 7. <https://doi.org/10.1186/s13742-015-0047-8>
- Chung, P. J., Patel, D. R., & Nizami, I. (2020). Disorder of written expression and dysgraphia: Definition, diagnosis, and management. *Translational Pediatrics*, 9(Suppl 1), S46-S54. <https://doi.org/10.21037/tp.2019.11.01>
- Coffey, E. B. J., Chepesiuk, A. M. P., Herholz, S. C., Baillet, S., & Zatorre, R. J. (2017). Neural Correlates of Early Sound Encoding and their Relationship to Speech-in-Noise Perception. *Frontiers in Neuroscience*, 11, 479. <https://doi.org/10.3389/fnins.2017.00479>

- Coffey, E. B. J., Herholz, S. C., Chepesiuk, A. M. P., Baillet, S., & Zatorre, R. J. (2016). Cortical contributions to the auditory frequency-following response revealed by MEG. *Nature Communications*, 7(1), 11070. <https://doi.org/10.1038/ncomms11070>
- Coffey, E. B. J., Mogilever, N. B., & Zatorre, R. J. (2017). Speech-in-noise perception in musicians: A review. *Hearing Research*, 352, 49-69. <https://doi.org/10.1016/j.heares.2017.02.006>
- Coffey, E. B. J., Nicol, T., White-Schwoch, T., Chandrasekaran, B., Krizman, J., Skoe, E., Zatorre, R. J., & Kraus, N. (2019). Evolving perspectives on the sources of the frequency-following response. *Nature Communications*, 10(1), 5036. <https://doi.org/10.1038/s41467-019-13003-w>
- Cuetos Vega, F., Arribas Águila, D., & Ramos Sánchez, J. L. (2016). *PROLEC-SE-R. Bateria de evaluación de los procesos lectores en Secundaria y Bachillerato, revisada: Manual* (TEA Ediciones, S.A.). TEA Ediciones. <http://hdl.handle.net/10651/54843>
- Cunningham, J., Nicol, T., King, C., Zecker, S. G., & Kraus, N. (2002). Effects of noise and cue enhancement on neural responses to speech in auditory midbrain, thalamus and cortex. *Hearing Research*, 169(1), 97-111. [https://doi.org/10.1016/S0378-5955\(02\)00344-1](https://doi.org/10.1016/S0378-5955(02)00344-1)
- Cunningham, J., Nicol, T., Zecker, S. G., Bradlow, A., & Kraus, N. (2001). Neurobiologic responses to speech in noise in children with learning problems: Deficits and strategies for improvement. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 112(5), 758-767. [https://doi.org/10.1016/s1388-2457\(01\)00465-5](https://doi.org/10.1016/s1388-2457(01)00465-5)
- Cutini, S., Szűcs, D., Mead, N., Huss, M., & Goswami, U. (2016). Atypical right hemisphere response to slow temporal modulations in children with developmental dyslexia. *Neuroimage*, 143, 40-49. <https://doi.org/10.1016/j.neuroimage.2016.08.012>

- David, O., Kilner, J. M., & Friston, K. J. (2006). Mechanisms of evoked and induced responses in MEG/EEG. *NeuroImage*, 31(4), 1580-1591.
<https://doi.org/10.1016/j.neuroimage.2006.02.034>
- De Vos, A., Vanvooren, S., Ghesquière, P., & Wouters, J. (2020). Subcortical auditory neural synchronization is deficient in pre-reading children who develop dyslexia. *Developmental Science*, 23(6), e12945. <https://doi.org/10.1111/desc.12945>
- De Vos, A., Vanvooren, S., Vanderauwera, J., Ghesquière, P., & Wouters, J. (2017a). A longitudinal study investigating neural processing of speech envelope modulation rates in children with (a family risk for) dyslexia. *Cortex*, 93, 206-219.
<https://doi.org/10.1016/j.cortex.2017.05.007>
- De Vos, A., Vanvooren, S., Vanderauwera, J., Ghesquière, P., & Wouters, J. (2017b). Atypical neural synchronization to speech envelope modulations in dyslexia. *Brain and Language*, 164, 106-117. <https://doi.org/10.1016/j.bandl.2016.10.002>
- Deutsch, D. (2013). Absolute pitch. En *The psychology of music*, 3rd ed (pp. 141-182). Elsevier Academic Press. <https://doi.org/10.1016/B978-0-12-381460-9.00005-5>
- Dias, B. G., & Ressler, K. J. (2014). Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nature Neuroscience*, 17(1), 89-96.
<https://doi.org/10.1038/nn.3594>
- Don, M., Ponton, C. W., Eggermont, J. J., & Masuda, A. (1993). Gender differences in cochlear response time: An explanation for gender amplitude differences in the unmasked auditory brain-stem response. *The Journal of the Acoustical Society of America*, 94(4), 2135-2148. <https://doi.org/10.1121/1.407485>
- Drayna, D., Manichaikul, A., Lange, M. D., Snieder, H., & Spector, T. (2001). Genetic Correlates of Musical Pitch Recognition in Humans. *Science*, 291(5510), 1969-1972.
<https://doi.org/10.1126/science.291.5510.1969>

- Drennan, W. R. (2008). Music perception in cochlear implant users and its relationship with psychophysical capabilities. *The Journal of Rehabilitation Research and Development*, 45(5), 779-790. <https://doi.org/10.1682/JRRD.2007.08.0118>
- EBF2 EBF transcription factor 2 [Homo sapiens (human)]—Gene—NCBI*. (s. f.). Recuperado 11 de febrero de 2024, de <https://www.ncbi.nlm.nih.gov/gene/64641>
- Eggermont, J. J., & Don, M. (1982). Analysis of click-evoked brainstem auditory electric potentials using high-pass noise masking and its clinical application. *Annals of the New York Academy of Sciences*, 388, 471-486. <https://doi.org/10.1111/j.1749-6632.1982.tb50810.x>
- Elhezzani, N. S. (2018). Improved estimation of SNP heritability using Bayesian multiple-phenotype models. *European Journal of Human Genetics*, 26(5), Article 5. <https://doi.org/10.1038/s41431-018-0100-z>
- Elhilali, M. (2019). Modulation Representations for Speech and Music. En K. Siedenburg, C. Saitis, S. McAdams, A. N. Popper, & R. R. Fay (Eds.), *Timbre: Acoustics, Perception, and Cognition* (Vol. 69, pp. 335-359). Springer International Publishing. https://doi.org/10.1007/978-3-030-14832-4_12
- Erb, J., Henry, M. J., Eisner, F., & Obleser, J. (2012). Auditory skills and brain morphology predict individual differences in adaptation to degraded speech. *Neuropsychologia*, 50(9), 2154-2164. <https://doi.org/10.1016/j.neuropsychologia.2012.05.013>
- Erb, J., Henry, M. J., Eisner, F., & Obleser, J. (2013). The brain dynamics of rapid perceptual adaptation to adverse listening conditions. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 33(26), 10688-10697. <https://doi.org/10.1523/JNEUROSCI.4596-12.2013>
- Farahani, E. D., Wouters, J., & van Wieringen, A. (2021). Brain mapping of auditory steady-state responses: A broad view of cortical and subcortical sources. *Human Brain Mapping*, 42(3), 780-796. <https://doi.org/10.1002/hbm.25262>

