Universidade do Minho
Escola de Psicologia

Margarida Fátima Gomes Vasconcelos

The effects of musical training on auditory statistical learning: electrophysiological and behavioral evidence.

Março 2019

UMinho | 2019

Margarida Fátima Gomes Vasconcelos

The effects of musical training on auditory statistical learning: electrophysiological and behavioral evidence.

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA

POCH
QREN QUADRO DE REFERÊNCIA ESTRATÉGICO NACIONAL PORTUGAL 2007.2013
GOVERNO DA REPÚBLICA PORTUGUESA
UNIÃO EUROPEIA
Fundo Social Europeu

Cofinanciado por:
COMPETE 2020
PORTUGAL 2020
UNIÃO EUROPEIA
Fundo Europeu de Desenvolvimento Regional
aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Universidade do Minho
Escola de Psicologia

Margarida Fátima Gomes Vasconcelos

The effects of musical training on auditory statistical learning: electrophysiological and behavioral evidence.

Tese de Doutoramento
em Psicologia Básica

Trabalho efetuado sob a orientação da
Professora Doutora Ana Patrícia Teixeira Pinheiro

Março 2019

Statement of Integrity

I hereby declare having conducted my thesis with integrity. I confirm that I have not used plagiarism or any form of falsification of results in the process of the thesis elaboration. I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

University of Minho, March 19th 2019

Full Name: Margarida Fátima Gomes Vasconcelos

Signature: Margarida F. Gomes Vasconcelos

To my beloved parents.

AGRADECIMENTOS

Este trabalho é fruto de um longo processo criativo e de investigação. Foi um processo de vários anos, que envolveu muita dedicação, questões e procura por soluções e que, por isso, se refletiu em crescimento académico e pessoal. Não teria sido, portanto, possível sem os contributos de inúmeras pessoas que comigo se cruzaram e o marcaram em definitivo. Faz, pois, sentido marcá-las a elas também neste espaço, agradecendo-lhes sinceramente.

À minha orientadora científica, Professora Ana Pinheiro: por me permitir aprender consigo há já quase uma década; por me desafiar a perseguir as minhas ideias; pela partilha de conhecimento; pelo estímulo ao rigor; pelo desafio ao pensamento crítico e pelo apoio académico e científico sem os quais este trabalho não se materializaria.

Ao Laboratório de Neurociências Psicológicas, na pessoa da Professora Adriana Sampaio, por me acolher, proporcionando-me as condições para desenvolver os meus estudos e permitindo-me partilhar espaço e debate com investigadores fantásticos. Agradeço-lhe a enorme disponibilidade e todo o apoio.

Aos meus colegas do Laboratório, pela partilha de ideias, discussão, aprendizagem e ambiente genial que fazem do Laboratório um local tão acolhedor. Um agradecimento especial à Diana, à Paula, à Carla e à Carol por serem parceiras incríveis ao longo destes anos e pelo apoio constante e alegre. À Carla, ao Alberto, ao Dudi e ao Diego, por todo o tempo e conhecimento partilhado que foi fundamental para mim.

Aos meus colegas de Doutoramento, àqueles que comigo se cruzaram no âmbito do doutoramento – Inês e Carlos – e aos Professores Óscar Gonçalves e Armando Machado, por serem pessoas inspiradoras com quem pude aprender muito durante este percurso académico.

À Fundação Portuguesa para a Ciência e a Tecnologia, pela bolsa de doutoramento que permitiu a concretização deste trabalho.

À Escola de Música da Universidade do Minho, na pessoa do Professor Ângelo Martingo, por ter sido uma importante aliada no processo de recrutamento de participantes e recolha de dados.

A todos os participantes dos meus estudos, em particular aos músicos, que muito alheios à investigação no domínio da Psicologia e das Neurociências acederam a colaborar.

To the Association for Psychological Science (APS), for awarding me a student grant that has contributed to the realisation of this project.

Aos meus amigos, Nocas, Tatiana, Rocky, Fábio, Ana e Diogo, por serem únicos e fundamentais (que sorte a minha!). Pelo vosso carinho enorme, pela vossa alegria e força, e por tantos momentos de partilha. Por saber-vos sempre lá.

Aos meus queridos pais: por serem uma presença constante e o meu porto seguro; pelo tanto que investem, acreditam e vibram comigo; por serem os meus pés na terra mas também a motivação para voar; pela compreensão infinita e o estímulo permanente; pelo amor constante que é o maior dos motores de concretização. Devo-vos o mais daquilo que hoje sou e nem estas nem outras palavras chegarão perto de transmitir o sentimento imensurável de gratidão que me acompanha por vos ter. Obrigada!

A ti, querido Marcelo, meu companheiro de viagem. Obrigada por ocupares esse lugar irrepetível e por, tornando-o feliz, contribuíres decisivamente para o que aqui se encerra. Obrigada pelo tanto que o teu amor, amizade, honestidade e inquietação me deram. Obrigada... Por seres uma imensa fonte de luz.

This work was funded by the Portuguese Foundation for Science and Technology (FCT, Portugal), with the doctoral Grant with the reference SFRH/BD/52400/2013 and by the Association for Psychological Science (APS, USA) with one of the 2017 student grants attributed in the realm of the student grant competition. The scientific research reported here was conducted at the Psychology Research Centre (UID/PSI/01662/2013), University of Minho, and supported by the Portuguese Foundation for Science and Technology and the Ministry of Science, Technology and Higher Education through national funds and co-funded by FEDER through COMPETE202 under the PT2020 Partnership Agreement (POCI-01-0145-FEDER-007653).

Cofinanciado por:

The effects of musical training on auditory statistical learning: electrophysiological and behavioral evidence.

ABSTRACT

Musical training is an experience-driven model of neuroplasticity that allows studying the consequences of long-term multi-modal training on the brain. Research has shown that training is associated with structural and functional brain changes that bring advantages to auditory processing and may enhance specific cognitive functions. The literature has showed transfer effects from musical training to speech processing, namely prosody processing, word learning and speech segmentation. Yet, it is still not clear which brain mechanisms are modified by musical training contributing to musicians' enhanced segmentation and learning of speech sounds. One hypothesis is that training promotes the ability to compute regularities and extract patterns from the acoustic environment – auditory statistical learning (ASL) –, which, in turn, may facilitate the processing of speech.

The studies reported in this Dissertation aimed at examining the effects of musical training on ASL abilities. We attempted to clarify if musical training facilitates the ASL of different types of auditory sequences (i.e., linguistic vs. musical), with or without musical characteristics (i.e., within vs. cross-domain effects), with distinct levels of complexity (e.g., simple vs. more complex structures), at different levels of processing (i.e., learning vs. recognition), with or without the influence of attention (i.e., attentively vs. pre-attentively) at neural and behavioral levels. In order to fully investigate the neural dynamics of ASL mechanisms with high temporal precision, the Event-Related Potentials (ERP) technique was applied.

Chapter 3 explored the neural correlates of the pre-attentive processing of statistical regularities in pure tones in musicians and non-musicians. There were no differences between groups in the mismatch negativity (MMN) amplitude to pitch deviants. These findings suggest that musicians and non-musicians have the same sensitivity to detect deviances based on pitch modulations, which indicate that, at least at a pre-attentive level, musicians and non-musicians compute simple statistical regularities similarly. Chapter 4 investigated the role of musical training on the attentive ASL of different types of auditory stimuli. Three ERP experiments, each composed of learning, implicit test and explicit behavioral test phases were

designed to evaluate the ASL of three types of auditory sequences: prosodic (i.e., words with melodic contour), non-prosodic (i.e., words with flat contour) and musical (i.e., tri-tone piano melodies). During learning, musicians evidenced an enhanced negativity in the 250-300 ms latency window in response to prosodic words and an increased positivity in the first 100 ms post-melodies' onset. With exception of the responses to the prosodic stream, musicians and non-musicians neural responses to violations of the streams' structure were similar. At the behavioral level, only musicians demonstrated successful learning across the three experiments. These findings suggest that training has an effect on the neural dynamics underlying ASL and on the capacity to learn from the regularities of distinct types of acoustic sequences, evidencing cross-domain neuroplasticity effects. Chapter 5 investigated the effects of training on ASL when the auditory stimulation is outside the focus of attention. The statistical learning of prosodic, non-prosodic and musical sequences was tested while participants were attending a visual task. While the learning unfolded, the groups did not differ in their ERP responses to both types of linguistic sequences but musicians showed larger positivities in two distinct latency windows while processing the musical stream. Musicians' and non-musicians' ERP responses to violations only differed when participants were processing melodies. Behaviorally, the groups showed similar learning performances. The results suggest that participants demonstrate statistical learning of speech regularities even when the auditory input is outside the focus of attention. Thus, when the auditory input is task-irrelevant, musical training is not advantageous to ASL but it affects the way the brain computes regularities among musical sequences (i.e., within-domain neuroplasticity).

Generally, the ERP and behavioral results of our studies clarify the extent and the conditions under which musical training has effects on ASL. The findings suggest that training modifies the attentive processing of regularities independently of stimulus-type, which is evidence for cross-domain neuroplasticity. However, they also indicate that when ASL proceeds outside the focus of attention training effects are only observable for musical streams. This Dissertation provides relevant evidence on basic auditory learning mechanisms and on how they can be modified by musical training, contributing to shed light on the brain dynamics of music-to-speech transfer effects.

Os efeitos do treino musical sobre a aprendizagem estatística auditiva: evidência electrofisiológica e comportamental.

RESUMO

O treino musical é considerado um modelo de neuroplasticidade que possibilita o estudo dos efeitos do treino de longo-prazo sobre o cérebro. A investigação tem mostrado que o treino musical está associado a mudanças na estrutura e função cerebrais que acarretam vantagens para o processamento auditivo e que podem beneficiar determinadas funções cognitivas. A literatura demonstrou que existem efeitos de transferência positivos do treino musical para o processamento de discurso, nomeadamente o processamento da prosódia, a aprendizagem de palavras e a segmentação de discurso. Contudo, ainda não são claros quais os mecanismos cerebrais que são modificados pelo treino e que contribuem para que os músicos evidenciem uma maior facilidade na aprendizagem e segmentação do discurso. Uma hipótese explicativa baseia-se na ideia de que o treino beneficia o cálculo de regularidades e a abstração de padrões do ambiente acústico – aprendizagem estatística auditiva (AEA) -, o que, em consequência, pode facilitar o processamento do discurso.

Os estudos desenvolvidos no âmbito desta Dissertação examinaram os efeitos do treino musical sobre a AEA. Procurou-se clarificar se o treino musical facilita a AEA de diferentes tipos de sequências auditivas (i.e., linguísticas vs. musicais), com ou sem características melódicas, com diferentes níveis de complexidade (i.e., estruturas simples vs. complexas), em diferentes níveis de processamento (i.e., aprendizagem vs. reconhecimento), com ou sem a influência da atenção, a nível comportamental e cerebral. Para investigar os mecanismos neuronais envolvidos na AEA com elevada precisão temporal foi usada a técnica dos potenciais evocados (PE).

O Capítulo 3 explorou os correlatos neuronais do processamento pré-atencional de regularidades estatísticas com base em sons puros. A amplitude da *mismatch negativity* não diferiu entre músicos e não-músicos embora a sua latência tendesse a ser mais precoce nos músicos. Estes resultados revelam que o treino musical parece não influenciar a sensibilidade para a detecção de desvios acústicos, mas parece influenciar a rapidez da resposta a estas desvios, o que indica que, a um nível pré-atencional, músicos e não-músicos percepcionam regularidades acústicas simples de forma similar. No Capítulo 4 investigou-se o papel do treino na AEA de diferentes tipos de estímulos auditivos. Três experiências distintas de PE, cada uma composta por fases de aprendizagem, de teste implícito da aprendizagem e de teste comportamental, foram criadas para avaliar a AEA de diferentes sequências auditivas: prosódicas (i.e., palavras pronunciadas de forma melodiosa), não-prosódicas (i.e., palavras sem melodia) e musicais (i.e., melodias). Durante a aprendizagem, os músicos evidenciaram um negatividade aumentada (250-300 ms) em resposta às palavras prosódicas e uma positividade aumentada nos primeiros 100 ms de processamento das melodias. Durante a fase de teste implícito, as respostas neuronais dos grupos só diferiram aquando do processamento das palavras prosódicas. Ao nível comportamental, apenas os músicos demonstraram ter aprendido os diferentes tipos de sequências. Estes resultados sugerem que o treino altera os mecanismos neuronais subjacentes à AEA e influencia positivamente a capacidade para aprender regularidades estatísticas, independentemente do tipo de estímulo. No Capítulo 5 examinaram-se os efeitos do treino sobre a AEA quando as sequências auditivas estavam fora do foco atencional dos participantes. Os resultados mostraram que não houve efeito do treino durante a aprendizagem das sequências linguísticas; contudo, os músicos evidenciaram maiores amplitudes em duas positividades enquanto aprendiam as melodias. No mesmo sentido, as respostas dos grupos durante as fases de teste implícito apenas diferiram durante a experiência com melodias. Ao nível comportamental, os grupos mostraram performances similares no reconhecimento das palavras prosódicas. Estes resultados indicam que a aprendizagem das regularidades estatísticas contidas em discurso com prosódia ocorre mesmo quando a atenção não está focada nesses estímulos. Assume-se então que, quando a estimulação auditiva não é relevante para a tarefa, possuir treino musical não é vantajoso para a AEA apesar de, ao nível cerebral, o treino ter efeito sobre a forma como as sequências auditivas são processadas.

Os resultados desta Dissertação clarificam a extensão e as condições sobre as quais possuir treino musical se traduz em efeitos positivos para a AEA. Sugere-se que o treino promove o processamento das regularidades acústicas, independentemente do tipo de estímulo, quando a aprendizagem ocorre ativamente (i.e., atenção focada nos estímulos) mas não quando os indivíduos estão focados numa tarefa distinta. A presente Dissertação fornece evidência relevante sobre os mecanismos básicos envolvidos na aprendizagem auditiva e sobre a forma como o treino musical pode modificá-los, contribuindo para o nosso conhecimento sobre as bases neuronais dos efeitos de transferência entre música e discurso.

Table of Contents

Abbreviations

SL – Statistical Learning
ASL – Auditory Statistical Learning
VSL – Visual Statistical Learning
AGL – Artificial Grammar Learning
TPs – Transitional Probabilities
EEG - Electroencephalogram
ERP – Event-Related Potential
MRI – Magnetic Resonance Imaging
fMRI – functional Magnetic Resonance Imaging
MEG – Magnetoencephalography
MMN – Mismatch Negativity
mMMN – magnetic Mismatch Negativity
LPC - Late Positive Component
LPPs – Late Positive Potentials
STD – Standard
DEV – Deviant
SSIRH – Shared Syntactic Integration Resource Hypothesis
f0 – Pitch
LDT – Lexical Decision Task
2AFC – Two Alternative Forced-Choice
ISI – Inter-Stimulus Interval
ROI – Region of Interest
M – Mean
SD – Standard Deviation
SEM – Standard Error of the Mean
min – Minute(s)
sec - Seconds
ms – Millisseconds
FFRs – Frequency-Following Responses

ns – Non-Significant

List of Figures

List of Tables

PART I

GENERAL INTRODUCTION

Chapter 1. Introduction

"I was delighted with the suggestion of sending a record for a different reason: we could send music. Our previous messages had contained information about what we perceive and how we think. But there is much more to human beings than perceiving and thinking. We are feeling creatures. However, our emotional life is more difficult to communicate, particularly to beings of very different biological make-up. Music, it seemed to me, was at least a creditable attempt to convey human emotions. Perhaps a sufficiently advanced civilization would have made an inventory of the music of species on many planets, and by comparing our music with such a library, might be able to deduce a great deal about us."

(Sagan, 1978)

In 1977, the Voyager spacecraft were launched. Within each craft was a golden phonograph record addressing extra-terrestrial life in the case of an encounter in some distant corner of the universe. The golden records had the mission to communicate the essence of the human civilization and, soon, Carl Sagan and his team realized that it had to contain more than scientific information. For that reason, besides including greetings in sixty human languages (and one whale language), photographs of the planet and an audio essay on "The Sounds of Earth", the records contained approximately 90 minutes of the world's greatest music. The team believed that music was essential to transmit the essence of mankind due to its unique ability to express human emotions and to its highly complex mathematical relationships (that the scientists thought would hold up for all cultures, species and even planets that Voyager would meet).

Music's impact on our societies' everyday life is indisputable, but Carl Sagan's words call our attention to the broader scope of music, granted by its special features. The characteristics of music make it a unique kind of *language.* Over the last decades, the attention of the scientific community to the impact of music on people's lives grew noticeably (Kraus &

Chandrasekaran, 2010; Münte, Altenmüller, & Jäncke, 2002; Parbery-Clark, Skoe, & Kraus, 2009; Strait & Kraus, 2011). An increasing number of studies have dedicated efforts to understand how long-term experience with a musical instrument shapes cortical and subcortical structures and circuitry. Musicians are experts in multi-sensory integration that have to combine the effective manipulation of their musical instrument, with the integration of constant auditory feedback and the adequate emotional/artistic expression of a musical piece (Pantev, Lappe, Herholz, & Trainor, 2009). For these reasons, musical training is considered an ideal experience-dependent model of brain plasticity.

The impact of musical training on cognition has been the subject of a vast number of studies revealing that training brings advantages for motor and auditory processing, attention and memory. In the last decades, research has revealed that the modifications caused by training can extend beyond the abovementioned domains and generalize to speech processing. Numerous studies support positive transfer effects from music to speech, showing that musicians have enhanced processing of speech both at neural and behavioral levels (Flaugnacco et al., 2015; François, Jaillet, Takerkart, & Schön, 2014; Intartaglia, White-Schwoch, Kraus, & Schön, 2017; Magne, Schön, & Besson, 2006; Marie, Delogu, Lampis, Belardinelli, & Besson, 2011; Schön, Magne, & Besson, 2004; Wong, Skoe, Russo, Dees, & Kraus, 2007). There are several aspects accounting for these effects, namely: (1) the acoustic commonalities between music and linguistic stimuli (Patel, 2003); (2) the organization of the elements of both domains (Patel & Morgan, 2017); (3) the partial overlap of the brain circuitry underlying music and speech processes (Abrams et al., 2011; Patel, 2003a); (4) a robust body of evidence showing the effects of musical training on language-related structures/projections such as the ventral and dorsal streams (Bengtsson et al., 2005; Dittinger et al., 2016; Dittinger, Valizadeh, Jäncke, Besson, & Elmer, 2018; Oechslin, 2010; Wan & Schlaug, 2010), and (5) the reports of enhanced cognitive abilities in musicians (Flaugnacco et al., 2015; Lima & Castro, 2011; Pallesen et al., 2010; Schlaug, Norton, Overy, & Winner, 2005). Existing literature shows that musicians are better at detecting pitch[1] changes in speech (Intartaglia et al., 2017; Martínez-Montes et al., 2013; Schön et al., 2004; Wong et al., 2007), have enhanced proficiency in a second language (Gomez-Dominguez, Fonseca-Mora, & Machancoses, 2018; Martínez-Montes et al., 2013; Milovanov, Huotilainen, Välimäki, Esquef, & Tervaniemi, 2008; Swaminathan & Schellenberg, 2017) and word learning (Dittinger et al.,

---

[1] Pitch refers to the fundamental frequency of sound or F0; in other words, whereas frequency is a one-dimensional feature of sound, pitch refers to the complex perception of frequency.

2016; Dittinger, Chobert, Ziegler, & Besson, 2017), and that musician children show better results in reading and verbal tasks compared to children without musical expertise (Flaugnacco et al., 2015; Gomez-Dominguez et al., 2018; Gordon, Fehd, & McCandliss, 2015; Ho, Cheung, & Chan, 2003; Moreno et al., 2009).

More recently, some reports suggest musical training to facilitate the segmentation of artificial sung languages (Francois & Schön, 2011; François & Schön, 2014). Segmentation refers to the capacity to extract meaningful units from continuous streams of information (e.g., to extract words from continuous speech, see Hay, Pelucchi, Estes, & Saffran, 2011; Keij, 1986; Mirman, Magnuson, Estes, & Dixon, 2008). Segmentation relies substantially on the ability to detect statistical regularities in auditory input that is referred to as Statistical Learning (SL) and operates based on the computation of transitional probabilities (TPs) between streams' units (e.g., the probability that a given syllable "Y" emerges after the occurrence of syllable "X") (Aslin, Saffran, & Newport, 1998; Kuhl, 2004; Saffran, Aslin, & Newport, 1996; Saffran, Johnson, Aslin, & Newport, 1999; Teinonen, Fellman, Näätänen, Alku, & Huotilainen, 2009).

The existing evidence leaves room to wonder if training-related plasticity enhances auditory statistical learning (ASL). Advantages in ASL may account for the benefits in speech processing observed in musical experts. Still, the role of musical training in ASL has so far been poorly investigated. The literature lacks a systematic investigation of the impact of musical training on the SL of distinct auditory stimuli (e.g., linguistic and musical), with various levels of complexity (e.g., simple vs. complex statistical sequences), at different stages of processing (e.g., learning vs. recognition), and at both neurophysiological and behavioral levels. Additionally, there is no published empirical evidence on the contribution of attention to the possible positive linkage between training and ASL abilities. Musicians exhibit increased selective auditory attention skills when compared with musically naïve subjects (hereafter non-musicians) (Patston, Hogg, & Tippett, 2007; Strait & Kraus, 2011; Strait, Kraus, Parbery-Clark, & Ashley, 2010; Strait, Slater, O'Connell, & Kraus, 2015). Yet, the question of if and how benefits in auditory attention contribute to neural alterations and higher percentages of correct responses in ASL tasks remains to be clarified, as well as if and how musical experts benefit from their enhanced attentional abilities in this context.

In the present Dissertation, we will address these open questions using neural and behavioral methods. We hope to contribute to the growing body of research examining the role of musical training in promoting brain plasticity and in developing complex cognitive skills such as speech processing.

1.1 Thesis Outline

The present dissertation is dedicated to examining the role of musical training on ASL. In this chapter, we already highlighted the universal relevance of music and shared some introductory notes on the growing interest in studying the impact of musical training on the brain, providing also a glimpse on the link between training, speech processing and SL.

Chapter 2 presents a review on the abovementioned topics. We will provide a summary of the existing evidence on musical training as a model of experience-driven plasticity that allows investigating the impact of instrumental training on the brain. This chapter offers a comprehensive view of the parallel between music and speech and summarizes behavioral and brain data supporting positive transfer effects from musical training to speech processing. Based on this theoretical background, SL will be introduced and framed as a possible explanatory mechanism accounting for the benefits that training brings to speech processing. We will explain the advantages demonstrated by musicians in the processing of linguistic stimuli and discuss to what extent can training-induced brain alterations enhance ASL mechanisms in general. Also, the role of attention as a mediator variable in the relationship between musical training and ASL will be approached. An introduction to the Event-related Potentials (ERPs) technique will be provided and the results of the existing studies on these matters will be reviewed, pointing its limitations and the need for further developments. Finally, we will provide a synthesis on the aims of the research presented in the chapters to follow.

Chapter 3 presents a study that explores the role of musical training in the pre-attentive processing of statistical regularities among pure tones using ERPs.

Chapter 4 reports a study composed of three experiments aiming to investigate the role of musical training on attentive ASL of different types of auditory stimuli. The experiments were designed to provide us with neural and behavioral data on the processing of statistical regularities in both linguistic (i.e., prosodic and non-prosodic words) and musical structures in musicians and non-musicians.

In Chapter 5, we present a study aiming to examine the role of musical training in the

SL of linguistic and musical structures under pre-attentive conditions.

Chapter 6 is dedicated to our concluding remarks. A summary of the present Dissertation's results will be provided, followed by a general discussion on the most relevant findings. The implications and limitations of the studies reported here will be addressed and prospects of future research on these topics will be presented.

1.2 References

Abrams, D. A., Bhatara, A., Ryali, S., Balaban, E., Levitin, D. J., & Menon, V. (2011). Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cerebral Cortex*, *21*(7), 1507–1518. http://doi.org/10.1093/cercor/bhq198

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8 month old infants. *Psychological Science*, *9*(4), 321–324.

Bengtsson, S. L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., & Ullén, F. (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nature Neuroscience*, *8*(9), 1148–1150. http://doi.org/10.1038/nn1516

Dittinger, E., Barbaroux, M., D'Imperio, M., Jäncke, L., Elmer, S., & Besson, M. (2016). Professional music training and novel word learning: from faster semantic encoding to longer-lasting word representations. *Journal of Cognitive Neuroscience*, *28*(10), 1584–1602. http://doi.org/10.1162/jocn

Dittinger, E., Chobert, J., Ziegler, J. C., & Besson, M. (2017). Fast Brain Plasticity during Word Learning in Musically-Trained Children. *Frontiers in Human Neuroscience*, *11*(May), 1–16. http://doi.org/10.3389/fnhum.2017.00233

Dittinger, E., Valizadeh, S. A., Jäncke, L., Besson, M., & Elmer, S. (2018). Increased functional connectivity in the ventral and dorsal streams during retrieval of novel words in professional musicians. *Human Brain Mapping*, *39*(2), 722–734. http://doi.org/10.1002/hbm.23877

Flaugnacco, E., Lopez, L., Terribili, C., Montico, M., Zoia, S., & Schön, D. (2015). Music training increases phonological awareness and reading skills in developmental dyslexia: A randomized control trial. *PLoS ONE*, *10*(9).

http://doi.org/10.1371/journal.pone.0138715

François, C., Jaillet, F., Takerkart, S., & Schön, D. (2014). Faster sound stream segmentation in musicians than in nonmusicians. *PLoS ONE*, *9*(7), 26–29. http://doi.org/10.1371/journal.pone.0101340

Francois, C., & Schön, D. (2011). Musical expertise boosts implicit learning of both musical and linguistic structures. *Cerebral Cortex*, *21*(10), 2357–2365. http://doi.org/10.1093/cercor/bhr022

François, C., & Schön, D. (2014). Neural sensitivity to statistical regularities as a fundamental biological process that underlies auditory learning: The role of musical practice. *Hearing Research*, *308*, 122–128. http://doi.org/10.1016/j.heares.2013.08.018

Gomez-Dominguez, M., Fonseca-Mora, M. C., & Machancoses, F. H. (2018). First and foreign language early reading abilities: The influence of musical perception. *Psychology of Music*, 030573561774673. http://doi.org/10.1177/0305735617746734

Gordon, R. L., Fehd, H. M., & McCandliss, B. D. (2015). Does music training enhance literacy skills? A meta-analysis. *Frontiers in Psychology*, *6*(DEC). http://doi.org/10.3389/fpsyg.2015.01777

Hay, J. F., Pelucchi, B., Estes, K. G., & Saffran, J. R. (2011). Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology*, *63*(2), 93–106. http://doi.org/10.1016/j.cogpsych.2011.06.002

Ho, Y.-C., Cheung, M.-C., & Chan, A. S. (2003). Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. *Neuropsychology*, *17*(3), 439–450. http://doi.org/10.1037/0894-4105.17.3.439

Intartaglia, B., White-Schwoch, T., Kraus, N., & Schön, D. (2017). Music training enhances the automatic neural processing of foreign speech sounds. *Scientific Reports*, *7*(1), 1–7. http://doi.org/10.1038/s41598-017-12575-1

Keij, B. M. (1986). *Rhythm & cues: rhythmic structure and segmentation in early language acquisition*.

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Science and Society*, *11*(August), 599–606. http://doi.org/10.1038/nrn2882

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*(11), 831–843. http://doi.org/10.1038/nrn1533

Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion*, *11*(5), 1021–1031. http://doi.org/10.1037/a0024521

Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, *18*(2), 199–211. http://doi.org/10.1162/jocn.2006.18.2.199

Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O., & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *Journal of Cognitive Neuroscience*, *23*(10), 2701–2715. http://doi.org/doi: 10.1162/jocn.2010.21585

Martínez-Montes, E., Hernández-Pérez, H., Chobert, J., Morgado-Rodríguez, L., Suárez-Murias, C., Valdés-Sosa, P. A., & Besson, M. (2013). Musical expertise and foreign speech perception. *Frontiers in Systems Neuroscience*, *7*(November), 1–11. http://doi.org/10.3389/fnsys.2013.00084

Milovanov, R., Huotilainen, M., Välimäki, V., Esquef, P. A. A., & Tervaniemi, M. (2008). Musical aptitude and second language pronunciation skills in school-aged children: Neural and behavioral evidence. *Brain Research*, *1194*, 81–89. http://doi.org/10.1016/j.brainres.2007.11.042

Mirman, D., Magnuson, J. S., Estes, K. G., & Dixon, J. A. (2008). The link between statistical segmentation and word learning in adults. *Cognition*, *108*(1), 271–280. http://doi.org/10.1016/j.cognition.2008.02.003

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. http://doi.org/10.1093/cercor/bhn120

Münte, T. F., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience*, *3*(6), 473–478. http://doi.org/10.1038/nrn843

Oechslin, M. S. (2010). The plasticity of the superior longitudinal fasciculus as a function of musical expertise: a diffusion tensor imaging study. *Frontiers in Human Neuroscience*, *3*(February), 1–12. http://doi.org/10.3389/neuro.09.076.2009

Pallesen, K. J., Brattico, E., Bailey, C. J., Korvenoja, A., Koivisto, J., Gjedde, A., & Carlson, S. (2010). Cognitive control in auditory working memory is enhanced in musicians. *PLoS*

ONE, *5*(6). http://doi.org/10.1371/journal.pone.0011120

Pantev, C., Lappe, C., Herholz, S. C., & Trainor, L. (2009). Auditory-somatosensory integration and cortical plasticity in musical training. *Annals of the New York Academy of Sciences*, *1169*(August), 143–150. http://doi.org/10.1111/j.1749-6632.2009.04588.x

Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical Experience Limits the Degradative Effects of Background Noise on the Neural Processing of Sound. *Journal of Neuroscience*, *29*(45), 14100–14107. http://doi.org/10.1523/JNEUROSCI.3256-09.2009

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, *6*(7), 674–681.

Patel, A., & Morgan, E. (2017). Exploring Cognitive Relations Between Prediction in Language and Music. *Cognitive Science*, *41*, 303–320. http://doi.org/10.1111/cogs.12411

Patston, L. L. M., Hogg, S. L., & Tippett, L. J. (2007). Attention in musicians is more bilateral than in non-musicians. *Laterality*, *12*(3), 262–272. http://doi.org/10.1080/13576500701251981

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27–52. http://doi.org/10.1016/S0010-0277(98)00075-4

Sagan, C. (1978). *Murmurs of earth: The Voyager interstellar record.* New York, NY: Ballantine Books.

Schlaug, G., Norton, A., Overy, K., & Winner, E. (2005). Effects of music training on the child's brain and cognitive development. *Annals of the New York Academy of Sciences*, *1060*, 219–230. http://doi.org/10.1196/annals.1360.015

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, *41*(3), 341–349. http://doi.org/10.1111/1469-8986.00172.x

Strait, D. L., & Kraus, N. (2011). Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Frontiers in Psychology*, *2*(JUN), 1–10. http://doi.org/10.3389/fpsyg.2011.00113

Strait, D. L., Kraus, N., Parbery-Clark, A., & Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: Evidence from masking and auditory attention performance. *Hearing Research*, *261*(1–2), 22–29. http://doi.org/10.1016/j.heares.2009.12.021

Strait, D. L., Slater, J., O'Connell, S., & Kraus, N. (2015). Music training relates to the development of neural mechanisms of selective auditory attention. *Developmental Cognitive Neuroscience*, *12*, 94–104. http://doi.org/10.1016/j.dcn.2015.01.001

Swaminathan, S., & Schellenberg, E. G. (2017). Musical competence and phoneme perception in a foreign language. *Psychonomic Bulletin and Review*, *24*(6), 1929–1934. http://doi.org/10.3758/s13423-017-1244-5

Teinonen, T., Fellman, V., Näätänen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience*, *10*. http://doi.org/10.1186/1471-2202-10-21

Wan, C. Y., & Schlaug, G. (2010). Music Making as a Tool for Promoting Plasticity across the Life Span. *Neuroscientist*, *16*(5), 566–577. http://doi.org/10.1177/1073858410377805.Music

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*(4), 420–422. http://doi.org/10.1038/nn1872

Chapter 2. Musical training, speech processing and ASL abilities: a review

2.1 Musical training as model of neuroplasticity

Playing a musical instrument is a highly demanding multimodal task. The practice of a musical instrument over many years can lead to neuroanatomical and neurophysiological changes in the brain (Hyde et al., 2009; Lappe, Herholz, Trainor, & Pantev, 2008; Schlaug, Jäncke, Huang, Staiger, & Steinmetz, 1995; Vaquero et al., 2016). It has been suggested that professional pianists and violinists might play around 7,500 hours before being 18 years old (Ericsson, Krampe, & Tesch-Römer, 1993). The amount of instrumental practice is a determinant factor for musical expertise (Jäncke, 2009). There is a substantial body of evidence demonstrating the positive influence of musical training[2] on specific behavioral outcomes, such as motor coordination and precision, working memory, attention or perception of speech-in-noise (see Okhrei, Kutsenko, & Makarchuk, 2017; Parbery-Clark et al., 2009; Strait & Kraus, 2011; Suárez, Elangovan, & Au, 2016; Vaquero et al., 2016). Studies in the last decades also converged in showing that musical training induces both functional and anatomical changes in sensory-motor systems, thereby showing that the musician's brain represents a unique model of neuroplasticity (Bidelman & Alain, 2015; Herdener et al., 2010; Münte, Altenmüller, & Jäncke, 2002). These changes have been reported in studies using techniques such as structural and functional magnetic resonance imaging (e.g., MRI and fMRI, respectively), ERPs and magnetoencephalography (MEG).

---

[2] Musical training can assume many forms: it can refer to singing-based training, nonverbal training (e.g., instrumental, that do not include voice) or theoretical-based training. In this Dissertation we will refer to musical training as nonverbal training (e.g. the practice of a musical instrument).