- Finn, E. S., Shen, X., Holahan, J. M., Scheinost, D., Lacadie, C., Papademetris, X., Shaywitz, S. E., Shaywitz, B. A., & Constable, R. T. (2014). Disruption of functional networks in dyslexia: A whole-brain, data-driven analysis of connectivity. *Biological psychiatry*, 76(5), 397-404. <https://doi.org/10.1016/j.biopsych.2013.08.031>
- Forte, A. E., Etard, O., & Reichenbach, T. (2017). The human auditory brainstem response to running speech reveals a subcortical mechanism for selective attention. *eLife*, 6, e27203. <https://doi.org/10.7554/eLife.27203>
- Foy, J. G., & Mann, V. A. (2014). Bilingual children show advantages in nonverbal auditory executive function task. *International Journal of Bilingualism*, 18(6), 717-729. <https://doi.org/10.1177/1367006912472263>
- François, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex (New York, N.Y.: 1991)*, 23(9), 2038-2043. <https://doi.org/10.1093/cercor/bhs180>
- Fujioka, T., Ross, B., Kakigi, R., Pantev, C., & Trainor, L. J. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain: A Journal of Neurology*, 129(Pt 10), 2593-2608. <https://doi.org/10.1093/brain/awl247>
- Ghitza, O., & Greenberg, S. (2009). On the possible role of brain rhythms in speech perception: Intelligibility of time-compressed speech with periodic and aperiodic insertions of silence. *Phonetica*, 66(1-2), 113-126. <https://doi.org/10.1159/000208934>
- Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511-517. <https://doi.org/10.1038/nn.3063>
- Goldstein, E. B. (2010). *Sensation and perception* (8. ed., internat. ed). Thomson Wadsworth Belmont, Calif.
- Gorina-Careta, N., Kurkela, J. L. O., Hämäläinen, J., Astikainen, P., & Escera, C. (2021). Neural generators of the frequency-following response elicited to stimuli of low and high

- frequency: A magnetoencephalographic (MEG) study. *NeuroImage*, 231, 117866.
<https://doi.org/10.1016/j.neuroimage.2021.117866>
- Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, 15(1), 3-10. <https://doi.org/10.1016/j.tics.2010.10.001>
- Goswami, U. (2015). Sensory theories of developmental dyslexia: Three challenges for research. *Nature Reviews Neuroscience*, 16(1), 43-54. <https://doi.org/10.1038/nrn3836>
- Goswami, U. (2019). Speech rhythm and language acquisition: An amplitude modulation phase hierarchy perspective. *Annals of the New York Academy of Sciences*, 1453(1), 67-78.
<https://doi.org/10.1111/nyas.14137>
- Goswami, U., Thomson, J., Richardson, U., Stainthorpe, R., Hughes, D., Rosen, S., & Scott, S. K. (2002). Amplitude envelope onsets and developmental dyslexia: A new hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, 99(16), 10911-10916. <https://doi.org/10.1073/pnas.122368599>
- Gourévitch, B., Martin, C., Postal, O., & Eggermont, J. J. (2020). Oscillations in the auditory system and their possible role. *Neuroscience & Biobehavioral Reviews*, 113, 507-528.
<https://doi.org/10.1016/j.neubiorev.2020.03.030>
- Gregersen, P. K., Kowalsky, E., Lee, A., Baron-Cohen, S., Fisher, S. E., Asher, J. E., Ballard, D., Freudenberg, J., & Li, W. (2013). Absolute pitch exhibits phenotypic and genetic overlap with synesthesia. *Human Molecular Genetics*, 22(10), 2097-2104.
<https://doi.org/10.1093/hmg/ddt059>
- Groussard, M., Viader, F., Landeau, B., Desgranges, B., Eustache, F., & Platel, H. (2014). The effects of musical practice on structural plasticity: The dynamics of grey matter changes. *Brain and Cognition*, 90, 174-180.
<https://doi.org/10.1016/j.bandc.2014.06.013>
- Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., Pallesen, J., Agerbo, E., Andreassen, O. A., Anney, R., Awashti, S., Belliveau, R., Bettella, F., Buxbaum, J. D.,

- Bybjerg-Grauholm, J., Bækvad-Hansen, M., Cerrato, F., Chambert, K., Christensen, J. H., ... Børglum, A. D. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics*, 51(3), Article 3. <https://doi.org/10.1038/s41588-019-0344-8>
- GTEx Portal*. (s. f.-a). Recuperado 11 de febrero de 2024, de <https://www.gtexportal.org/home/gene/EBF2>
- GTEx Portal*. (s. f.-b). Recuperado 11 de febrero de 2024, de <https://www.gtexportal.org/home/gene/RBFOX1>
- GTEx Portal*. (s. f.-c). Recuperado 11 de febrero de 2024, de <https://www.gtexportal.org/home/gene/OPALIN>
- Gui, Y., Zhou, X., Wang, Z., Zhang, Y., Wang, Z., Zhou, G., Zhao, Y., Liu, M., Lu, H., & Zhao, H. (2022). Sex-specific genetic association between psychiatric disorders and cognition, behavior and brain imaging in children and adults. *Translational Psychiatry*, 12(1), Article 1. <https://doi.org/10.1038/s41398-022-02041-6>
- Guo, L., Li, P. hui, Li, H., Colicino, E., Colicino, S., Wen, Y., Zhang, R., Feng, X., Barrow, T. M., Cayir, A., Baccarelli, A. A., & Byun, H. M. (2017). Effects of environmental noise exposure on DNA methylation in the brain and metabolic health. *Environmental Research*, 153, 73-82. <https://doi.org/10.1016/j.envres.2016.11.017>
- Gustavson, D. E., Coleman, P. L., Wang, Y., Nitin, R., Petty, L. E., Bush, C. T., Mosing, M. A., Wesseldijk, L. W., Ullén, F., 23 and Me Research Team, Below, J. E., Cox, N. J., & Gordon, R. L. (2023). Exploring the genetics of rhythmic perception and musical engagement in the Vanderbilt Online Musicality Study. *Annals of the New York Academy of Sciences*, 1521(1), 140-154. <https://doi.org/10.1111/nyas.14964>
- GWAS Catalog*. (s. f.). Recuperado 11 de febrero de 2024, de <https://www.ebi.ac.uk/gwas/genes/RBFOX1>

- Hagenbeek, F. A., Hirzinger, J. S., Breunig, S., Bruins, S., Kuznetsov, D. V., Schut, K., Odintsova, V. V., & Boomsma, D. I. (2023). Maximizing the value of twin studies in health and behaviour. *Nature Human Behaviour*, 7(6), 849-860. <https://doi.org/10.1038/s41562-023-01609-6>
- Handel, S. (1995). Chapter 12—Timbre Perception and Auditory Object Identification. En B. C. J. Moore (Ed.), *Hearing* (pp. 425-461). Academic Press. <https://doi.org/10.1016/B978-012505626-7/50014-5>
- Hatoum, A. S., Morrison, C. L., Mitchell, E. C., Lam, M., Benca-Bachman, C. E., Reineberg, A. E., Palmer, R. H. C., Evans, L. M., Keller, M. C., & Friedman, N. P. (2023). Genome-wide Association Study Shows That Executive Functioning Is Influenced by GABAergic Processes and Is a Neurocognitive Genetic Correlate of Psychiatric Disorders. *Biological Psychiatry*, 93(1), 59-70. <https://doi.org/10.1016/j.biopsych.2022.06.034>
- Hornickel, J., & Kraus, N. (2013). Unstable Representation of Sound: A Biological Marker of Dyslexia. *Journal of Neuroscience*, 33(8), 3500-3504. <https://doi.org/10.1523/JNEUROSCI.4205-12.2013>
- Hornickel, J., Skoe, E., Nicol, T., Zecker, S., & Kraus, N. (2009). Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *Proceedings of the National Academy of Sciences of the United States of America*, 106(31), 13022-13027. <https://doi.org/10.1073/pnas.0901123106>
- Hu, X., Oroszi, G., Chun, J., Smith, T. L., Goldman, D., & Schuckit, M. A. (2005). An expanded evaluation of the relationship of four alleles to the level of response to alcohol and the alcoholism risk. *Alcoholism, Clinical and Experimental Research*, 29(1), 8-16. <https://doi.org/10.1097/01.alc.0000150008.68473.62>
- Hurley, L. M., & Hall, I. C. (2011). Context-dependent modulation of auditory processing by serotonin. *Hearing Research*, 279(1), 74-84. <https://doi.org/10.1016/j.heares.2010.12.015>

- Ignatious, E., Azam, S., Jonkman, M., & De Boer, F. (2023). Frequency and Time Domain Analysis of EEG Based Auditory Evoked Potentials to Detect Binaural Hearing in Noise. *Journal of Clinical Medicine*, 12(13), 4487. <https://doi.org/10.3390/jcm12134487>
- Islam, T. U., & Abbas, E. (2022). *Validity of ANOVA under Non-normality & Heterogeneity* [Preprint]. In Review. <https://doi.org/10.21203/rs.3.rs-2071136/v1>
- Iurescia, S., Seripa, D., & Rinaldi, M. (2017). Looking Beyond the 5-HTTLPR Polymorphism: Genetic and Epigenetic Layers of Regulation Affecting the Serotonin Transporter Gene Expression. *Molecular Neurobiology*, 54(10), 8386-8403. <https://doi.org/10.1007/s12035-016-0304-6>
- Jafari, Z., Villeneuve, C., Thompson, J., & Koravand, A. (2023). Subcortical plasticity and enhanced neural synchrony in multilingual adults. *Bilingualism: Language and Cognition*, 26(2), 307-316. <https://doi.org/10.1017/S1366728922000645>
- James, C. E., Altenmüller, E., Kliegel, M., Krüger, T. H. C., Van De Ville, D., Worschech, F., Abdili, L., Scholz, D. S., Jünemann, K., Hering, A., Grouiller, F., Sinke, C., & Marie, D. (2020). Train the brain with music (TBM): Brain plasticity and cognitive benefits induced by musical training in elderly people in Germany and Switzerland, a study protocol for an RCT comparing musical instrumental practice to sensitization to music. *BMC Geriatrics*, 20(1), 418. <https://doi.org/10.1186/s12877-020-01761-y>
- James, C. E., Oechslin, M. S., Van De Ville, D., Hauert, C.-A., Descloux, C., & Lazeyras, F. (2014). Musical training intensity yields opposite effects on grey matter density in cognitive versus sensorimotor networks. *Brain Structure and Function*, 219(1), 353-366. <https://doi.org/10.1007/s00429-013-0504-z>
- Jeng, F.-C., Peris, K. S., Hu, J., & Lin, C.-D. (2013). Evaluation of an Automated Procedure for Detecting Frequency-Following Responses in American and Chinese Neonates. *Perceptual and Motor Skills*, 116(2), 456-465. <https://doi.org/10.2466/24.10.PMS.116.2.456-465>

- Joris, P. X., Schreiner, C. E., & Rees, A. (2004). Neural Processing of Amplitude-Modulated Sounds. *Physiological Reviews*, 84(2), 541-577.
<https://doi.org/10.1152/physrev.00029.2003>
- Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *The Journal of the Acoustical Society of America*, 61(5), 1337-1351.
<https://doi.org/10.1121/1.381436>
- Kay, L. M., Beshel, J., Brea, J., Martin, C., Rojas-Líbano, D., & Kopell, N. (2009). Olfactory oscillations: The what, how and what for. *Trends in Neurosciences*, 32(4), 207-214.
<https://doi.org/10.1016/j.tins.2008.11.008>
- King, C., Warrier, C. M., Hayes, E., & Kraus, N. (2002). Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. *Neuroscience Letters*, 319(2), 111-115. [https://doi.org/10.1016/S0304-3940\(01\)02556-3](https://doi.org/10.1016/S0304-3940(01)02556-3)
- Koelsch, S. (2011). Toward a Neural Basis of Music Perception – A Review and Updated Model. *Frontier in Psychology*, 2. <https://doi.org/10.3389/fpsyg.2011.00110>
- Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, 11(8), Article 8. <https://doi.org/10.1038/nrn2882>
- Kraus, N., Skoe, E., Parbery-Clark, A., & Ashley, R. (2009). Experience-induced Malleability in Neural Encoding of Pitch, Timbre, and Timing. *Annals of the New York Academy of Sciences*, 1169, 543-557. <https://doi.org/10.1111/j.1749-6632.2009.04549.x>
- Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research*, 25(1), 161-168.
<https://doi.org/10.1016/j.cogbrainres.2005.05.004>
- Krizman, J., Bonacina, S., & Kraus, N. (2019). Sex differences in subcortical auditory processing emerge across development. *Hearing Research*, 380, 166-174.
<https://doi.org/10.1016/j.heares.2019.07.002>

- Krizman, J., Bonacina, S., & Kraus, N. (2020). Sex differences in subcortical auditory processing only partially explain higher prevalence of language disorders in males. *Hearing Research, 398*, 108075. <https://doi.org/10.1016/j.heares.2020.108075>
- Krizman, J., & Kraus, N. (2019). Analyzing the FFR: A tutorial for decoding the richness of auditory function. *Hearing Research, 382*, 107779. <https://doi.org/10.1016/j.heares.2019.107779>
- Krizman, J., Marian, V., Shook, A., Skoe, E., & Kraus, N. (2012a). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of the National Academy of Sciences, 109*(20), 7877-7881. <https://doi.org/10.1073/pnas.1201575109>
- Krizman, J., Marian, V., Shook, A., Skoe, E., & Kraus, N. (2012b). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of the National Academy of Sciences, 109*(20), Article 20. <https://doi.org/10.1073/pnas.1201575109>
- Krizman, J., Marian, V., Shook, A., Skoe, E., & Kraus, N. (2012c). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of the National Academy of Sciences, 109*(20), 7877-7881. <https://doi.org/10.1073/pnas.1201575109>
- Krizman, J., Skoe, E., & Kraus, N. (2012). Sex differences in auditory subcortical function. *Clinical Neurophysiology, 123*(3), 590-597. <https://doi.org/10.1016/j.clinph.2011.07.037>
- Krizman, J., Skoe, E., Marian, V., & Kraus, N. (2014). Bilingualism increases neural response consistency and attentional control: Evidence for sensory and cognitive coupling. *Brain and Language, 128*(1), 34-40. <https://doi.org/10.1016/j.bandl.2013.11.006>
- Krizman, J., Tierney, A., Nicol, T., & Kraus, N. (2021). Listening in the Moment: How Bilingualism Interacts With Task Demands to Shape Active Listening. *Frontiers in Neuroscience, 15*, 717572. <https://doi.org/10.3389/fnins.2021.717572>

- Kunert, R., Willems, R. M., & Hagoort, P. (2016). An Independent Psychometric Evaluation of the PROMS Measure of Music Perception Skills. *PLOS ONE*, *11*(7), e0159103.
<https://doi.org/10.1371/journal.pone.0159103>
- Lahey, B. B., Tong, L., Pierce, B., Hedeker, D., Berman, M. G., Cardenas-Iniguez, C., Moore, T. M., Applegate, B., Tiemeier, H., & Kaczkurkin, A. N. (2022). Associations of polygenic risk for attention-deficit/hyperactivity disorder with general and specific dimensions of childhood psychological problems and facets of impulsivity. *Journal of Psychiatric Research*, *152*, 187-193. <https://doi.org/10.1016/j.jpsychires.2022.06.019>
- Lam, S. S.-Y., White-Schwoch, T., Zecker, S. G., Hornickel, J., & Kraus, N. (2017). Neural stability: A reflection of automaticity in reading. *Neuropsychologia*, *103*, 162-167.
<https://doi.org/10.1016/j.neuropsychologia.2017.07.023>
- Lambert, S. A., Gil, L., Jupp, S., Ritchie, S. C., Xu, Y., Buniello, A., McMahon, A., Abraham, G., Chapman, M., Parkinson, H., Danesh, J., MacArthur, J. A. L., & Inouye, M. (2021). The Polygenic Score Catalog as an open database for reproducibility and systematic evaluation. *Nature Genetics*, *53*(4), Article 4. <https://doi.org/10.1038/s41588-021-00783-5>
- Lamminmäki, S., Parkkonen, L., & Hari, R. (2014). Human Neuromagnetic Steady-State Responses to Amplitude-Modulated Tones, Speech, and Music. *Ear & Hearing*, *35*(4), 461-467. <https://doi.org/10.1097/AUD.0000000000000033>
- Lavezzi, A. M., Pusioli, T., & Matturri, L. (2015). Cytoarchitectural and functional abnormalities of the inferior colliculus in sudden unexplained perinatal death. *Medicine*, *94*(6), e487. <https://doi.org/10.1097/MD.0000000000000487>
- Law, L. N. C., & Zentner, M. (2012). Assessing Musical Abilities Objectively: Construction and Validation of the Profile of Music Perception Skills. *PLoS ONE*, *7*(12), e52508.
<https://doi.org/10.1371/journal.pone.0052508>