Evidence for a positive relationship between musical training and electric and magnetic responses in the auditory cortex has received increased support (Brown et al., 2017; Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004; Pantev et al., 1998). Research has been showing the profound impact of musical training on auditory processing by promoting functional and structural changes in the auditory neural circuitry. This evidence is reinforced by studies in which neural responses were found to be strengthened for musicians who started training at an early age (Pantev et al., 1998; Wong, Skoe, Russo, Dees, & Kraus, 2007). Compared with non-musicians, musicians show increased volume of Heschl's gyrus (Schneider et al., 2002) and increased grey matter in the superior temporal gyrus (Bermudez & Zatorre, 2005) as well as changes in specific white-matter pathways (Bengtsson et al., 2005; Steele, Bailey, Zatorre, & Penhune, 2013). Besides the auditory cortex, changes resulting from musical training also include motor areas. Pianists reveal a reduction of grey matter volume in areas related to auditory processing, sensorimotor control and musical score reading, whereas a set of regions implicated in reinforcement learning (e.g., bilateral putamen and basal ganglia) show increased grey matter volume (Vaquero et al., 2016). Similar changes were found in lip and tongue-related areas (e.g., anterior and posterior parts of the post-central gyrus, respectively) in wind musicians, specifically at the level of cortical thickness and resting-state neuronal networks (Choi, Sung, Hong, Chung, & Ogawa, 2015). Professional pianists also show higher hemodynamic activity during acoustic and mute-motion tasks compared to non-musicians, in regions of the frontal cortex (involving Broca's area), the Wernicke's area, the premotor and supplementary motor areas, and the supramarginal gyrus (Bangert et al., 2006). In the same sense, previous studies showed evidence of increased cortical representation of the fingers of the left hand in string players (Elbert, Pantev, Wienbruch, Rockstroh, & Taub, 1995) and of increased intrasulcal length of the precentral gyrus in keyboard players, which was negatively correlated with the onset of musical training (Amunts et al., 1997). Gaser and Schlaug (Gaser & Schlaug, 2003) also found grey matter volume differences between keyboard players and amateur musicians/non-musicians in motor, auditory and visual/spatial areas. The authors attributed these differences to musical practice since they were correlated with musician status (e.g., professional musicians, amateur musicians or non-musicians) and practice intensity. This type of brain alterations are already observed in six-year old children that underwent fifteen months of musical training showing structural changes in a brain area involved in motor control (e.g., right part of the pre-central gyrus) along with auditory cortex and corpus callosum

(Hyde et al., 2009). Schneider and colleagues (Schneider et al., 2002) found musicians to have 130% larger Heschl's gyrus volume (anteromedial part) and a 100% larger neurophysiology-magnetic response to tone onset in the auditory cortex comparatively to non-musicians. Schneider and colleagues (Schneider et al., 2005) also found a relationship between tone-processing strategies (e.g., the preference for fundamental or spectral pitch analysis in a psychometric pitch test) and anatomical features of Heschl's gyrus: musicians that gave preference to spectral analysis of complex tones revealed a rightward asymmetry of grey matter; those who preferred fundamental pitch analysis showed leftward asymmetry of grey matter. This study revealed that, in nearly 90% of musicians, Heschl's gyrus showed multiplications (e.g., morphological alterations such as common stem duplication, complete posterior duplication or multiple duplications) in one or both hemispheres of the brain. There is also evidence of altered white matter integrity in musicians. Han and colleagues (Han et al., 2009) found increased fractional anisotropy in the right posterior limb of the internal capsule along with higher grey matter density in pianists relative to non-musicians. Thus, musical training fine-tunes the auditory system to process acoustic information.

Training effects extend beyond auditory areas, leading to alterations in a wide-ranging network of brain areas spanning from frontal to parietal regions (Chen, Penhune, & Zatorre, 2008; Koelsch, Fritz, Schulze, Alsop, & Schlaug, 2005; Margulis, Mlsna, Uppunda, Parrish, & Wong, 2009; Zarate & Zatorre, 2008). There is evidence of increased volume and microstructural complexity in white matter fibers connecting auditory cortical areas to inferior frontal cortex in singers, which has been taken as neural support for their enhanced sound perception and production (Halwani, Loui, Rüber, & Schlaug, 2011). Moreover, aging research revealed that orchestral musical training reduced or stalled the decrease in grey matter density over frontal brain regions such as the dorsolateral prefrontal cortex (bilaterally) and left inferior frontal gyrus (Sluming et al., 2002).

Evidence for the impact of musical training on brain function comes also from studies investigating brainstem responses during auditory processing. Musicians show increased Frequency Following Responses (FFRs) to speech and music samples (Musacchia, Sams, Skoe, & Kraus, 2007). The FFRs are neural responses generated in the auditory brainstem (inferior colliculus) that encode the fundamental frequency or pitch (f0) of sounds with high precision. By means of the auditory corticofugal pathway, music and linguistic long-term experience can alter the sensory auditory encoding in the brainstem, modifying the auditory functions in either domain (e.g., music and speech). A study from Bidelman and colleagues (Bidelman, Gandour, & Krishnan, 2011b) showed musicians and tone language experts (e.g., natives of Mandarin Chinese) to have enhanced brainstem encoding of pitch patterns when compared with non-musicians, which is also in agreement with the findings of Wong and collaborators (Wong et al., 2007) of more robust brainstem encoding of linguistic pitch patterns in musicians. Interestingly, Wong and collaborators (2007) also found the FFRs to be positively correlated with the age of onset of musical training and the amount of practice. These findings revealed the impact of musical training on speech processing by indicating a strong association between extensive musically driven pitch training and the modulation of subcortical responses to pitch variations in speech stimuli. Changes in neurophysiological responses were also confirmed by MEG studies: Lappe and collaborators (2008) demonstrated that participants who learned to play piano during two weeks exhibited enhanced amplitude of the Mismatch Negativity (MMN) component (an ERP component reflecting deviance detection) to piano sequences comparatively to a group of participants who did not receive sensorimotor-auditory training. More recently, a study from the same group (Lappe, Trainor, Herholz, & Pantev, 2011) showed the MMN to be more negative in a group of subjects who received instrumental musical training compared to a group of subjects who were only trained to listen and evaluate rhythmic sequences. In an attempt to overcome the difficulties in developing longitudinal studies with a random assignment design, a growing body of studies has correlated the age of onset of musical training with neuroplasticity changes (Steele et al., 2013; Vaquero et al., 2016; Wong et al., 2007). The results indicate a positive relationship between earlier age of onset and the magnitude of brain changes: for example, earlier piano training was related with smaller grey matter volume in the right putamen and greater temporal precision in piano performance (Vaquero et al., 2016).

Overall, the studies reviewed above demonstrate that musical training has a deep impact on brain function and on different cognitive processes. The instrumental training provided by music modifies not only the behavioral responses to musical sounds, but also its underlying brain mechanisms (for a review see Herholz & Zatorre, 2012). Thus, musical training is associated with within-domain auditory plasticity (Patel, 2014).

2.2 Music and speech as unique aspects of human cognition

Together, music and language represent two of the most complex human creations. Defined as the one of the greatest mysteries of humanity by Darwin (Darwin, 1871), music is seen by many as an universal language (Higgins, 2012). Music represents the harmonious combination of sounds, elapsing during a temporal window of seconds to hours, generated by musical instruments such as voice and/or mechanical objects, which have the capacity to convey meaning and emotion. Its key features are melody, rhythm, timbre and harmony. Melody pertains to the combination of pitches; harmony refers to the progression of chords – e.g., groups of simultaneously presented tones; rhythm relates to the temporal structure of sounds; and timbre refers to the quality of the sound resulting from the combination of spectral and temporal features of the acoustic signal (Loui & Wessel, 2007). As a song unfolds over time, our expectations about the upcoming acoustic information is established by chord progressions and tempo: the chords inform us about what combination of tones could match the harmony of the song, whether its temporal structure provides a sense of the rhythm and enables the anticipation of new chords at specific moments (Lappe et al., 2011).

Language, on the other hand, is defined as “the method of human communication, either spoken or written, consisting of the use of words in a structured and conventional way” (“Language", Definition 1, n.d.). Language comprises several processing domains that include grammar or syntax, semantics, pragmatics and phonology (Bruner, 1974; Cromer, 1981). The most basic entities of spoken language are phonemes that can be organized into words when assembled together according to syntax rules. Words are the most important constituents of speech. Speech relies on the functional organization of words into meaningful sentences according to the rules established by grammar. Grammar is responsible for the elasticity of speech, allowing sentences composed of the same words to have different meanings according to word order and punctuation (Chomsky, 1957, 1965; Liberman, 1970; Seidenberg, 2014).

Above all, music and speech are considered universal since both are present in all human cultures (Cohen, 2008). Music[3] and speech are inherently human auditory-motor functions (Zatorre, 2013) that comprise several on demanding modulations of acoustic

---

[3] Of note, the literature reviewed here regarding music processing will mainly consider studies focusing on Western music. The choice for western music is due both to the need to select a specific type of music, considering variables such as culture, period and style, but also due to the greater body of research informing on its perception and processing (Deutsch, 1999).

parameters used to produce meaning. They are both made of discrete elements such as tones or syllables, have a complex and hierarchically organized structure, involve the modulation of acoustic features such as pitch, intensity and rhythm to communicate, are rooted in the auditory system and involve the decoding of complex acoustic patterns (Abrams et al., 2011; Patel, 1998, 2003a, 2008, 2012; Patel & Morgan, 2017). Songs, for example, result from the combination of music and speech: musical rhythms and scales are combined with phonemes, syllables, words and prosody. Speech, in turn, makes use of a constellation of acoustic parameters to create melodic patterns or intonations that convey informational content and emotion. According to Brandt, Gebrian and Slevc (Brandt, Gebrian, & Slevc, 2012, p. 4) "*from a musical perspective, speech is a concert of phonemes and syllables, melodically inflected by prosody*". Speech prosody is critical for vocal communication (Thönnessen et al., 2010) because it allows the punctuation of verbal messages, which plays an important role in speech segmentation and word learning (Cutler, Dahan, & Van Donselaar, 1997; de Diego-Balaguer, Rodríguez-Fornells, & Bachoud-Lévi, 2015; Mattys, 2004).

Much of the resemblance between the musical and linguistic domains has its roots in the generative nature that both share: they are the result of rule-based arrangements of a finite number of elementary sounds that are organized to generate meaningful sequences that can, by its turn, be reorganized into increasingly complex structures according to syntactic rules (Krumhansl, 1990). Syntax can be defined as the set of rules that dictate the functional organization of the basic elements of language and music as sequences (Jackendoff, 2002). Words or musical tones need to be combined in specific ways so that the sequences of elements they integrate, such as sentences or chord progressions, make sense as a whole. In speech, for example, syntax sets the stage for understanding the subject to which the predication relates to (e.g., 'Daisy owns a beautiful cat': "a beautiful cat" is the predication that refers to the subject "Daisy"). In music, it works the same way: musical phrases are expected to have peaks of tension/resolution at specific times. Speech is built upon complex linguistic expressions composed of clauses that relate to each other in different ways and that unfold over time, whereas music relies on variations of chord progressions occurring at specific moments of the song (Lerdahl & Jackendoff, 1983; Patel, 2003). Human beings acquire syntactic knowledge early in development and apply it implicitly to the many musical and linguistic forms they are exposed to everyday (Patel, 1998). Our knowledge of syntax is used constantly when we speak or when we listen to songs and it is important for humans to be capable of perceiving deviations in the structure of the incoming input. Without this ability, a sequence of sounds, either linguistic or melodic, would be only the juxtaposition of different elements, one after the other. Thus, syntax is then at the core of our capacity to derive meaning from auditory patterns and to communicate, and it is also one of the critical elements in the overlap between music and speech (Patel, 2003; Everaert et al., 2017).

Additionally, one of the most powerful parallels between music and speech deals with timing. Music's transformative power may reside precisely on how we process time by means of rhythm. Rhythm is defined as "a strong, regular repeated pattern of movement or sound" ("Rhythm," Definition 1, n.d.). The only possible way to find out the rhythm of a song, the pace of a person or a cardiac rhythm is to track how a series of events unfold over time. Essentially, the ability to perceive the regular recurrence of events over time is fundamental for our understanding of the world. It is possible that to perceive time in music and speech the brain relies on the same neurofunctional mechanisms (Patel, 2012) and, specifically, on the auditory working memory brain circuitry (Hickok, Buchsbaum, Humphries, & Muftuler, 2003; Joseph et al., 2015). There is a close link between rhythmic and linguistic abilities. Listeners rely on predominant timing and stress patterns to segment syllables and words from a continuous speech signal. That is, humans need to perceive the rhythm of speech in order to decode it (Patel, 2008). For this reason, recent studies have suggested a positive relationship between higher performance in rhythmic tasks and phonological processing (Bhide, 2013; Goswami, 2011). These studies are complemented by others indicating a link between musical training - in which the perception of rhythm is at the core -, and phonological skills (Chobert, François, Velay, & Besson, 2012; Slevc & Miyake, 2006; Tierney & Kraus, 2013).

Despite the considerable set of commonalities, music and speech are also clearly distinct. One of the main differences lies in the instruments used to produce sound: whereas speech is only the product of the human voice - resulting from the orchestration between the lips, the nostrils, the pharynges and both nasal and oral cavities -, instrumental music is generated by the manipulation of objects. The temporal characteristics of speech and music also differ. The rate at which speech is delivered is faster relative to music. For example, there is evidence that syllables are almost 50% shorter in duration than musical notes (Greenberg, 1996; Watt & Quinn, 2006). The change in the spectral shape of speech and melodies is also a distinctive characteristic of these sounds: whereas the amount of change in the spectral shape varies considerably in syllables, it varies much less in musical notes (Patel, 2014).

Music and speech are two complex forms of auditory input whose perception and production have been intriguing scientists in the neuroscience and cognitive psychology fields for centuries. In these fields, musical training and linguistic expertise have gained an increasing interest from the scientific community as experience-dependent models of brain plasticity. These models are important tools to examine music and speech's influence on the brain architecture and function. The parallel between both domains has been a matter of interest in psychology and neuroscience greatly due to the fact that both play a key role in human existence and involve detailed/complex sequences of sounds (Patel, 2008). Moreover, a more comprehensive understanding of the parallel between music and speech can provide outstanding contributions to cognitive research by unraveling essential mechanisms behind human communication.

In summary, there are several aspects that may account for transfer effects from musical training to speech processing, namely: (1) the acoustic commonalities between music and speech; (2) the hierarchical organization of both domains; (3) the neural overlap between music and speech processes, and (4) the evidence of training-related alterations in speech-related structures/projections such as the ventral and dorsal streams.

2.3 Transfer effects between music and speech

The comparative analysis of speech and music provides an explanatory and integrative view of the mind while simultaneously elucidating the neural architecture of both domains. In this realm, a growing body of studies have provided evidence for shared perceptual and cognitive mechanisms during the processing of musical and speech input (Patel & Iversen, 2007). Based on this evidence and on musicians' enhanced sensitivity to process acoustic cues, researchers have devoted efforts to investigate positive transfer of training effects from music to speech (Patel, 2011).

The studies of Aniruddh Patel have provided evidence for an overlap between speech and music processing, namely due to its intrinsic syntactic structure. As mentioned before, syntax is defined as a set of principles that govern the way perceptually unique elements are combined to form structured sequences (Jackendoff, 2002). These sequences are hierarchically organized so that listeners can decode meaning from auditory input. Despite some contradictory findings, the majority of studies claim that there is a significant overlap in the neural resources used to process syntax in speech and in music (Jung, Sontag, Park, & Loui, 2015; Koelsch, Gunter, Wittfoth, & Sammler, 2005; Patel, Gibson, Ratner, Besson, & Holcomb, 1998). Patel (Patel, 1998) offers an integrative explanation – the Shared Syntactic Integration Resource Hypothesis (SSIRH) - for the apparently inconsistent evidence. Based on findings from previous studies (Haarmann & Kolk, 1991; Kaan & Swaab, 2002), Patel suggests that the center for syntax processing across domains (e.g., music and speech) may be located in "*frontal regions, which do not themselves contain syntactic representations but which provide resources for computations in posterior regions where syntactic representations reside*" (Patel, 2003, p. 678). The SSIRH hypothesis expects interference to follow whenever musical and linguistic syntactic integration co-occur. The study of Koelsch and collaborators (2005) corroborates this hypothesis: the authors found that the left anterior negativity - a correlate of initial syntactic structure building (Friederici, 2002) - to syntactically incorrect words was reduced when words were paired with syntactically incorrect chords, which suggests that the processing of syntax in speech interacts with the processing of syntax in music.

The studies that probed how musicians process music and speech stimuli have revealed that musical training results in an enhanced sensitivity to perceive supra-segmental pitch variations (e.g., in native language, see Schön, Magne, & Besson, 2004; and in foreign language, see Marques, Moreno, Castro, & Besson, 2007) along with segmental and tonal manipulations of pitch (Marie, Delogu, Lampis, Belardinelli, & Besson, 2011) as well as alterations in speech's metrical structure (Marie, Magne, & Besson, 2011). Recent research argues that training develops musicians' capacity to process more abstract and complex auditory input. For example, Bidelman and colleagues (Bidelman, Gandour, & Krishnan, 2011) studied the influence of linguistic and musical training on the capacity to represent pitch patterns in the auditory brainstem. The study revealed that native Chinese participants and English musicians showed higher pitch tracking-accuracy and strength (computed from brainstem FFRs) compared to English non-musicians when processing lexical tones and musical intervals. This study suggested that pitch pattern encoding in the brainstem depends not only on the specific characteristics of the sounds but also on the musical expertise of the listeners. Moreover, there is evidence of enhanced phonological processing (Anvari, Trainor, Woodside, & Levy, 2002; Moreno et al., 2009; Slevc & Miyake, 2006) and higher sensitivity to vowel/consonant changes in words (Marie, Delogu, et al., 2011) in musicians. These findings were also extended to children: Chobert and collaborators (Chobert et al., 2012) found that children receiving twelve-months of active musical training demonstrate more negative MMNs and smaller error rates in response to variations in syllabic duration and voice onset time in syllables when compared with non-musician children receiving painting training. Nevertheless, we should highlight that only half of the training given to these children was purely instrumental (part of the training time was dedicated to singing). Singing training is a speech-based intervention, not being considered uniquely instrumental. In conclusion, these findings are in line with the notion that musical training increases the low-level processing of acoustic information, which might enhance related cognitive processes such as the building of phonological representations (and their associated meaning) of linguistic materials.

More evidence of positive transfer of training effects comes from studies investigating emotional prosody. Some studies suggest that musicians are more accurate than non-musicians in the recognition of emotions in speech (Lima & Castro, 2011; Thompson, Schellenberg, & Husain, 2004), being more accurate than non-musicians in the recognition of angry prosody even when sentences are devoid of intelligible semantic content (Pinheiro, Vasconcelos, Dias, Arrais, & Gonçalves, 2015). There is also evidence of increased speech prosody sensitivity in musically-trained children (Thompson et al., 2004). Additionally, musical training has been linked to increased reading and speech perception skills (see, for example, Moreno et al., 2009). The study of Boll-Avetisyan and collaborators (Boll-Avetisyan, Bhatara, & Höhle, 2017) investigated the relationship between musicality (or musical aptitude) and rhythmic grouping of speech based on the premise that rhythmic perception is based on cognitive processes that are shared by both music and speech. The results showed that rhythmic grouping preferences were predicted by musical rhythm receptivity (i.e., the higher the scores on the rhythm component of a musical test the greater the consistency of participants in grouping rhythmic patterns from speech). The perception of speech-in-noise has also been examined in musicians and non-musicians. Studies with children (Strait, Parbery-Clark, Hittner, & Kraus, 2012) and adults (Parbery-Clark et al., 2009; Strait et al., 2012; Strait & Kraus, 2011) consistently reveal musical training to be associated with enhanced perceptual (i.e., less degradation of auditory percepts) and subcortical (i.e., more robust representations) encoding of speech in noisy environments. However, there is also evidence for null training effects: Ruggles and colleagues (Ruggles, Freyman, & Oxenham, 2014) found that musicians did not show enhanced perception of voice or whispered sentences produced in noisy backgrounds when compared to non-musicians, despite showing enhanced pitch discrimination skills.

Mostly due to time, costs and difficulties in participants' recruitment there are no more than a dozen longitudinal studies addressing the relationship between musical training and speech processing skills (Chobert et al., 2012; François, Chobert, Besson, & Schön, 2013; Moreno et al., 2009; Roden, Kreutz, & Bongard, 2012; Thompson et al., 2004). In a study administering musical vs. painting training to children, Moreno and collaborators (Moreno et al., 2009) found that six months of musical training increased reading skills and the detection of small pitch variations in words in eight-year old children, which was reflected in both behavioral and electrophysiological (i.e., larger positivity – 300-900 milliseconds (ms) latency window – in musicians for weak incongruities) measures. The effects of training extended to emotional prosody: six-year old children receiving keyboard lessons, when compared with children who did not receive any training were better at identifying negative emotions (i.e., anger or fear) in spoken sentences and tone utterances (Thompson et al., 2004). The effects of training were also observed in verbal memory in children who underwent 18 months of instrumental sessions compared with groups of 'control' children who either received no training or training in natural sciences (Roden et al., 2012). Interestingly, the authors found no effects of training in visual memory tests, which supports within-domain neuroplasticity. Another longitudinal study investigated speech segmentation in children receiving musical or painting training using a test-training-retest procedure (François et al., 2013). After training, only musician children showed improved segmentation of artificially created speech sounds observed both in electrophysiological (i.e., larger negativity in the 450-550 ms latency window) and behavioral (i.e., performance on post-learning 2-alternative forced-choice [2AFC] tasks) measures.

The findings of studies using longitudinal designs are critical since they rule out possible preexisting individual differences in perceptual and cognitive skills that could be responsible for the positive results associated with musical training. Moreover, the reviewed research points to more than common acoustic processing between music and speech. It shows that long-term training in the *musical* domain influences the development of specific and distinctive cognitive representations in the *speech* domain, which is taken as evidence for transfer effects (for a review, see Besson, Chobert, & Marie, 2011).

2.4 Sensitivity to statistical regularities: A fundamental aspect of auditory processing

In their daily life, humans are exposed to a plethora of stimuli. Stimulation comes in different forms, with distinct spatial and temporal characteristics, conveying essential information about the world around us. The constant and varied auditory stimulation to which humans are exposed to could have a chaotic effect if the brain did not have the capacity to efficiently and rapidly disentangle the complex information contained in the sensory input. Yet, we are able to segment a stream of sounds based only on the regularities of the auditory events presented, and based on that, to infer semantic and lexical meaning (e.g., speech streams) (Hay, Pelucchi, Estes, & Saffran, 2011; Romberg & Saffran, 2011; Saffran, Johnson, Aslin, & Newport, 1999).

Perceiving regularities in the acoustic environment is highly dependent on the capacity to perceive time, which makes time a fundamental variable in auditory perception. The stimuli inhabiting our auditory landscape are temporal in nature and unfold over different time scales. For example, pure tones take up to tens of ms, phonemes take a bit longer, syllables and words even more but often less than a second (sec); phrases occur for hundreds of ms to secs, while musical phrases and sentences last secs to minutes (min) (Teki & Griffiths, 2016). Before sentences can be recognized as such, they are merely sequences of words that form acoustic patterns. The acoustic patterns characterizing speech and also music are based on complex statistical structures.

Decades ago, researchers became interested in studying how the brain disentangles words from continuous speech. Artificial language learning paradigms showed that listeners could learn the statistical structure of a speech stream and recognize its constituent words based on the mere exposure to that stream. Dating back to 1996, Saffran and collaborators (Saffran, Aslin, & Newport, 1996) made an important discovery regarding speech segmentation: the authors found that infants were able to segment continuous streams of pseudo-words (word-like letter strings) based only on the TPs between syllable pairs. In that study, subjects were exposed to a familiarization phase in which syllables were presented continuously. Subjects were not aware that the stream's structure was organized into triplets of syllables. Nevertheless, the implicit capacity to extract regularities from the auditory input would suggest they were sensitive to the triplets. A habituation test, designed to measure subjects' responses to triplets and novel sequences, showed that infants were successful at segmenting the continuous stream. This was the first report showing that the segmentation of artificial speech could be achieved by decoding its statistical structure.

Artificial languages, such as the one used in Saffran and collaborators study (1996), mirror the complex structure of natural languages while removing the confounding effect of semantic information. The absence of semantic information allows researchers to examine the learning of grammar in isolation. The learning of artificial languages started to be investigated using artificial grammar learning (AGL) paradigms. AGL paradigms examine the ability to encode an artificial grammar designed in an experimental context. The grammar is composed by a fixed set of symbols (that can assume may forms, for example, consonants, vowels, numbers and figures), whose combination can form a finite set of entities according to the grammars' rules. In the context of language research, these entities are pseudo-words that form a small-scale language, which is highly controlled. These languages present low external validity, as their complexity is reduced compared to natural languages. Nevertheless, they provide the means to study the detection of patterns and the processing of statistical regularities in continuous speech streams.

These paradigms have been applied to different topics of research (e.g., language learning, see Gómez & Gerken, 2000; Petersson, Folia, & Hagoort, 2012; visual SL, see Stobbe, Westphal-Fitch, Aust, & Tecumseh Fitch, 2012; structural priming, see Brooks & Kempe, 2013) and with stimuli of different sensory modalities. These paradigms suffered modifications over time (see the classical studies of Reber, Reber, 1967, 1976, 1989; Reber & Squire, 1994). AGL paradigms were proposed to represent a more adequate experimental designs to test the learning of complex linguistic sequences (Christiansen, Dale, Ellefson, & Conway, 2002). Generally, the paradigm is composed of two parts: learning and test phases. The learning phase (also referred to as *exposure* or *familiarization* phase in the literature) relies on the presentation of a stream of stimuli according to a set of rules. Each stimulus is repeated numerous times and concatenated to form a continuous (without pauses) stream. Typically, the same stimulus does not occur twice in a row. Participants are not aware of the streams' structure and not informed how to decode it. Therefore, the decoding of input patterns, such as pseudo-words (hereafter called 'words') is only made possible by the computation of TPs between syllable pairs (see next section). The common instruction is to pay attention to the stream. The output of the learning phase is then assessed with behavioral tests such auditory 2AFC tasks (Batterink & Paller, 2017; Cunillera et al., 2009; François et al., 2013), lexical decision tasks (LDTs) (François, Cunillera, et al., 2017), among others. The test phase aims to evaluate if participants are able to correctly recognize words from the artificial language they were exposed to. In some studies, researchers test word segmentation by presenting the stream's items (e.g., words) among new items (e.g., non-words) and asking participants to perform LDTs (i.e., to decide whether the item presented belongs to the artificial language or not; François, Cunillera, et al., 2017a); in other studies, researchers opt for 2AFC tests in which participants answer to pairs of test-items (e.g., word - non-word) (Cunillera et al., 2009; de Diego-Balaguer, Toro, Rodriguez-Fornells, & Bachoud-Lévi, 2007; Francois & Schön, 2011).

Researchers have struggled with the limitations imposed by behavioral test phases. The responses to explicit behavioral tests might be biased by evaluative processes and memory skills, not providing the most accurate measures of learning (Noam Siegelman, Bogaerts, Kronenfeld, & Frost, 2017). Using neuroimaging methodologies, such as the ERP technique, it is possible to implicitly evaluate learning outcomes without the interference of evaluative judgments. Therefore, some studies introduced an intermediate phase between learning and test aiming to evaluate the participants' responses to violations of the streams' structure (de Diego-Balaguer et al., 2007; François, Cunillera, et al., 2017). This intermediate phase - that probes the response to stream violations - consists on the presentation of the same stimuli (previously heard) among new stimuli (test items) randomly inserted. The new stimuli can be new arrangements of the original item's components or completely new items. For example, in an artificial language paradigm, the new items of the intermediate phase (also known as violation phase or implicit test phase) could be words whose syllables were previously presented but are organized in the wrong order, or new words formed by syllables that were not heard before. Participants are not informed about the presence of test items. Thus, the stream presented in this phase is heard as if it was a continuation of the learning phase. Since the new items violate the statistical structure of the language, the violation phase allows researchers to infer participants' learning implicitly, without overt responses. Not only it is possible to evaluate if participants had learnt the language, as it is also possible to assess how the brain responds to violations of syllable' sequences (see Fig. 1).

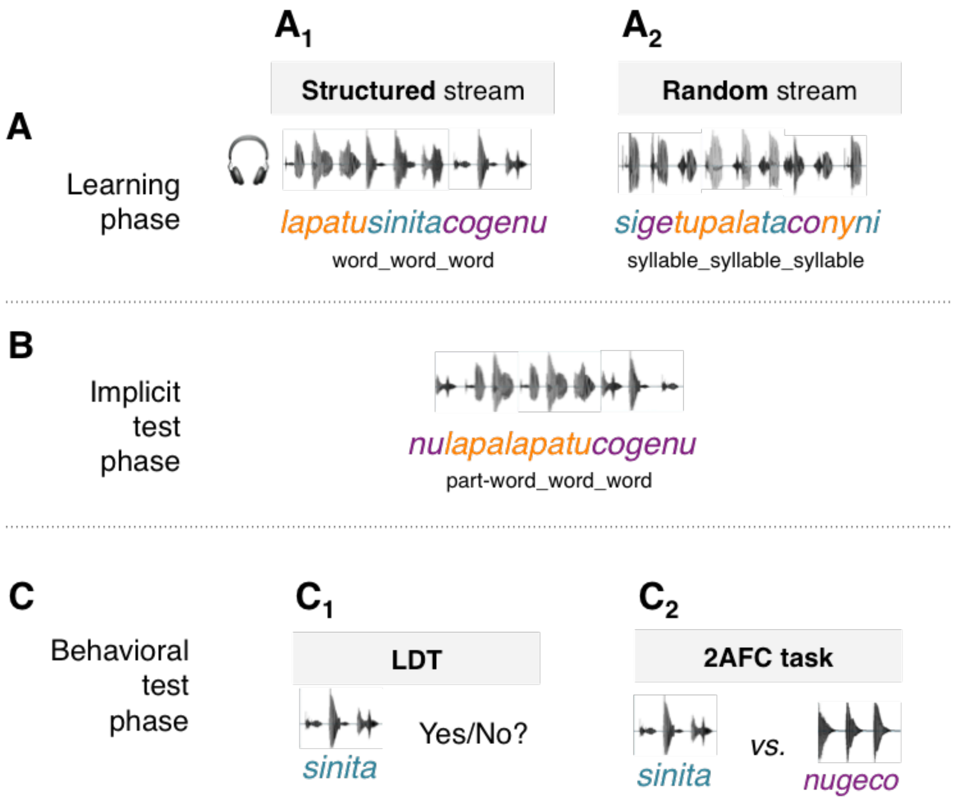

FIGURE 1. Representation of a typical learning paradigm. **(A)** The paradigm starts with the learning phase in which participants are presented with the language stream via headphones **(A1)**; in some studies (Batterink & Paller, 2017; Buiatti, Peña, & Dehaene-Lambertz, 2009; Cunillera, Toro, Sebastián-Gallés, & Rodríguez-Fornells, 2006; de Diego-Balaguer et al., 2015), researchers additionally include a random stream of syllables **(A2)** which allow the comparison of ERP responses, for example, between structured and unstructured language streams. **(B)** Recent studies (François, Teixidó, et al., 2017; François, Cunillera, Garcia, Laine, & Rodriguez-Fornells, 2017) aiming to obtain an implicit measure of learning have inserted an implicit test phase after the learning phase; the words consisting on structural violations can result from different combinations of syllables [e.g., part-words – combination of the last syllable of a word with the first syllable-pair of another word; non-words – words shown in the reversed form (i.e., *lapatu* – *tupala*); new words]. **(C)** The paradigm ends with the evaluation of participants' learning outcomes typically by means of one of two tasks: **(C1)** a LDT, in which each trial presents an item (e.g., words, part-word, non-word), or **(C2)** a 2AFC task in which pairs of words (e.g., word vs. non-word, part-word vs. word) are presented and participants are asked to choose, by pressing one of two buttons, the word that seems the most familiar to them.

2.5 Encoding of regularities by statistical learning

The ability to decode the statistical structure of linguistic streams has been demonstrated in adults (Abla & Okanoya, 2009; Mirman, Magnuson, Estes, & Dixon, 2008), in children (Kristin McNealy, Mazziotta, & Dapretto, 2010; Saffran et al., 1999; Vasuki, Sharma, Ibrahim, & Arciuli, 2017a), even in 8-month-old infants (Aslin, Saffran, & Newport, 1998; Saffran, Aslin, et al., 1996) and also in other species (e.g., nonhuman primates, see Hauser, Newport, & Aslin, 2001; Newport, Hauser, Spaepen, & Aslin, 2004; rodents, see Toro & Trobalón, 2005). The evidence stemming from these studies point to a general regularity-processing mechanism defined in the literature as SL that allows decoding the units of rapid and continuous streams of auditory information since early months of age (Perruchet & Pacton, 2006; Saffran, 2003). SL is one of the mechanisms involved in the segmentation of continuous streams of speech stimuli that allow babies to acquire their mother language and allow individuals in general to learn new languages (Hay et al., 2011; Romberg & Saffran, 2011; Saffran, 2003).

SL has great relevance for the perception and encoding of environmental regularities. It allows the extraction of patterns from different sources of environmental input and the prediction of upcoming events (Arciuli, 2017; Hasson, 2017). The term *statistical learning* has been defined after the abovementioned language studies due to its power to explain the detection of word boundaries (Onnis, Waterfall, & Edelman, 2008), the computation of phonotactic or orthographic patterns (Chambers, Onishi, & Fisher, 2003; Pacton, Perruchet, Fayol, & Cleeremans, 2001), allowing the detection of adjacent (Endress & Mehler, 2009) and non-adjacent dependencies (i.e., interdependence relationships between adjacent (contiguous) or non-adjacent (intercalated) elements in a stream) (Newport & Aslin, 2004) along with syntax learning and comprehension (Gomez, 2002; Saffran & Wilson, 2003). SL operates by segmenting streams of information into chunks taking into account the TPs between the streams' units (Saffran, 2003). TPs refer to the probability of X given the occurrence of Y. Within a stream of continuous events made up of stimulus sequences, the TPs within sequences are higher than across sequences. For example, in the speech stream 'what/a/nice/song", the TPs between the syllables 'ni' and 'ce' are higher than between the syllables 'ce' and 'song', because 'ni' and 'ce' belong to a word while 'ce' and 'song' refer to word boundaries. By computing TPs between the streams' simplest units (e.g., syllable-pairs in the case of speech; tone-pairs in the case of melodies), the brain progressively develops temporal correlations and creates chunks of information (Franco & Destrebecqz, 2012; Perruchet, Poulin-charronnat, Tillmann, & Peereman, 2014). These computations occur implicitly and provide the brain with information regarding possible upcoming events, allowing it to predict future input. Since the seminal paper of Saffran and collaborators (1996), the role of TPs in the segmentation of artificial speech streams has been demonstrated in several studies (Aslin et al., 1998; Kuhl, 2004; Saffran, Aslin, et al., 1996; Saffran, Newport, & Aslin, 1996).