- Lee, D. J., Chen, Y., & Schlaug, G. (2003). Corpus callosum: Musician and gender effects. *NeuroReport*, 14(2), 205.
- Lee, K. M., Skoe, E., Kraus, N., & Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 29(18), 5832-5840.
<https://doi.org/10.1523/JNEUROSCI.6133-08.2009>
- Leeuw, C. A. de, Mooij, J. M., Heskes, T., & Posthuma, D. (2015). MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLOS Computational Biology*, 11(4), e1004219.
<https://doi.org/10.1371/journal.pcbi.1004219>
- Lehongre, K., Morillon, B., Giraud, A.-L., & Ramus, F. (2013). Impaired auditory sampling in dyslexia: Further evidence from combined fMRI and EEG. *Frontiers in Human Neuroscience*, 7. <https://www.frontiersin.org/articles/10.3389/fnhum.2013.00454>
- Lehongre, K., Ramus, F., Villiermet, N., Schwartz, D., & Giraud, A.-L. (2011). Altered low- γ sampling in auditory cortex accounts for the three main facets of dyslexia. *Neuron*, 72(6), 1080-1090. <https://doi.org/10.1016/j.neuron.2011.11.002>
- Lenzenweger, M. F. (2013). Endophenotype, intermediate phenotype, biomarker: Definitions, concept comparisons, clarifications. *Depression and Anxiety*, 30(3), 185-189.
<https://doi.org/10.1002/da.22042>
- Li, M. Y. C., Braze, D., Kukona, A., Johns, C. L., Tabor, W., Van Dyke, J. A., Mencl, W. E., Shankweiler, D. P., Pugh, K. R., & Magnuson, J. S. (2019). Individual differences in subphonemic sensitivity and phonological skills. *Journal of Memory and Language*, 107, 195-215. <https://doi.org/10.1016/j.jml.2019.03.008>
- Li, P., Legault, J., & Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: Anatomical changes in the human brain. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 58, 301-324.
<https://doi.org/10.1016/j.cortex.2014.05.001>

- Li, X., & Jeng, F.-C. (2011). Noise tolerance in human frequency-following responses to voice pitch. *The Journal of the Acoustical Society of America*, 129(1), EL21-26.
<https://doi.org/10.1121/1.3528775>
- Lizarazu, M., Lallier, M., Molinaro, N., Bourguignon, M., Paz-Alonso, P. M., Lerma-Usabiaga, G., & Carreiras, M. (2015). Developmental evaluation of atypical auditory sampling in dyslexia: Functional and structural evidence. *Human Brain Mapping*, 36(12), 4986-5002. <https://doi.org/10.1002/hbm.22986>
- Lorenzi, C., Dumont, A., & Füllgrabe, C. (2000). Use of temporal envelope cues by children with developmental dyslexia. *Journal of Speech, Language, and Hearing Research: JSLHR*, 43(6), 1367-1379. <https://doi.org/10.1044/jslhr.4306.1367>
- Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, 54(6), 1001-1010.
<https://doi.org/10.1016/j.neuron.2007.06.004>
- Macaskill, M., Omidvar, S., & Koravand, A. (2022). Long Latency Auditory Evoked Responses in the Identification of Children With Central Auditory Processing Disorders: A Scoping Review. *Journal of Speech, Language, and Hearing Research: JSLHR*, 65(9), 3595-3619.
https://doi.org/10.1044/2022_JSLHR-21-00544
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research*, 50(4), 940-967.
[https://doi.org/10.1044/1092-4388\(2007/067\)](https://doi.org/10.1044/1092-4388(2007/067))
- Mayhew, A. J., & Meyre, D. (2017). Assessing the Heritability of Complex Traits in Humans: Methodological Challenges and Opportunities. *Current Genomics*, 18(4).
<https://doi.org/10.2174/1389202918666170307161450>

- McGrath, L. M., & Stoodley, C. J. (2019). Are there shared neural correlates between dyslexia and ADHD? A meta-analysis of voxel-based morphometry studies. *Journal of Neurodevelopmental Disorders*, 11(1), 31. <https://doi.org/10.1186/s11689-019-9287-8>
- Mohades, S. G., Struys, E., Van Schuerbeek, P., Mondt, K., Van De Craen, P., & Luypaert, R. (2012). DTI reveals structural differences in white matter tracts between bilingual and monolingual children. *Brain Research*, 1435, 72-80. <https://doi.org/10.1016/j.brainres.2011.12.005>
- Moore, B. C. J. (2019). The roles of temporal envelope and fine structure information in auditory perception. *Acoustical Science and Technology*, 40(2), 61-83. <https://doi.org/10.1250/ast.40.61>
- Morell, R. J., Brewer, C. C., Ge, D., Snieder, H., Zalewski, C. K., King, K. A., Drayna, D., & Friedman, T. B. (2007). A twin study of auditory processing indicates that dichotic listening ability is a strongly heritable trait. *Human Genetics*, 122(1), 103-111. <https://doi.org/10.1007/s00439-007-0384-5>
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-Term Music Training Enhances Verbal Intelligence and Executive Function. *Psychological science*, 22(11), 1425-1433. <https://doi.org/10.1177/0956797611416999>
- Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical Training Influences Linguistic Abilities in 8-Year-Old Children: More Evidence for Brain Plasticity. *Cerebral Cortex*, 19(3), 712-723. <https://doi.org/10.1093/cercor/bhn120>
- Morillon, B., & Schroeder, C. E. (2015). Neuronal oscillations as a mechanistic substrate of auditory temporal prediction. *Annals of the New York Academy of Sciences*, 1337(1), 26-31. <https://doi.org/10.1111/nyas.12629>
- Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National*

Academy of Sciences, 104(40), 15894-15898.

<https://doi.org/10.1073/pnas.0701498104>

Musacchia, G., Strait, D., & Kraus, N. (2008). Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hearing Research*, 241(1), 34-42. <https://doi.org/10.1016/j.heares.2008.04.013>

Navarro, L., Martín-Torres, F., & Salas, A. (2021). Sensogenomics and the Biological Background Underlying Musical Stimuli: Perspectives for a New Era of Musical Research. *Genes*, 12(9), 1454. <https://doi.org/10.3390/genes12091454>

Nayak, M., Das, D., Pradhan, J., Ahmed, R. G., Laureano-Melo, R., & Dandapat, J. (2022). Epigenetic signature in neural plasticity: The journey so far and journey ahead. *Heliyon*, 8(12), e12292. <https://doi.org/10.1016/j.heliyon.2022.e12292>

Niarchou, M., Gustavson, D. E., Sathirapongsasuti, J. F., Anglada-Tort, M., Eising, E., Bell, E., McArthur, E., Straub, P., 23andMe Research Team, Aslibekyan, S., Auton, A., Bell, R. K., Bryc, K., Clark, S. K., Elson, S. L., Fletez-Brant, K., Fontanillas, P., Furlotte, N. A., Gandhi, P. M., ... Gordon, R. L. (2022). Genome-wide association study of musical beat synchronization demonstrates high polygenicity. *Nature Human Behaviour*, 6(9), 1292-1309. <https://doi.org/10.1038/s41562-022-01359-x>

Nourski, K. V., & Brugge, J. F. (2011). Representation of temporal sound features in the human auditory cortex. *Revneuro*, 22(2), 187-203. <https://doi.org/10.1515/rns.2011.016>

Obara, N., Kamiya, H., & Fukuda, S. (2014). Serotonergic modulation of inhibitory synaptic transmission in mouse inferior colliculus. *Biomedical Research*, 35(1), 81-84. <https://doi.org/10.2220/biomedres.35.81>

Oikkonen, J., Huang, Y., Onkamo, P., Ukkola-Vuoti, L., Raijas, P., Karma, K., Vieland, V. J., & Järvelä, I. (2015). A genome-wide linkage and association study of musical aptitude identifies loci containing genes related to inner ear development and neurocognitive functions. *Molecular Psychiatry*, 20(2), 275-282. <https://doi.org/10.1038/mp.2014.8>

- Olszewska, A. M., Gaca, M., Herman, A. M., Jednoróg, K., & Marchewka, A. (2021). How Musical Training Shapes the Adult Brain: Predispositions and Neuroplasticity. *Frontiers in Neuroscience*, 15. <https://www.frontiersin.org/articles/10.3389/fnins.2021.630829>
- Otto-Meyer, S., Krizman, J., White-Schwoch, T., & Kraus, N. (2018). Children with autism spectrum disorder have unstable neural responses to sound. *Experimental Brain Research*, 236(3), 733-743. <https://doi.org/10.1007/s00221-017-5164-4>
- Oxenham, A. J. (2018). How We Hear: The Perception and Neural Coding of Sound. *Annual review of psychology*, 69, 27-50. <https://doi.org/10.1146/annurev-psych-122216-011635>
- Parbery-Clark, A., Anderson, S., Hittner, E., & Kraus, N. (2012). Musical experience strengthens the neural representation of sounds important for communication in middle-aged adults. *Frontiers in Aging Neuroscience*, 4. <https://doi.org/10.3389/fnagi.2012.00030>
- Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical Experience Limits the Degradative Effects of Background Noise on the Neural Processing of Sound. *The Journal of Neuroscience*, 29(45), 14100-14107. <https://doi.org/10.1523/JNEUROSCI.3256-09.2009>
- Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician Enhancement for Speech-In-Noise. *Ear & Hearing*, 30(6), 653-661. <https://doi.org/10.1097/AUD.0b013e3181b412e9>
- Parbery-Clark, A., Strait, D. L., & Kraus, N. (2011). Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia*, 49(12), 3338-3345. <https://doi.org/10.1016/j.neuropsychologia.2011.08.007>
- Park, H., Lee, S., Kim, H.-J., Ju, Y. S., Shin, J.-Y., Hong, D., Von Grotthuss, M., Lee, D.-S., Park, C., Kim, J. H., Kim, B., Yoo, Y. J., Cho, S.-I., Sung, J., Lee, C., Kim, J.-I., & Seo, J.-S. (2012). Comprehensive genomic analyses associate *UGT8* variants with musical ability in a

- Mongolian population. *Journal of Medical Genetics*, 49(12), 747-752.
<https://doi.org/10.1136/jmedgenet-2012-101209>
- Paulraj, M. P., Subramaniam, K., Yaccob, S. B., Adom, A. H. B., & Hema, C. R. (2015). Auditory Evoked Potential Response and Hearing Loss: A Review. *The Open Biomedical Engineering Journal*, 9(1), 17-24. <https://doi.org/10.2174/1874120701509010017>
- Pelham, S. D., & Abrams, L. (2014). Cognitive advantages and disadvantages in early and late bilinguals. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 40(2), 313-325. <https://doi.org/10.1037/a0035224>
- Plack, C. J. (2005). *The Sense of Hearing* (1.^a ed.). Psychology Press.
<https://doi.org/10.4324/9781410613745>
- Plourde, G. (2006). Auditory evoked potentials. *Best Practice & Research Clinical Anaesthesiology*, 20(1), 129-139. <https://doi.org/10.1016/j.bpa.2005.07.012>
- Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time'. *Speech Communication*, 41(1), 245-255.
[https://doi.org/10.1016/S0167-6393\(02\)00107-3](https://doi.org/10.1016/S0167-6393(02)00107-3)
- Poeppel, D., Idsardi, W. J., & van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1493), 1071-1086. <https://doi.org/10.1098/rstb.2007.2160>
- Poncet, A., Courvoisier, D. S., Combescure, C., & Perneger, T. V. (2016). Normality and Sample Size Do Not Matter for the Selection of an Appropriate Statistical Test for Two-Group Comparisons. *Methodology*, 12(2), 61-71. <https://doi.org/10.1027/1614-2241/a000110>
- Prévost, F., Laroche, M., Marcoux, A. M., & Dajani, H. R. (2013). Objective measurement of physiological signal-to-noise gain in the brainstem response to a synthetic vowel. *Clinical Neurophysiology*, 124(1), 52-60. <https://doi.org/10.1016/j.clinph.2012.05.009>

- Privé, F., Aschard, H., Carmi, S., Folkersen, L., Hoggart, C., O'Reilly, P. F., & Vilhjálmsdóttir, B. J. (2022). Portability of 245 polygenic scores when derived from the UK Biobank and applied to 9 ancestry groups from the same cohort. *The American Journal of Human Genetics*, 109(1), 12-23. <https://doi.org/10.1016/j.ajhg.2021.11.008>
- Protopapas, A. (2014). From temporal processing to developmental language disorders: Mind the gap. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1634), 20130090. <https://doi.org/10.1098/rstb.2013.0090>
- Rajendran, V. G., Teki, S., & Schnupp, J. W. H. (2018). Temporal Processing in Audition: Insights from Music. *Neuroscience*, 389, 4-18. <https://doi.org/10.1016/j.neuroscience.2017.10.041>
- Rasmussen, K. M. B., West, N. C., Bille, M., Sandvej, M. G., & Cayé-Thomasen, P. (2022). Cochlear Implantation Improves Both Speech Perception and Patient-Reported Outcomes: A Prospective Follow-Up Study of Treatment Benefits among Adult Cochlear Implant Recipients. *Journal of Clinical Medicine*, 11(8), 2257. <https://doi.org/10.3390/jcm11082257>
- RBFOX1* RNA binding fox-1 homolog 1 [*Homo sapiens* (human)]—Gene—NCBI. (s. f.). Recuperado 11 de febrero de 2024, de <https://www.ncbi.nlm.nih.gov/gene/54715>
- Ressel, V., Pallier, C., Ventura-Campos, N., Díaz, B., Roessler, A., Ávila, C., & Sebastián-Gallés, N. (2012). An Effect of Bilingualism on the Auditory Cortex. *The Journal of Neuroscience*, 32(47), 16597-16601. <https://doi.org/10.1523/JNEUROSCI.1996-12.2012>
- Rey, V., De Martino, S., Espesser, R., & Habib, M. (2002). Temporal processing and phonological impairment in dyslexia: Effect of phoneme lengthening on order judgment of two consonants. *Brain and Language*, 80(3), 576-591. <https://doi.org/10.1006/brln.2001.2618>
- Ribas-Prats, T., Almeida, L., Costa-Faidella, J., Plana, M., Corral, M. J., Gómez-Roig, M. D., & Escera, C. (2019). The frequency-following response (FFR) to speech stimuli: A