SL represents one of the quintessential learning abilities that have been linked to word learning and auditory stream segmentation. SL is conceptualized as a mechanism of regularity processing that applies to different sensory modalities and cognitive processes as a domain-general learning mechanism that exhibits great inter-individual variability (Armstrong, Frost, & Christiansen, 2017; Frost, Armstrong, Siegelman, & Christiansen, 2015; N. Siegelman & Frost, 2015). SL is considered to be domain-general because it is "*a general mechanism for learning and processing any type of sensory input that unfolds across time and space*" (Frost et al., 2015, p. 117). Recent studies have suggested that SL is stimulus-specific since the output of statistical computations is different depending on the stimulus nature: for example, the combined results of the studies of Shufaniya and Arnon (2018) and Raviv and Arnon (2018) show that SL improves with age for non-linguistic stimuli (e.g., bell ringing) but it does not for linguistic sounds (e.g., syllables), despite both being auditory in nature. Besides, studies have revealed that language learning by SL is intact in subjects with congenital amusia whereas musical learning (which is based on the same statistical structures) is not (Peretz, Saffran, Schön, & Gosselin, 2012). This is revealing of distinct SL processing systems depending on the input nature (e.g., syllables or tones), also suggesting that SL occurs differently within the same sensory modality (e.g., auditory) (Shufaniya & Arnon, 2018; Raviv & Arnon, 2018).

Concrete and abstract features of stimuli can be encoded during SL: abstract structures relate to distributional properties of the stimuli; concrete structures relate to concrete features of stimuli such as shapes or colors (Daltrozzo & Conway, 2014). Since the seminal paper of Saffran and colleagues (Saffran, Aslin & Newport, 1996), SL has been investigated with different stimuli (e.g., auditory – tones, see Creel, Newport, & Aslin, 2004; Saffran et al., 1999; chords, see Daikoku, Yatomi, & Yumoto, 2016; speech sounds, see Saffran, Newport, et al., 1996; Toro, Sinnett, & Soto-Faraco, 2005; morse code, see Shook, Marian, Bartolotti, & Schroeder, 2013; visual – shapes, see Fiser & Aslin, 2002a; Kirkham, Slemmer, & Johnson, 2002; figures, see Arciuli & Simpson, 2012; Kidd & Arciuli, 2016) in human (e.g., newborns, see Bulf, Johnson, & Valenza, 2011; François, Teixidó, et al., 2017;

children, see Arciuli & Simpson, 2012; Kidd & Arciuli, 2016; adults, see Batterink, 2017; Cunillera et al., 2009; de Diego-Balaguer et al., 2015; François, Cunillera, et al., 2017; Mirman et al., 2008) and non-human populations (e.g., tamarins, see Hauser et al., 2001; Newport et al., 2004; rodents, see Toro & Trobalón, 2005).

A recent study (Batterink & Paller, 2017) offered evidence of two distinct components in linguistic ASL. The first component, referred to as *word identification component,* concerns the perceptual encoding of streams' structure. It is perceptual in its nature and refers to the process by which the streams' most basic units (e.g., syllables) are integrated into more complex and larger items (e.g., words). The other component refers to the storage of the mental representations of units in long-term memory. As Batterink and Paller (2017) state, this *memory storage component* is 'peripheral' to the main processes underlying the SL of speech streams but it is nonetheless of great importance since it is a requirement for further processing, and simultaneously dictates the successful performance of subjects on post-learning tasks. This double component framework for SL offers relevant insights because it dissociates the formulation of word-like entities from the successful memory storage of those entities' representations. Frequently, in SL research, the authors examine learning through behavioral, post-learning tasks without stating that those tasks are only testing learning outcomes. Research shows that around one third of a sample fails to outperform chance levels in SL tasks (Frost et al., 2015; N. Siegelman & Frost, 2015) and, therefore, the performance in these tasks is highly variable among individuals mainly due to individual differences in memory capacity (Bors & MacLeod, 1996). Most of the previous studies did not make a clear distinction between the two components of this framework, frequently conceptualizing in the same way the results of SL tasks that test the learning process instead of the learning outcomes, or vice-versa. Thereby, it is critical to carefully address this dissociation in order to provide data on these two components of learning separately.

The implicit nature of SL processes lends itself to inquiry by neuroimaging techniques. The ERPs are particularly well suited to probe the fast, dynamic brain processes underlying SL, specifically due to its excellent temporal resolution (Siegelman, Bogaerts, Kronenfeld, et al., 2017; for a review on the topic, see Daltrozzo & Conway, 2014). As such, evidence stemming from ERP studies has brought important contributions for the understanding of the temporal course of the neural responses in SL tasks.

2.6 The ERP technique

Neuroimaging is the set of techniques employed to measure brain activity that can be directly or indirectly associated with neural mechanisms or anatomical aspects of the brain (Haynes & Rees, 2006). The ERP technique derived from the electroencephalogram (EEG) is one of those techniques. The EEG signal consists on the electrical brain activity captured at the scalp-level. This type of brain activity emerges from the exchange of information occurring at the level of the synaptic cleft (the gap between axons and dendrites of neurons) (Luck, 2005). The neural communication established by neurotransmitters triggers post-synaptic potentials (electrical signals) that travel across neurons. Since there is an unbalance in the chemical and electrical states between the inner and outer neuron, the potential originates a current outside the cell (Buzsáki, Anastassiou, & Koch, 2012). When multiple pyramidal neurons – that have a perpendicular orientation in reference to the scalp - act synchronously, their currents generate a large electric field that can be detected at the scalp by the EEG (Allison, Wood, & McCarthy, 1986).

The ERP technique, derived from the EEG, is a non-invasive and safe method that measures the online electrophysiological responses to specific events. The ERP technique is known in the cognitive neuroscience domain for its excellent temporal resolution, deriving from the fact that the recorded signals originate directly from electrical brain activity (Cohen, 2017; Luck, 2005). It offers a unique window to study how information is processed in the brain and to characterize the temporal dynamics of the brains' responses to stimuli or actions. Conversely, the spatial resolution of the technique is low mainly due to the unavoidable interactions between distant electric fields and volume conduction (Cohen, 2017; Luck, 2005). Therefore, interpretations of the spatial location of specific electrical activity need careful consideration and the help of source localization methods.

The ERPs are based on the analysis of voltage deflections extracted from the scalp-recorded EEG by means of signal averaging (Duncan et al., 2009). The voltage fluctuations embedded in the EEG signal reflect sensory (low-level) and/or cognitive (high-level) brain processes generated in response to specific events, physical or mental (Duncan et al., 2009; Luck, 2005; Picton, Lins, & Scherg, 1995). The neural manifestations of these processes are categorized as ERP components with specific polarities (e.g., positive or negative), latencies, scalp topographies and with specific associations to their generating events (Cohen, 2017; Luck, 2005). The latency of the components refers to the timing of a given process, while its amplitude relates to the neural resources allocated to that process (Cohen, 2017; Luck, 2005). The high temporal resolution of this methodology is providing researchers with a dynamic view of the cognitive operations happening in the brain while tracking them in time with high precision. Hence, ERPs are a well-suited method to probe how the extraction of statistical regularities from acoustic streams dynamically occurs in the brain.

SL can be examined in different ways using ERPs. Electrophysiological data can be recorded during the learning phase providing online, neurophysiological measures of learning and allowing the investigation of the temporal dynamics of SL with ms precision. EEG can also be recorded during test phases and the ERPs obtained during implicit test phases mixing legal (i.e., items whose structure is in accordance with the structure of the items presented in the learning phase) with new/illegal (e.g., items that violate previous items' structure) items, are a relevant implicit measure of learning outcomes, revealing the electrophysiological responses to statistical violations of the input structure (for a similar procedure, see de Diego-Balaguer et al., 2007; François, Cunillera, et al., 2017). If obtained during behavioral post-learning tasks, ERPs allow the investigation of the neural signatures of familiar vs. unfamiliar stimulus processing (for a similar procedure, see Chobert et al., 2012; de Diego-Balaguer et al., 2007; François et al., 2014; Francois & Schön, 2011).

Several ERP components have been related to the processing of simple statistical contingencies and to the establishment of more complex statistical relationships between the constituents of a stream. The MMN, the N1/P2 complex and the N400 are such findings. The characteristics and relevance of these ERP components to the SL research field are addressed more thoroughly in the following section.

2.6.1 Electrophysiological signatures of the processing of simple auditory regularities: The example of the MMN

The SL of simple auditory regularities has been studied mainly through oddball tasks. Oddball paradigms are characterized by the intermixed presentation of two stimuli, a standard

(STD) and a deviant (DEV), that occur at a high probability and at a low probability rate, respectively (Näätänen, Simpson, & Loveless, 1982; Nyman et al., 1990). The study of the neural responses to this sequence of stimuli allows researchers to understand how very simple statistical structures are learned and how the brain responds to deviations in the auditory background. When participants are instructed to perform an unrelated task (e.g., to watch a movie) while the oddball is presented -, the MMN ERP component is elicited (Näätänen, 2002; Näätänen, Paavilainen, Rinne, & Alho, 2007). The MMN is frequently followed by the P3a component – a sub-component of the P300 - reflecting an attention orienting response. The fact that the MMN can be elicited even when individuals are not actively attending to the stimuli makes it useful in studies with infants and children, along with populations with clinical disorders. When individuals are paying attention to the stimuli of the oddball task, DEV stimuli typically elicit the P300 (Courchesne, Hillyard, & Courchesne, 1977; Debener, Kranczioch, Herrmann, & Engel, 2002; Escera & Corral, 2007; Hillyard, Squires, Bauer, & Lindsay, 1971; Simons, Graham, Miles, & Chen, 2001; Sutton, Braren, Zubin, & John, 1965). Whereas the MMN is a fronto-central negativity peaking at approximately 100-250 ms, P300 is a centro-parietal positivity peaking approximately between 300-400 ms after stimulus' onset (Duncan et al., 2009). The MMN and the P300 reflect deviance detection mechanisms that are typically triggered by discriminable changes in an otherwise regular continuum of sounds (Polich, 2003). It is a well-established finding that both the MMN and P300 amplitude to DEVs are increased comparatively to STDs, and thus both ERP components are considered indices of deviance detection.

It should be noted that most studies reporting the occurrence of SL under none/minimal attentional demands used very simple input structures. Under these procedures and using these types of stimuli, researchers frequently find MMN to emerge as a consequence of a pre-attentive regularity extraction (Gottselig, Brandeis, Hofer-Tinguely, Borbély, & Achermann, 2004; Van Zuijen, Sussman, Winkler, Näätänen, & Tervaniemi, 2005). The input invariance can have distinct degrees of complexity, varying from simple rules with single tones to more abstract rules with complex sounds (Winkler, 2003). According to Schröger (1997), MMN is the reflection of a cascade of processes that lead to attentive auditory perception: the characteristics of every incoming stimulus are compared to the ones established by previous STD stimuli; if there is a mismatch between an incoming and an expected stimulus, a MMN is elicited, and if the MMN signal exceeds a given threshold, the detection of deviance reaches consciousness prompting a shift in attention towards DEV stimuli. This attention shift is indexed by the P3a ERP component. MMN also indexes the neural representation of a mechanism of expectancy violation detection when there is no actual stimulation: Herholz and colleagues (Herholz, Lappe, Knief, & Pantev, 2009) found that MMN could be elicited in the absence of physical sensory input/sensory memory trace. These results challenged previous conceptions about the component (e.g., the *memory trace theory*, see Risto Näätänen, Paavilainen, & Reinikainen, 1989). They suggest MMN elicitation not to be only based on local regularities extracted from sequences of physical stimuli; MMN also emerges from the violation of a prediction - that can be based on more abstract, long time-range global regularities that govern the auditory input - even in the absence of physical auditory stimulation (e.g., omission MMN, Yabe, Tervaniemi, Reinikainen, & Näätänen, 1997).

In 1999, Koeslch and colleagues (Koelsch, Schröger, & Tervaniemi, 1999b) provided the first evidence of enhanced pre-attentive auditory processing mechanisms in musicians. The authors tested the ERP responses of violinists and musical novices to chord and simple-frequency DEVs using a standard MMN paradigm. While a MMN to frequency-DEVs was elicited in experts and non-experts, a similar response to major chord deviations only occurred in musicians, showing how long-term auditory training modulated the neural pre-attentive discrimination of complex sounds. Besides, the authors found that only musicians exhibited an N200b-P300 complex in response to the task of actively detecting DEV chords. This finding indicates that musical training not only modulates neural mechanisms underlying sensory memory but it also interferes with higher-order cognitive processes. The detection of DEV melodic patterns is also modulated by musical expertise and the musical background of experts: the MMN is more negative to DEVs in melody contour in jazz and rock musicians than in classical musicians (Tervaniemi, Janhunen, Kruck, Putkinen, & Huotilainen, 2016).

MMN was also studied using the MEG technique. Consistently with the findings of electrophysiological studies (Herholz, Lappe, & Pantev, 2009; Van Zuijen et al., 2005), the MMNm - the magnetic equivalent of the MMN – was not modulated by musical expertise when it concerns to simple tones discrimination (e.g., frequency (Hz) DEVs) (Paraskevopoulos, Kuchenbuch, Herholz, & Pantev, 2012). Yet, the discrimination of contour and interval information (Fujioka et al., 2004) and the processing of high/low “voice” and in-key/out-of-key deviances (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2005) elicited increased MMNm in musicians compared to non-musicians. Other studies reveal an effect of training on MMN latency but not MMN amplitude: in the work of Nikjeh and collaborators (Nikjeh, Lister, & Frisch, 2008), musicians (i.e., vocalists) exhibited an earlier MMN compared to non-musicians. Of note, in this study the authors did not find statistically significant differences between instrumentalists and non-musicians, only between vocalists and non-musicians. This finding was extended later in 2009 in a study investigating preattentive ERP responses to different types of acoustic stimuli (Nikjeh, Lister, & Frisch, 2009): pure tones, harmonic sequences and speech syllables. Again, musical training modulated the latency but not the amplitude of the MMN response to all types of stimuli (i.e., musicians elicited an earlier MMN in all conditions).

The existing studies indicate that musicians are more sensitive, even at a preattentive level, to changes in the structure of sequences of sounds. At the behavioral level, there is evidence of faster and more accurate responses to pitch DEVs in musical experts (Boh, Herholz, Lappe, & Pantev, 2011; Nikjeh et al., 2008; Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005). At the neural level, musicians show an enhanced sensitivity to the metrical structure of sounds reflected in a more negative MMN to metre-incongruent DEVs (Geiser, Sandmann, Jäncke, & Meyer, 2010), which was accompanied by an improved performance (i.e., higher hit rates) in a behavioral metre perception test. Of note, being a musician seems to only represent an advantage when the task of detecting changes is not simple, such when pitch deviances are very subtle (e.g., 0.8% and 2% of change between STD and DEV stimuli, see Tervaniemi et al., 2005; one eight of a difference in a tone, see Nikjeh et al., 2008). The MMN can be elicited in response to DEVs, irrespective of musical training, when auditory regularities rely on pitch similarity (Zuijen, Sussman, Winkler, Näätänen, & Tervaniemi, 2004). In contrast, when regularities are derived from good continuation of pitch (i.e., tone-groups composed of ascending pitches instead of tone-group marked by ascending and descending pitches), a MMN response is only observed in musicians (Zuijen et al., 2004). Van Zuijen and collaborators (Van Zuijen et al., 2005) also found that whereas a violation of tempo could elicit a MMN irrespective of subjects' musical background, only musicians exhibited a MMN response to altered numerical regularities.

Koelsch and colleagues (Koelsch, Busch, Jentschke, & Rohrmeier, 2016) tested the neurophysiological responses to timbre sequences with varying TPs. The authors showed subjects triplets whose first two items were equiprobable but the third could occur with low, intermediate or high probability. The results revealed an early right anterior negativity in response to the onset of the third item, maximal at frontal electrodes, whose amplitude increased linearly as the probability of the occurrence of the item decreased. This was the first time a statistical MMN response was reported. However, this study did not compare musicians with non-musicians. Herholz and colleagues (Herholz, Lappe, Knief, et al., 2009) tested whether a MMN emerged if the expectation of the imagined continuations of well-known songs were violated and if its amplitude was modulated by musical expertise. Indeed, only musicians showed a MMN response to DEV tones. Behaviorally, musicians outperformed non-musicians (i.e., participants had to judge if a presented tone was a good continuation for the melody). This study showed that musical expertise modulates imagery of music and modulates MMN amplitude when it is elicited based on the global statistical knowledge of auditory regularities instead of sensory memory traces (Herholz, Lappe, Knief, et al., 2009). Musical background along with training practice can account for some variability in the way musicians' brains process acoustic information. Musicians who practice based on improvisation, on playing by the ear and on listening to records (e.g., aural practice strategies) show enhanced discrimination of contour and interval changes in melody-like patterns (Seppänen, Brattico, & Tervaniemi, 2007).

Whereas musical training enhances the ability to automatically process abstract characteristics/rules governing melodic sequences, it seems to not affect the sensory processing of simple tones (Fujioka et al., 2004; Nikjeh et al., 2008, 2009; Paraskevopoulos et al., 2012; Tervaniemi et al., 2005). This suggests that non-musicians and musicians show similar responses when processing deviations in streams of simple pure tones, whereas musicians have advantages in the processing of more abstract features in more complex auditory sequences. Koelsch and colleagues (Koelsch et al., 1999b) suggested that whereas the sensory memory traces of musicians for auditory stimuli were rich in various dimensions of acoustic information, these traces were simpler in non-musicians. This notion was supported by studies showing no MMN differences between musical experts and non-experts in auditory discrimination tasks based on simple pitch changes (i.e., simple, not complex sounds) (Fujioka et al., 2004; Nikjeh et al., 2008; Paraskevopoulos et al., 2012; Tervaniemi et al., 2005). The human brain may be tuned to process frequency information (i.e., pitch) with high efficiency because it is essential to accurately process verbal messages and, not less importantly, to perceive music. For that reason, non-musicians might be as good as musicians in detecting pitch changes.

In the context of SL research, MMN has been found in response to the SL of tone sequences (Paraskevopoulos et al., 2012) and to structural violations of speech streams (de Diego-Balaguer et al., 2007; François, Cunillera, et al., 2017). MMN is increased for stimuli that violate the structure of an artificial language (e.g., DEV words that violate the statistical structure of an artificial language) compared to stimuli that follow the structure of that language. This finding is taken as evidence that the brain is able to implicitly detect mismatches between stimuli that conform or do not conform to recently acquired knowledge about a streams' structure. To the best of our knowledge, only one study addressed the influence of musical training on the MMN response to statistical violations of tone sequences, showing no amplitude/latency differences between musicians and non-musicians (Paraskevopoulos et al., 2012).

2.6.2 N1-P2 complex

The N1 and P2 are two distinct ERP components that frequently emerge together, following one another in time (Ostroff, Martin, & Boothroyd, 1998; Whiting, Martin, & Stapells, 1998). They represent early components elicited by any perceptual event (externally or internally generated), that are associated with sensory/low-level perception mechanisms and that can be modulated by attention (Beres, 2017). The N1 indexes the automatic evaluation of the sensory features of a stimulus by the auditory cortex (Luck, 2005). The N1 has been deeply investigated in the domain of speech studies and it is considered to reflect also a predictive mechanism occurring during speech perception and production (Heinks-Maldonado, Mathalon, Gray, & Ford, 2005). The P2 component has been associated with auditory discrimination processes (Näätänen, 1992; Sheehan, McArthur, & Bishop, 2005) and, in the field of SL studies, with perceptual changes involved in the computation of dependencies (de Diego-Balaguer et al., 2007; Snyder, Alain, & Picton, 2006). Individuals highly trained in auditory discrimination (e.g., training in the discrimination of small frequency changes in pure tones, Bosnyak, Eaton, & Roberts, 2004; training in the discrimination of vowels sounds, Reinke, He, Wang, & Alain, 2003; musical training, Shahin, Bosnyak, Trainor, & Roberts, 2003; training in the identificiation of synthetic speech variants of a syllable, Tremblay & Kraus, 2002) exhibit increased P2 amplitude in distinct tasks. These studies suggest that the

P2 represents a marker of auditory neuroplasticity that is modulated by attention and perceptual learning (Hillyard, Hink, Schwent, & Picton, 1973; Reinke et al., 2003).

Recently, studies examining the neurophysiological responses during on-line SL of speech sounds reported N1 and P2 modulations in response to words' onset (Cunillera et al., 2006; de Diego-Balaguer et al., 2007; Sanders, Newport, & Neville, 2002). These early ERP components are not specific to language research but, since they index the sensory processing of a stimulus, they are expected to emerge in response to linguistic materials. Cunillera and colleagues (Cunillera, Laine, & Rodríguez-Fornells, 2016) found that the N1 and the P2 were elicited in response to word onsets, but the components' amplitude did not differ between experimental conditions (e.g., anchor [a language stream containing two words that were previously learned by the participants] vs. non-anchor [a language stream composed only of new words] vs. random [a unstructured stream made by the random juxtaposition of syllables] conditions). Before, Cunillera and collaborators (Cunillera et al., 2009) have also found no differences in the N1 component between language (structured stream) and random streams or across blocks during the learning phase. Larger N1 onset effects in response to words have been associated with improved performance (i.e., higher number of correct responses) in a offline SL test (Sanders et al., 2002). Similarly, in the study of Astheimer and Sanders (Astheimer & Sanders, 2011), N1 effects were increased for unpredicted vs. predicted words. In this study participants could predict words based on the cumulative presentations of their pairs. In another study, de Diego-Balaguer and colleagues (de Diego-Balaguer et al., 2015) presented participants with language streams with or without pauses between words in order to examine if pauses facilitated learning by eliminating the need to process adjacent computations. Both N1 and P2 were elicited in response to words. However, the insertion of pauses before words' onset led to N1 attenuation, independently of whether learning of non-adjacent dependencies was possible or not. It also led to a P2 enhancement (observed only when learning occurred), which corroborated the role of pauses as a factor modulating online segmentation mechanisms.

A study investigating the influence of statistical and stress (e.g., pitch accentuation) cues in speech segmentation also found N1 and P2 effects in response to words' onset (Cunillera et al., 2006). The results revealed differences in both ERP components between structured and random streams with stress cues, with amplitude increases observed for the stream combining both statistical and stress patterns. However, when the language was devoid of pitch accentuations, N1 was similar between structured and random streams. Only the P2 (with a frontal and medial distribution) differed between structured streams as a function of the presence of stress cues (e.g., higher amplitudes for stressed words). Additionally, a study from de Diego-Balaguer and colleagues (de Diego-Balaguer et al., 2007) with rule-words (i.e., the first syllable of a trissillabic word predicted the last one) showed the P2 to increase throughout the blocks of the learning phase. Since the mean amplitude of the P2 during the 3rd block (e.g., block with maximal amplitude) was correlated with participants' performance in a rule-learning behavioral task, the authors considered it a marker of rule learning. Based on that evidence, the authors investigated P2 modulations dividing the sample in groups of good and poor learners. The P2 was more positive in good learners, exhibiting a linear increase that was maximal at right fronto-central electrodes. This finding corroborated previous evidence of a positive correlation between the P2 amplitude and the scores on the perpetual segregation of a continuous stream of sounds in two distinct streams (Snyder et al., 2006).

The modulation of the N1 and P2 components during online SL has also been observed for non-linguistic stimuli. Koelsch and colleagues (Koelsch et al., 2016) found a more negative N1 (right lateralized) amplitude for the first relative to the second item of timbre triplets, which increased linearly as a function of exposure to the sounds, a result that the authors took as evidence of developing word segmentation. A similar result was found for pure tri-tone sequences in a study that divided the sample in three groups according to their level of performance in a behavioral post-learning task (Abla, Katahira, & Okanoya, 2008): high learners showed larger N1 onset effects (maximal at middle frontal sites) during the first learning session (among three), while this effect only emerged later (i.e., subsequent learning sessions) for middle learners and did not emerge at all for participants with a lower performance. These findings indicate that, during exposure to the sounds, participants adopt distinct learning strategies that are reflected in N1 amplitude modulations. Therefore, the temporal dynamics of N1 might be informative about the participants' learning abilities.

From the few studies that investigated SL in musicians, two reported modulations of the N1 that interacted with musical expertise. Despite the emergence of the N1 component in both musicians and non-musicians, the learning of musical tone-triplets elicited N1 onset effects only in musicians during the early part of the learning phase (e.g., first stream) (Vasuki, Sharma, Ibrahim, & Arciuli, 2017b). It is worth mentioning that both groups learnt the statistical structure of the auditory stream, as demonstrated by behavioral findings. Also, when examining the learning of a sung language, François and colleagues (François & Schön, 2014) found the N1 to emerge in both groups without significant differences, despite musicians showing an increased N1 during the first minute of exposure. The authors interpreted these divergent results as consequence of discrepancies between the stimulus features of their study and previous reports, namely the variable attack times of the consonants composing the words that could have led to larger variability in the latency of the components.

The accumulating evidence indicates the N1 onset effect, typically found for triplets, to be mainly observed in participants exhibiting the higher number of correct responses in behavioral SL tasks (Abla et al., 2008; Sanders, Ameral, & Sayles, 2009; Sanders et al., 2002; Vasuki et al., 2017b). This finding might be explained both by the use of more effective learning strategies in high learners and/or the successful recruitment of additional resources, such as selective attention, that might aid the processing of regularities (Daltrozzo & Conway, 2014). The P2 amplitude increase observed in SL paradigms seems to indicate that the P2 reflects a perceptual change (due to the capture of attention by the features of the streams) that aid in perceptual grouping of adjacent and non-adjacent dependencies. The evidence agrees with reports suggesting that, in order to efficiently manage the incoming speech information during communicative situations, individuals need to dynamically attend to different acoustic cues over time (Astheimer & Sanders, 2011; Toro et al., 2005). Increased levels of attention are needed at the initial stages of exposure to the input. As repetitions of stimuli accumulate and learning progresses attention decreases. Attentional resources need to be reoriented during speech segmentation so that the brain maximizes the computation of dependencies. Therefore, attention seems to play a relevant role in ASL, which might be reflected in N1 and P2 amplitude modulations. Nonetheless, there is a lack of studies confirming it: more SL studies are required to corroborate the involvement of these early ERP components in the online learning process and in its interactions with musical expertise.

2.6.3 The N400

The N400 was first reported almost forty years ago by Kutas and Hillyard (1980). In 1980, the authors employed a modified oddball paradigm to investigate the role of sentence context in word recognition. They presented participants with 75% of sentences whose 7-words were semantically congruous, and with 25% of sentences whose endings could be highly improbable or invalid (this manipulation is referred to as the *anomalous sentence paradigm*). The incongruous endings elicited a parietally-distributed, large negativity peaking at 400 ms. Since then, the N400 has been the focus of approximately a thousand studies on a myriad of topics from language processing, to object and face processing, to semantic memory and clinical disorders (Kutas & Federmeier, 2011).

A negative component emerging from subtraction procedures typically studied in SL research has been referred to as N400-like component. In this case, it tends to peak between 200 and 600 ms and to be distributed over centro-parietal regions (Daltrozzo & Conway, 2014). Years of studies addressing the functional significance of the N400 with different experimental manipulations and stimuli from different types/modalities have led to the conclusion that the N400 does not merely reflect the violation of a given pattern. Instead, its elicitation relates to the processing of meaning, to stimulus probability, to processes involved in lexical search, and semantic and recognition memory (Kutas & Federmeier, 2010).

Typically, it is the amplitude of the component that is more prone to manipulation, with higher amplitudes being associated with low expectations and increased difficulty in stimulus processing (Kutas & Federmeier, 2009), whereas the latency seems to remain stable. The N400 is sensitive to specific linguistic manipulations, but they do not include syntactic ones, which typically elicit the P600 or the left anterior negativity (Osterhout & Mobley, 1995). When elicited by auditory stimuli, the N400 tends to be longer (i.e., to take more time), appear earlier and to be larger at frontal electrodes (Kutas & Van Petten, 1994) as opposed to when it is elicited by visual stimuli. The N400 exhibits modality sensitivity but not modality specificity: there are differences in the scalp topography of the N400 response to distinct meaningful stimuli, along with similarities regarding the shape and timing of the component (Ganis, Kutas, & Sereno, 1984; Olivares, Iglesias, & Antonieta Bobes, 1999; van Petten & Rheinfelder, 1995). Regarding the influence of attention on the component, most studies converge to show that N400 elicitation is modulated by the individuals' selective attention to the stimuli: it is not fully automatic but also not fully dependent on high levels of awareness (Kutas & Federmeier, 2011).

A study investigating second language word learning compared learners of French (with approximately 14 hours of instruction) with non-learners (i.e., people not learning/without knowledge of French), across three sessions, in a task where word-pairs were presented (McLaughlin, Osterhout, & Kim, 2004). The pairs could be composed of semantically related words (e.g., chien-chat), semantically unrelated words (e.g., maison-soif) or words and pseudo-words (mot-nasier). Whereas non-learners did not show any N400 modulation in response to the pairs, learners show progressively increasing N400 amplitude across sessions to pseudo-words in comparison with related and unrelated words. McLaughlin and colleagues (2004) found that learners' changes in the N400 amplitude were accompanied by chance-level responses in a LDT (i.e., word/non-word) occurring in parallel with the EEG recording, which supports the role of the N400 as a relevant implicit measure of learning.

The N400 has also been associated with predictive processing. N400 amplitude was found to be less negative in response to words that were predicted given a preceding context as opposed to words that mismatched the participants' predictions (Van Petten, Coulson, Rubin, Plante, & Parks, 1999). This evidence also shows that the brain is able to differentiate between predictable and unpredictable words even before the words are behaviorally recognized, that is, as soon as perceptual mismatches between predictions and incoming input are detected.

The role of N400 as a neural index of word learning has been vastly explored in recent studies (e.g., Batterink & Neville, 2011; Borovsky, Kutas, & Elman, 2010; Mestres-Missé, Rodriguez-Fornells, & Münte, 2007; Perfetti, Wlotko, & Hart, 2005). Studies with adults have shown that the N400 is associated with semantic processing since it is characterized by more negative amplitude for stimuli that are novel or unexpected given their context (King & Kutas, 1995; Kutas & Federmeier, 2011; Neville, Mills, & Lawson, 1992; Pulvermüller, Lutzenberger, & Birbaumer, 1995) and it is less negative when words are predictable due to high cloze probability (Federmeier & Kutas, 1999; Federmeier, McLennan, Ochoa, & Kutas, 2002). Generally, known words elicit reduced N400 compared to unknown words (Batterink & Neville, 2011; Borovsky et al., 2010; McLaughlin et al., 2004; Mestres-Missé et al., 2007). Similar findings were observed in studies with children, who show less negative N400 for words to which they were exposed to while having to associate meanings, compared to words with no meanings associated (Abel, Schneider, & Maguire, 2018).

Despite the huge amount of research investigating the N400 due to its high relevance to our understanding of language processing, the functional significance of the component is still a matter of debate. The difficulty in defining the functional significance of the N400 comes from the myriad of theoretical accounts on language processing that a great amount of data have generated. A review on the topic suggests the N400 effect to be the output of facilitated lexical access (Lau, Phillips, & Poeppel, 2008). This hypothesis is supported by a study from Szewczyk and Schriefers (2018) that found that the N400 is eliminated when incongruent target words are predicted by previous words, which suggests that the N400 component is a marker of prediction and supports the *preactivation theory* of N400. This conceptual account postulates N400 attenuation when a given word is predicted based on its preceding context (Kutas & Federmeier, 2011; Lau et al., 2008; Van Berkum, 2008). If the word is predictable, its representation is pre-activated and its subsequent processing is easier, then resulting in a reduced or inexistent N400 effect; when it is not, lexical access has to occur in full and a N400 is elicited and/or larger. More recently, what seems to be a consensual computational account has emerged: it suggests that the N400 component reflects "*the input-driven update of a representation of sentence meaning - one that implicitly and probabilistically represents all aspects of meaning as it evolves in real time during comprehension*" (Rabovsky, Hansen, & McClelland, 2018, p. 693). Based on the available evidence, a growing body of research has probed the N400 during word learning tasks. By doing that it is possible to examine when the mental representations of words-to-be-learned become so robust that the brain responds to them as if they were part of the lexicon.

In the context of SL studies, an N400-like component has been consistently found and proposed to represent a neurophysiological correlate of speech segmentation (Abla et al., 2008; Buiatti et al., 2009; Cunillera et al., 2009). In the realm of word learning studies with adult and children samples (for a review, see Rodriguez-Fornells, Cunillera, Mestres-Misse, & de Diego-Balaguer, 2009), the component typically emerges in the 200-500 ms latency range with a fronto-central distribution, which explains why it is named an N400-like effect. N400-like amplitude increases (i.e., more negative amplitude) were proposed to represent an index of the segmentation of acoustic streams, either speech or non-linguistic streams (Abla et al., 2008; Buiatti et al., 2009; Cunillera et al., 2006; Francois & Schön, 2011; Sanders et al., 2002).

In AGL experiments, the TPs within words are higher than between word boundaries. For that reason, the onset of a new word is difficult to predict. Sanders and colleagues (2002) found the N400 to have increased amplitude for the initial syllables of words when participants were exposed to an artificial speech stream. This modulation of the N400 is known as the word onset effect, a correlate of successful stream segmentation (Abla et al., 2008; Cunillera et al., 2006; Sanders et al., 2002; Vasuki et al., 2017b). More recently, François and colleagues (2017a) and Batterink and Paller (2017) implemented different AGL paradigms using speech stimuli by presenting participants with random and structured streams of syllables' sequences. Both studies found an N400-like component to be larger in response to the sequences in the structured stream, despite differences in the topographical distribution and latency of the component (e.g., early FN400 [200-350 m], maximal at fronto-central electrodes in François et al., 2016; N400-like [300-500 ms], maximal at midline sites in Batterink & Paller, 2017). This evidence is also supported by the findings of Cunillera and collaborators (Cunillera et al., 2016, 2006). The authors taught participants two new words (anchor words) and then presented them with speech streams in which those words were present (anchor word) or absent (non-anchor condition) (Cunillera et al., 2016). More negative N400 amplitudes to words embedded in the non-anchor condition were found, which the authors took as evidence of more difficult word learning in the non-anchor condition due to the absence of anchor words. Similarly, de Diego-Balaguer and colleagues (2007) found N400 to have increased amplitude in response to words in 'good learners' (participants with higher number of correct responses in post-learning tests) during the first min of the learning phase.

Similar results were found for non-linguistic stimuli. Abla and collaborators (2008) presented three continuous streams of pure tone triplets to which participants should attend without intentionally analyzing the sounds. This learning phase was followed by a test phase (e.g., behavioral task) and, according to their performance level, participants were divided into three groups: high, middle and low learners. High learners showed more negative N400 – a triplet onset effect - during the presentation of the first stream of sounds. Moreover, N400 amplitude decreased across sessions for the high learners, increased for middle learners and showed no modulations in the low learners' group. A similar effect was also found in an experiment testing the SL of geometric shapes for the first streams and only in high learners (Abla & Okanoya, 2009). Previously, in the de Diego-Balaguer and colleagues' (2007) study, the modulation of the N400 was only observed during the first minute of stream exposure (maximal at the second block of the learning phase at central sites) and vanished after this point. Hence, it is important to highlight that the N400 effect tends to be observed early in the learning phase, diminishing over time as learning takes place (for examples, see Abla et al., 2008; Cunillera et al., 2009; de Diego-Balaguer et al., 2007; Vasuki et al., 2017).