- normative dataset in healthy newborns. *Hearing Research*, 371, 28-39.
<https://doi.org/10.1016/j.heares.2018.11.001>
- Robles, L., & Ruggero, M. A. (2001). Mechanics of the mammalian cochlea. *Physiological Reviews*, 81(3), 1305-1352. <https://doi.org/10.1152/physrev.2001.81.3.1305>
- Román-Caballero, R., Arnedo, M., Triviño, M., & Lupiáñez, J. (2018). Musical practice as an enhancer of cognitive function in healthy aging—A systematic review and meta-analysis. *PLOS ONE*, 13(11), e0207957. <https://doi.org/10.1371/journal.pone.0207957>
- Rosen, S. (1992). Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 336(1278), 367-373. <https://doi.org/10.1098/rstb.1992.0070>
- Rosenthal, M. A. (2020). A systematic review of the voice-tagging hypothesis of speech-in-noise perception. *Neuropsychologia*, 136, 107256.
<https://doi.org/10.1016/j.neuropsychologia.2019.107256>
- Russo, N. M., Skoe, E., Trommer, B., Nicol, T., Zecker, S., Bradlow, A., & Kraus, N. (2008). Deficient brainstem encoding of pitch in children with Autism Spectrum Disorders. *Clinical Neurophysiology*, 119(8), 1720-1731.
<https://doi.org/10.1016/j.clinph.2008.01.108>
- Russo, N., Nicol, T., Musacchia, G., & Kraus, N. (2004). Brainstem responses to speech syllables. *Clinical Neurophysiology*, 115(9), 2021-2030.
<https://doi.org/10.1016/j.clinph.2004.04.003>
- Sachs, M. B. (1984). Neural Coding of Complex Sounds: Speech. *Annual Review of Physiology*, 46(1), 261-273. <https://doi.org/10.1146/annurev.ph.46.030184.001401>
- Sachs, M. B., Young, E. D., Winslow, R. L., & Shofner, W. P. (1986). Some Aspects of Rate Coding in the Auditory Nerve. En B. C. J. Moore & R. D. Patterson (Eds.), *Auditory Frequency Selectivity* (pp. 121-128). Springer US. https://doi.org/10.1007/978-1-4613-2247-4_14

- Sato, H., Sando, I., & Takahashi, H. (1991). Sexual dimorphism and development of the human cochlea. Computer 3-D measurement. *Acta Oto-Laryngologica*, 111(6), 1037-1040.
<https://doi.org/10.3109/00016489109100753>
- Satterlee, J. S., Beckel-Mitchener, A., Little, A. R., Procaccini, D., Rutter, J. L., & Lossie, A. C. (2015). Neuroepigenomics: Resources, obstacles, and opportunities. *Neuroepigenetics*, 1, 2-13. <https://doi.org/10.1016/j.nepig.2014.10.001>
- Schlaug, G., Jäncke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia*, 33(8), 1047-1055.
[https://doi.org/10.1016/0028-3932\(95\)00045-5](https://doi.org/10.1016/0028-3932(95)00045-5)
- Schlaug, G., Jäncke, L., Huang, Y., & Steinmetz, H. (1995). In Vivo Evidence of Structural Brain Asymmetry in Musicians. *Science*, 267(5198), 699-701.
<https://doi.org/10.1126/science.7839149>
- Schochat, E., Rocha-Muniz, C. N., & Filippini, R. (2017). Understanding Auditory Processing Disorder Through the FFR. En N. Kraus, S. Anderson, T. White-Schwoch, R. R. Fay, & A. N. Popper (Eds.), *The Frequency-Following Response: A Window into Human Communication* (pp. 225-250). Springer International Publishing.
https://doi.org/10.1007/978-3-319-47944-6_9
- Sebastián-Gallés, N., & Soto-Faraco, S. (1999). Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition*, 72(2), 111-123.
[https://doi.org/10.1016/S0010-0277\(99\)00024-4](https://doi.org/10.1016/S0010-0277(99)00024-4)
- Selinger, L., Zarnowiec, K., Via, M., Clemente, I. C., & Escera, C. (2016a). Involvement of the Serotonin Transporter Gene in Accurate Subcortical Speech Encoding. *The Journal of Neuroscience*, 36(42), 10782-10790. <https://doi.org/10.1523/JNEUROSCI.1595-16.2016>
- Selinger, L., Zarnowiec, K., Via, M., Clemente, I. C., & Escera, C. (2016b). Involvement of the Serotonin Transporter Gene in Accurate Subcortical Speech Encoding. *The Journal of Neuroscience*, 36(42), Article 42. <https://doi.org/10.1523/JNEUROSCI.1595-16.2016>

- Shettigar, N., Yang, C.-L., Tu, K.-C., & Suh, C. S. (2022). On The Biophysical Complexity of Brain Dynamics: An Outlook. *Dynamics*, 2(2), Article 2.
<https://doi.org/10.3390/dynamics2020006>
- Shi, L.-F. (2010). Perception of Acoustically Degraded Sentences in Bilingual Listeners Who Differ in Age of English Acquisition. *Journal of Speech, Language, and Hearing Research*, 53(4), 821-835. [https://doi.org/10.1044/1092-4388\(2010/09-0081\)](https://doi.org/10.1044/1092-4388(2010/09-0081))
- Skipper, J. I. (2022). A voice without a mouth no more: The neurobiology of language and consciousness. *Neuroscience and Biobehavioral Reviews*, 140, 104772.
<https://doi.org/10.1016/j.neubiorev.2022.104772>
- 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.
<https://doi.org/10.1016/j.neuroscience.2017.02.049>
- Skoe, E., García-Sierra, A., Ramírez-Esparza, N., & Jiang, S. (2022). Automatic sound encoding is sensitive to language familiarity: Evidence from English monolinguals and Spanish-English bilinguals. *Neuroscience Letters*, 777, 136582.
<https://doi.org/10.1016/j.neulet.2022.136582>
- Slater, J., Skoe, E., Strait, D. L., O'Connell, S., Thompson, E., & Kraus, N. (2015). Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program. *Behavioural Brain Research*, 291, 244-252.
<https://doi.org/10.1016/j.bbr.2015.05.026>
- Smalt, C. J., Krishnan, A., Bidelman, G. M., Ananthakrishnan, S., & Gandour, J. T. (2012). Distortion products and their influence on representation of pitch-relevant information in the human brainstem for unresolved harmonic complex tones. *Hearing research*, 292(1-2), 26-34. <https://doi.org/10.1016/j.heares.2012.08.001>
- Smit, D. J. A., Boomsma, D. I., Schnack, H. G., Hulshoff Pol, H. E., & De Geus, E. J. C. (2012). Individual Differences in EEG Spectral Power Reflect Genetic Variance in Gray and

- White Matter Volumes. *Twin Research and Human Genetics*, 15(3), 384-392.
<https://doi.org/10.1017/thg.2012.6>
- Smit, D. J. A., Wright, M. J., Meyers, J. L., Martin, N. G., Ho, Y. Y. W., Malone, S. M., Zhang, J., Burwell, S. J., Chorlian, D. B., De Geus, E. J. C., Denys, D., Hansell, N. K., Hottenga, J., McGue, M., Van Beijsterveldt, C. E. M., Jahanshad, N., Thompson, P. M., Whelan, C. D., Medland, S. E., ... Boomsma, D. I. (2018). Genome-wide association analysis links multiple psychiatric liability genes to oscillatory brain activity. *Human Brain Mapping*, 39(11), 4183-4195. <https://doi.org/10.1002/hbm.24238>
- Smith, A. R., Kwon, J. H., Navarro, M., & Hurley, L. M. (2014). Acoustic trauma triggers upregulation of serotonin receptor genes. *Hearing research*, 315, 40-48.
<https://doi.org/10.1016/j.heares.2014.06.004>
- Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 416(6876), 87-90. <https://doi.org/10.1038/416087a>
- Snowling, M. J., Nash, H. M., Gooch, D. C., Hayiou-Thomas, M. E., Hulme, C., & Wellcome Language and Reading Project Team. (2019). Developmental Outcomes for Children at High Risk of Dyslexia and Children With Developmental Language Disorder. *Child Development*, 90(5), e548-e564. <https://doi.org/10.1111/cdev.13216>
- Sonck, R., Vanthornhout, J., Bonin, E., & Francart, T. (2023). *Auditory steady-state responses: Multiplexed amplitude modulation frequencies to reduce recording time* [Preprint]. Neuroscience. <https://doi.org/10.1101/2023.06.01.543200>
- Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2011). Perception of Speech in Noise: Neural Correlates. *Journal of Cognitive Neuroscience*, 23(9), 2268-2279.
<https://doi.org/10.1162/jocn.2010.21556>
- Song, J. H., Skoe, E., Wong, P. C. M., & Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of cognitive neuroscience*, 20(10), 1892-1902. <https://doi.org/10.1162/jocn.2008.20131>

- Spanoudis, G. C., Papadopoulos, T. C., & Spyrou, S. (2019). Specific Language Impairment and Reading Disability: Categorical Distinction or Continuum? *Journal of Learning Disabilities*, 52(1), 3-14. <https://doi.org/10.1177/0022219418775111>
- Su, Y., & Delgutte, B. (2020). Robust Rate-Place Coding of Resolved Components in Harmonic and Inharmonic Complex Tones in Auditory Midbrain. *The Journal of Neuroscience*, 40(10), 2080-2093. <https://doi.org/10.1523/JNEUROSCI.2337-19.2020>
- Suresh, C. H., Krishnan, A., & Luo, X. (2019). Auditory Frequency-Following Responses to Vocoded Speech: Amplitude Modulation Versus Amplitude Plus Frequency Modulation. En *Encyclopedia of Computational Neuroscience* (pp. 1-13). Springer. https://doi.org/10.1007/978-1-4614-7320-6_100689-1
- Swaminathan, J., & Heinz, M. G. (2012). Psychophysiological Analyses Demonstrate the Importance of Neural Envelope Coding for Speech Perception in Noise. *Journal of Neuroscience*, 32(5), 1747-1756. <https://doi.org/10.1523/JNEUROSCI.4493-11.2012>
- Talge, N. M., Adkins, M., Kileny, P. R., & Frownfelter, I. (2022). Click-evoked auditory brainstem responses and autism spectrum disorder: A meta-analytic investigation of disorder specificity. *Pediatric Research*, 92(1), Article 1. <https://doi.org/10.1038/s41390-021-01730-0>
- Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language*, 9(2), 182-198. [https://doi.org/10.1016/0093-934x\(80\)90139-x](https://doi.org/10.1016/0093-934x(80)90139-x)
- Tallal, P. (2004). Improving language and literacy is a matter of time. *Nature Reviews Neuroscience*, 5(9), 721-728. <https://doi.org/10.1038/nrn1499>
- Tallal, P., & Gaab, N. (2006). Dynamic auditory processing, musical experience and language development. *Trends in Neurosciences*, 29(7), 382-390. <https://doi.org/10.1016/j.tins.2006.06.003>

- Tallal, P., Miller, S., & Fitch, R. H. (1993). Neurobiological Basis of Speech: A Case for the Preeminence of Temporal Processing. *Annals of the New York Academy of Sciences*, 682(1 Temporal Info), 27-47. <https://doi.org/10.1111/j.1749-6632.1993.tb22957.x>
- Tanigawa, Y., Qian, J., Venkataraman, G., Justesen, J. M., Li, R., Tibshirani, R., Hastie, T., & Rivas, M. A. (2022). Significant sparse polygenic risk scores across 813 traits in UK Biobank. *PLOS Genetics*, 18(3), e1010105. <https://doi.org/10.1371/journal.pgen.1010105>
- Theusch, E., Basu, A., & Gitschier, J. (2009). Genome-wide Study of Families with Absolute Pitch Reveals Linkage to 8q24.21 and Locus Heterogeneity. *The American Journal of Human Genetics*, 85(1), 112-119. <https://doi.org/10.1016/j.ajhg.2009.06.010>
- Theusch, E., & Gitschier, J. (2011). Absolute Pitch Twin Study and Segregation Analysis. *Twin Research and Human Genetics*, 14(2), 173-178. <https://doi.org/10.1375/twin.14.2.173>
- Thompson, E. C., Carr, K. W., White-Schwoch, T., Otto-Meyer, S., & Kraus, N. (2017). Individual differences in speech-in-noise perception parallel neural speech processing and attention in preschoolers. *Hearing research*, 344, 148-157. <https://doi.org/10.1016/j.heares.2016.11.007>
- Tschentscher, N., Ruisinger, A., Blank, H., Díaz, B., & Von Kriegstein, K. (2019). Reduced structural connectivity between left auditory thalamus and the motion-sensitive planum temporale in developmental dyslexia. *The Journal of Neuroscience*, 1435-18. <https://doi.org/10.1523/JNEUROSCI.1435-18.2018>
- Ukkola, L. T., Onkamo, P., Raijas, P., Karma, K., & Järvelä, I. (2009). Musical Aptitude Is Associated with AVPR1A-Haplotypes. *PLoS ONE*, 4(5), e5534. <https://doi.org/10.1371/journal.pone.0005534>
- Ullén, F., Mosing, M. A., Holm, L., Eriksson, H., & Madison, G. (2014). Psychometric properties and heritability of a new online test for musicality, the Swedish Musical Discrimination Test. *Personality and Individual Differences*, 63, 87-93. <https://doi.org/10.1016/j.paid.2014.01.057>

- Vanden Bempt, F., Van Herck, S., Economou, M., Vanderauwera, J., Vandermosten, M., Wouters, J., & Ghesquière, P. (2022). Speech perception deficits and the effect of envelope-enhanced story listening combined with phonics intervention in pre-readers at risk for dyslexia. *Frontiers in Psychology, 13*, 1021767.
<https://doi.org/10.3389/fpsyg.2022.1021767>
- Vassoler, F. M., White, S. L., Schmidt, H. D., Sadri-Vakili, G., & Pierce, R. C. (2013). Epigenetic inheritance of a cocaine-resistance phenotype. *Nature Neuroscience, 16*(1), 42-47.
<https://doi.org/10.1038/nn.3280>
- Vihla, M., Kiviniemi, K., & Salmelin, R. (2002). Auditory cortical activation in Finnish and Swedish speaking Finns: A magnetoencephalographic study. *Neuroscience Letters, 322*(3), 141-144. [https://doi.org/10.1016/S0304-3940\(01\)02517-4](https://doi.org/10.1016/S0304-3940(01)02517-4)
- Vuust, P., Heggli, O. A., Friston, K. J., & Kringelbach, M. L. (2022). Music in the brain. *Nature Reviews Neuroscience, 23*(5), Article 5. <https://doi.org/10.1038/s41583-022-00578-5>
- Wang, X.-J. (2010). Neurophysiological and Computational Principles of Cortical Rhythms in Cognition. *Physiological Reviews, 90*(3), 1195-1268.
<https://doi.org/10.1152/physrev.00035.2008>
- Wankerl, M., Miller, R., Kirschbaum, C., Hennig, J., Stalder, T., & Alexander, N. (2014). Effects of genetic and early environmental risk factors for depression on serotonin transporter expression and methylation profiles. *Translational Psychiatry, 4*(6), e402.
<https://doi.org/10.1038/tp.2014.37>
- Ward, L. M. (2003). Synchronous neural oscillations and cognitive processes. *Trends in Cognitive Sciences, 7*(12), 553-559. <https://doi.org/10.1016/j.tics.2003.10.012>
- Warrier, C. M., Abrams, D. A., Nicol, T. G., & Kraus, N. (2011). Inferior colliculus contributions to phase encoding of stop consonants in an animal model. *Hearing Research, 282*(1-2), 108-118. <https://doi.org/10.1016/j.heares.2011.09.001>