More recent studies showed the N400 effect to be modulated by musical training, both when in response to artificial speech and non-linguistic stimuli, and also to fluctuate over time (François et al., 2014; Vasuki et al., 2017b). A study from Vasuki and collaborators (2017) tested ASL with pure tone triplets and found an N400-like component (with centro-parietal distribution) in both musicians and non-musicians. Nevertheless, whereas in musicians the effect was stable across the three streams composing the learning phase, in non-musicians it was only observed from the second stream onwards. François and collaborators (2014) presented participants with a continuous stream of trisyllabic sung words – e.g., words whose syllables had varying pitches – and analyzed the ERPs as learning unfolded over time during the four blocks of familiarization. The fronto-central N400-like effect exhibited an inverted u-shaped learning curve in musicians: i.e., an amplitude increase from the first to second block – blocks stand for time bins of 1'20'' duration - followed by an asymptote between the second and third blocks and a decrease from then until the end of the phase. In non-musicians it increased linearly. The inverted U-shaped learning curve was considered to reflect the translation of syllable sequences into word units, as well as the consolidation of the learned units due to repetition (familiarity effect). This was taken as evidence for faster speech segmentation in musicians relative to non-musicians. Interestingly, the level of performance in a behavioral linguistic task was higher in individuals who showed maximal N400 amplitude early on during the learning phase (e.g., second block). These findings are not only suggestive of differences in the neural dynamics underlying speech segmentation depending on the individuals' musical background, but also of the utility of the N400-like as an implicit index of successful SL.

The reviewed electrophysiological evidence indicates that ASL involves the formation of predictions whose violation or confirmation is associated with specific patterns of neural activity. Altogether, N1, P2 and N400-like components represent the ERP signatures that have been most frequently associated with the online SL of speech and non-linguistic streams (Abla et al., 2008; Cunillera et al., 2009, 2006; de Diego-Balaguer et al., 2015; de Diego-Balaguer et al., 2007; Sanders et al., 2002; Vasuki et al., 2017b). Until now, only a few studies have examined the effects of musical training on ASL using both neurophysiological and behavioral measures (see Table 1.). Nevertheless, most of the reviewed studies ignore the contribution of top-down influences to the modulation of these ERP effects. There are no reports providing an examination of the online neural dynamics underlying ASL when participants are actively attending to the input vs. when they are not, an important issue that needs to be addressed more thoroughly.

2.7 The role of attention in ASL

Research probing artificial language learning has proposed that artificial grammars are learned implicitly since, in most studies, participants are not provided with information regarding the grammars' structure and tend to perform above chance in behavioral tests (Batterink & Neville, 2013; Batterink, Reber, & Paller, 2015; Reber, 1976). However, participants are usually told to focus their attention on the speech streams. Would they perform above chance in the SL tests if attention was focused elsewhere? Despite evidence pointing to implicit learning mechanisms underlying speech segmentation, the existing studies are not clear regarding the influence of attention on these mechanisms, not clarifying if learning would still occur if there was not an active effort to focus attention on the auditory input.

The bulk of SL studies suggest that SL is incidental (Fiser & Aslin, 2005) and automatic (Fiser & Aslin, 2002a; Pierre Perruchet & Pacton, 2006; Saffran, Aslin, et al., 1996; Turk-Browne, Jungé, & Scholl, 2005). Research on both AGL and natural languages has reported effects on early ERP components in response to syntactic violations (de Diego-Balaguer et al., 2007; Friederici, 2002; Mueller, Bahlmann, & Friederici, 2008). This evidence suggests that the processing of syntactic relationships occurs under minimal attentional conditions, which in turn indicates that this processing involves implicit mechanisms. Yet, the processing of statistical regularities is different from syntactic processing. The impact of specific experimental manipulations, such as changes in attentional load or lack of selective attention on SL, has just recently started to be investigated. There are claims that attention is a prerequisite for SL (Ahissar & Hochstein, 1993; Gilbert, Ito, Kapadia, & Westheimer, 2000). Research comparing dual task settings that manipulate attentional load (e.g., high vs. low attentional load) indicate that SL performance decreases as a function of diverted attention and increased attentional load, both in visual and auditory tasks (Toro et al., 2005; Turk-Browne et al., 2005).

Table 1

Studies Investigating ASL as a Function of Musical Training Using Linguistic Sounds.

Linguistic sounds						
Authors	Study Design	Sample size and features	Stimuli	Procedure	Neural Results	Behavioral Results
François & Schon, 2011	Cross-sectional	16 M & 20 NM	sung words	- Learning Phase - Test Phases [Linguistic & Musical (2AFC tasks)]	Linguistic Test: N1: M > NM Musical Test: N1/P2: M > NM	Linguistic Test: - M (↑ 50%) = NM (↑ 50%) Musical Test: - M (=50%) = NM (=50%)
François, Chobert, Besson & Schon, 2013	Longitudinal [3 evaluation moments (i.e., before training (T0), after 1 year (T1), and 2 years (T2)]	12 M & 12 NM (8-yo children randomly assigned to music or painting training groups)	sung words	- Learning Phase - Test Phase [Linguistic test (2AFC task)]	Test Phase Familiarity effect (N450-550): T0: M = NM; T1: no data; T2: M > NM.	T0: M (=50%) = NM (=50%) T1: M (↑ 50%) > NM (=50%) T2: M (↑ 50%) > NM (=50%)
François, Jaillet, Takerkart & Schon, 2014	Cross-sectional	13 M & 13 NM	sung words	- Learning Phase - Test Phases [Linguistic & Musical (2AFC tasks)]	Learning Phase: N1/P2: no effects; N400: M – inverted U-shape learning curve; NM – linear increase.	Linguistic Test: - M (↑ 50%) = NM; - ↑ performance, ↑ N400 during the 2nd time bin. Musical Test: - M = NM

Tonal sounds (continuation of Table 1)

Authors	**Study Design**	**Sample size and features**	**Stimuli**	**Procedure**	**Neural Results**	**Behavioral Results**
Paraskevopoulos, Kuchenbuch, Herholz & Pantev, 2012	Cross-sectional	15 M & 15 NM	Tri-tone sequences	- Learning Phase (oddball paradigm) - Test phase (2AFC; pairs of STD and DEV sequences)	Learning Phase: mMMN: M = NM mP50: M > NM	M (=50%) = NM (=50%)
Emerson, Daltrozzo & Conway, 2014	Cross-sectional	6 M & 7 NM[a]	Complex tones	Predictor-target paradigm[b]	P300-like: M = NM; CNV: M = NM; Late FCN: M = NM.	Reaction times: M = NM
Vasuki, Sharma, Ibrahim & Arciuli, 2017b	Cross-sectional	17 M & 18 NM	Tone triplets	- Learning Phase (3 streams paired with oddball detection task) - Test phase (2AFC)	Learning phase: N1 TOE: M > NM (1st stream); P2: M = NM; N400 TOE: M > NM (1st stream).	M (↑ 50%) > NM (↑ 50%)
Vasuki, Sharma, Ibrahim, Arciuli, 2017a	Cross-sectional	24 M & 24 NM (9-11 yo children)	Tone triplets	- Learning Phase (paired with oddball detection task) - Test phase (2AFC)	Learning phase: P1: M = NM; N250 TOE: M > NM.	M (↑ 50%) > NM (↑ 50%)

Note. M = Musicians; NM = Non-Musicians; 2AFC = Two-alternative forced-choice task; (↑ 50%) = significantly higher than chance level; (=50%) = not different from chance level; N450-550 = negativity elicited in the 450-550 ms latency range; yo = years old; STD = Standard stimulus; DEV = Deviant stimulus; mMMN = magnetic MMN; mP50 = Magnetic P50; CNV- Contingent Negative Variation; FCN- Fronto-central Negativity; TOE – triplet onset effect.

[a] The authors compared participants with low vs. high musical aptitude. [b] Modified version of the standard oddball paradigm composed of standard, predictor and target tones; participant's task was to press a button whenever a target occurred.

Turk-Browne and colleagues (Turk-Browne et al., 2005) manipulated attention in five different behavioral experiments. The authors found that visual statistical learning (VSL) of temporal sequences is modulated by the allocation of attention, suggesting that selective attention is required for VSL to occur. The same was found when testing SL of shapes' sequences (Baker, Olson, & Behrmann, 2004). Thus, the evidence supports the idea that VSL is not passive: both attention and perceptual grouping are needed along with statistical regularities' processing so that unitary object representations are formed. Besides, visual location studies have found that manipulating the attentional load by means of dual-task conditions have detrimental effects on participants' performance on serial reaction time tasks – i.e., paradigms designed to evaluate sequence learning (Shanks, Rowland, & Ranger, 2005). Still, similar studies that systematically investigate the role played by attention in the SL of auditory materials are lacking.

Only a few auditory studies have attempted to clarify this issue (Paraskevopoulos et al., 2012; Toro & Trobalón, 2005). Toro, Sinnett and Soto-Faraco (2005) investigated the impact of high attentional load on ASL using concurrent tasks in parallel with speech segmentation, namely: discriminating alterations in a concurrent noise stream; discriminating pitch deviations in the actual speech stream; and, indicating picture repetitions in a stream of images. With these manipulations, the authors were able to test the impact of attentional load across sensory modalities and within the same auditory modality using different tasks. Behavioral performance in recognition tests was at chance level as a function of diverted attention, independently of the experimental manipulation. On the contrary, performance was significantly higher than chance for those participants asked to attend to the speech stream. Recent electrophysiological evidence (Ding et al., 2018) support these results indicating that, for the brain to chunk syllables into complex sequences (e.g., words), individuals need to actively attend to the sounds. Therefore, the existing evidence suggests that, at the neural level, attention is needed for the brain to extract the statistical structure governing speech streams. These results agree with the limited evidence suggesting that attentional resources need to be actively focused on the SL task so that knowledge of the statistical structure of the input reaches awareness, i.e. attention is needed for standard and deviant patterns to be detected behaviorally (Cunillera et al., 2009; Toro et al., 2005; Turk-Browne, Scholl, Johnson, & Chun, 2010). This is not surprising: SL seems to be enhanced when individuals are actively attending to the input which leads sensory, low–level processing to become attention-driven and, simultaneously, more efficient (Daikoku & Yumoto, 2017; Ding et al., 2018; López-Barroso, Cucurell, Rodríguez-Fornells, & de Diego-Balaguer, 2016).

While research investigating the relationship between attention and ASL is limited, studies dedicated to the understanding of the mediating role of attention in the effects of musical training on ASL are almost inexistent. Despite some preliminary evidence of faster speech segmentation mechanisms and enhanced ASL in musicians compared to non-musicians (François et al., 2014; Francois & Schön, 2011; Vasuki et al., 2017b), the literature lacks reports testing the associations between attention, musical training and ASL. Previous research showed that musicians exhibit increased top-down modulatory effects of attention on ERPs (e.g., P2, P3) emerging in response to auditory stimuli (Tervaniemi et al., 2009). This is consistent with findings of participants with vast formal musical training recruiting more extensively brain networks involved in cognitive control and sustained attention compared to non-musicians (Pallesen et al., 2010). By making a more efficient use of attentional and cognitive control resources and strategies, the decoding of the statistical information that defines the input's structure might be facilitated in musicians. These advantages might work together to facilitate auditory learning in musicians. An increase in selective attention to sounds in musicians might then be at the core of recent claims of enhanced speech and tone-sequences segmentation in musicians (Strait & Kraus, 2011; Zendel & Alain, 2009, 2014). For that reason, it is necessary to rule out if attention mediates the relationship between musical training and speech segmentation and auditory learning in general. To the best of our knowledge, there is only one MEG report examining the influence of musical training on the SL of tone sequences under minimal attentional conditions (e.g., while watching a silent movie; see Paraskevopoulos et al., 2012). Despite the absence of group differences in MMN amplitude, the study showed that musical training modulated the magnitude of the STD-DEV difference in the magnetic P50 - an index of sensory gating that occurs approximately at 50 ms post-stimulus onset (Boutros et al., 1995; Boutros & Belger, 1999). This study is novel in two aspects: (1) it shows that the P50 response was altered by the statistical structure of varying-pitch sequences, which suggests TPs to be computed very early (at approximately 50 ms) in the course of processing, and (2) it demonstrates an effect of musical training on the P50 amplitude. Even when sounds are task-irrelevant, the brain makes predictions about incoming input and responds differently to sound transitions with varying probabilities. This is corroborated by a more recent MMN study (Mittag, Takegata, & Winkler, 2016) that provided evidence that the auditory system prioritizes the calculation of TPs of tone-pairs over probabilities of the occurrence of single sounds. Moreover, the study of Paraskevopoulos and colleagues (2012) also showed that neither musicians nor non-musicians explicitly learned the tonal sequences, which suggests that participants only memorize the statistical structure of the tonal stream if attention is allocated to the task. Nevertheless, the study also indicated that attention is not a prerequisite for the computation of TPs between simple tones to occur, but it is required for the sequences of sounds to be retrieved as unified, three-tone-patterns.

2.8 Considerations on the reviewed literature

The research reviewed in the present chapter explored the impact of musical training on brain plasticity and musicians' advantages in auditory perception. The existing evidence suggests that these advantages might impact upon high cognitive functions and generalize to speech processing. Yet, the mechanisms by which musical training changes musicians' brains and makes them better able to process speech are still a matter of debate. We reviewed evidence demonstrating that ASL, benefiting from long-term auditory training, is at the core of musicians' enhanced abilities to process speech. After discussing behavioral and electrophysiological data on these topics, some considerations emerged. Below, we present these considerations that represent a critical approach to the existing evidence, aiming to provide a theoretical framework for the experimental studies reported in the present Dissertation.

2.8.1 Is ASL the key to understand musicians' advantages in speech processing?

Understanding how musical training fosters learning-related plasticity and develops high cognitive abilities, namely speech processing skills, is of great importance considering both the scientific interest in musical training as a neuroplasticity model and its societal implications. One of the hypotheses is that the lifelong auditory training of musicians results in improved pitch perception skills. Since spoken language relies on the modulation of acoustic parameters for communicative purposes (e.g., speech prosody) and the decoding of this information is highly dependent on pitch processing, musicians' enhanced pitch skills might explain their overall advantages in the processing of speech input. This hypothesis would explain several differences between musically-trained and naïve subjects, particularly: differences in the processing of speech prosody and pitch violations, both in mother and foreign languages (Bidelman et al., 2011b; Kraus & Chandrasekaran, 2010; Marques et al., 2007) and, differences in the brainstem encoding of linguistic pitch patterns as well as in hearing speech in noise (Strait & Kraus, 2011) and reading abilities (Parbery-Clark et al., 2009). However, this explanation is less suitable to explain the facilitated speech segmentation and language learning in musicians. SL, as a domain-general learning mechanism is common to both music and language (Bergstrom, Howard, & Howard, 2012; Conway & Christiansen, 2005; Frost et al., 2015). The ability to perceive and encode sequences of syllables is essential for speech segmentation and is one of the key elements for the learning of new languages to be achieved (Mirman et al., 2008; Pierre Perruchet & Tillmann, 2010; Saffran et al., 1999; Saffran & Wilson, 2003). The long-term study of musical pieces implicates deep perceptual learning of metre, rhythm and western music contingencies (Loui & Wessel, 2007; Virtala, Huotilainen, Partanen, & Tervaniemi, 2014; Zhao, Lam, Sohi, & Kuhl, 2017), which contributes to a great knowledge about auditory input's structure and might facilitate the encoding and storage of representations about statistical regularities. Expert musicians might strongly develop the capacity to process TPs in music, which in turn, generalizes to the processing of statistical regularities in other types auditory input. Moreover, besides practice-related changes in primary and associative auditory cortex areas, musical training was associated with changes in the brain's language network, namely in frontal areas involved in speech segmentation (e.g., prefrontal cortex, see Cunillera et al., 2009; McNealy, Mazziotta, & Dapretto, 2006; Kristin McNealy, Mazziotta, & Dapretto, 2010) and in fiber bundles connecting auditory to motor areas (Bengtsson et al., 2005; Oechslin, Meyer, & Jäncke, 2010; Wan & Schlaug, 2010) implicated in word learning (López-Barroso et al., 2013). This evidence supports the idea that the possible advantages of musicians in the processing of speech emerge from enhanced (generalized and modality-specific) abilities to compute statistical regularities. According to this hypothesis, musicians would differ from non-musicians in the learning of artificial spoken languages, with or without melodic structure, and in the learning of sound sequences, generically. Unfortunately, there are no studies addressing this hypothesis with a systematic approach.

2.8.2 Musicians' statistical learning of simple and complex auditory regularities

Oddball and SL paradigms are two approaches used to test an individual's capacity to extract regularities from environmental input. Standard oddball paradigms intermix the presentation of a (high frequency) STD stimulus and of a (low frequency) DEV stimulus. These paradigms are useful for testing behavioral and brain responses to auditory streams with basic statistical structures, consisting of high and low probability stimuli. Since the input's structure is defined by the number of repetitions of two stimuli, the extraction of patterns is relatively easy. On the other hand, traditional SL paradigms allow the examination of how more sophisticated regularities are computed, namely TPs between streams' units, which is an essential ability for the accurate prediction of complex incoming input.

MMN studies have shown that musicians, compared with non-musicians, have an enhanced sensitivity to detect violations in chords (Koelsch, Schröger, & Tervaniemi, 1999a), in contour, interval, "voice" and "key" information (Fujioka et al., 2004, 2005) and in the numerical and metrical structure of a melody (Geiser et al., 2010; Van Zuijen et al., 2005). Conversely, musicians seem to be no different from non-musicians in the automatic processing of simple irregularities stemming from basic acoustic features (e.g., Fujioka et al., 2004; Nikjeh et al., 2008, 2009; Paraskevopoulos, Kuchenbuch, Herholz, & Pantev, 2012; Tervaniemi et al., 2005). These findings may suggest that, when it comes to the pre-attentive processing of auditory input, training only makes difference when the sounds are complex, such as when there are changes in multiple acoustic cues or in abstract features of complex-music sequences. Based on that assumption, musical training would bring changes to the learning of auditory sequences only when they are characterized by complex statistical structures (e.g., speech). Musical training might develop the ability of musicians to implicitly calculate TPs between the units of an auditory stream. If so, behavioral and neural advantages in the SL of sounds would only be observed in musicians when the task of abstracting patterns from a stream is challenging, but this hypothesis needs further investigation. Therefore, a systematic examination of musicians' and non-musicians' SL of simple vs. complex auditory regularities is lacking.

2.8.3 The need for studying the neural dynamics of learning

Much of the SL research has addressed how individuals learn different materials by testing their performance with 2AFC tasks (Batterink & Paller, 2017; Cunillera et al., 2009; François et al., 2013), LDTs (Batterink & Neville, 2011; François, Cunillera, et al., 2017; van Petten & Rheinfelder, 1995), serial reaction time tasks (Eimer, Goschke, Schlaghecken, & Stürmer, 1996; Schultz, Stevens, Keller, & Tillmann, 2013; Shanks et al., 2005; Terry, Stevens, Weidemann, & Tillmann, 2016), among others. The findings derived from these studies provide information on the behavioral outcomes of learning. These findings are intrinsically related to the integrated representations of the input's structure stored in the long-term memory and are dependent on conscious retrieval and judgment processes (Batterink & Paller, 2017; Siegelman, Bogaerts, Kronenfeld, et al., 2018). For those reasons, behavioral results can represent biased measures of SL. An illustration of this problem would be the situation in which an individual presents poor SL outcomes, despite having intact SL abilities, due to impaired memory storage. Besides, the main shortcoming of these behavioral approaches resides in their unfeasibility to uncover what happens during the learning process *per se*. Therefore, studying the mechanisms involved in SL and how they change over the learning process is necessary for a deeper understanding of the phenomenon. Neuroimaging techniques are well-suited approaches to shed light on the neural underpinnings of SL. ERPs, more specifically, have had an important role in the understanding of how cognitive processes unfold in the brain with a milisecond precision. Applied to SL research, the use of electrophysiology is particularly well suited for two reasons: on the one hand, it provides a unique window into the electrophysiological mechanisms associated with the processing of regularities and, on the other hand, by studying the temporal dynamics of ERPs, it is possible to observe how the brain correlates of SL change over time. When applied to the study of the relationship between musical training and ASL, ERPs are of great importance because they allow both to test whether expertise can modify the on-line processes implicated in auditory segmentation and to examine whether changes in the neural dynamics are stimulus-specific, i.e., vary as a function of stimulus type.

2.8.4 The need to combine implicit and explicit measures of learning

The bulk of studies on SL (except for some reports, e.g., Cunillera et al., 2009, 2006; Karuza et al., 2013; McNealy et al., 2006; Kristin McNealy et al., 2010; Turk-browne, Scholl, Chun, & Johnson, 2008) have followed a common procedure: a learning phase followed by an offline test during which behavioral measures are collected. This approach is limited in two ways: first, as seen above, it does not allow the examination of the time course of learning (i.e., it is focused on the learning outcomes rather than on the learning process *per se*); and second, the outcomes are frequently assessed only through post-learning, explicit behavioral tasks that can be biased by previously mentioned factors. The behavioral tasks provide important information on a listener's capacity to store representations of the streams' units in long-term memory and to retrieve them. However, they do not allow the evaluation of the neural mechanisms underlying SL *per se*. Furthermore, with such methods, subjects' responses are not dissociable from their memory capacity, which might compromise how SL mechanisms are examined.

The claim for implicit measures of learning comes from various research reports suggesting that explicit behavioral responses are not sensitive enough to measure speech segmentation or sequence learning in general (Batterink, Reber, Neville, & Paller, 2015; Cleeremans, 2006; Francois & Schön, 2010; McLaughlin et al., 2004; Tremblay, Kraus, & McGee, 1998). Therefore, the literature lacks studies that collect implicit measures (e.g., ERP) of learning before explicitly examining the outcomes of the learning phase with behavioral tasks. These implicit measures will inform on the neural dynamics underlying the processing of stimuli that violate a pattern or conform to a previously learned statistical structure, while providing more objective measures of the learning outcomes. Combining ERPs and behavior as implicit and explicit measures of learning, respectively, will offer a more comprehensive understanding of the cognitive mechanisms involved in auditory perception and integration.

2.8.5 The role of stimulus type in ASL and the need to investigate within vs. cross-domain neuroplasticity

SL is a domain-general mechanism that operates on different types of stimuli, despite the classical association of this learning system with language learning. It has been tested in children (Arciuli & Simpson, 2012; Kidd & Arciuli, 2016) and adults (Batterink, 2017; Cunillera et al., 2009; de Diego-Balaguer et al., 2015; François, Cunillera, et al., 2017a; Mirman et al., 2008) and with a myriad of stimuli (tones, see Creel et al., 2004; Saffran et al., 1999; chords, see Daikoku et al., 2016; speech sounds, see Saffran, Newport, et al., 1996; Toro et al., 2005; morse code, see Shook et al., 2013; shapes, see Fiser & Aslin, 2002a; Kirkham et al., 2002; or figures, see Arciuli & Simpson, 2012; Kidd & Arciuli, 2016). Despite the numerous studies, a concerted investigation that evaluates if ASL in adults is stimulus-specific is lacking. In children, for example, very recent studies (Raviv & Arnon, 2018; Shufaniya & Arnon, 2018) found that SL changes with age and depends on the stimulus: the SL of linguistic material is age-invariant while the SL of non-linguistic stimuli was found to increase with age. A similar investigation that sheds light on the neural dynamics of SL as a function of stimulus type is warranted.

Ultimately, examining how musicians learn different types of sounds will provide information on the extent of the impact of training on ASL, clarifying whether training promotes ASL in general or if training-related alterations are stimulus-specific. This investigation would provide relevant insights regarding the extent of the effects of musical training on auditory processing by showing if musicians exhibit within-domain (i.e., specific to melodic sounds) or cross-domain (i.e., extensive to non-melodic sounds) neuroplasticity.

2.8.6 Speech segmentation: sung vs. flat languages

The vast majority of research papers from the existent literature on ASL have investigated the involvement of SL, as a learning mechanism, in the segmentation of speech. Due to this theoretical motif, linguistic sounds - from natural to synthetic speech sounds – have been vastly used and studied. Previous research investigating artificial language learning has used both syllables with varying pitches (i.e., with prosodic or melodic characteristics) and syllables with stable pitches. Words stemming from pitch intonated syllables are characterized by specific melodic contours, which is why they form “sung” languages. These languages, as opposed to spoken languages (e.g., with flat contour), facilitate word segmentation since their underlying musical structure and associated motivational characteristics may positively influence the parsing out of speech streams (Schön et al., 2008). Also, the combination of statistical cues with prosodic information may prompt an increased and more orchestrated response of neurons in the auditory cortex (Cunillera et al., 2006). To the best of our knowledge, research addressing the influence of musical training on speech segmentation has always resorted to sung languages. Since musical training is known to alter pitch-processing skills, and sung languages are characterized by pitch modulations, the eventual benefit of musicians in speech segmentation might be attributed to a better capacity to process the melodic structure of the linguistic streams. This methodological option precludes the possibility of accurately testing whether musical training promotes the processing of regularities in speech materials, independently of having or not a melodic structure. Therefore, in order to test if the benefits in speech segmentation observed in musicians are due to enhanced ASL mechanisms, a direct contrast of sung with flat languages is needed.

Moreover, previous studies on this topic created artificially pseudo-words whose phonotactic constraints were the same as the ones of the participant’s mother tongue (François et al., 2014; Francois & Schön, 2011). A recent study showed that learners’ prior knowledge about speech co-occurrences on their mother tongue had a positive impact on their performance on auditory verbal tasks with novel input (Noam Siegelman et al., 2018). In visual SL studies, participants have no prior knowledge on the co-occurrences of units embedded in the visual streams; however, in research using artificial languages participants’ performance in the SL tasks interact with long-standing representations of the statistical structure of their native language (Noam Siegelman et al., 2018). As listeners are not *tabula rasa*, language-learning tasks might not only evaluate participants’ effective SL skills but also their level of linguistic entrenchment. Therefore, an investigation of the role of musical training in speech segmentation, using verbal material with distinct phonotactic constraints from those of the participants’ native language, is essential to avoid the potential confounds raised by linguistic entrenchment.

2.8.7 The mediating role of attention on the effects of training in ASL

The existing studies claim that SL proceeds automatically, even without a conscious effort or explicit instructions that would facilitate the extraction of patterns (Dienes, Broadbent, & Berry, 1991; Dulany, Carlson, & Dewey, 1984; Reber, 1976). SL occurs both when listeners are passively attending to the stimuli (e.g., without having been asked to perform a task, see Fiser & Aslin, 2002b; Saffran et al., 1999; Toro et al., 2005), when they are concurrently addressing a task that can be or not be related to the stimuli to be learned (e.g., Turk-Browne et al., 2005; Turk-browne et al., 2008) and independently of the type of instructions (Arciuli, von Koss Torkildsen, Stevens, & Simpson, 2014; Batterink, Reber, & Paller, 2015). However, the literature is somewhat inconclusive regarding the role of attention in ASL. Most of the studies claiming that SL is implicit tested participants under passive listening conditions that do not rule out the influence of selective attention on SL. Recent evidence stemming from auditory and visual studies indicate that attention needs to be actively focused on the input in order for stream segmentation to be achieved (Daikoku & Yumoto, 2017; Ding et al., 2018; Toro et al., 2005). Nevertheless, most of the existing studies (with the exception of Paraskevopoulos et al., 2012) manipulated attentional load, which only inform us about what happens to performance in behavioral SL tasks when multiple sources of stimulation are competing for attention. These studies do not clarify how SL proceeds under pre-attentive conditions. Only by testing ASL when participants are strictly focused on other, unrelated tasks would it be possible to investigate if the processing of regularities is truly implicit or dependent on attention.

In addition to the lack of studies investigating SL under pre-attentive conditions, there is also no systematic investigation on the role of musical training in ASL when the stimulation is task-irrelevant. Neither are there reports investigating speech segmentation in musicians under pre-attentive conditions, nor studies examining ASL in general (using stimuli other than artificial languages) under those conditions. Besides, there is need to provide more than behavioral data regarding the effects of musical training on ASL when subjects are not attending to the auditory input: first, a broad understanding of the neural dynamics characterizing ASL in the “absence” of attention is needed along with more evidence on the relationship between these neural phenomena and behavioral measures; and ultimately, it is important to clarify if attention mediates the effects of musical training in ASL.

2.9 Aims of the present Dissertation

The literature review reported in this chapter led to some considerations regarding gaps in the ASL and musical training research (described in the previous section) and raised important questions that will be addressed in the experimental studies composing this Dissertation. The following questions were the motto for our experimental work:

1. Does musical training enhance ASL of simple and complex auditory regularities?
2. Are the effects of musical training on ASL dependent on the type of auditory input?
3. What are the neural dynamics underlying ASL in musicians and non-musicians?
4. Does musical training influence how the brain responds to statistical violations in auditory input?
5. Are there effects of musical training on ASL when subjects are not paying attention to sounds?
6. Can better ASL abilities account for musicians' enhanced speech processing skills?

In order to answer these queries, three experimental studies were conducted (chapters[4] 3, 4 and 5). The use of the ERP technique was common to all studies since we aimed to provide evidence on the neural dynamics of SL processes with high temporal precision.

The study described in chapter 3 examined the role of musical training in the SL of simple auditory regularities. Musician and non-musician participants underwent an oddball task composed of pure tones of varying pitches that tested the processing of simple acoustic regularities under pre-attentive task requirements. Thereafter, chapter 4 pertains to the influence of musical training

---

[4] The chapters reporting experimental studies are organized in the format of scientific papers that were submitted for publication or that are in preparation for submission. For that reason, we advise in advance the reader that specific segments of this dissertation might seem redundant.

on the ASL of distinct types of auditory input under attentive conditions. For that reason, chapter 4 includes three experiments aiming to study the neural and behavioral correlates of the learning of linguistic (e.g., prosodic and non-prosodic – flat - words) and musical sounds. Using ERPs, we investigated the temporal dynamics of the learning process *per se*, testing the influence of musical training on the learning process. To provide a more comprehensive account on the neural mechanisms involved in ASL, electrophysiological responses were recorded during the period in which participants were exposed to the auditory streams. To assess the implicit learning of the streams, ERPs were recorded during implicit test phases (often referred to as violation phases). These phases combined the presentation of previously learned items with new legal and illegal items, setting the stage for the evaluation of the neural responses to violations of statistical regularities. The explicit assessment of learning was provided by LDTs in which behavioral data were collected. With this manipulation, we were able to assess the learning outcomes both implicitly (e.g., without participant's awareness) and explicitly. The combination of ERP recordings and behavioral measures during both online and offline learning tasks provided complementary measures on the processes underlying ASL abilities in both groups. Chapter 5 addresses the question of what happens at the neural and behavioral levels when the ASL of distinct types of sounds proceeds pre-attentively and what is the role of musical training in ASL under these experimental conditions. Musicians and non-musicians underwent similar experiments as the ones reported in chapter 4 while watching a movie while being instructed to pay attention to it while ignoring the sounds. Therefore, ERPs were collected during ASL tasks with linguistic and non-linguistic stimuli under pre-attentive task requirements.

The present Dissertation provides important insights linking basic learning mechanisms to aspects of auditory processing and its interaction with experience-driven neuronal changes resulting from musical training. We believe that it contributes to a broader understanding of brain and behavioral consequences of long-term musical training. Simultaneously, our findings may provide new insights about the societal and educational implications of musical training and also help to inform the research on speech-related disorders.

2.10 References

Abel, A. D., Schneider, J., & Maguire, M. J. (2018). N400 Response Indexes Word Learning from Linguistic Context in Children. *Language Learning and Development*, *14*(1), 61–71. http://doi.org/10.1080/15475441.2017.1362347

Abla, D., Katahira, K., & Okanoya, K. (2008). On-line assessment of statistical learning by event-related potentials. *Journal of Cognitive Neuroscience*, *20*(6), 952–964. http://doi.org/10.1162/jocn.2008.20058

Abla, D., & Okanoya, K. (2009). Visual statistical learning of shape sequences: An ERP study. *Neuroscience Research*, *64*(2), 185–190. http://doi.org/10.1016/j.neures.2009.02.013

Abrams, D. A., Bhatara, A., Ryali, S., Balaban, E., Levitin, D. J., & Menon, V. (2011). Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cerebral Cortex*, *21*(7), 1507–1518. http://doi.org/10.1093/cercor/bhq198

Ahissar, M., & Hochstein, S. (1993). Attentional Control of Early Perceptual Learning. *Proceedings of the National Academy of Sciences*, *90*(12), 5718–5722.

Allison, T., Wood, C. C., & McCarthy, G. (1986). The central nervous system. In *Coles, M.G.H., Donchin, E., Porges, S. (Eds.), Psychophysiology: Systems, Processes, and Applications.* (pp. 5–25). New York,: Guilford Press.

Amunts, K., Schlaug, G., Jäncke, L., Steinmetz, H., Schleicher, A., Dabringhaus, A., & Zilles, K. (1997). Motor cortex and hand motor skills: Structural compliance in the human brain. *Human Brain Mapping*, *5*(3), 206–215. http://doi.org/10.1002/(SICI)1097-0193(1997)5:3<206::AID-HBM5>3.0.CO;2-7

Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology*, *83*(2), 111–130. http://doi.org/10.1016/S0022-0965(02)00124-8

Arciuli, J. (2017). The multi-component nature of statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160058. http://doi.org/10.1098/rstb.2016.0058

Arciuli, J., & Simpson, I. C. (2012). Statistical learning is lasting and consistent over time. *Neuroscience Letters*, *517*(2), 133–135. http://doi.org/10.1016/j.neulet.2012.04.045

Arciuli, J., von Koss Torkildsen, J., Stevens, D. J., & Simpson, I. C. (2014). Statistical learning under incidental versus intentional conditions. *Frontiers in Psychology*, *5*(JUL), 1–8. http://doi.org/10.3389/fpsyg.2014.00747

Armstrong, B. C., Frost, R., & Christiansen, M. H. (2017). The long road of statistical learning research: Past, present and future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711). http://doi.org/10.1098/rstb.2016.0047

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8 month old infants. *Psychological Science*, *9*(4), 321–324. http://doi.org/http://doi.org/10.1111/1467-9280.00063

Astheimer, L. B., & Sanders, L. D. (2011). Predictability affects early perceptual processing of word onsets in continuous speech. *Neuropsychologia*, *49*(12), 3512–3516. http://doi.org/10.1016/j.neuropsychologia.2011.08.014

Atienza, M., Cantero, J. L., Grau, C., Gomez, C., Dominguez-Marin, E., & Escera, C. (2003). Effects of temporal encoding on auditory object formation: A mismatch negativity study. *Cognitive Brain Research*, *16*(3), 359–371. http://doi.org/10.1016/S0926-6410(02)00304-X

Baker, C. I., Olson, C. R., & Behrmann, M. (2004). Visual Statistical Learning Role of Attention and Perceptual Grouping in Visual Statistical Learning. *Psychological Science*, *15*(7), 1–8.

Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., … Altenmüller, E. (2006). Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. *NeuroImage*, *30*(3), 917–926. http://doi.org/10.1016/j.neuroimage.2005.10.044

Batterink, L. J. (2017). Rapid Statistical Learning Supporting Word Extraction From Continuous Speech. *Psychological Science*, *28*(7), 921–928. http://doi.org/10.1177/0956797617698226

Batterink, L. J., & Paller, K. A. (2017). Online neural monitoring of statistical learning. *Cortex*, *90*, 31–45. http://doi.org/10.1016/j.cortex.2017.02.004

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language*, *83*, 62–78.

http://doi.org/10.1016/j.jml.2015.04.004

Batterink, L. J., Reber, P. J., & Paller, K. A. (2015). Functional differences between statistical learning with and without explicit training. *Learning and Memory*, *22*(11), 544–556. http://doi.org/10.1101/lm.037986.114

Batterink, L., & Neville, H. (2011). Implicit and explicit mechanism of word learning in a narrative context: An event-related potential study. *Journal of Cognitive Neuroscience*, *23*(11), 3181–3196. http://doi.org/10.1162/jocn_a_00013

Batterink, L., & Neville, H. J. (2013). The Human Brain Processes Syntax in the Absence of Conscious Awareness. *Journal of Neuroscience*, *33*(19), 8528–8533. http://doi.org/10.1523/JNEUROSCI.0618-13.2013

Bengtsson, S. L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., & Ullén, F. (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nature Neuroscience*, *8*(9), 1148–1150. http://doi.org/10.1038/nn1516

Beres, A. M. (2017). Time is of the Essence: A Review of Electroencephalography (EEG) and Event-Related Brain Potentials (ERPs) in Language Research. *Applied Psychophysiology Biofeedback*, *42*(4), 247–255. http://doi.org/10.1007/s10484-017-9371-3

Bergstrom, J. C., Howard, J. H., & Howard, D. V. (2012). Enhanced Implicit Sequence Learning in College-age Video Game Players and Musicians. *Applied Cognitive Psychology*, *26*(1), 91–96. http://doi.org/10.1002/acp.1800

Bermudez, P., & Zatorre, R. J. (2005). Differences in gray matter between musicians and nonmusicians. *Annals of the New York Academy of Sciences*, *1060*(2005), 395–399. http://doi.org/10.1196/annals.1360.057

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology*, *2*(MAY), 1–12. http://doi.org/10.3389/fpsyg.2011.00094

Bhide, A. (2013). A Rhythmic Musical Intervention for Poor Readers: A Comparison of Efficacy with a Letter-Based Intervention, *7*(2), 113–123. Retrieved from https://www.cne.psychol.cam.ac.uk/pdfs/phds/ABMphil.pdf

Bidelman, G. M., & Alain, C. (2015). Musical Training Orchestrates Coordinated Neuroplasticity in Auditory Brainstem and Cortex to Counteract Age-Related Declines in Categorical Vowel

Perception. *Journal of Neuroscience*, *35*(3), 1240–1249. http://doi.org/10.1523/JNEUROSCI.3292-14.2015

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011a). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience*, *23*(2), 425–434. http://doi.org/10.1162/jocn.2009.21362

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011b). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain and Cognition*, *77*(1), 1–10. http://doi.org/10.1016/j.bandc.2011.07.006

Bidelman, G. M., & Krishnan, A. (2010). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Research*, *1355*, 112–125. http://doi.org/10.1016/j.brainres.2010.07.100

Boh, B., Herholz, S. C., Lappe, C., & Pantev, C. (2011). Processing of complex auditory patterns in musicians and nonmusicians. *PLoS ONE*, *6*(7). http://doi.org/10.1371/journal.pone.0021458

Boll-Avetisyan, N., Bhatara, A., & Höhle, B. (2017). Effects of Musicality on the Perception of Rhythmic Structure in Speech. *Laboratory Phonology Journal of the Association for Laboratory Phonology*, *8*(9), 1–16. http://doi.org/10.5334/labphon.91

Borovsky, A., Kutas, M., & Elman, J. (2010). Learning to use words: Event-related potentials index single-shot contextual word learning. *Cognition*, *116*(2), 289–296. http://doi.org/10.1016/j.cognition.2010.05.004

Bors, D. A., & MacLeod, C. M. (1996). Individual differences in memory. In E. L. Bjork & R. A. Bjork (Eds.), *Memory: Handbook of Perception and Cognition* (pp. 411–441). San Diego: Academic Press, Inc.

Bosnyak, D. J., Eaton, R. A., & Roberts, L. E. (2004). Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. *Cerebral Cortex*, *14*(10), 1088–1099. http://doi.org/10.1093/cercor/bhh068

Boutros, N. N., & Belger, A. (1999). Midlatency evoked potentials attenuation and augmentation reflect different aspects of sensory gating. *Biological Psychiatry*, *45*(7), 917–922. http://doi.org/10.1016/S0006-3223(98)00253-4

Boutros, N. N., Torello, M. W., Barker, B. A., Tueting, P. A., Wu, S. C., & Nasrallah, H. A. (1995). The P50 evoked potential component and mismatch detection in normal volunteers: implications for the study of sensory gating. *Psychiatry Research*, *57*(1), 83–88. http://doi.org/10.1016/0165-1781(95)02637-C

Brandt, A., Gebrian, M., & Slevc, L. R. (2012). Music and early language acquisition. *Frontiers in Psychology*, *3*(SEP), 1–17. http://doi.org/10.3389/fpsyg.2012.00327

Brooks, P. J., & Kempe, V. (2013). Individual differences in adult foreign language learning: The mediating effect of metalinguistic awareness. *Memory & Cognition*, *41*(2), 281–296. http://doi.org/10.3758/s13421-012-0262-9

Brown, C. J., Jeon, E.-K., Driscoll, V., Mussoi, B., Deshpande, S. B., Gfeller, K., & Abbas, P. J. (2017). Effects of Long-Term Musical Training on Cortical Auditory Evoked Potentials. *Ear and Hearing*, *38*(2), e74–e84. http://doi.org/10.1097/AUD.0000000000000375

Bruner, J. S. (1974). From communication to language—A psychological perspective. *Cognition*, *3*(3), 255–287. http://doi.org/10.1016/0010-0277(74)90012-2

Buiatti, M., Peña, M., & Dehaene-Lambertz, G. (2009). Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses. *NeuroImage*, *44*(2), 509–519. http://doi.org/10.1016/j.neuroimage.2008.09.015

Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition*, *121*(1), 127–132. http://doi.org/10.1016/j.cognition.2011.06.010

Buzsáki, G., Anastassiou, C. A., & Koch, C. (2012). The origin of extracellular fields and currents-EEG, ECoG, LFP and spikes. *Nature Reviews Neuroscience*, *13*(6), 407–420. http://doi.org/10.1038/nrn3241

Chambers, K. E., Onishi, K. H., & Fisher, C. (2003). Infants learn phonotactic regularities from brief auditory experience. *Cognition*, *87*(2), B69–B77. http://doi.org/10.1016/S0

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Moving on time: Brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience*, *20*(2), 226–239. http://doi.org/10.1162/jocn.2008.20018

Chobert, J., François, C., Velay, J. L., & Besson, M. (2012). Twelve months of active musical training in 8-to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral Cortex*, *24*(4), 956–967.

http://doi.org/10.1093/cercor/bhs377

Choi, U.-S., Sung, Y.-W., Hong, S., Chung, J.-Y., & Ogawa, S. (2015). Structural and functional plasticity specific to musical training with wind instruments. *Frontiers in Human Neuroscience*, *9*(October), 1–8. http://doi.org/10.3389/fnhum.2015.00597

Chomsky, N. (1957). *Syntactic structures.* Berlin: Walter de Gruyter, GmbH & Co. KG.

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* Cambrige, Massachussetts: MIT Press.

Christiansen, M. H., Dale, R. a C., Ellefson, M. R., & Conway, C. M. (2002). The role of sequential learning in language evolution: Computational and experimental studies. In *Simulating the Evolution of Language* (pp. 165–187). London: Springer. http://doi.org/10.1007/978-1-4471-0663-0_8

Cleeremans, A. X. (2006). Conscious and unconscious cognition: A graded, dynamic perspective. In *Q. Jing, M. Rosenzweig, G. d'Ydewalle, H. Zhang, H.-C. Chen, & K. Zhang (Eds.), Progress in Psychological Science around the World. Vol 1 Neural, Cognitive and Developmental Issues* (pp. 401–418). Hove, England: Psychology Press. http://doi.org/10.4324/9780203783122

Cohen, C. (2008). Music: A universal language? In I. B. Tauris (Ed.), *Music and conflict transformation. Harmonies and dissonances in geopolitics* (Olivier Ur, pp. 26–39). London and New York.

Cohen, M. X. (2017). Where Does EEG Come From and What Does It Mean? *Trends in Neurosciences*, *40*(4), 208–218. http://doi.org/10.1016/j.tins.2017.02.004

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning Memory and Cognition*, *31*(1), 24–39. http://doi.org/10.1037/0278-7393.31.1.24

Courchesne, E., Hillyard, S. A., & Courchesne, R. Y. (1977). P3 Waves to the Discrimination of Targets in Homogeneous and Heterogeneous Stimulus Sequences. *Psychophysiology*. http://doi.org/10.1111/j.1469-8986.1977.tb01206.x

Creel, S. C., Newport, E. L., & Aslin, R. N. (2004). Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning Memory and Cognition*, *30*(5), 1119–1130. http://doi.org/10.1037/0278-7393.30.5.1119

Cromer, R. F. (1981). Developmental language disorders: Cognitive processes, semantics, pragmatics, phonology, and syntax. *Journal of Autism and Developmental Disorders*, *11*(1), 57–74. http://doi.org/10.1016/S1364-6613(02)01947-2

Cunillera, T., Càmara, E., Toro, J. M., Marco-Pallares, J., Sebastián-Galles, N., Ortiz, H., ... Rodríguez-Fornells, A. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *NeuroImage*, *48*(3), 541–553. http://doi.org/10.1016/j.neuroimage.2009.06.069

Cunillera, T., Laine, M., & Rodríguez-Fornells, A. (2016). Headstart for speech segmentation: A neural signature for the anchor word effect. *Neuropsychologia*, *82*, 189–199. http://doi.org/10.1016/j.neuropsychologia.2016.01.011

Cunillera, T., Toro, J. M., Sebastián-Gallés, N., & Rodríguez-Fornells, A. (2006). The effects of stress and statistical cues on continuous speech segmentation: An event-related brain potential study. *Brain Research*, *1123*(1), 168–178. http://doi.org/10.1016/j.brainres.2006.09.046

Cutler, A., Dahan, D., & Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, *40*(2), 141–201. http://doi.org/10.1016/j.jadohealth.2005.06.003

Daikoku, T., Yatomi, Y., & Yumoto, M. (2016). Pitch-class distribution modulates the statistical learning of atonal chord sequences. *Brain and Cognition*, *108*, 1–10. http://doi.org/10.1016/j.bandc.2016.06.008

Daikoku, T., & Yumoto, M. (2017). Single, but not dual, attention facilitates statistical learning of two concurrent auditory sequences. *Scientific Reports*, *7*(1), 1–10. http://doi.org/10.1038/s41598-017-10476-x

Daltrozzo, J., & Conway, C. M. (2014). Neurocognitive mechanisms of statistical-sequential learning: what do event-related potentials tell us? *Frontiers in Human Neuroscience*, *8*(June), 1–22. http://doi.org/10.3389/fnhum.2014.00437

Darwin, C. R. (1871). *The Descent of Man, and Selection in Relation to Sex.* (John Murray, Ed.). London.

de Diego-Balaguer, R., Rodríguez-Fornells, A., & Bachoud-Lévi, A.-C. (2015). Prosodic cues enhance rule learning by changing speech segmentation mechanisms. *Frontiers in Psychology*, *6*(September). http://doi.org/10.3389/fpsyg.2015.01478

de Diego-Balaguer, R., Toro, J. M., Rodriguez-Fornells, A., & Bachoud-Lévi, A. C. (2007). Different neurophysiological mechanisms underlying word and rule extraction from speech. *PLoS ONE*, *2*(11). http://doi.org/10.1371/journal.pone.0001175

Debener, S., Kranczioch, C., Herrmann, C. S., & Engel, A. K. (2002). Auditory novelty oddball allows reliable distinction of top-down and bottom-up processes of attention. *International Journal of Psychophysiology*, *46*(1), 77–84. http://doi.org/10.1016/S0167-8760(02)00072-7

Deutsch, D. (1999). The processing of pitch combinations. In *The Psychology of Music* ((Second Ed, pp. 349–411).

Dienes, Z., Broadbent, D., & Berry, D. C. (1991). Implicit and explicit knowledge bases in artificial grammar. *Journal of Experimental Psychology, Learning, Memory and Cognition*, *17*, 875–887. http://doi.org/10.1037/0278-7393.17.5.875

Ding, N., Pan, X., Luo, C., Su, N., Zhang, W., & Zhang, J. (2018). Attention is required for knowledge-based sequential grouping: Insights from the integration of syllables into words. *The Journal of Neuroscience*, *38*(5), 1178–1188. http://doi.org/10.1523/JNEUROSCI.2606-17.2017

Dulany, D. E., Carlson, R. A., & Dewey, G. I. (1984). A case of syntactical learning and judgment: how conscious and how abstract? *Journal of Experimental Psychology: General*, *113*, 541–555. http://doi.org/10.1037/0096-3445.113.4.541

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., ... Van Petten, C. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, *120*(11), 1883–1908. http://doi.org/10.1016/j.clinph.2009.07.045

Eimer, M., Goschke, T., Schlaghecken, F., & Stürmer, B. (1996). Explicit and implicit learning of event sequences: Evidence from event-related brain potentials. *Journal of Experimental Psychology: Learning Memory and Cognition*, *22*(4), 970–987. http://doi.org/10.1037/0278-7393.22.4.970

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., & Taub, E. (1995). Increased Cortical Representation of the Fingers of the Left Hand in String Players. *Science*, *270*(5234), 305–307. http://doi.org/10.1126/science.270.5234.305

Endress, A. D., & Mehler, J. (2009). Primitive computations in speech processing. *Quarterly*

Journal of Experimental Psychology, *62*(11), 2187–2209. http://doi.org/10.1080/17470210902783646

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406. http://doi.org/10.1037/0033-295X.100.3.363

Escera, C., & Corral, M. J. (2007). Role of mismatch negativity and novelty-P3 in involuntary auditory attention. *Journal of Psychophysiology*, *21*(3–4), 251–264. http://doi.org/10.1027/0269-8803.21.3.xxx

Everaert, M. B. H., Huybregts, M. A. C., Berwick, R. C., Chomsky, N., Tattersall, I., Moro, A., & Bolhuis, J. J. (2017). What is Language and How Could it Have Evolved? *Trends in Cognitive Sciences*, *21*(8), 569–571. http://doi.org/10.1016/j.tics.2017.05.007

Federmeier, K. D., & Kutas, M. (1999). A Rose by Any Other N ame : Long-Term Memory Structure and Sentence Processing. *Journal of Memory and Language*, *41*, 469–495.

Federmeier, K. D., McLennan, D. B., Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, *39*(2), 133–146. http://doi.org/10.1111/1469-8986.3920133

Fiser, J., & Aslin, R. N. (2002a). Statistical Learning of Higher-Order Temporal Structure from Visual Shape Sequences. *Journal of Experimental Psychology: Learning Memory and Cognition*, *28*(3), 458–467. http://doi.org/10.1037/0278-7393.28.3.458

Fiser, J., & Aslin, R. N. (2002b). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(24), 15822–6. http://doi.org/10.1073/pnas.232472899

Fiser, J., & Aslin, R. N. (2005). Encoding multielement scenes: Statistical learning of visual feature hierarchies. *Journal of Experimental Psychology: General*, *134*(4), 521–537. http://doi.org/10.1037/0096-3445.134.4.521

Franco, A., & Destrebecqz, A. (2012). Chunking or not chunking? How do we find words in artificial language learning? *Advances in Cognitive Psychology*, *8*(2), 144–154. http://doi.org/10.2478/v10053-008-0111-3

François, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex*, *23*(9), 2038–2043. http://doi.org/10.1093/cercor/bhs180

François, C., Cunillera, T., Garcia, E., Laine, M., & Rodriguez-Fornells, A. (2017). Neurophysiological evidence for the interplay of speech segmentation and word-referent mapping during novel word learning. *Neuropsychologia*, *98*(October), 56–67. http://doi.org/10.1016/j.neuropsychologia.2016.10.006

François, C., Jaillet, F., Takerkart, S., & Schön, D. (2014). Faster sound stream segmentation in musicians than in nonmusicians. *PLoS ONE*, *9*(7), 26–29. http://doi.org/10.1371/journal.pone.0101340

Francois, C., & Schön, D. (2010). Learning of musical and linguistic structures: Comparing event-related potentials and behavior. *NeuroReport*, *21*(14), 928–932. http://doi.org/10.1097/WNR.0b013e32833ddd5e

Francois, C., & Schön, D. (2011). Musical expertise boosts implicit learning of both musical and linguistic structures. *Cerebral Cortex*, *21*(10), 2357–2365. http://doi.org/10.1093/cercor/bhr022

François, C., & Schön, D. (2014). Neural sensitivity to statistical regularities as a fundamental biological process that underlies auditory learning: The role of musical practice. *Hearing Research*, *308*, 122–128. http://doi.org/10.1016/j.heares.2013.08.018

François, C., Teixidó, M., Takerkart, S., Agut, T., Bosch, L., & Rodriguez-Fornells, A. (2017). Enhanced Neonatal Brain Responses to Sung Streams Predict Vocabulary Outcomes by Age 18 Months. *Scientific Reports*, *7*(1), 1–13. http://doi.org/10.1038/s41598-017-12798-2

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78–84. http://doi.org/10.1016/s1364-6613(00)01839-8

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, *19*(3), 117–125. http://doi.org/10.1016/j.tics.2014.12.010

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience*, *16*(6), 1010–1021. http://doi.org/10.1162/0898929041502706

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2005). Automatic encoding of melodies in musicians and non musicians. *Journal of Cognitive Neuroscience*, *Vol. 17*(10), 1578–1592. http://doi.org/10.1162/089892905774597263

Ganis, G., Kutas, M., & Sereno, M. I. (1984). The Search for " Common Sense ": A n Electrophysiological Study of the Comprehension of ' words and Pictures in Reading, 89–106.

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *The Journal of Neuroscience*, *23*(27), 9240–9245. http://doi.org/23/27/9240 [pii]

Geiser, E., Sandmann, P., Jäncke, L., & Meyer, M. (2010). Refinement of metre perception - training increases hierarchical metre processing. *European Journal of Neuroscience*, *32*(11), 1979–1985. http://doi.org/10.1111/j.1460-9568.2010.07462.x

Gilbert, C., Ito, M., Kapadia, M., & Westheimer, G. (2000). Interactions between attention, context and learning in primary visual cortex. *Vision Research*, *40*(10–12), 1217–1226. http://doi.org/10.1016/S0042-6989(99)00234-5

Gómez, R., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, *4*(5), 178–186. http://doi.org/10.1016/S1364-6613(00)01467-4

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431–436. http://doi.org/10.7897/2277-4343.04136

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences*, *15*(1), 3–10. http://doi.org/10.1016/j.tics.2010.10.001

Gottselig, J. M., Brandeis, D., Hofer-Tinguely, G., Borbély, A. A., & Achermann, P. (2004). Human Central Auditory Plasticity Associated with Tone Sequence Learning. *Learning and Memory*, *11*(2), 162–171. http://doi.org/10.1101/lm.63304

Greenberg, S. (1996). Understanding speech understanding: Towards a unified theory of speech perception. *ESCA Tutorial and Advanced Research Workshop on the Auditory Basis of Speech Perception*, 1–8. Retrieved from http://www1.icsi.berkeley.edu/ftp/global/pub/speech/papers/keele96-usu.pdf

Haarmann, H. J., & Kolk, H. H. J. (1991). Syntactic priming in broca's aphasics: Evidence for slow activation. *Aphasiology*, *5*(3), 247–263. http://doi.org/10.1080/02687039108248527

Halwani, G. F., Loui, P., Rüber, T., & Schlaug, G. (2011). Effects of practice and experience on the arcuate fasciculus: Comparing singers, instrumentalists, and non-musicians. *Frontiers in Psychology*, *2*(156), 1–9. http://doi.org/10.3389/fpsyg.2011.00156

Han, Y., Yang, H., Lv, Y. T., Zhu, C. Z., He, Y., Tang, H. H., ... Dong, Q. (2009). Gray matter density and white matter integrity in pianists' brain: A combined structural and diffusion tensor MRI study. *Neuroscience Letters*, *459*(1), 3–6. http://doi.org/10.1016/j.neulet.2008.07.056

Hasson, U. (2017). The neurobiology of uncertainty : implications for statistical learning Neurobiological approaches to uncertainty and statistical learning : Interfaces and. *Phil. Trans. R. Soc. B.*, *372*(20160048), 1–12. http://doi.org/10.1098/not

Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition*, *78*(3), 53–64. http://doi.org/10.1016/S0010-0277(00)00132-3

Hay, J. F., Pelucchi, B., Estes, K. G., & Saffran, J. R. (2011). Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology*, *63*(2), 93–106. http://doi.org/10.1016/j.cogpsych.2011.06.002

Haynes, J. D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, *7*(7), 523–534. http://doi.org/10.1038/nrn1931

Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., & Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology*, *42*(2), 180–190. http://doi.org/10.1111/j.1469-8986.2005.00272.x

Herdener, M., Esposito, F., di Salle, F., Boller, C., Hilti, C. C., Habermeyer, B., ... Cattapan-Ludewig, K. (2010). Musical Training Induces Functional Plasticity in Human Hippocampus. *Journal of Neuroscience*, *30*(4), 1377–1384. http://doi.org/10.1523/JNEUROSCI.4513-09.2010

Herholz, S. C., Lappe, C., Knief, A., & Pantev, C. (2009). Imagery mismatch negativity in musicians. *Annals of the New York Academy of Sciences*, *1169*, 173–177. http://doi.org/10.1111/j.1749-6632.2009.04782.x

Herholz, S. C., Lappe, C., & Pantev, C. (2009). Looking for a pattern: An MEG study on the abstract mismatch negativity in musicians and nonmusicians. *BMC Neuroscience*, *10*, 1–10. http://doi.org/10.1186/1471-2202-10-42

Herholz, S. C., & Zatorre, R. J. (2012). Musical Training as a Framework for Brain Plasticity: Behavior, Function, and Structure. *Neuron*, *76*(3), 486–502. http://doi.org/10.1016/j.neuron.2012.10.011

Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience*, *15*(5), 673–682. http://doi.org/10.1162/089892903322307393

Higgins, K. M. (2012). *The music between us: Is music a universal language?.* Press., University of Chicago.

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*(4108), 177–180.

Hillyard, S. A., Squires, K. C., Bauer, J. W., & Lindsay, P. H. (1971). Evoked potential correlates of auditory signal detection. *Science*, *172*(3990), 1357–1360.

Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). The effects of musical training on structural brain development: A longitudinal study. *Annals of the New York Academy of Sciences*, *1169*, 182–186. http://doi.org/10.1111/j.1749-6632.2009.04852.x

Jackendoff, R. (2002). English particle constructions, the lexicon, and the autonomy of syntax. *Verb-Particle Explorations*, 67–94.

Jäncke, L. (2009). Music drives brain plasticity Major recent advances Structural brain plasticity, (October). http://doi.org/10.3410/B1-78

Joseph, S., Iverson, P., Manohar, S., Fox, Z., Scott, S. K., & Husain, M. (2015). Precision of working memory for speech sounds. *Quarterly Journal of Experimental Psychology*, *68*(10), 2022–2040. http://doi.org/10.1080/17470218.2014.1002799

Jung, H., Sontag, S., Park, Y. S., & Loui, P. (2015). Rhythmic effects of syntax processing in music and language. *Frontiers in Psychology*, *6*(NOV), 1–11. http://doi.org/10.3389/fpsyg.2015.01762

Kaan, E., & Swaab, T. Y. (2002). The brain circuitry of syntactic comprehension. *Trends in Cognitive Sciences*, *6*(8), 350–356. http://doi.org/10.1016/S1364-6613(02)01947-2

Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain and*

Language, *127*(1), 46–54. http://doi.org/10.1016/j.bandl.2012.11.007

Kidd, E., & Arciuli, J. (2016). Individual Differences in Statistical Learning Predict Children's Comprehension of Syntax. *Child Development*, *87*(1), 184–193. http://doi.org/10.1111/cdev.12461

King, J. W., & Kutas, M. (1995). Who did what and when – using word level and clause-level ERPs to monitor working-memory usage in reading. *Journal of Cognitive Neuroscience*, *7*(1986), 376–395.

Kirkham, N. Z., Slemmer, J. a, & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain-general learning mechanism. *Cognition*, *83*, 35–42. http://doi.org/10.1016/S0010-0277(02)00004-5

Koelsch, S., Busch, T., Jentschke, S., & Rohrmeier, M. (2016). Under the hood of statistical learning: A statistical MMN reflects the magnitude of transitional probabilities in auditory sequences. *Scientific Reports*, *6*, 1–11. http://doi.org/10.1038/srep19741

Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: An fMRI study. *NeuroImage*, *25*(4), 1068–1076. http://doi.org/10.1016/j.neuroimage.2004.12.050

Koelsch, S., Gunter, T. C., Wittfoth, M., & Sammler, D. (2005). Interaction between syntax processing in language and in music: An ERP study. *Journal of Cognitive Neuroscience*, *17*(10), 1565–1577. http://doi.org/10.1162/089892905774597290

Koelsch, S., Schröger, C. A. E., & Tervaniemi, M. (1999a). Superior pre-attentive auditory processing in musicians, *10*(6), 1309–1313.

Koelsch, S., Schröger, E., & Tervaniemi, M. (1999b). Superior pre-attentive auditory processing in musicians. *NeuroReport*, *10*(6), 1309–1313. http://doi.org/10.1097/00001756-199904260-00029

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Science and Society*, *11*(August), 599–606. http://doi.org/10.1038/nrn2882

Krumhansl, C. L. (1990). Tonal Hierarchies and Rare Intervals in Music Cognition. *Music Perception: An Interdisciplinary Journal*, *7*(3), 309–324. Retrieved from http://www.jstor.org/stable/40285467

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews*

Neuroscience, *5*(11), 831–843. http://doi.org/10.1038/nrn1533

Kutas, M., & Federmeier, K. D. (2009). Event-related brain potential (ERP) studies of sentence processing. In *The Oxford Handbook of Psycholinguistics* (pp. 1–40). Oxford University Press. http://doi.org/10.1093/oxfordhb/9780198568971.013.0023

Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology*. http://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*. http://doi.org/10.1126/science.7350657

Kutas, M., & Van Petten, C. (1994). ERP Psycholinguistics electrified: Event-related brain potential investigations. *Handbook of Psycholinguistics*, 83–143. http://doi.org/10.1016/B978-012369374-7/50018-3

Language. Def. 1. (n.d.). In Oxford English Dictionary Online. Retrieved from https://en.oxforddictionaries.com/definition/language. (n.d.). Retrieved from https://en.oxforddictionaries.com/definition/language

Lappe, C., Herholz, S. C., Trainor, L. J., & Pantev, C. (2008). Cortical Plasticity Induced by Short-Term Unimodal and Multimodal Musical Training. *Journal of Neuroscience*, *28*(39), 9632–9639. http://doi.org/10.1523/JNEUROSCI.2254-08.2008

Lappe, C., Trainor, L. J., Herholz, S. C., & Pantev, C. (2011). Cortical plasticity induced by short-term multimodal musical rhythm training. *PLoS ONE*, *6*(6). http://doi.org/10.1371/journal.pone.0021493

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Nature Reviews Neuroscience*, *9*(12), 920–933. http://doi.org/10.1038/nrn2532

Lerdahl, F., & Jackendoff, R. S. (1983). *A Generative Theory of Tonal Music.* MIT Press. http://doi.org/10.1525/mts.1985.7.1.02a00120

Liberman, A. M. (1970). The Grammars of Speech and Language. *Cognitive Psychology*, *323*, 301–323.

Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion*, *11*(5), 1021–1031. http://doi.org/10.1037/a0024521

López-Barroso, D., Catani, M., Ripolles, P., Dell'Acqua, F., Rodriguez-Fornells, A., & de Diego-Balaguer, R. (2013). Word learning is mediated by the left arcuate fasciculus. *Proceedings of the National Academy of Sciences*, *110*(32), 13168–13173. http://doi.org/10.1073/pnas.1301696110

López-Barroso, D., Cucurell, D., Rodríguez-Fornells, A., & de Diego-Balaguer, R. (2016). Attentional effects on rule extraction and consolidation from speech. *Cognition*, *152*, 61–69. http://doi.org/10.1016/j.cognition.2016.03.016

Loui, P., & Wessel, D. (2007). Harmonic expectation and affect in Western music: Effects of attention and training. *Perception and Psychophysics*, *69*(7), 1084–1092. http://doi.org/10.3758/BF03193946

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique* (Second, Vol. 78). Cambridge, MA: MIT Press. http://doi.org/10.1118/1.4736938

Margulis, E. H., Mlsna, L. M., Uppunda, A. K., Parrish, T. B., & Wong, P. C. M. (2009). Selective neyrophysiologic responses to mysic jn instruinentalists with different listening biographies. *Human Brain Mapping*, *30*(1), 267–275. http://doi.org/10.1002/hbm.20503

Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O., & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *Journal of Cognitive Neuroscience*, *23*(10), 2701–2715. http://doi.org/doi: 10.1162/jocn.2010.21585

Marie, C., Magne, C., & Besson, M. (2011). Musicians and the metric structure of words. *Journal of Cognitive Neuroscience*, *23*(2), 294–305. http://doi.org/10.1162/jocn.2010.21413

Marques, C., Moreno, S., Castro, S. L., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience*, *19*(9), 1453–1463. http://doi.org/10.1162/jocn.2007.19.9.1453

Mattys, S. L. (2004). Stress Versus Coarticulation: Toward an Integrated Approach to Explicit Speech Segmentation. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(2), 397–408. http://doi.org/10.1037/0096-1523.30.2.397

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience*, *7*(7), 703–704. http://doi.org/10.1038/nn1264

McNealy, K., Mazziotta, J. C., & Dapretto, M. (2006). Cracking the Language Code: Neural Mechanisms Underlying Speech Parsing. *Journal of Neuroscience*, *26*(29), 7629–7639. http://doi.org/10.1523/JNEUROSCI.5501-05.2006

McNealy, K., Mazziotta, J. C., & Dapretto, M. (2010). The neural basis of speech parsing in children and adults. *Developmental Science*, *13*(2), 385–406. http://doi.org/10.1111/j.1467-7687.2009.00895.x

Mestres-Missé, A., Rodriguez-Fornells, A., & Münte, T. F. (2007). Watching the brain during meaning acquisition. *Cerebral Cortex*, *17*(8), 1858–1866. http://doi.org/10.1093/cercor/bhl094

Mirman, D., Magnuson, J. S., Estes, K. G., & Dixon, J. A. (2008). The link between statistical segmentation and word learning in adults. *Cognition*, *108*(1), 271–280. http://doi.org/10.1016/j.cognition.2008.02.003

Mittag, M., Takegata, R., & Winkler, I. (2016). Transitional Probabilities Are Prioritized over Stimulus/Pattern Probabilities in Auditory Deviance Detection: Memory Basis for Predictive Sound Processing. *Journal of Neuroscience*, *36*(37), 9572–9579. http://doi.org/10.1523/JNEUROSCI.1041-16.2016

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. http://doi.org/10.1093/cercor/bhn120

Mueller, J. L., Bahlmann, J., & Friederici, A. D. (2008). The role of pause cues in language learning: The emergence of event-related potentials related to sequence processing. *Journal of Cognitive Neuroscience*, *20*(5), 892–905. http://doi.org/10.1162/jocn.2008.20511

Münte, T. F., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience*, *3*(6), 473–478. http://doi.org/10.1038/nrn843

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. http://doi.org/10.1073/pnas.0701498104

Näätänen, R. (1992). *Attention and Brain Function.* Lawrence Erlbaum Associates, Inc.

Näätänen, R. (2002). The perception of speech sounds by the human brain as reflected by the mismatch negativity brain response. *International Congress Series*, *1232*, 97–105. http://doi.org/10.1016/S0531-5131(01)00795-6

Näätänen, R., Paavilainen, P., & Reinikainen, K. (1989). Do event-related potentials to infrequent decrements in duration of auditory stimuli demonstrate a memory trace in man? *Neuroscience Letters*, *107*(1–3), 347–352. http://doi.org/10.1016/0304-3940(89)90844-6

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*, *118*(12), 2544–2590. http://doi.org/10.1016/j.clinph.2007.04.026

Näätänen, R., Simpson, M., & Loveless, N. E. (1982). Stimulus deviance and evoked potentials. *Biological Psychology*, *14*(1–2), 53–98. http://doi.org/10.1016/0301-0511(82)90017-5

Neville, H. J., Mills, D. L., & Lawson, D. S. (1992). Fractionating language: Different neural subsytems with different sensitive periods. *Cerebral Cortex*, *2*(June), 244–258.

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*(2), 127–162. http://doi.org/10.1016/S0010-0285(03)00128-2

Newport, E. L., Hauser, M. D., Spaepen, G., & Aslin, R. N. (2004). Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology*, *49*(2), 85–117. http://doi.org/10.1016/j.cogpsych.2003.12.002

Nikjeh, D. A., Lister, J. J., & Frisch, S. A. (2008). Hearing of note: An electrophysiologic and psychoacoustic comparison of pitch discrimination between vocal and instrumental musicians. *Psychophysiology*, *45*(6), 994–1007. http://doi.org/10.1111/j.1469-8986.2008.00689.x

Nikjeh, D. A., Lister, J. J., & Frisch, S. A. (2009). Preattentive cortical-evoked responses to pure tones, harmonic tones, and speech: Influence of music training. *Ear and Hearing*, *30*(4), 432–446. http://doi.org/10.1097/AUD.0b013e3181a61bf2

Nyman, G., Alho, K., Laurinen, P., Paavilainen, P., Radil, T., Reinikainen, K., … Näätänen. (1990). Mismatch negativity (MMN) for sequences of auditory and visual stimuli: evidence for a mechanism specific to the auditory modality. *Electroencephalography and Clinical Neurophysiology*, *77*(6), 436–444. http://doi.org/10.1016/0168-5597(90)90004-w

Oechslin, M. S., Meyer, M., & Jäncke, L. (2010). Absolute pitch-functional evidence of speech-relevant auditory acuity. *Cerebral Cortex*, *20*(2), 447–455. http://doi.org/10.1093/cercor/bhp113

Okhrei, A., Kutsenko, T., & Makarchuk, M. (2017). Performance of working memory of musicians and non-musicians in tests with letters, digits, and geometrical shapes. *Biologija*, *62*(4), 207–215. http://doi.org/10.6001/biologija.v62i4.3408

Olivares, E. I., Iglesias, J., & Antonieta Bobes, M. (1999). Searching for face-specific long latency ERPs: A topographic study of effects associated with mismatching features. *Cognitive Brain Research*, *7*(3), 343–356. http://doi.org/10.1016/S0926-6410(98)00038-X

Onnis, L., Waterfall, H. R., & Edelman, S. (2008). Learn locally, act globally: Learning language from variation set cues. *Cognition*, *109*(3), 423–430. http://doi.org/10.1016/j.cognition.2008.10.004

Osterhout, L., & Mobley, L. A. (1995). Event-Related Brain Potentials Elicited by Failure to Agree. *Journal of Memory and Language*. http://doi.org/10.1006/jmla.1995.1033

Ostroff, J. M., Martin, B. A., & Boothroyd, A. (1998). Cortical evoked response to acoustic change within a syllable. *Ear and Hearing*, *19*(4), 290–297.

Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General*, *130*(3), 401–426. http://doi.org/10.1037/0096-3445.130.3.401

Pallesen, K. J., Brattico, E., Bailey, C. J., Korvenoja, A., Koivisto, J., Gjedde, A., & Carlson, S. (2010). Cognitive control in auditory working memory is enhanced in musicians. *PLoS ONE*, *5*(6). http://doi.org/10.1371/journal.pone.0011120

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, *392*(6678), 811–814. http://doi.org/10.1038/33918

Paraskevopoulos, E., Kuchenbuch, A., Herholz, S. C., & Pantev, C. (2012). Statistical learning effects in musicians and non-musicians: An MEG study. *Neuropsychologia*, *50*(2), 341–349. http://doi.org/10.1016/j.neuropsychologia.2011.12.007

Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical Experience Limits the Degradative Effects of Background Noise on the Neural Processing of Sound. *Journal of Neuroscience*, *29*(45), 14100–14107. http://doi.org/10.1523/JNEUROSCI.3256-09.2009

Patel, A. D. (1998). Syntactic Processing in Language and Music: Different Cognitive Operations, Similar Neural Resources? *Music Perception: An Interdisciplinary Journal*, *16*(1), 27–42. http://doi.org/10.2307/40285775

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, *6*(7), 674–681. http://doi.org/10.1038/nn1082

Patel, A. D. (2008). *Music, language, and the brain.* Oxford university press.

Patel, A. D. (2008). Science & music: Talk of the tone. *Nature*, *453*(7196), 726–727. http://doi.org/10.1038/453726a

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology*, *2*(JUN), 1–14. http://doi.org/10.3389/fpsyg.2011.00142

Patel, A. D. (2012). Language, music, and the brain: a resourche-sharing framework. *Language and Music as Cognitive Systems*, 204–223. http://doi.org/10.1111/j.1756-8765.2009.01080.x

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research*, *308*, 98–108. http://doi.org/10.1016/j.heares.2013.08.011

Patel, A. D., & Iversen, J. R. (2007). The linguistic benefits of musical abilities. *Trends in Cognitive Sciences*, *11*(9), 369–372. http://doi.org/10.1016/j.tics.2007.06.007

Patel, A., & Morgan, E. (2017). Exploring Cognitive Relations Between Prediction in Language and Music. *Cognitive Science*, *41*, 303–320. http://doi.org/10.1111/cogs.12411

Peretz, I., Saffran, J., Schön, D., & Gosselin, N. (2012). Statistical learning of speech, not music, in congenital amusia. *Annals of the New York Academy of Sciences*, *1252*(1), 361–366. http://doi.org/10.1111/j.1749-6632.2011.06429.x

Perfetti, C. A., Wlotko, E. W., & Hart, L. A. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology: Learning Memory and Cognition*, *31*(6), 1281–1292. http://doi.org/10.1037/0278-7393.31.6.1281

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233–238.

http://doi.org/10.1016/j.tics.2006.03.006

Perruchet, P., Poulin-charronnat, B., Tillmann, B., & Peereman, R. (2014). New evidence for chunk-based models in word segmentation. *Acta Psychologica*, *149*, 1–8.

Perruchet, P., & Tillmann, B. (2010). Exploiting multiple sources of information in learning an artificial language: Human data and modeling. *Cognitive Science*, *34*(2), 255–285. http://doi.org/10.1111/j.1551-6709.2009.01074.x

Petersson, K. M., Folia, V., & Hagoort, P. (2012). What artificial grammar learning reveals about the neurobiology of syntax. *Brain and Language*, *120*(2), 83–95. http://doi.org/10.1016/j.bandl.2010.08.003

Picton, T. W., Lins, O., & Scherg, M. (1995). The recording and analysis of event-related potentials. In *Handbook of Neuropsychology* (Vol. 10, pp. 3–73).

Pinheiro, A. P., Vasconcelos, M., Dias, M., Arrais, N., & Gonçalves, Ó. F. (2015). The music of language: An ERP investigation of the effects of musical training on emotional prosody processing. *Brain and Language*, *140*, 24–34. http://doi.org/10.1016/j.bandl.2014.10.009

Polich, J. (2003). *Detection of Change: event-related potential and fMRI findings.* (J. Polich, Ed.). Springer Science & Business Media. http://doi.org/10.1007/978-1-4615-0294-4

Pulvermüller, F., Lutzenberger, W., & Birbaumer, N. (1995). Electrocortical distinction of vocabulary types. *Electroencephalography and Clinical Neurophysiology*, *94*(5), 357–370. http://doi.org/10.1016/0013-4694(94)00291-R

Rabovsky, M., Hansen, S. S., & McClelland, J. L. (2018). Modelling the N400 brain potential as change in a probabilistic representation of meaning. *Nature Human Behaviour*. http://doi.org/10.1038/s41562-018-0406-4

Raviv, L., & Arnon, I. (2018). The developmental trajectory of children's auditory and visual statistical learning abilities: modality-based differences in the effect of age. *Developmental Science*, *21*(4), 1–13. http://doi.org/10.1111/desc.12593

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*(6), 855–863. http://doi.org/10.1016/S0022-5371(67)80149-X

Reber, A. S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory*, *2*(1), 88–94. http://doi.org/10.1037/0278-7393.2.1.88

Reber, A. S. (1989). Implicit Learning and Tacit Knowledge. *Journal of Experimental Psychology: General*, *118*(3), 219–235. http://doi.org/10.1037/0096-3445.118.3.219

Reber, P. J., & Squire, L. R. (1994). Parallel brain systems for learning with and without awareness. *Learning & Memory*, *1*, 217–229. http://doi.org/10.1101/lm.1.4.217

Reinke, K. S., He, Y., Wang, C., & Alain, C. (2003). Perceptual learning modulates sensory evoked response during vowel segregation. *Cognitive Brain Research*, *17*(3), 781–791. http://doi.org/10.1016/S0926-6410(03)00202-7

Rhythm. Def. 1. (n.d.). In Oxford English Dictionary Online. Retrieved fromhttps://en.oxforddictionaries.com/definition/rhythm. (n.d.).

Roden, I., Kreutz, G., & Bongard, S. (2012). Effects of a school-based instrumental music program on verbal and visual memory in primary school children: A longitudinal study. *Frontiers in Psychology*, *3*(DEC), 1–9. http://doi.org/10.3389/fpsyg.2012.00572

Rodriguez-Fornells, A., Cunillera, T., Mestres-Misse, A., & de Diego-Balaguer, R. (2009). Neurophysiological mechanisms involved in language learning in adults. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536), 3711–3735. http://doi.org/10.1098/rstb.2009.0130

Romberg, A. R., & Saffran, J. R. (2011). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews. Cognitive Science*, *1*(6), 906–914. http://doi.org/10.1002/wcs.78.Statistical

Ruggles, D. R., Freyman, R. L., & Oxenham, A. J. (2014). Influence of musical training on understanding voiced and whispered speech in noise. *PLoS ONE*, *9*(1). http://doi.org/10.1371/journal.pone.0086980

Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*, *12*(4), 110–114. http://doi.org/10.1111/1467-8721.01243

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. http://doi.org/10.1126/science.274.5294.1926

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27–52. http://doi.org/10.1016/S0010-0277(98)00075-4

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*(4), 606–621. http://doi.org/10.1006/jmla.1996.0032

Saffran, J. R., & Wilson, D. P. (2003). From syllables to syntax: Multilevel statistical learning by 12-month-old infants. *Infancy*, *4*(2), 273–284. http://doi.org/10.1207/S15327078IN0402_07

Sanders, L. D., Ameral, V., & Sayles, K. (2009). Event-related potentials index segmentation of nonsense sounds. *Neuropsychologia*, *47*(4), 1183–1186. http://doi.org/10.1016/j.neuropsychologia.2008.11.005

Sanders, L. D., Newport, E. L., & Neville, H. J. (2002). Segmenting nonsense: An event-related potential index of perceived onsets in continuous speech. *Nature Neuroscience*, *5*(7), 700–703. http://doi.org/10.1038/nn873

Schlaug, G., Jäncke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia*, *33*(8), 1047–1055. http://doi.org/10.1016/0028-3932(95)00045-5

Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience*, *5*(7), 688–694. http://doi.org/10.1038/nn871

Schneider, P., Sluming, V., Roberts, N., Scherg, M., Goebel, R., Specht, H. J., ... Rupp, A. (2005). Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nature Neuroscience*, *8*(9), 1241–1247. http://doi.org/10.1038/nn1530

Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition*, *106*(2), 975–983. http://doi.org/10.1016/j.cognition.2007.03.005

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, *41*(3), 341–349. http://doi.org/10.1111/1469-8986.00172.x

Schröger, E. (1997). On the detection of auditory deviations: A pre-attentive activation model. *Psychophysiology*, *34*(3), 245–257. http://doi.org/10.1111/j.1469-8986.1997.tb02395.x

Schultz, B. G., Stevens, C. J., Keller, P. E., & Tillmann, B. (2013). A Sequence Identification Measurement Model to Investigate the Implicit Learning of Metrical Temporal Patterns. *PLoS*

ONE, *8*(9), 1–13. http://doi.org/10.1371/journal.pone.0075163

Seidenberg, M. S. (2014). Language Acquisition and Use : Learning and Applying Probabilistic Constraints, (April 1997). http://doi.org/10.1126/science.275.5306.1599

Seppänen, M., Brattico, E., & Tervaniemi, M. (2007). Practice strategies of musicians modulate neural processing and the learning of sound-patterns. *Neurobiology of Learning and Memory*, *87*(2), 236–247. http://doi.org/10.1016/j.nlm.2006.08.011

Shahin, A., Bosnyak, D. J., Trainor, L. J., & Roberts, L. E. (2003). Enhancement of Neuroplastic P2 and N1c Auditory Evoked Potentials in Musicians. *Journal of Neuroscience*, *23*(13), 5545–5552. http://doi.org/10.1523/JNEUROSCI.23-13-05545.2003

Shanks, D. R., Rowland, L. A., & Ranger, M. S. (2005). Attentional load and implicit sequence learning. *Psychological Research*, *69*(5–6), 369–382. http://doi.org/10.1007/s00426-004-0211-8

Sheehan, K. A., McArthur, G. M., & Bishop, D. V. M. (2005). Is discrimination training necessary to cause changes in the P2 auditory event-related brain potential to speech sounds? *Cognitive Brain Research*, *25*(2), 547–553. http://doi.org/10.1016/j.cogbrainres.2005.08.007

Shook, A., Marian, V., Bartolotti, J., & Schroeder, S. (2013). Musical experience influences novel language learning. *The American Journal of Psychology*, *126*(1), 95–104. http://doi.org/10.5406/amerjpsyc.126.1.0095

Shufaniya, A., & Arnon, I. (2018). Statistical Learning Is Not Age-Invariant During Childhood: Performance Improves With Age Across Modality. *Cognitive Science*, 1–16. http://doi.org/10.1111/COGS.12692

Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., & Frost, R. (2018). Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition*, *177*(April), 198–213. http://doi.org/10.1016/j.cognition.2018.04.011

Siegelman, N., Bogaerts, L., Kronenfeld, O., & Frost, R. (2017). Redefining "Learning" in Statistical Learning: What Does an Online Measure Reveal About the Assimilation of Visual Regularities? *Cognitive Science*, *42*, 692–727. http://doi.org/10.1111/cogs.12556

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, *81*, 105–120. http://doi.org/10.1016/j.jml.2015.02.001

Simons, R. F., Graham, F. K., Miles, M. A., & Chen, X. (2001). On the relationship of P3a and the Novelty-P3. *Biological Psychology*, *56*(3), 207–218. Retrieved from papers3://publication/uuid/32F1F356-4F45-4D03-A28F-1F0F872E0679

Slevc, L. R., & Miyake, A. (2006). Individual Differences in Second- Language Proficiency, *17*(8), 675–681. http://doi.org/10.1111/j.1467-9280.2006.01765.x

Sluming, V., Barrick, T., Howard, M., Cezayirli, E., Mayes, A., & Roberts, N. (2002). Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. *NeuroImage*, *17*(3), 1613–1622. http://doi.org/10.1006/nimg.2002.1288

Snyder, J. S., Alain, C., & Picton, T. W. (2006). Effects of attention on neuroelectric correlates of auditory stream segregation. *Journal of Cognitive Neuroscience*, *18*(1), 1–13. http://doi.org/10.1162/089892906775250021

Steele, C. J., Bailey, J. A., Zatorre, R. J., & Penhune, V. B. (2013). Early Musical Training and White-Matter Plasticity in the Corpus Callosum: Evidence for a Sensitive Period. *Journal of Neuroscience*, *33*(3), 1282–1290. http://doi.org/10.1523/JNEUROSCI.3578-12.2013

Stobbe, N., Westphal-Fitch, G., Aust, U., & Tecumseh Fitch, W. (2012). Visual artificial grammar learning: Comparative research on humans, kea (Nestor notabilis) and pigeons (Columba livia). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1598), 1995–2006. http://doi.org/10.1098/rstb.2012.0096

Strait, D. L., & Kraus, N. (2011). Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Frontiers in Psychology*, *2*(113), 1–10. http://doi.org/10.3389/fpsyg.2011.00113

Strait, D. L., Parbery-Clark, A., Hittner, E., & Kraus, N. (2012). Musical training during early childhood enhances the neural encoding of speech in noise. *Brain and Language*, *123*(3), 191–201. http://doi.org/10.1016/j.bandl.2012.09.001.Musical

Suárez, L., Elangovan, S., & Au, A. (2016). Cross-sectional study on the relationship between music training and working memory in adults. *Australian Journal of Psychology*, *68*(1), 38–46. http://doi.org/10.1111/ajpy.12087

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*(3700), 1187–1188.

http://doi.org/10.1126/science.150.3700.1187

Szewczyk, J. M., & Schriefers, H. (2018). The N400 as an index of lexical preactivation and its implications for prediction in language comprehension. *Language, Cognition and Neuroscience*, *33*(6), 665–686. http://doi.org/10.1080/23273798.2017.1401101

Teki, S., & Griffiths, T. D. (2016). Brain bases of working memory for time intervals in rhythmic sequences. *Frontiers in Neuroscience*, *10*(JUN). http://doi.org/10.3389/fnins.2016.00239

Terry, J., Stevens, C. J., Weidemann, G., & Tillmann, B. (2016). Implicit learning of between-group intervals in auditory temporal structures. *Attention, Perception, and Psychophysics*, *78*(6), 1728–1743. http://doi.org/10.3758/s13414-016-1148-x

Tervaniemi, M., Janhunen, L., Kruck, S., Putkinen, V., & Huotilainen, M. (2016). Auditory profiles of classical, jazz, and rock musicians: Genre-specific sensitivity to musical sound features. *Frontiers in Psychology*, *6*(JAN), 1–11. http://doi.org/10.3389/fpsyg.2015.01900

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: An event-related potential and behavioral study. *Experimental Brain Research*, *161*(1), 1–10. http://doi.org/10.1007/s00221-004-2044-5

Tervaniemi, M., Kruck, S., De Baene, W., Schröger, E., Alter, K., & Friederici, A. D. (2009). Top-down modulation of auditory processing: Effects of sound context, musical expertise and attentional focus. *European Journal of Neuroscience*, *30*(8), 1636–1642. http://doi.org/10.1111/j.1460-9568.2009.06955.x

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding Speech Prosody: Do Music Lessons Help? *Emotion*, *4*(1), 46–64. http://doi.org/10.1037/1528-3542.4.1.46

Thönnessen, H., Boers, F., Dammers, J., Chen, Y. H., Norra, C., & Mathiak, K. (2010). Early sensory encoding of affective prosody: Neuromagnetic tomography of emotional category changes. *NeuroImage*, *50*(1), 250–259. http://doi.org/10.1016/j.neuroimage.2009.11.082

Tierney, A., & Kraus, N. (2013). *Music training for the development of reading skills*. *Progress in Brain Research* (1st ed., Vol. 207). Elsevier B.V. http://doi.org/10.1016/B978-0-444-63327-9.00008-4

Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, *97*(2), 25–34. http://doi.org/10.1016/j.cognition.2005.01.006

Toro, J. M., & Trobalón, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception and Psychophysics*, *67*(5), 867–875. http://doi.org/10.3758/BF03193539

Tremblay, K., Kraus, N., & McGee, T. (1998). The time course of auditory perceptual learning: Neurophysiological changes during speech-sound training. *NeuroReport*, *9*(16), 3557–3560. http://doi.org/10.1097/00001756-199811160-00003

Tremblay, K. L., & Kraus, N. (2002). Auditory training induces asymmetrical changes in cortical neural activity. *Journal of Speech, Language, and Hearing Research*, *45*(3), 564–572.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*(4), 552–564. http://doi.org/10.1037/0096-3445.134.4.552

Turk-browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2008). Neural Evidence of Statistical Learning : Efficient Detection of Visual Regularities Without Awareness. *Journal of Cognitive Neuroscience*, 1934–1945. http://doi.org/10.1162/jocn.2009.21131

Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit Perceptual Anticipation Triggered by Statistical Learning. *Journal of Neuroscience*, *30*(33), 11177–11187. http://doi.org/10.1523/JNEUROSCI.0858-10.2010

Van Berkum, J. J. A. (2008). Understanding sentences in context: What brain waves can tell us. *Current Directions in Psychological Science*, *17*(6), 376–380. http://doi.org/10.1111/j.1467-8721.2008.00609.x

Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time Course of Word Identification and Semantic Integration in Spoken Language. *Journal of Experimental Psychology: Learning Memory and Cognition*, *25*(2), 394–417. http://doi.org/10.1037/0278-7393.25.2.394

van Petten, C., & Rheinfelder, H. (1995). Conceptual relationships between spoken words and environmental sounds: Event-related brain potential measures. *Neuropsychologia*, *33*(4), 485–508. http://doi.org/10.1016/0028-3932(94)00133-A

Van Zuijen, T. L., Sussman, E., Winkler, I., Näätänen, R., & Tervaniemi, M. (2005). Auditory organization of sound sequences by a temporal or numerical regularity - A mismatch negativity study comparing musicians and non-musicians. *Cognitive Brain Research*, *23*(2–3), 270–276. http://doi.org/10.1016/j.cogbrainres.2004.10.007

Vaquero, L., Hartmann, K., Ripollés, P., Rojo, N., Sierpowska, J., François, C., ... Altenmüller, E. (2016). Structural neuroplasticity in expert pianists depends on the age of musical training onset. *NeuroImage*, *126*(November), 106–119. http://doi.org/10.1016/j.neuroimage.2015.11.008

Vasuki, P. R. M., Sharma, M., Ibrahim, R., & Arciuli, J. (2017a). Statistical learning and auditory processing in children with music training: An ERP study. *Clinical Neurophysiology*, *128*(7), 1270–1281. http://doi.org/10.1016/j.clinph.2017.04.010

Vasuki, P. R. M., Sharma, M., Ibrahim, R. K., & Arciuli, J. (2017b). Musicians' Online Performance during Auditory and Visual Statistical Learning Tasks. *Frontiers in Human Neuroscience*, *11*(114), 1–12. http://doi.org/10.3389/fnhum.2017.00114

Virtala, P., Huotilainen, M., Partanen, E., & Tervaniemi, M. (2014). Musicianship facilitates the processing of Western music chords-An ERP and behavioral study. *Neuropsychologia*, *61*(1), 247–258. http://doi.org/10.1016/j.neuropsychologia.2014.06.028

Wan, C. Y., & Schlaug, G. (2010). Music Making as a Tool for Promoting Plasticity across the Life Span. *The Neuroscientist*, *16*(5), 566–577. http://doi.org/10.1177/1073858410377805.Music

Watt, R., & Quinn, S. (2006). Analysis of local and global timing and pitch change in or- dinary melodies. *Proceedings of the 9th International Conference on Music Perception and Cognition (ICMPC) and 6th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM)*.

Whiting, K. A., Martin, B. A., & Stapells, D. R. (1998). The Effects of Broadband Noise Masking on Cortical Event-Related Potentials to Speech Sounds /ba/ and /da/. *Ear and Hearing*, *19*(3), 218–231. http://doi.org/10.1097/00003446-199806000-00005

Winkler, I. (2003). Change detection in complex auditory environment: beyond the oddball paradigm. In *Detection of Change* (p. (pp. 61-81).). Boston, MA: Springer. http://doi.org/10.1007/978-1-4615-0294-4

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*(4), 420–422. http://doi.org/10.1038/nn1872

Yabe, H., Tervaniemi, M., Reinikainen, K., & Näätänen, R. (1997). Temporal window of integration revealed by MMN to sound omission. *NeuroReport*, *8*(8), 1971–1974. http://doi.org/10.1097/00001756-199705260-0003

Zarate, J. M., & Zatorre, R. J. (2008). Experience-dependent neural substrates involved in vocal pitch regulation during singing. *NeuroImage*, *40*(4), 1871–1887. http://doi.org/10.1016/j.neuroimage.2008.01.026

Zatorre, R. J. (2013). Predispositions and Plasticity in Music and Speech Learning: Neural Correlates and Implications. *Science*, *342*(November), 585–589. http://doi.org/10.1126/science.1238414

Zendel, B. R., & Alain, C. (2009). Concurrent sound segregation is enhanced in musicians. *Journal of Cognitive Neuroscience*, *21*(8), 1488–1498. http://doi.org/10.1162/jocn.2009.21140

Zendel, B. R., & Alain, C. (2014). Enhanced attention-dependent activity in the auditory cortex of older musicians. *Neurobiology of Aging*, *35*(1), 55–63. http://doi.org/10.1016/j.neurobiolaging.2013.06.022

Zhao, T. C., Lam, H. T. G., Sohi, H., & Kuhl, P. K. (2017). Neural processing of musical meter in musicians and non-musicians. *Neuropsychologia*, *106*, 289–297. http://doi.org/10.1016/j.neuropsychologia.2017.10.007

Zuijen, T. L. van, Sussman, E., Winkler, I., Näätänen, R., & Tervaniemi, M. (2004). Grouping of Sequential Sounds—An Event-Related Potential Study Comparing Musicians and Nonmusicians. *Journal of Cognitive Neuroscience*, *16*(2), 331–338. http://doi.org/10.1162/089892904322984607

PART II

EMPIRICAL STUDIES

The work presented in Chapter 4 and Chapter 5 was submitted for publication to international peer-reviewed journals:

Vasconcelos, M. & Pinheiro, P. (submitted). Musical training modulates ASL as a function of stimulus type: combined behavioral and EEG insights.

Vasconcelos, M. & Pinheiro, P. (under review). Does attention affect ASL in musicians? EEG and behavioral insights.

The empirical studies reported in this part of the Dissertation were also publicly presented at national and international scientific meetings:

Vasconcelos, M & Pinheiro, A. P. (2018, April). *Neural and behavioral correlates of auditory statistical learning in musicians.* Poster presentation at APPE 2018, Braga, Portugal.

Vasconcelos, M. & Pinheiro, A. P. (2017, August). *An electrophysiological investigation on the role of musical training on speech prosody processing and segmentation.* Oral presentation at ESCOM 2017, Ghent, Belgium.

Vasconcelos, M & Pinheiro, A. P. (2017, June). *Musicians' brains are tuned for speech: behavioral and neural evidence of distinctive statistical learning abilities in musicians.* Oral presentation at Interdisciplinary Advances in Statistical learning, Bilbao, Spain.

Vasconcelos, M & Pinheiro, A. P. (2016, November). *How music changes the linguistic brain: from auditory perception to auditory prediction.* Oral presentation at II Colloquium in Musical Psychophysiology, Braga, Portugal.

Chapter 3. The effects of musical training on the pre-attentive processing of simple statistical regularities

3.1 Abstract

The ability to extract regularities from the auditory landscape makes humans able to predict and comprehend highly relevant auditory percepts such as speech and music. Musical training has been studied as a model of neuroplasticity that benefits auditory perception and may have a positive impact in the processing of regularities in musical and speech input. The MMN is an ERP component of the EEG that is elicited in response to violations of predictions about sound properties (e.g., pitch, duration) and arrangements (e.g., harmonic/ rhythmic structure). By means of ERPs, the present study investigated if the MMN to violations of simple regularities (i.e., pitch changes) is modulated by musical expertise. EEG data were recorded from 19 musicians and 23 non-musicians. Participants were presented with high-probability (1000 Hz) and low-probability (1200 Hz) pure tones while attending to a movie with neutral content. A MMN was observed in response to pitch changes. There were no differences between musicians and non-musicians regarding MMN amplitude or latency. These findings indicate that musical expertise does not affect the pre-attentive processing of sound regularities based on pitch information. Previous research has shown that musical training enhances abstract regularity processing fostering the computation of complex statistical contingencies, which might suggest a broad and increased capacity of musicians to learn regularities and predict events in the auditory input. By failing to show neurophysiological differences in the processing of simple auditory regularities between musical experts and naïve subjects our findings challenge previous studies and suggest that musical training brings no positive advantages to auditory learning when the task of tracking regularities relies on the perception of pitch changes.

Keywords: Musical training, irregularity detection, MMN, pitch processing.

3.2 Introduction

For the brain to attribute meaning to input proceeding from the acoustic environment the auditory system needs to process the acoustic input as temporally structured patterns (Bendixen, Schroger, & Winkler, 2009; Winkler, Denham, & Nelken, 2009). This process is based on the integration of the acoustic features of sound percepts over time. The perception of regularities is an important aspect of auditory perception that simultaneously allows the detection of changes in the environment: It is difficult to detect gaps in temporally irregular sequences whereas there is the need of less cognitive effort to notice silent gaps within a regular sound sequence (Mizuno, Schwartz, & Cazals, 1994). The study of these processes is highly relevant since the processing of events that violate predictions based on learned statistical regularities is an important aspect of everyday life.

Musical training has been associated with enhanced auditory perceptual abilities in musicians (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004; Koelsch, Schmidt, & Kansok, 2002; Koelsch, Schröger, & Tervaniemi, 1999; Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005). For that reason, musical training is a model of neuroplasticity that provides a unique window to probe the brain mechanisms involved in auditory perception in general (Gaser & Schlaug, 2003; Münte, Kohlmetz, Nager, & Altenmüller, 2001; Nager, Kohlmetz, Altenmüller, Rodriguez-Fornells, & Münte, 2003) and regularity detection in particular (Bidelman & Alain, 2015; François, Jaillet, Takerkart, & Schön, 2014; Kraus & Chandrasekaran, 2010). For instance, musicians detect DEV chords and melodic patterns (Koelsch, Schröger & Tervaniemi, 1999; Tervaniemi et al., 2001), as well as small pitch changes between tones (Tervaniemi et al., 2005) more easily when compared to musically-naïve subjects. Pitch is related to the fundamental frequency of the lowest frequency partial of a sound and refers to the subjective perception of that sound (Gelfand, 1998). Pitch perception is very important in the context of music perception, representing also a critical acoustic feature for the decoding of vocal emotions (Hevner, 1937; Schön, Magne, & Besson, 2004). In the English language, for example, pitch contours that sign important information are typically larger than one-half an octave, whereas in Western music pitch changes are much smaller (e.g., approximately 1/6th to 1/12th of an octave; see Ayotte, Peretz, & Hyde, 2002; Sundberg, 1994).

The capacity to extract statistical regularities from continuous streams of stimuli has been studied in the realm of SL research (Saffran, Aslin, & Newport, 1996). SL represents the process by which statistical regularities of any input are learned allowing the prediction of upcoming events (Saffran, Aslin, et al., 1996; Siegelman & Frost, 2015). The SL framework has been used to study the processing of TPs in the context of artificial grammars and word learning (Aslin, Saffran, & Newport, 1998; Perruchet & Pacton, 2006; Saffran, Newport, & Aslin, 1996). Beyond the typical SL paradigms, oddball tasks offer efficient experimental setups to investigate predictive mechanisms such as those underlying the processing of basic auditory contingencies. Oddball tasks are characterized by the intermixed presentation of STD and DEV stimuli that occur at a high and at a low probability rate, respectively. When stimuli are task-irrelevant, occasional changes – i.e., DEVs – in one or several physical or abstract aspects of an otherwise regular input – i.e., STDs - elicit the MMN ERP component (Näätänen, 2002; Näätänen, Paavilainen, Rinne, & Alho, 2007). The MMN is a fronto-central negativity originating mainly from the auditory cortex (Alho et al., 1996; Alho, 1995), which peaks at approximately 100-250 ms after stimulus onset (Duncan et al., 2009). It reflects deviance detection mechanisms that are typically triggered by discriminable changes in an otherwise regular stream of sounds (Näätänen et al., 1978; Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001; Picton, Alain, Otten, Ritter, & Achim, 2000). The MMN - elicited even in the absence of conscious attention (Näätänen et al., 2007, 2001) - is a pre-attentive neurophysiological signature of sensory predictive processes that are essential for subsequent stimulus discrimination. When the attention of the listeners is automatically oriented to the acoustic changes, MMN might be followed by the N2b and P3a ERP components that index sound discrimination and the conscious shift of attention towards the target stimuli (Näätänen, Simpson, & Loveless, 1982; Ritter et al., 1992). MMN has been elicited by different forms of auditory violations: changes in tonal sequences (Sussman, Ritter, & Vaughan, 1999) or tonal relationships within a melody (Tervaniemi, Rytkönen, Schröger, Ilmoniemi, & Näätänen, 2001; Trainor, McDonald, & Alain, 2002), tone alternations (Nordby, Roth, & Pfefferbaum, 1998), changes in stimulus intensity and duration (Escera, Corral, & Yago, 2002), tone omissions (Herholz, Lappe, Knief, & Pantev, 2009; Herholz, Lappe, & Pantev, 2009), among others. In a recent study, the MMN was elicited by tone triplets whose endings (i.e., the third tone) could have different TPs (Koelsch, Busch, Jentschke, & Rohrmeier, 2016). Low probability endings elicited an enhanced anterior negativity (that was referred to by the authors as statistical MMN) compared to intermediate and high probability endings. These findings add to the evidence of

MMN as a neurophysiological index of learned regularities, revealing that the component also reflects the SL of distributional cues, such as TPs. Therefore, the MMN is a crucial tool to examine regularity processing and deviance detection in musically trained and naïve participants without the contamination of attentional or cognitive load demands (Tervaniemi & Huotilainen, 2003).

The existing studies have provided evidence for enhanced pre-attentive auditory processing mechanisms in musicians (Besson, Faïta, & Requin, 1994; Münte et al., 2001; Pantev et al., 1998). Research indicates that pitch violations in pure tones elicit a MMN both in musicians and non-musicians (Koelsch et al., 1999). This is true only for easily noticeable pitch changes. However, subtle pitch changes that are difficult to perceive (see, for example, Nikjeh, Lister, & Frisch, 2008; Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005) and acoustic deviations in more complex auditory patterns such as chords (Koelsch et al., 1999), melodies (Tervaniemi et al., 2001), rule-base sequences (e.g., melodic contour and interval sequences, see Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004; Pantev et al., 2003) and metre-incongruent DEVs (Geiser, Sandmann, Jäncke, & Meyer, 2010) lead to increased MMN responses (e.g., larger amplitude) only in participants with musical expertise. Musicians are able to group sequences of tones based on good-continuation-of-pitch (i.e., tone sequences composed of ascending pitches instead of sequences marked by ascending and descending pitch), whereas non-musicians are capable of auditory grouping based only on pitch similarity (Zuijen, Sussman, Winkler, Näätänen, & Tervaniemi, 2004). The amplitude of the MMNm – the magnetic counterpart of the MMN – was also found not to be modulated by musical expertise when it comes to simple tones discrimination (e.g., frequency (Hz) DEVs) (Paraskevopoulos, Kuchenbuch, Herholz, & Pantev, 2012). Notwithstanding, musicians show increased MMNm in response to high/low "voice" and in-key/out-of-key changes (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2005). Herholz and colleagues (Herholz, Lappe, & Pantev, 2009) tested whether a MMN emerged if the expectation of the imagined continuations of well-known songs were violated and if its amplitude was modulated by musical expertise. Only musicians showed a MMN for DEV tones, which suggests that musical expertise modulates music imagery and modulates MMN amplitude even when MMN is elicited based on the global statistical knowledge of auditory input regularities instead of sensory memory traces (Herholz, Lappe, & Pantev, 2009). Other studies indicate that whereas the magnitude of the MMN response (i.e., its amplitude) to simple pitch deviances may not differ between musicians and non-musicians, its latency does: earlier MMN to pure tones, harmonic sequences and speech syllables was observed in musicians compared to non-musicians (Nikjeh et al., 2008; Nikjeh, Lister, & Frisch, 2009).

3.2.1 The current study and hypotheses

In the current study, we examined the capacity of musically trained and musically-naïve participants to encode simple statistical regularities. We aimed to investigate whether musical training influences the pre-attentive processing of simple statistical contingencies in the auditory domain. Therefore, participants were presented with a standard oddball paradigm composed of pure tones of distinct frequencies (Hz) - 1000 Hz and 1200 Hz – while watching a soundless movie with neutral content. The oddball task probed the capacity to compute auditory regularities based on a unique acoustic feature – pitch – which is relevant for both music and speech processing (Heinks-Maldonado, Mathalon, Gray, & Ford, 2005; Kishon-Rabin, Amir, Vexler, & Zaltz, 2001; Tang, Xiong, Zhang, Dong, & Nan, 2016; Wong, Skoe, Russo, Dees, & Kraus, 2007). Since pitch perception is an essential skill to disentangle the multiple and successive objects that compose auditory landscapes (Bregman, 1990), we chose to evaluate the influence of musical training on auditory deviance detection based on pitch modulations. We ensured that the musicians integrating this study had strong musical expertise marked by several years of uninterrupted instrumental training and confirmed by the high performance on a test of music audiation (please see Table 2). Moreover, the group of musicians was heterogeneous regarding the type of musical background (i.e., they were expert in nine different types of musical instruments, see Table 3), which in relation to early studies, ensured that the processing of the pitch tones would not be enhanced or compromised *a priori* by participants specific technical expertise. Based on previous findings (Näätänen, 2002; Näätänen et al., 2007; Nyman et al., 1990), MMN elicitation would indicate that participants implicitly learned the statistical structure of the auditory stream. If musical training leads to a facilitated and faster processing of acoustic contingencies in general (i.e., independently of the degree of acoustic complexity of the sounds), we expected musicians to show an increased (i.e., more negative) and earlier MMN relative to non-musicians in response to pitch changes. Conversely, if musical training only represented an advantage for the encoding of auditory statistical regularities when listeners are processing more complex acoustic stimuli (i.e., involving the detection of changes in a set of acoustic features), no differences in the MMN response would be expected between groups.

3.3 Material and Methods

3.3.1 Participants

A group of 19 musicians (10 women; 21.05 ± 2.88 years [*M* ± *SD*]) and 23 non-musicians (13 women; 23.09 ± 4.53 years) participated in this study. The groups did not differ significantly in age, $t_{(38)}$ = 1.83, *p* = .075, (see Table 2). Participants in the musician group had more than 8 years of musical experience (years of training = 13.05 ± 3.21 years; training onset = 7.79 ± 2.88 years), were currently practicing at least one musical instrument and were teaching or attending musical classes at college (e.g., Music Bachelors of the University of Minho) or cultural groups (see Table 3 for a detailed description of musicians' musical background). The group of non-musicians was composed by individuals who did not play any musical instrument currently or in the past, and have never had formal musical training besides attending the two-year musical classes that are mandatory in the Portuguese school curricula for fifth and sixth grade students. Participants had normal hearing, normal or corrected-to-normal vision, reported no history of psychopathological/neurological disorders or substance abuse and were not taking medication. All participants were right-handed (Edinburgh Handedness inventory, Oldfield, 1971), native speakers of European Portuguese. The Advanced Measures of Music Audiation (AMMA; Gordon, 1989) were used to guarantee that musicians had higher musical aptitude abilities (see Table 2). Furthermore, participants were assessed using the Graffar Scale (Graffar, 1956) and The Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005; Simões et al., 2008) in order to control for differences in socioeconomic status and cognitive functioning, respectively (see Table 2). The study was conducted under the approval and the ethical oversight of the Ethics Committee for Life and Health Sciences of the University of Minho. Written consent was provided according to the updated version of the Declaration of Helsinki. Participants who were college students received course credit for their participation while the remaining participants received gift vouchers.

TABLE 2.

Socio-Demographic and Other Relevant Characteristics of The Sample.

	Musicians (*N* = 19)	Non-Musicians (*N* = 23)	Statistics	
	M (SE)	*M (SE)*	t	*p*
Age	21.05 (0.66)	23.09 (0.95)	1.69	0.098
Graffar Scale	10.95 (0.56)	12.48 (0.59)	1.85	0.072
MoCA	29.26 (0.19)	28.91 (0.22)	-1.19	0.238
AMMA Total	70.53 (3.32)	53.96 (4.20)	**-3.00**	**0.005**
AMMA Tonal	69.74 (3.71)	55.09 (4.27)	**-2.53**	**0.015**
AMMA Rhythm	68.58 (3.29)	53.52 (3.81)	**-2.92**	**0.006**
Musical training (years)	13.05 (0.74)	N.A.	N.A.	
Onset of training (years)	7.79 (0.66)	N.A.	N.A.	

Note. The MoCA is from Nasreddine et al. (2005) and the MoCA adaptation to the Portuguese population is from Simões et al. (2008); The AMMA is from Gordon (1989); M (Mean); SE – Standard Error; N.A. – Non-Applicable. Significant differences are highlighted in bold.

3.3.2 Stimuli

The stimuli were 1000 Hz (STDs) and 1200 Hz (DEVs) pure tones with 100 ms duration. The sounds were normalized in intensity (70 dB) using Praat software (Boersma & Weenink, 2013). The MMN experiment was composed of one block containing 1050 stimuli: 950 STD (probability [p] ≈ 0.905) and 100 DEV (p ≈ 0.095) pure tones. The inter-stimulus interval (ISI) was 300 ms. The presentation of stimuli was pseudo-randomized with at least two STDs occurring between each DEV. Stimuli were presented with Sennheiser earphones, using Presentation® software (version 16.3; Neurobehavioral Systems, Albany, CA). Figure 2 displays a schematic representation of the MMN experiment.

3.3.3 Procedure

The EEG was conducted in a sound-attenuated and electrically shielded room where participants sat comfortably in a reclining chair at a distance of 100 cm from a desktop screen. Participants were instructed to watch a silent movie – a nature documentary with neutral content - and to ignore the auditory stimuli so they would be able to answer a brief questionnaire at the end of the session. The movie was presented on a LG ACPI x86-based computer. The experiment took approximately 15 minutes.

TABLE 3.

Characterization of Musicians' Musical Background and Training Routines.

Participant Code	Musical training[1]	Onset of Training[1]	Instrument[2]	Practice days/week	Practice hours/day
M1	15	6	violin	4	2
M2	16	6	piano	7	2
M3	8	11	clarinet	7	2.5
M4	9	11	saxophone	7	6
M5	10	9	clarinet	7	3
M6	13	6	violoncello	7	6.5
M7	10	15	trumpet	7	4
M8	15	5	violoncello	5	5.5
M9	11	12	percussion	7	2
M10	13	6	percussion	5	3
M11	15	11	piano	5	3
M12	14	5	piano	6	4
M13	20	7	violin	7	2
M14	13	5	violin	7	4
M15	17	7	piano	7	3.5
M16	17	6	transverse flute	4	2
M17	10	6	violin	4	1.5
M18	12	8	piano	7	2
M19	10	6	guitar	7	1

Note. [1] Measured in years; [2] The column displays the participant's main instrument despite several participants reporting expertise in more than one musical instrument.

3.3.4 EEG data acquisition and processing

EEG data were recorded from a 64 channels Active-Two system (Biosemi, Amsterdam, The Netherlands) according to the extended international 10-20 system (Jasper, 1958). The data was analyzed using Brain Vision Analyzer software (Brain Products, Munich, Germany).

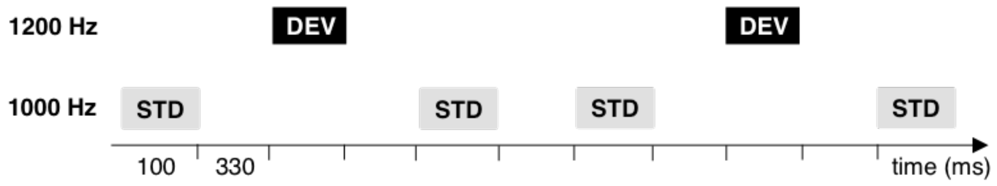

FIGURE 2. Schematic illustration of the MMN paradigm. DEV sounds (black rectangles) of 1200 Hz were intermixed with STD sounds (grey rectangles) of 1000 Hz. *Note:* ISI = 300 ms.

Eye movements were monitored with 4 electrodes: horizontal movements were measured with 2 electrodes placed on the outer canthi of each eye, and vertical movements were monitored with one electrode placed at the infra-orbital ridge of the left eye and with FP1. The electrode offset was kept below 20 mV. The EEG was digitized at a rate of 512 Hz and filtered offline using a 0.1-30 Hz Infinite Impulse Response filter. The method developed by Graton and collaborators (Gratton, Coles, & Donchin, 1983) was implemented for ocular correction. Artifacts were detected based on eye blinks and movement artifacts exceeding ±100 μV.

For ERP analyses, the signal was re-referenced offline to the algebraic average of the left and right mastoids. ERPs were time-locked to tone onset in the two conditions and averaged for epochs of 400 ms (e.g., 100 ms pre-stimulus and 300 ms post-stimulus). The baseline was defined as the 100 ms period preceding stimulus onset. STD and DEV segments were averaged for each participant, separately, at each electrode site. The MMN was obtained by calculating difference waveforms, i.e., the subtraction of ERP activity of STD sounds to that elicited by DEV sounds. Individual amplitudes were extracted based on mean amplitude voltage and the latencies were extracted based on peak latency measures (see Luck, 2005). Amplitudes and latencies were calculated from a 100 ms window (160-260 ms) approximately centered on the MMN peak latency based on the difference waveforms of all participants and all target electrode locations (i.e., frontal, fronto-central and central topographical regions). The option for this time window was also based on previous studies (Duncan et al., 2009; Garrido, Kilner, Stephan, & Friston, 2009; Näätänen et al., 2007).

3.3.5 Statistical analysis

Since MMN is know to be largest at fronto-central sites (Duncan et al., 2009; Näätänen et al., 2001; Näätänen, 1990), a region of interest (ROI) analysis was performed using three different ROIs: frontal (Fz, F3, F4), fronto-central (FCz, FC3, FC4) and central (Cz, C3, C4). Mean amplitude and peak latency were analyzed separately using repeated-measures analysis of variance (ANOVA) including ROI (3 levels: frontal, fronto-central and central) as within-subjects factor and Expertise (2 levels: Musicians vs. Non-musicians) as between-subjects factor.

3.4 Results

The EEG results are illustrated in Figures 3 and 4. Figure 3 shows the grand average waveforms obtained for STDs and DEVs (panel A), the grand average difference waveforms (panel B) and the topographical maps (panel C) of the MMN effect in musicians and non-musicians. Figure 4 illustrates amplitude and latency differences between groups and across topographical regions. Table 4 reports means and standard error values of MMN amplitudes and latencies.

3.4.1 MMN amplitude

The amplitude analysis revealed a ROI effect, F (2, 80) = 10.461, $p < .001$, $\eta^2_p = .207$, showing increased (i.e., more negative) amplitude at frontal and fronto-central ROIs relative to the central ROI ($p = .001$, for both comparisons) for all participants (see Figs. 3B and 4B and Table 4). This analysis revealed no significant differences between musicians and non-musicians and no interaction effects ($p > .05$).

3.4.2 MMN latency

MMN tended to peak earlier in musicians compared to non-musicians even though the effect of Expertise did not reach statistical significance, F (1, 40) = 2.993; $p = .095$, $\eta^2_p = .068$ (see Figs. 3B and 4A and Table 4). There were no other significant main effects or interactions regarding MMN latency ($p > .05$).

TABLE 4.

Mean and Standard Error Values of MMN Amplitudes and Latencies Separately for Each ROI and Group at the Time Window of Interest (160-260 ms).

		Amplitude	Latency
		M (SE)	*M (SE)*
Musicians			
	Frontal	-3.153 (.421)	199.664 (4.652)
	Fronto-central	-2.898 (.434)	203.125 (4.974)
	Central	-2.355 (.419)	204.804 (5.165)
Non-Musicians			
	Frontal	-3.021 (.463)	214.136 (4.228)
	Fronto-central	-2.641 (.478)	214.447 (4.521)
	Central	-2.301 (.461)	211.362 (4.695)

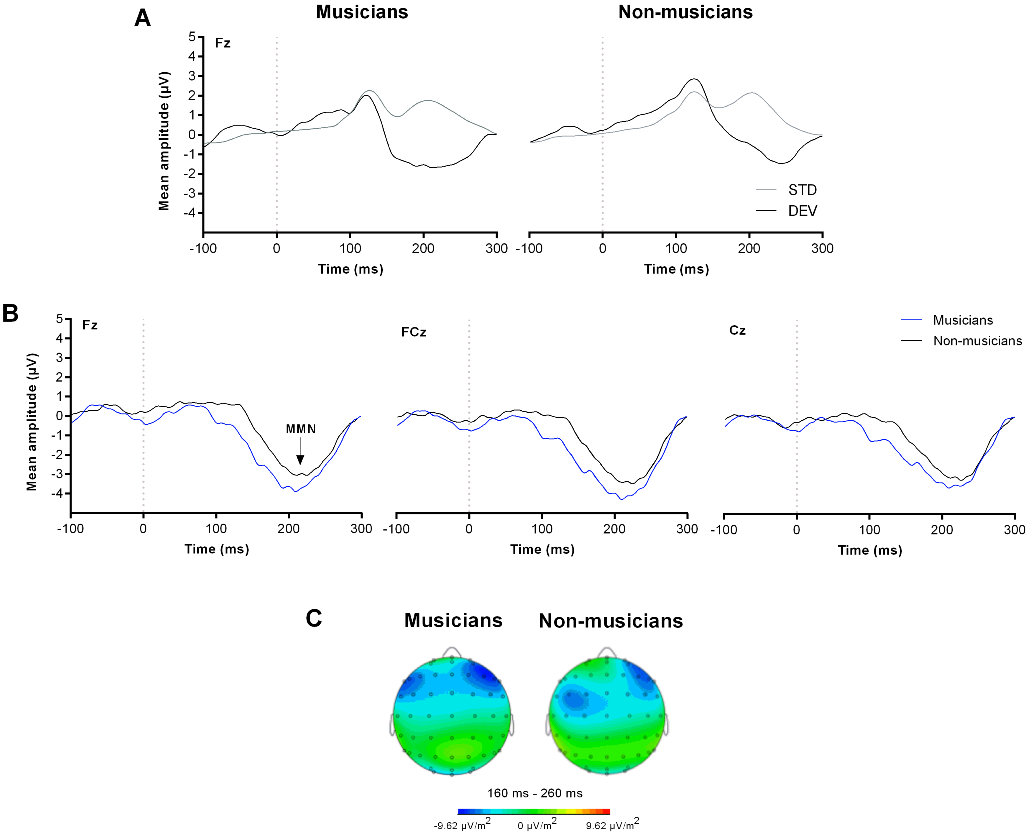

FIGURE 3. Grand-averaged waveforms obtained for STD/DEV tones and difference waveforms and topographical maps of the MMN effect. **(A)** Grand-averaged ERP waveforms obtained for STD (grey line) and DEV (black line) sounds in musicians (left) and non-musicians (right). **(B)** Grand-averaged difference waveforms

recorded in musicians (blue line) and non-musicians (black line) at frontal (Fz), fronto-central (FCz) and central (Cz) sites. **(C)** Topographical maps are illustrated for each group at the latency window (160-260 ms) used to extract mean amplitude and peak latency values.

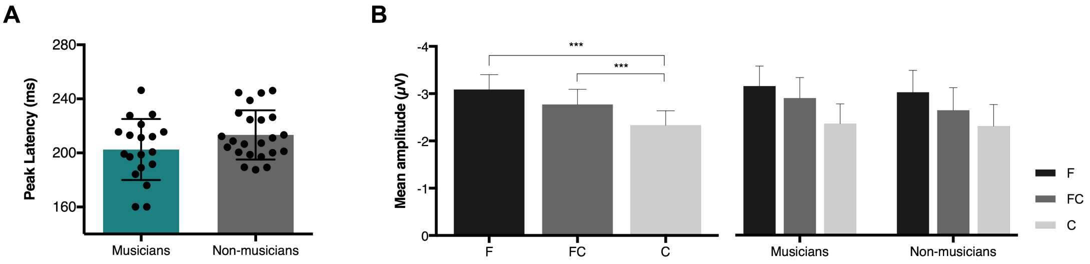

FIGURE 4. Means and Standard Deviations (SD) are plotted for latency **(A)** and amplitude **(B)** measures. **(A)** Individual MMN peak latencies (black circles) are displayed for musicians and non-musicians: the means are plotted per group in green (musicians) and grey (non-musicians) bars. **(B)** The bar plot on the left displays the means an SDs amplitudes at frontal (F, black), fronto-central (FC, dark grey) and central (C, light grey) regions; the amplitude was increased at F and FC regions relatively to the C region. The plots on the right illustrate means and SDs amplitudes for F, FC, and C regions separately for each group. * $p < .05$; ** $p < .01$; *** $p < .001$

3.5. Discussion

The current study probed the effects of musical expertise on the pre-attentive detection of pitch changes. Participants were expected to automatically track the probabilistic occurrence of two types of pure tones (1000 Hz vs. 12000 Hz) and to build a predictive model of the auditory environment while attending to an unrelated, concurrent task. Deviance processing elicited a MMN response in both groups, distributed mainly over frontal and fronto-central scalp regions, which was maximal between approximately 160 to 260 ms after sound onset. The MMN response of musicians and non-musicians did not differ significantly in amplitude or latency, despite a tendency for musicians to exhibit an early MMN response. These findings indicate that the magnitude of the electrical brain response to pitch deviance is not affected by musical expertise but the speed at which this response emerges might be.

The MMN is elicited when the incoming input deviates from a sensory memory template that was created based on environmental regularities. Here, the regularities were established based on the invariance of sounds' features (e.g., duration, intensity) and the MMN was elicited in response to the irregular occurrence of pitch changes (1200 Hz sounds). MMN was elicited in both groups, which indicates that participants were able to establish representations of the statistical structure of the auditory stream based purely on its acoustical nature (e.g., frequency information), even though their attention was focused on concurrent visual stimuli. Detecting pitch violations is critical as changes in this acoustic feature are at the core of both music and speech processing (Patel, 2003).

However, the current findings indicate that musical training does not translate into a facilitated pre-attentive detection of pitch changes. This finding agrees with previous studies showing that musicians do not differ from non-musicians in the elicitation of MMN responses to changes in physical features (e.g., intensity, pitch or duration) of the sounds (see Fujioka et al., 2004; Nikjeh et al., 2008, 2009; Paraskevopoulos, Kuchenbuch, Herholz, & Pantev, 2012; Tervaniemi et al., 2005). The studies of Tervaniemi and colleagues (Tervaniemi et al., 2005), Fujioka and colleagues (Fujioka et al., 2004) and Nikjeh and colleagues (Nikjeh et al., 2008) found no differences between musical experts and non-experts in auditory discrimination tasks that relied on simple pitch changes. Fujioka and collaborators (Fujioka et al., 2004) also investigated whether musical training influenced the automatic discrimination of contour and interval features of melodies using MEG. The authors found a mMMN effect in both musicians and non-musicians that was significantly increased in the former both when testing contour and interval discrimination. These findings indicate that while musical training enhances the ability to automatically process abstract characteristics/rules governing melodic sequences, it does not affect deviance detection in pitch-modulated sequences. Tervaniemi and colleagues (Tervaniemi et al., 2005) tested pitch discrimination with oddball paradigms when attention was or was not directed to the sounds. Similarly to the current study, they found that the MMN elicited by DEV sounds in the unattended condition did not differ in amplitude as a function of musical expertise. Conversely, at a subsequent processing stage in which subjects' attention was automatically oriented towards the sounds, musicians exhibited enhanced N2b and P3 responses. Another study found musicians to exhibit an earlier P3a response to syllabic and harmonic sequences following the MMN, which suggests that the neural mechanisms engaged in shifting attention towards relevant changes in speech or harmonics is faster in musicians (Nikjeh et al., 2008). Our results are also corroborated by a MEG study (Paraskevopoulos et al., 2012) that found no differences between groups in the MMNm elicited to pure tone triplets. Nevertheless, the impact of musical expertise over regularity processing can be observed at later stages of processing when cognitive/top-down processes come into play (Tervaniemi et al., 2005).

The acoustic complexity of the stimuli may account for the apparently discrepant findings of studies that investigated pre-attentive responses to deviations in simple or more acoustically complex sounds as a function of musical expertise. Previously, musicians showed an enhanced abstract-feature MMN – a MMN elicited in response to multi-feature sounds, such as sounds presenting variations in a set of acoustic features (e.g., chords, Koelsch et al., 1999; melodies, Tervaniemi et al., 2001). This suggests that musical training may only represent an advantage in change detection when stimuli are acoustically complex in nature. When changes in auditory stimulation rely on changes on a unique acoustic feature (e.g., only pitch modulation) as opposed to changes in multiple acoustic features, or when they require abstract rule processing, musical expertise does not make a difference in the brain's response to task-irrelevant auditory irregularities. We argue that while non-musicians and musicians show a similar sensitivity to process deviations in simple pure tones, musical training may bring advantages to the detection of more abstract deviations in auditory patterns.

We should also note that, in the present study, the magnitude of the difference between tones was considerable: DEVs were increased in 200 Hz relative to STDs, which is a 20% frequency difference. Musicians and non-musicians are equally accurate in behaviorally detecting 4% pitch deviances (musicians are only faster and more accurate responding to pitch changes with 0.8% and 2% of magnitude) (Tervaniemi et al., 2005). The fact that the discrepancy between sounds was considerable, even considering that participant's attention was focused on a concurrent task, might also have contributed to the similar brain responses found between groups.

We found a trend towards an earlier MMN response in musicians. Previous studies have found that musicians show earlier pre-attentive change detection in response to pure tones and speech syllables (Nikjeh et al., 2008), in harmonic tone complexes (Tervaniemi et al. 2005; Nikjeh et al., 2009) and in pitch alterations in familiar vs. unfamiliar interval patterns (Brattico, Näätänen, & Tervaniemi, 2001). The existing studies suggest that the higher the difference between the STD and DEV sounds, the shorter the MMN latency (Brattico et al., 2001; Koelsch et al., 1999; Novitski et al., 2004). In the case of the current study, there were no multiple DEVs and so it was not possible to compare latencies in response to distinct differences between STDs and DEVs. Nevertheless, the trend for earlier MMN responses in musicians goes along with the findings of previous studies and might suggest that musicians are faster in the detection of pitch deviances.

It is relevant to mention that, at least in one study, an earlier MMN effect to pure tones (Nikjeh et al., 2008) occurred only in the musicians who were both vocalists and instrumentalists; the MMN response of instrumentalists did not differ from non-musicians. This finding highlights the potential impact of musicians' technical background on the MMN response to pitch DEVs. Previous research has shown that instrumental musicians encode acoustic parameters in a different way depending on their musical genre (Nager et al., 2003; Pantev et al., 2003; Seppänen, Brattico, & Tervaniemi, 2007), type of training (Pantev et al., 2003), or training strategies (Seppänen et al., 2007). In the present study, musicians were all instrumentalists, with expertise in different instruments, with some of them being particularly efficient in the processing of frequency information (e.g., the musicians who have to attune their own instrument). Future studies should explore whether the musical background of musicians (e.g., type of instrument played, type of training, etc.) contributes to brain changes in the processing of acoustic deviance. Also, future research should examine auditory discrimination by testing other type of sounds, along with pitch manipulations. Synthetic timbres that lack a pre-existing hierarchical organization (see, for example, Koelsch et al., 2016) might bring important insights to the study of auditory perception and integration and extend the literature on the effects of musical training on these processes (Chomsky, 1957; Lerdahl & Jackendoff, 1983; Rohrmeier, 2011).

3.6 Conclusion

In summary, electrophysiological measures were used to investigate automatic (pre-attentive) pitch change detection in musicians *vs.* non-musicians. All participants evidenced a MMN response to pitch deviance, which represents a neurophysiological index of the efficient learning of statistical regularities. Musicians did not show enhanced pre-attentive processing of pitch compared with non-musicians, despite musicians showing a tendency for a faster MMN response, which is in agreement with early studies. These findings suggest that musical expertise does not influence the sensitivity of the brain's pre-attentive response to basic regularities processing when they are based on the manipulations of a single acoustic feature.

3.7 References

Alho, K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNM) elicited by sound changes. *Ear and Hearing*, *16*(1), 38–51. http://doi.org/10.1097/00003446-199502000-00004

Alho, Tervaniemi, M., Huotilainen, M., Lavikainen, J., Tiitinen, H., Ilmoniemi, R. J., ...

Näätänen, R. (1996). Processing of complex sounds in the human auditory cortex as revealed by magnetic brain responses. *Psychophysiology*. http://doi.org/10.1111/j.1469-8986.1996.tb01061.x

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8 month old infants. *Psychological Science*, *9*(4), 321–324. http://doi.org/http://doi.org/10.1111/1467-9280.00063

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia. A group study of adults afflicted with a music-specific disorder. *Brain*, *125*(2), 238–251. http://doi.org/10.1093/brain/awf028

Bendixen, A., Schroger, E., & Winkler, I. (2009). I Heard That Coming: Event-Related Potential Evidence for Stimulus-Driven Prediction in the Auditory System. *Journal of Neuroscience*, *29*(26), 8447–8451. http://doi.org/10.1523/JNEUROSCI.1493-09.2009

Besson, M., Faïta, F., & Requin, J. (1994). Brain waves associated with musical incongruities differ for musicians and non-musicians. *Neuroscience Letters*, *168*(1–2), 101–105. http://doi.org/10.1016/0304-3940(94)90426-X

Bidelman, G. M., & Alain, C. (2015). Musical Training Orchestrates Coordinated Neuroplasticity in Auditory Brainstem and Cortex to Counteract Age-Related Declines in Categorical Vowel Perception. *Journal of Neuroscience*, *35*(3), 1240–1249. http://doi.org/10.1523/JNEUROSCI.3292-14.2015

Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer. Amsterdam: Institute of Phonetic Sciences/University of Amsterdam.

Brattico, E., Näääätäänen, R., & Tervaniemi, M. (2001). Context effects on pitch perception in musicians and nonmusicians: Evidence from event-related-potential recordings. *Music Perception: An Interdisciplinary Journal*, *19*(2), 199–222. http://doi.org/10.1525/mp.2001.19.2.199

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge: MIT Press.

Chomsky, N. (1957). *Syntactic structures.* Berlin: Walter de Gruyter, GmbH & Co. KG.

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., ... Van Petten, C. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, *120*(11), 1883–1908. http://doi.org/10.1016/j.clinph.2009.07.045

Escera, C., Corral, M. J., & Yago, E. (2002). An electrophysiological and behavioral investigation of involuntary attention towards auditory frequency, duration and intensity changes. *Cognitive Brain Research*, *14*(3), 325–332. http://doi.org/10.1016/S0926-6410(02)00135-0

François, C., Jaillet, F., Takerkart, S., & Schön, D. (2014). Faster sound stream segmentation in musicians than in nonmusicians. *PLoS ONE*, *9*(7), 26–29. http://doi.org/10.1371/journal.pone.0101340

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience*, *16*(6), 1010–1021. http://doi.org/10.1162/0898929041502706

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2005). Automatic encoding of melodies in musicians and non musicians. *Journal of Cognitive Neuroscience*, *Vol. 17*(10), 1578–1592. http://doi.org/10.1162/089892905774597263

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology*, *120*(3), 453–463. http://doi.org/10.1016/j.clinph.2008.11.029

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *The Journal of Neuroscience*, *23*(27), 9240–9245. http://doi.org/23/27/9240 [pii]

Geiser, E., Sandmann, P., Jäncke, L., & Meyer, M. (2010). Refinement of metre perception - training increases hierarchical metre processing. *European Journal of Neuroscience*, *32*(11), 1979–1985. http://doi.org/10.1111/j.1460-9568.2010.07462.x

Gelfand, S. (1998). *Hearing: An introduction to psychological and phys- iological acoustics.* New York: Marcel Dekker, Inc.

Gordon, E. E. (1989). *Advanced measures of music audiation.* Chicago, IL: G.I.A. Publications.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468–484.

Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., & Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology*, *42*(2), 180–190. http://doi.org/10.1111/j.1469-8986.2005.00272.x

Herholz, S. C., Lappe, C., Knief, A., & Pantev, C. (2009). Imagery mismatch negativity in musicians. *Annals of the New York Academy of Sciences*, *1169*, 173–177. http://doi.org/10.1111/j.1749-6632.2009.04782.x

Herholz, S. C., Lappe, C., & Pantev, C. (2009). Looking for a pattern: An MEG study on the abstract mismatch negativity in musicians and nonmusicians. *BMC Neuroscience*, *10*, 1–

10. http://doi.org/10.1186/1471-2202-10-42

Hevner, K. (1937). The affective value of pitch and tempo in music. *The American Journal of Psychology*, *49*(4), 621–630.

Jasper, H. H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 370–375.

Kishon-Rabin, L., Amir, O., Vexler, Y., & Zaltz, Y. (2001). Pitch discrimination: Are professional musicians better than non-musicians? *Journal of Basic and Clinical Physiology and Pharmacology*, *12*(2), 125–144. http://doi.org/10.1515/JBCPP.2001.12.2.125

Koelsch, S., Busch, T., Jentschke, S., & Rohrmeier, M. (2016). Under the hood of statistical learning: A statistical MMN reflects the magnitude of transitional probabilities in auditory sequences. *Scientific Reports*, *6*, 1–11. http://doi.org/10.1038/srep19741

Koelsch, S., Schmidt, B. H., & Kansok, J. (2002). Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. *Psychophysiology*, *39*(5), 657–663. http://doi.org/10.1017/S0048577202010508

Koelsch, S., Schröger, E., & Tervaniemi, M. (1999). Superior pre-attentive auditory processing in musicians. *NeuroReport*, *10*(6), 1309–1313. http://doi.org/10.1097/00001756-199904260-00029

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Science and Society*, *11*(August), 599–606. http://doi.org/10.1038/nrn2882

Lerdahl, F., & Jackendoff, R. S. (1983). *A Generative Theory of Tonal Music.* MIT Press. http://doi.org/10.1525/mts.1985.7.1.02a00120

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique* (Second, Vol. 78). Cambridge, MA: MIT Press. http://doi.org/10.1118/1.4736938

Mizuno, C., Schwartz, J. L., & Cazals, Y. (1994). Periodicity of long-term context can influence gap detection. *Hearing Research*, *78*(1), 41–48. http://doi.org/10.1016/0378-5955(94)90042-6

Münte, T. F., Kohlmetz, C., Nager, W., & Altenmüller, E. (2001). Superior auditory spatial tuning in conductors. *Nature*, *409*(6820), 580. http://doi.org/10.1038/35054668

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*, *13*(2), 201–233. http://doi.org/10.1017/S0140525X00078407

Näätänen, R. (2002). The perception of speech sounds by the human brain as reflected by the mismatch negativity brain response. *International Congress Series*, *1232*, 97–105.

http://doi.org/10.1016/S0531-5131(01)00795-6

Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, *42*(4), 313–329. http://doi.org/http://doi.org/10.1016/0001-6918(78)90006-9

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*, *118*(12), 2544–2590. http://doi.org/10.1016/j.clinph.2007.04.026

Näätänen, R., Simpson, M., & Loveless, N. E. (1982). Stimulus deviance and evoked potentials. *Biological Psychology*, *14*(1–2), 53–98. http://doi.org/10.1016/0301-0511(82)90017-5

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001a). “Primitive intelligence” in the auditory cortex. *Trends in Neurosciences*, *24*(5), 283–288. http://doi.org/10.1016/S0166-2236(00)01790-2

Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001b). “Primitive intelligence” in the auditory cortex. *Trends in Neurosciences*, *24*(5), 283–288. http://doi.org/10.1016/S0166-2236(00)01790-2

Nager, W., Kohlmetz, C., Altenmüller, E., Rodriguez-Fornells, A., & Münte, T. F. (2003). The fate of sounds in conductors’ brains: An ERP study. *Cognitive Brain Research*, *17*(1), 83–93. http://doi.org/10.1016/S0926-6410(03)00083-1

Nasreddine, Z., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., & Collin, I. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *American Geriatrics Society*, *53*(4), 695–699. http://doi.org/10.1111/j.1532-5415.2005.53221.x

Nikjeh, D. A., Lister, J. J., & Frisch, S. A. (2008). Hearing of note: An electrophysiologic and psychoacoustic comparison of pitch discrimination between vocal and instrumental musicians. *Psychophysiology*, *45*(6), 994–1007. http://doi.org/10.1111/j.1469-8986.2008.00689.x

Nikjeh, D. A., Lister, J. J., & Frisch, S. A. (2009). Preattentive cortical-evoked responses to pure tones, harmonic tones, and speech: Influence of music training. *Ear and Hearing*, *30*(4), 432–446. http://doi.org/10.1097/AUD.0b013e3181a61bf2

Nordby, H., Roth, W. T., & Pfefferbaum, A. (1998). Event-related potentials to breaks in sequences of alternating pitches or interstimulus intervals. *Psychophysiology*, *25*, 262 – 268.

Novitski, N., Tervaniemi, M., Huotilainen, M., & Näätänen, R. (2004). Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures. *Cognitive Brain Research*, *20*(1), 26–36. http://doi.org/10.1016/j.cogbrainres.2003.12.011

Nyman, G., Alho, K., Laurinen, P., Paavilainen, P., Radil, T., Reinikainen, K., … Näätänen. (1990). Mismatch negativity (MMN) for sequences of auditory and visual stimuli: evidence for a mechanism specific to the auditory modality. *Electroencephalography and Clinical Neurophysiology*, *77*(6), 436–444. http://doi.org/10.1016/0168-5597(90)90004-w

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. http://doi.org/10.1016/0028-3932(71)90067-4

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, *392*(6678), 811–814. http://doi.org/10.1038/33918

Pantev, C., Ross, B., Fujioka, T., Trainor, L., Schulte, M., & Schulz, M. (2003). Music and learning-induced cortical plasticity. *Annals of the New York Academy of Sciences*, *999*, 438–450.

Paraskevopoulos, E., Kuchenbuch, A., Herholz, S. C., & Pantev, C. (2012). Statistical learning effects in musicians and non-musicians: An MEG study. *Neuropsychologia*, *50*(2), 341–349. http://doi.org/10.1016/j.neuropsychologia.2011.12.007

Patel, A. D. (2003). Music, language, syntax and the brain. *Nature Neuroscience*, *6*(7), 674–681.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233–238. http://doi.org/10.1016/j.tics.2006.03.006

Picton, T. W., Alain, C., Otten, L., Ritter, W., & Achim, A. (2000). Mismatch Negativity: Different Water in the Same River. *Audiology & Neuro-Otology*, *5*, 111–139. Retrieved from http://www.scielo.cl/scielo.php?pid=S0718-48162008000200011&script=sci_arttext

Ritter, W., Paavilainen, P., Lavikainen, J., Reinikainen, K., Alho, K., Sams, M., & Näätänen, R. (1992). Event-related potentials to repetition and change of auditory stimuli, *83*, 306–321.