- Watanabe, K., Taskesen, E., van Bochoven, A., & Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nature Communications*, 8(1), Article 1. <https://doi.org/10.1038/s41467-017-01261-5>
- Watanabe, K., Umičević Mirkov, M., de Leeuw, C. A., van den Heuvel, M. P., & Posthuma, D. (2019). Genetic mapping of cell type specificity for complex traits. *Nature Communications*, 10(1), Article 1. <https://doi.org/10.1038/s41467-019-11181-1>
- Wesseldijk, L. W., Abdellaoui, A., Gordon, R. L., 23andMe Research Team, Aslibekyan, S., Auton, A., Babalola, E., Bell, R. K., Bielenberg, J., Bryc, K., Bullis, E., Coker, D., Partida, G. C., Dhamija, D., Das, S., Elson, S. L., Filshtein, T., Fletez-Brant, K., Fontanillas, P., ... Mosing, M. A. (2022). Using a polygenic score in a family design to understand genetic influences on musicality. *Scientific Reports*, 12(1), 14658. <https://doi.org/10.1038/s41598-022-18703-w>
- Wesseldijk, L. W., Ullén, F., & Mosing, M. A. (2023). Music and Genetics. *Neuroscience & Biobehavioral Reviews*, 152, 105302. <https://doi.org/10.1016/j.neubiorev.2023.105302>
- White-Schwoch, T., & Kraus, N. (2013). Physiologic discrimination of stop consonants relates to phonological skills in pre-readers: A biomarker for subsequent reading ability? *Frontiers in Human Neuroscience*, 7, 899. <https://doi.org/10.3389/fnhum.2013.00899>
- White-Schwoch, T., Woodruff Carr, K., Thompson, E. C., Anderson, S., Nicol, T., Bradlow, A. R., Zecker, S. G., & Kraus, N. (2015). Auditory Processing in Noise: A Preschool Biomarker for Literacy. *PLOS Biology*, 13(7), e1002196. <https://doi.org/10.1371/journal.pbio.1002196>
- Willcutt, E. G., McGrath, L. M., Pennington, B. F., Keenan, J. M., DeFries, J. C., Olson, R. K., & Wadsworth, S. J. (2019). Understanding Comorbidity Between Specific Learning Disabilities. *New Directions for Child and Adolescent Development*, 2019(165), 91-109. <https://doi.org/10.1002/cad.20291>

- Wilsch, A., & Obleser, J. (2016). What works in auditory working memory? A neural oscillations perspective. *Brain Research*, 1640, 193-207.
<https://doi.org/10.1016/j.brainres.2015.10.054>
- Xu, Z., Chen, Z., Shen, T., Chen, L., Tan, T., Gao, C., Chen, B., Yuan, Y., & Zhang, Z. (2022). The impact of HTR1A and HTR1B methylation combined with stress/genotype on early antidepressant efficacy. *Psychiatry and Clinical Neurosciences*, 76(2), 51-57.
<https://doi.org/10.1111/pcn.13314>
- Yang, J., Zeng, J., Goddard, M. E., Wray, N. R., & Visscher, P. M. (2017). Concepts, estimation and interpretation of SNP-based heritability. *Nature Genetics*, 49(9), 1304-1310.
<https://doi.org/10.1038/ng.3941>
- Yengo, L., Vedantam, S., Marouli, E., Sidorenko, J., Bartell, E., Sakaue, S., Graff, M., Eliassen, A. U., Jiang, Y., Raghavan, S., Miao, J., Arias, J. D., Graham, S. E., Mukamel, R. E., Spracklen, C. N., Yin, X., Chen, S.-H., Ferreira, T., Highland, H. H., ... Hirschhorn, J. N. (2022). A saturated map of common genetic variants associated with human height. *Nature*, 610(7933), Article 7933. <https://doi.org/10.1038/s41586-022-05275-y>
- Yuan, H., Voelcker-Rehage, C., Huebner, L., & Godde, B. (2017). *Resting State EEG Classification for Motor Learning Skills Using Echo State Networks*.
<https://doi.org/10.13140/RG.2.2.13507.50726>
- Zentner, M., & Strauss, H. (2017). Assessing musical ability quickly and objectively: Development and validation of the Short-PROMS and the Mini-PROMS: Assessing musical ability. *Annals of the New York Academy of Sciences*, 1400(1), 33-45.
<https://doi.org/10.1111/nyas.13410>
- Zhang, X., & Gong, Q. (2019). Frequency-Following Responses to Complex Tones at Different Frequencies Reflect Different Source Configurations. *Frontiers in Neuroscience*, 13, 130.
<https://doi.org/10.3389/fnins.2019.00130>

Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental Science*, 12(5), 732-745.
<https://doi.org/10.1111/j.1467-7687.2009.00817.x>

ANNEX

Table Supplementary 1: Spearman's correlations between FFR parameters

	F0 SpecAmp	F0 SNR	F1 /o/ SpecAmp	F1 /o/ SNR	F1 /a/ SpecAmp	F1 /a/ SNR	Cross /a/ flat	Cross /a/ rising	PitchS /a/ flat	PitchS /a/ rising	PitchE /a/ flat	PitchE /a/ rising	Neural lag
F0 SpecAmp	-												
F0 SNR	0.805***	-											
F1 /o/ SpecAmp	-0.072	-0.175	-										
F1 /o/ SNR	-0.06	-0.004	0.680***	-									
F1 /a/ SpecAmp	-0.004	-0.023	0.111	0.008	-								
F1 /a/ SNR	-0.044	0.019	0.097	0.146	0.726***	-							
Cross /a/ flat	0.348***	0.357***	-0.0817	0.043	0.108	0.053	-						
Cross /a/ rising	-0.271**	-0.152	0.026	-0.039	0.031	-0.061	0.052	-					
PitchS /a/ flat	0.467***	0.479***	0.151	0.243**	0.134	0.202*	0.274**	-0.149	-				
PitchS /a/ rising	0.0468***	0.478***	0.1557	0.244**	0.142	0.208*	0.273**	-0.152	0.998***	-			
PitchE /a/ flat	-0.176	-0.251**	-0.314***	-0.252**	-0.065	-0.186*	-0.152	-0.034	-0.752***	-0.749***	-		
PitchE /a/ rising	-0.177	-0.241**	-0.318***	-0.256**	-0.067	-0.182	-0.152	-0.052	-0.751***	-0.748***	0.997***	-	
Neural lag	-0.218*	-0.128	0.136	-0.197*	0.002	0.095	-0.094	0.078	0.143	0.144	-0.164	-0.158	-

*P<0.05, ** p<0.01, ***p<0.001



UNIVERSITAT DE
BARCELONA

Oficina de Gestió de la Recerca
Pavelló Rosa (recinte Maternitat) primer pis
Travessera de les Corts, 131-159 93-4035398
08028 Barcelona

**Comisión de Bioética
de la Universitat de Barcelona**

**Certificado de aprobación
Sobre experimentación en humanos o en muestras de origen humano**

Don Albert Royes Qui, Secretario de la Comisión de Bioética de la Universitat de Barcelona.

CERTIFICA:

Que el Dr. **Marc Via García** presentó el proyecto titulado "Paper del Gens i de factors epigenètics en la modulació de potencials cerebrals associats al processament del llenguatge i de la música (GEPLEM)".

La Comisión de Bioética de la Universitat de Barcelona analizó toda la documentación presentada por el Dr. **Marc Via García** y, por acuerdo de fecha 16 de noviembre de 2018, aprobó informar favorablemente desde el punto de vista bioético el proyecto de investigación de referencia.

Y para que conste y a los efectos que corresponda, firmo este documento con el visto bueno del presidente de la Comisión en Barcelona a 16 de noviembre de 2018


Universitat de Barcelona
Comissió de Bioètica

Vº Bº El presidente de la Comisión de Bioética de la
Universitat de Barcelona



**UNIVERSITAT DE
BARCELONA**
Oficina de Gestió de la Recerca
Domènec Espriu Climent
Vicepresident de Recerca