Rohrmeier, M. (2011). Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, *5*(1), 35–53. http://doi.org/10.1080/17459737.2011.573676

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. http://doi.org/10.1126/science.274.5294.1926

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*(4), 606–621. http://doi.org/10.1006/jmla.1996.0032

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, *41*(3), 341–349. http://doi.org/10.1111/1469-8986.00172.x

Seppänen, M., Brattico, E., & Tervaniemi, M. (2007). Practice strategies of musicians modulate neural processing and the learning of sound-patterns. *Neurobiology of Learning and Memory*, *87*(2), 236–247. http://doi.org/10.1016/j.nlm.2006.08.011

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, *81*, 105–120. http://doi.org/10.1016/j.jml.2015.02.001

Simões, M. R., Freitas, S., Santana, I., Firmino, H., Martins, C., Nasreddine, Z., & Vilar, M. (2008). Montreal cognitive assessment (MoCA): versão final portuguesa. *Coimbra: Serviço de Avaliação Psicológica, Faculdade de Psicologia e de Ciências Da Educação Da Universidade de Coimbra.*

Sundberg, J. (1994). Perceptual aspects of singing. *Journal of Voice*, *8*(2), 106–122. http://doi.org/10.1016/S0892-1997(05)80303-0

Sussman, E., Ritter, W., & Vaughan, H. G. J. (1999). An investigation of the auditory streaming effect using event-related brain potentials. *Psychophysiology*, *36*(1), 22–34. http://doi.org/10.1017/S0048577299971056

Tang, W., Xiong, W., Zhang, Y. xuan, Dong, Q., & Nan, Y. (2016). Musical experience facilitates lexical tone processing among Mandarin speakers: Behavioral and neural evidence. *Neuropsychologia*, *91*, 247–253. http://doi.org/10.1016/j.neuropsychologia.2016.08.003

Tervaniemi, M., & Huotilainen, M. (2003). The promises of change-related brain potentials in cognitive neuroscience of music. *Annals of the New York Academy of Sciences*, *999*(1), 29–39.

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: An event-related potential and behavioral study. *Experimental Brain Research*, *161*(1), 1–10. http://doi.org/10.1007/s00221-004-2044-

5

Tervaniemi, M., Rytkönen, M., Schröger, E., Ilmoniemi, R. J., & Näätänen, R. (2001). Superior formation of cortical memory traces for melodic patterns in musicians. *Learning and Memory*, *8*(5), 295–300. http://doi.org/10.1101/lm.39501

Trainor, L. J., McDonald, K. L., & Alain, C. (2002). Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. *Journal of Cognitive Neuroscience*, *14*(3), 430–442. http://doi.org/10.1162/089892902317361949

Winkler, I., Denham, S. L., & Nelken, I. (2009). Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences*, *13*(12), 532–540. http://doi.org/10.1016/j.tics.2009.09.003

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*(4), 420–422. http://doi.org/10.1038/nn1872

Zuijen, T. L. van, Sussman, E., Winkler, I., Näätänen, R., & Tervaniemi, M. (2004). Grouping of Sequential Sounds—An Event-Related Potential Study Comparing Musicians and Nonmusicians. *Journal of Cognitive Neuroscience*, *16*(2), 331–338. http://doi.org/10.1162/089892904322984607

Chapter 4. Musical training modulates ASL as a function of stimulus type: combined behavioral and EEG insights.

4.1 Abstract

The current study probed whether and how musical training is associated with enhanced SL of distinct types of auditory streams. Implicit (i.e., ERPs of the EEG) and explicit (i.e., behavioral responses) measures were combined to examine SL *per se* and its outcomes. Learning outcomes considered both the online neural responses to violations in the stream's statistical structure and the offline behavioral responses to recognition tasks. ERPs were recorded from 18 musicians and 22 non-musicians in three experiments examining the processing of statistical regularities in prosodic (i.e., words with pitch contour), non-prosodic (i.e., words with flat contour) and musical streams. Compared to non-musicians, musicians showed an enhanced negativity in the 250-300 ms latency window when learning prosodic words and an enhanced positivity in the first 100 ms post-stimulus onset when learning melodies. Only structural violations to prosodic words elicited different ERP responses between groups. Behaviorally, only musicians demonstrated successful learning of the items contained in the auditory streams. Overall, these findings suggest that the neurofunctional mechanisms underlying ASL are modulated both by the musical background of participants and by stimulus type. Additionally, they provide evidence for cross-domain neuroplasticity effects of training, shedding light on the processes underling positive transfer effects from music to speech.

Keywords: musical training, ASL, speech segmentation, stimulus type, ERPs

4.2 Introduction

Musical training has been considered a model of neuroplasticity (Bidelman & Alain, 2015; Hyde et al., 2009; Kraus & Chandrasekaran, 2010; Moreno et al., 2009; Münte, Altenmüller, & Jäncke, 2002; Strait & Kraus, 2011; Vaquero, Ramos-Escobar, François, Penhune, & Rodríguez-Fornells, 2018). Long-term musical training was associated with structural and functional changes in the brain that are mostly related to advantages in auditory processing: increased gray matter concentration in the auditory cortex (Bermudez & Zatorre, 2005), increased volume of the planum temporale (Schlaug, Norton, Overy, & Winner, 2005) and increased gray matter volume in Broca's area (Sluming et al., 2002). Musical expertise was also found to affect both structure and function of brain regions involved in phonological processing, such as the superior temporal sulcus (Hickok, 2012; Hickok & Poeppel, 2007) and the inferior frontal gyrus (Gelfand & Bookheimer, 2003). These changes may underlie behavioral differences between musicians and non-musicians in the processing of acoustic features such as frequency (pitch) and tempo (Ehrlé & Samson, 2005; Kishon-Rabin, Amir, Vexler, & Zaltz, 2001; Micheyl, Delhommeau, Perrot, & Oxenham, 2006). At the electrophysiological level, musical expertise has been shown to result in increased amplitude of the N1 and P2 ERP components in response to auditory stimuli such as melodies (Shahin, Bosnyak, Trainor, & Roberts, 2003) and speech (Cunillera et al., 2009; Cunillera, Toro, Sebastián-Gallés, & Rodríguez-Fornells, 2006; de Diego-Balaguer, Toro, Rodriguez-Fornells, & Bachoud-Lévi, 2007), which reflects an enhanced processing of the acoustic features of sounds.

4.2.1 Transfer effects between music and speech

The existing studies have revealed positive cross-domain links between musical and speech skills (Besson, Chobert, & Marie, 2011; Moreno et al., 2009; Wong, Skoe, Russo, Dees, & Kraus, 2007) both in adults (Milovanov, Huotilainen, Välimäki, Esquef, & Tervaniemi, 2008; Schön, Magne, & Besson, 2004) and children (Magne, Schön, & Besson, 2006; Moreno et al., 2009). Specifically, musical training was associated with a facilitated encoding of the acoustic features of speech (e.g., duration, periodicity) (Besson et al., 2011; Strait & Kraus, 2011). A facilitated processing of pitch patterns in speech has also been found in musicians, reflected in a more robust encoding of speech stimuli at the brainstem level (as shown by the f_0 amplitude of FFRs to prosodic stimuli, see Wong et al., 2007), and enhanced discrimination of pitch incongruities (Magne et al., 2006; Moreno et al., 2009; Schön et al., 2004; Wong et al., 2007). Additionally, there is evidence that musical training influences the earliest stages of emotional prosody processing (e.g., reduced amplitude of the P50 ERP component in musicians compared to non-musicians, see Pinheiro, Vasconcelos, Dias, Arrais, & Gonçalves, 2015), being also associated with increased accuracy in the recognition of vocal emotions (Lima & Castro, 2011; Thompson, Schellenberg, & Husain, 2004). Longitudinal studies with random assignment of children to musical programs found that speech segmentation – the ability to parse out a speech stream into its constituent units (which is critical in language development, see Aslin, Saffran, & Newport, 1998; Saffran, Newport, & Aslin, 1996) - was enhanced in musician children (François, Chobert, Besson, & Schön, 2013), who also showed enhanced word learning (Dittinger, Chobert, Ziegler, & Besson, 2017) and increased preattentive processing of duration and voice onset time manipulations (Chobert, François, Velay, & Besson, 2012). These findings suggest that musical training may benefit the processing of speech at various levels by enhancing low-level acoustic discrimination abilities (Strait & Kraus, 2011), by improving speech prosody perception and recognition (Lima & Castro, 2011; Pinheiro et al., 2015), and by possibly enhancing the detection of acoustic regularities in linguistic input (François et al., 2013; François & Schön, 2014).

4.2.2 SL as a key aspect of auditory learning

Considering the evidence supporting the strong link between music and speech (Besson et al., 2011; Dittinger et al., 2016; Elmer & Jäncke, 2018; Moreno et al., 2009; Patel, 1998, 2012; Patel & Morgan, 2017; Strait & Kraus, 2011), it is critical to clarify whether and how musical training facilitates the cognitive processes involved in speech processing. One of such candidate processes is the ability to detect statistical regularities in auditory input that is critical for the effective expression and comprehension of communication signals (Newport, 2016; Saffran, Newport, et al., 1996; Saffran, Senghas, & Trueswell, 2001). One of the mechanisms responsible for the efficient segmentation of continuous streams of speech is SL. SL relies on the analysis of the distributional cues present in streams of information in order to extract input regularities (Newport, 2016; Perruchet & Pacton, 2006; Turk-Browne, Jungé, &

Scholl, 2005). Studies of SL in the context of language research date back to 1996 when Saffran and collaborators (Saffran, Aslin, et al., 1996) found that infants were able to segment continuous streams of pseudo-words (hereafter words[5]) based on the TPs (i.e., the probability of X given the occurrence of Y) between syllable pairs. The role of TPs in the segmentation of artificial speech streams has been demonstrated in several studies with children and adults (Aslin et al., 1998; Kuhl, 2004; Saffran, Aslin, et al., 1996; Saffran, Newport, et al., 1996). The study by Saffran and collaborators (Saffran, Aslin, et al., 1996) provided the first evidence indicating that the segmentation of artificial speech is possible based on its statistical structure exclusively.

Recently, enhanced SL skills were proposed to represent the mechanism accounting for the facilitated speech processing in musicians (François et al., 2013; François, Jaillet, Takerkart, & Schön, 2014; Francois & Schön, 2011). Musical training imvolves the implicit learning of musical structures such as rhythmic patterns or chord progressions (Jonaitis & Saffran, 2009; Schultz, Stevens, Keller, & Tillmann, 2013). The ability to implicitly learn musical structures is critical for the prediction of incoming acoustic input and for its accurate segmentation (Loui, Wessel, & Kam, 2010; Saffran, Johnson, Aslin, & Newport, 1999). Further, music and speech have both a complex and hierarchically organized structure: While music is composed of variations in chord progressions occurring at specific moments of a melody, speech often entails the use of clauses that may relate to each other in different ways and that unfold over time (Patel, 2003; Patel, Gibson, Ratner, Besson, & Holcomb, 1998). Additionally, musicians were found to excel non-musicians in the learning of a novel Morse-code language (Shook, Marian, Bartolotti, & Schroeder, 2013), in the detection of pitch and melodic variations in a foreign language (Marie, Delogu, Lampis, Belardinelli, & Besson, 2011) and in foreign language learning (Chobert & Besson, 2013; Dittinger, Valizadeh, Jäncke, Besson, & Elmer, 2018). Based on this evidence, one could argue that musical training influences positively the processing of speech structure.

4.2.3 Detecting auditory regularities – insights from ERPs

The ERP methodology is a well-suited technique to investigate the time course of SL with millisecond precision. Studies investigating the SL of speech sequences using ERPs have

---

[5] In the context of the current work, ‘words’ refer to the constituent units of artificial speech streams.

been somewhat consistent in revealing that amplitude modulations of the N1, P2 and N400-like components represent learning indices. The auditory N1 and the P2 represent early sensory components typically elicited by any perceptual auditory stimulus (Beres, 2017). N1 and P2 are conceptualized as sensory gating mechanisms (Boutros, Korzyukov, Jansen, Feingold, & Bell, 2004; Lijffijt et al., 2009) with a distinct functional significance. The N1 reflects the processing of a stimulus' sensory features (Luck, 2005) and may also reflect predictive mechanisms involved in speech processing (Heinks-Maldonado, Mathalon, Gray, & Ford, 2005). The P2 indexes processes related to auditory discrimination and early stimulus categorization (Kramer & Donchin, 1987; Ritter et al., 1992; Sheehan, McArthur, & Bishop, 2005) or perceptual changes related to the computation of statistical regularities (de Diego-Balaguer et al., 2007; Snyder, Alain, & Picton, 2006). The auditory N1 peaks at approximately 100 ms (Abla, Katahira, & Okanoya, 2008; Sanders, Newport, & Neville, 2002; Vasuki, Sharma, Ibrahim, & Arciuli, 2017a), whereas the P2 component peaks approximately at 200 ms (Abla et al., 2008; Abla & Okanoya, 2009; Cunillera et al., 2009, 2006; de Diego-Balaguer, Rodríguez-Fornells, & Bachoud-Lévi, 2015; Reinke, He, Wang, & Alain, 2003; Vasuki, Sharma, Ibrahim, & Arciuli, 2017b), both with a fronto-central topographical distribution. These components are not exclusively elicited in speech paradigms; nevertheless, as they index the sensory processing of a stimulus, they are frequently found in response to speech stimuli (Tremblay, Kraus, & McGee, 1998). Additionally, N400-like amplitude modulations are a consistent finding in artificial language learning research and have been proposed to reflect processes related to lexical search that emerge in response to the successful segmentation of a stream into its units (Abla, Katahira, & Okanoya, 2008; Cunillera et al., 2009, 2006; de Diego-Balaguer, Toro, Rodriguez-Fornells, & Bachoud-Levi, 2007; François, Cunillera, Garcia, Laine, & Rodriguez-Fornells, 2017a; François et al., 2014; Vasuki, Sharma, Ibrahim, & Arciuli, 2017b; Sanders, Newport, & Neville, 2002). The N400-like component elicited in artificial language learning tasks typically emerges between 200 and 500 ms and is more pronounced over fronto-central electrodes (Cunillera et al., 2009; Dittinger et al., 2017; François, Cunillera, Garcia, Laine, & Rodriguez-Fornells, 2017; François et al., 2014; Sanders et al., 2002; Vasuki et al., 2017b).

Generally, SL studies involve a learning phase (e.g., when participants are familiarized with a stimulus stream whose structure is unknown to them) and a test phase (e.g., the time when the segmentation/learning of stimulus sequences is tested). To the best of our knowledge, François and colleagues were the first to investigate the influence of musical expertise on speech segmentation and to examine the neurophysiological responses during both learning and test phases (François et al., 2014; Francois & Schön, 2011). In a first study (Francois & Schön, 2011), the authors used an artificial language paradigm with sung words. The results of the test phase revealed increased P2 amplitude and an increased negativity at approximately 450 ms for unfamiliar items (i.e., new words) in musicians. In a latter study (François et al., 2014), the authors used the same type of stimuli but examined ERP responses during the learning phase. The learning phase was analyzed in four consecutive blocks to define learning curves reflecting the brain dynamics during learning. While non-musicians revealed a linear increase in the N400-like component, musicians showed an inverted U-shaped learning curve (e.g., an increase in the N400-like component until the third block, followed by a decrease). The authors took these results as evidence that musical training leads to faster speech segmentation. It is important to highlight that the materials used in these studies were nonsense sung words. Sung words are words with pitch modulations whose characteristics resemble the prosody that is specific of human speech. The pitch contours of these stimuli might benefit language learning by at least three mechanisms (Schön et al., 2008): (1) by increasing attention to the sounds; (2) by emphasizing phonological boundaries which promotes phonological discrimination, and; (3) by globally enhancing the computation of TPs through the parallel mapping of musical and linguistic information. A study from Vasuki and colleagues (Vasuki et al., 2017b) also compared SL in musicians and non-musicians using both auditory (pure tones) and visual (cartoon-like figures) streams. The N1 and N400-like components elicited in response to tone triplets were modulated by musical training and were associated with increased accuracy in post-learning behavioral ASL tasks. Nevertheless, musicians did not differ from non-musicians in visual SL. These results suggest that the benefits of musical training in SL tasks are specific of the auditory modality. Nonetheless, the above-mentioned studies do not clarify whether musical training benefits the processing of the statistical structure of speech input. First, speech segmentation was tested with sung materials that, as mentioned above, facilitate phonological discrimination and the computation of TPs. Second, these speech stimuli contain melodic variations whose processing previous research has shown to benefit from musical expertise (Francois & Schön, 2011). Therefore, does musical training facilitate speech segmentation if stimuli are devoid of prosodic modulations? This question warrants clarification in order to understand if previous claims of enhanced speech segmentation skills in musicians are due to a positive and broad transfer effect between music and speech (i.e., cross-domain neuroplasticity), or if they arise from a specific advantage in the processing of auditory streams with melodic features, such as sung languages (i.e., within-domain neuroplasticity). A direct comparison of the neurofunctional mechanisms underpinning how speech streams with or without prosodic modulations vs. pure musical streams are learned could dissociate the contributions of speech vs. melodic information to speech segmentation and clarify whether musical training benefits ASL irrespective of stimulus type.

Studies testing the impact of musical training on SL are scarce and, so far, most studies have evaluated learners through offline tests such 2AFC tasks (Batterink & Paller, 2017; Cunillera et al., 2009; François et al., 2013) or LDTs (François et al., 2017). With the exception of few reports (Cunillera et al., 2009, 2006; François et al., 2017; McNealy, Mazziotta, & Dapretto, 2010; Turk-Browne, Scholl, Johnson, & Chun, 2010), the existing studies have addressed SL using the same approach: experiments start by exposing participants to a structured stream and finish with post-learning offline tasks evaluating the participants' capacity to recognize the streams' items. By using this approach, studies are limited in their capacity to empirically examine the time course of online SL. Moreover, these studies have examined learning outcomes based on behavioral, post-learning tasks (Vasuki et al., 2017; François et al., 2014) that only evaluate the participants' ability to retrieve encoded representations of the streams' units from long-term memory. These tasks do not allow examining what happens at the brain level when participants are exposed both to items that conform to the statistical structure of the streams and to items that violate that structure. To the best of our knowledge, only a few studies have done this comparison before (de Diego-Balaguer et al., 2007; François et al., 2017) by introducing violations to the statistical structure of the streams in an intermediate test phase in between the learning phase and the offline test. During this intermediate phase (hereafter called 'implicit test phase'), illegal non-words were intermixed with legal words and elicit MMN/N200 components. Specifically, when compared to legal words, illegal words elicited larger negativities at the first and third syllabic positions (François et al., 2017). The MMN is an automatic, pre-attentive component elicited in response to physical or abstract changes in an otherwise repetitive sequence of sounds (Näätänen, 2002; Näätänen, Paavilainen, Rinne, & Alho, 2007). It reflects deviance detection mechanisms (Näätänen, Jacobsen, & Winkler, 2005) in response to both simple and complex auditory stimulation (Boh, Herholz, Lappe, & Pantev, 2011; Chobert et al., 2012; Herholz, Lappe, & Pantev, 2009; Näätänen, Gaillard, & Mäntysalo, 1978). The N200 is superimposed on the MMN (Näätänen & Gaillard, 1983) and is thought to reflect implicit knowledge mechanisms (related to working memory processes) in response to mismatches between incoming input and previous memory templates (Sams, Alho, & Näätänen, 1983). In the field of AGL studies, ungrammatical sequences elicited larger N200 amplitudes when compared to grammatical sequences (Carrión & Bly, 2007; Selchenkova et al., 2014). The MMN/N200 components found in response to structural violations typically have a fronto-central topography, peaking approximately at 100-250 ms after stimulus onset (François et al., 2017).

Therefore, a combined examination of brain responses during exposure to the streams (e.g., learning phase) and the evaluation of the learning outcomes by means of both neural and behavioral responses (Daltrozzo & Conway, 2014) is lacking. Additionally, it remains to be clarified whether the effects of musical training on ASL are stimulus-specific or whether ASL is modulated by musical training irrespective of stimulus type, making it clear whether musicians benefit from within and/or cross-domain neuroplasticity. Moreover, the relationship between the cognitive mechanisms characterizing the learning phase and the higher-order evaluative processes characterizing the overt responses in post-learning tasks (Paulmann & Pell, 2010) requires further investigation.

4.2.4 The current study and hypotheses

The current study addresses the need for a systematic investigation of the effects of musical training on the SL of different types of auditory stimuli, i.e. speech vs. musical streams. For that purpose, a group of participants with vast formal musical training and a group of participants without musical training took part in three experiments in which the SL of sung words (prosodic experiment: Experiment 1), non-prosodic words (linguistic experiment: Experiment 2) and tri-tone melodies (musical experiment: Experiment 3) was compared. EEG data was collected during the learning and implicit test phases of each experiment. We expected N1-P2 (de Diego-Balaguer et al., 2007; Sanders et al., 2002) and N400-like ERP components to be elicited in response to words' and melodies' onsets (Cunillera et al., 2009; de Diego-Balaguer et al., 2007; François et al., 2017, 2014; Vasuki et al., 2017b). We hypothesized that musicians would exhibit larger (i.e., more negative) N400 amplitude relative to non-musicians, which would corroborate a positive effect of musical expertise on the online segmentation of auditory streams. Also, we expected to observe changes in N400-like amplitude with an increased exposure time, i.e. an inverted U-shape learning curve in musicians as opposed to a linear N400 increase in non-musicians (see François et al., 2014).

The implicit test phase was designed to examine the brain responses to structural violations in the streams' statistical structure. Thus, after the learning phase, an implicit test phase was introduced in which the streams' items with ABC forms (e.g., A – first syllable/tone; B – second; C – third) were pseudo-randomly intermixed with new items (e.g., CBA, BCA or CAB forms) that represented structural violations of the stream. With this manipulation, we evaluated two different aspects: first, we obtained a sensitive, online measure of the neural response to violations, which provides an additional measure of learning; second, we investigated to what extent musical training influences the neural mechanisms involved in the detection of deviations in the statistical structure of the auditory input (for a similar procedure, see de Diego-Balaguer et al., 2007; François et al., 2017). During the implicit test, structural violations were expected to elicit modulations of the MMN/N200 components (according to François et al., 2017). Illegal new-items (i.e., non-words/melodies) with CBA forms were expected to trigger MMN/N200 components at the first and third syllabic/tone positions, while items with BCA or CAB forms (i.e., part-words/melodies) were expected to elicit negative responses at the second and third syllabic/tone positions. When comparing musicians with non-musicians, we expected musicians to show increased MMN/N200 amplitude reflecting enhanced SL and facilitated detection of structural violations.

The behavioral outcomes were examined using post-learning LDT tasks (i.e., explicit test phase). Behaviorally, we expected musicians to be more accurate (i.e., higher number of correct responses) in the LDTs than non-musicians (Francois & Schön, 2011; Vasuki et al., 2017a, 2017b). In order to investigate if significant electrophysiological effects in the learning phase (that reflect relevant neurocognitive processes underpinning ASL) are associated with explicit measures of learning (i.e., SL performance on LDTs), correlations between neural and behavioral measures were tested (see for example, Bidelman, Gandour, & Krishnan, 2011). Our prediction was that individual differences in the ERP components elicited during the learning phase would be correlated with the performance in the explicit test phase.

4.3. Material and Methods

4.3.1 Participants

To ensure that our statistical analysis had sufficient power, we performed an *a priori* power analysis with G*Power software (Faul, Erdfelder, Buchner, & Lang, 2009). We chose the *F* tests family, specifically the repeated-measures ANOVA, with a between-subjects design, considering 5 measures and a nonsphericity correction (ε) criteria of 0.25. G*Power estimated a minimum required sample size of 24 participants, 12 per group, to detect a medium effect size of 0.5 at a significance level of α= 0.05 with power of 1 - β = 0.95.

A group of 18 musicians (10 women; 20.78 ± 2.88 years [*M* ± *SD*]) and 22 non-musicians (12 women; 22.96 ± 4.32 years) participated in this study. The groups did not differ significantly in age, $t(38) = 1.83$, $p = .076$ (see Table 5). Participants who had more than 8 years of musical experience (years of training: 12.40 ± 3.09; training onset: 8.17 ± 3.08 years old) composed the musicians' group (see Table 5). These participants practiced regularly at least one musical instrument and were teaching or attending musical classes in college (e.g., Bachelor of Music) or cultural groups (see Table 6). The control participants included in the non-musicians group did not play any musical instrument in the present or in the past, and have never had formal musical training besides attending two-year musical classes that are mandatory in the Portuguese school curricula for fifth and sixth grade students. Participants had normal hearing, normal or corrected-to-normal vision, no history of psychopathological/neurological disorders or substance abuse, and were not taking psychotropic medication at the time of the assessment. All participants were right-handed (Edinburgh Handedness inventory, Oldfield, 1971), monolingual, native speakers of European Portuguese and none of them was fluent in German. Participants were assessed with the Advanced Measures of Music Audiation (AMMA; Gordon, 1989) in order to ensure the two groups had distinct musical aptitude abilities (see Table 1). Additionally, to control for differences between groups in socioeconomic status and cognitive functioning, participants were evaluated using the Graffar Scale (Graffar, 1956) and The Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005; Simões et al., 2008), respectively (see Table 5). The present study was approved by the Ethics Committee for Life and Health Sciences of the University of Minho. Written consent was provided according to the updated version of the Declaration of Helsinki. Participants who were college students received course credit for their participation, while the remaining participants received gift vouchers.

Table 5

Socio-demographic and Other Relevant Characteristics of the Sample

	Musicians (*N* = 18)	Non-Musicians (*N* = 22)	Statistics	
	M (SE)	*M (SE)*	*t*	*p*
Age	20.78 (0.68)	22.96 (0.92)	1.827	.076
Graffar Scale	11.61 (0.74)	12.46 (0.60)	.895	.376
MoCA	29.06 (0.21)	28.96 (0.22)	-.327	.745
AMMA Total	70.28 (3.34)	52.27 (4.22)	**-3.241**	**.002**
AMMA Tonal	68.94 (3.93)	53.72 (4.33)	**-2.551**	**.015**
AMMA Rhythm	68.56 (3.10)	51.86 (3.82)	**-3.290**	**.002**
Musical training (years)	12.40 (0.73)	N.A.	N.A.	
Training onset (years)	8.17 (0.70)	N.A.	N.A.	

Note. The MoCA is from Nasreddine et al. (2005) and the MoCA adaptation to the Portuguese population is from Simões et al. (2008); The AMMA is from Gordon (1989); SE – Standard Error; N.A. – Non-Applicable. Significant differences are highlighted in bold.

4.3.2 Stimuli and Tasks

Learning Phase: Sung words were used in order to create an artificial language stream with prosodic characteristics for the prosodic experiment (Experiment 1). The new artificial sung words resulted from the combination of four consonants (K, L, R, D) and three vowels (I, O, U) into a set of 10 different consonant-vowel syllables ('ku', 'li', 'ri', 'lu', 'lo', 'ru', 'do', 'ki', 'di', 'ro') according to the procedure reported by François and colleagues (François et al., 2014). The syllables were synthesized using the Mbrola speech synthesizer (http://tcts.fpms.ac.be/synthesis/mbrola.html) with the German male diphone database (i.e., ge1). German is a language that has distinct phonotactic constraints compared to the

European Portuguese. None of the participants were fluent in German.

Table 6

Characterization of Musicians' Musical Background and Training Routines

Participant Code	Musical training[1]	Onset of Training[1]	Instrument[2]	Practice days/week	Practice hours/day
M1	13	6	piano	7	4
M2	8	12	piano	6	2
M3	11	10	clarinet	7	4
M4	15	6	violin	4	2
M5	8	11	clarinet	7	2.5
M6	10	9	clarinet	7	3
M7	13	6	violoncello	7	6.5
M8	10	15	trumpet	7	4
M9	11	12	percussion	7	2
M10	13	6	percussion	5	3
M11	15	11	piano	5	3
M12	14	5	piano	6	4
M13	20	7	violin	7	2
M14	13	5	violin	7	4
M15	17	6	transverse flute	4	2
M16	10	6	violin	4	1.5
M17	12	8	piano	7	2
M18	10	6	guitar	7	1

Note. [1] Measured in years; [2] The column displays the participant's main instrument despite several participants reporting expertise in more than one musical instrument.

Each syllable was then associated with a specific pitch from a range of ten different pitches (i.e., B4, C4, D4, F4, G4, C5, D5, Db5, E5, F5, according to musical letter notation). Syllables were 200 ms long with 10 ms of rise and decay times. Syllables were combined using the Audacity® software (version 2.1.1) in order to form five trisyllabic sung words (kuliri [C4 D4

F4], luloru [E5 Db5 G4], dokiru [D5 C5 G4], diluro [B4 E5 F5], rudiki [G4 B4 C5]) that were always sung with the same melodic contour.

The characteristics of these stimuli are shown in Table 7. A pause of 50 ms was introduced between each syllable leading to 700 ms-long words. The five words were concatenated in a pseudo-random order to form a continuous stream in which the same word was not repeated twice consecutively. There was an ISI of 50 ms in between words. The pause between syllables and the ISI between words aimed to create 50 ms silent periods that could work as baseline periods (for EEG analyses) preceding each syllable. This was an important aspect as ERP epochs were time-locked to words' onset. Each of the five words was repeated 100 times to reach a total of 500 Items in the stream, at a total duration of 6 min and 15 sec. The pitch-contour changes present in the words could not be used to segment the stream because they took place 50% of the time only at word boundaries. TPs within words ranged from 0.5 to 1 (M = 0.7) and TPs across word boundaries ranged from 0.07 to 0.31 (M = 0.19). The only cues to word boundaries were the TPs between syllable pairs. This stream was presented in the learning phase of the prosodic experiment.

Implicit Test Phase: An implicit test stream was also presented. This stream was designed to test the brain responses to online mismatches between expected (e.g., previously learned items) vs. unexpected items (e.g., new items). The implicit stream was composed of the words that were previously presented and of two categories of foils: 'part-words' (that resulted from the combination of the two last syllables of a word plus the first of another, or vice-versa) and 'non-words' (i.e., each word in the reverse order – for example, the word 'kuliri' gave rise to the non-word 'riliku'). There were five part-words and five non-words. The part-words were legal items of the sung language, yet they were heard five times less than the words that composed them. On the contrary, non-words were illegal items that were never presented to the participants (see Table 7). Each word-item was pseudo-randomly presented in the auditory stream with no immediate repetition. Each word was repeated 90 times and each part-word and non-word were repeated eight times. Thus, the implicit test stream was composed of 530 items and lasted approximately 6 min and 38 sec. In this stream, the words were classified as STD stimuli, whereas the part and non-words were classified as DEV stimuli. The pitch-contour changes present in all items of the stream could not be used to segment it. Again, the only cues to word boundaries were the TPs between syllable pairs.

Table 7

Characterization of the Sounds Used in the Learning and Implicit Test Phases of Each Experiment

Experimental Phase	Category	Experiment		
		Prosodic	Linguistic	Musical
Learning phase	Words/melodies	kuliri (C4 D4 F4)	xemuvu	D4 E4 G4
		luloru (E5 Db5 G4)	memive	F5 D5# A4
		dokiru (D5 C5 G4)	zixuve	E5 D5 A4
		diluro (B4 E5 F5)	zumevi	C4 F5 G5
		rudiki (G4 B4 C5)	vezuxu	A5 C4 D5
Implicit Test Phase	Words/melodies	kuliri (C4 D4 F4)	xemuvu	D4 E4 G4
		luloru (E5 Db5 G4)	memive	F5 D5# A4
		dokiru (D5 C5 G4)	zixuve	E5 D5 A4
		diluro (B4 E5 F5)	zumevi	C4 F5 G5
		rudiki (G4 B4 C5)	vezuxu	A5 C4 D5
	Part-words/melodies (new legal items)	kirulu (C5 G4 E5)	mivevu	C4 D5 F5
		loruku (Db5 G4 C4)	vixemu	D#5 A4 D4
		dikilu (B4 C5 E5)	vuzixu	D5 A4 F5
		rokuli (F5 C4 D4)	xuveme	G4 A5 D5
		ridoki (F4 D5 C5)	zuxume	G5 D4 E4
	Non-words/melodies (new illegal items)	kidiru (C5 B4 G4)	vumuxe	G4 E4 D4
		riliku (F4 D4 C4)	vemime	A4 D5# F5
		roludi (F5 E5 B4)	vexuzi	A4 D5 E5
		rukido (G4 C5 D5)	vimezu	G5 F5 C4
		rulolu (G4 Db5 E5)	xuzuve	D5 C4 A5

Note. The pitch (Hz) of the notes that compose each melody is reported in musical letter notation according to the English convention. Each note is named by a scientific pitch notation with a letter-name and a number identifying pitch's octave; additionally, "#" stands for *sharp*, "b" stands for *flat*.

In order to ensure that the stimuli used in this experiment were truly unfamiliar to the participants, a validation study with 31 first-year college students (29 women; 18.6 ± 1.85 years) was conducted. Participants were Portuguese speakers, with no fluency in more than two languages (German excluded) and no history of auditory impairment. The words were presented via loudspeakers and participants were asked to rate in a paper sheet how familiar each word sounded to them, using a scale that varied between 1 and 7 (1 – absolutely unfamiliar; 7 – highly familiar). Participants were also asked to provide a confidence judgment about their previous response using a scale from 1 to 7 (1 – not confident; 7 – absolutely confident). The mean perceived familiarity of words was 1.95 (*SD* = 0.38) indicating that the words were consistently classified as unfamiliar. This finding was supported by high confidence rates (6.15 ± 0.18).

For the linguistic (Experiment 2) and musical (Experiment 3) experiments, new words and melodies were developed to create two new streams of sounds to be presented during the learning phases of linguistic and musical experiments, respectively (see Table 7 for details). The same procedure described above was used to create both types of stimuli and streams. The stimuli used in these two additional experiments have the exact same duration, rise and decay times, and interval between syllables or notes of the sung words used previously. The words of the linguistic experiment (Experiment 2) were non-prosodic (i.e., all syllables had the same pitch [1000 Hz]). The melodies (Experiment 3) resulted from the combination of 10 different midi piano tones (D4, E4, G4, A4, C4, F5, D5#, E5, G5, A5) created using Piano FxStudio Software. Using Audacity® software (version 2.1.1), tones were combined into five tri-tone sequences (D4 E4 G4, F5 D5# A4, E5 D5 A4, C4 F5 G5, A5 C4 D5). The words and melodies were concatenated separately to form two distinct streams of sounds in which the same stimulus was not repeated consecutively. The ISI was 50 ms. These streams had the same characteristics of the stream used in the learning phase of the prosodic experiment regarding stimulus repetitions, duration and TPs. Again, the only cues to words' and melodies' boundaries were the TPs between syllables and tone pairs, respectively. The frequency changes that characterized the melodies could not be used to segment the musical stream. The implicit test phases of the linguistic (Experiment 2) and musical (Experiment 3) experiments were designed using the same procedure adopted for the implicit test phase of the prosodic experiment. As in the prosodic experiment, an implicit test stream was presented after the learning phases of the linguistic and musical experiments. The unfamiliarity of linguistic stimuli was examined in a validation study conducted with the sample described above. Participants were fully confident on their responses (6.29 ± 0.17 [*M* ± *SD*]), which confirmed the subjective unfamiliarity of the words (1.72 ± 0.25 [*M* ± *SD*]).

4.3.3 Procedure

The prosodic experiment was composed of three distinct phases in the following order: learning phase, implicit test phase and explicit behavioral test phase (see Fig. 5). During the learning and the implicit test phases, participants were told they would listen to an unknown artificial language and were instructed to listen carefully to it, making the effort to discover the words of that language. They were also told about a task at the end of stimuli presentation in which their guesses about the words of the unknown language would be tested. Thus, both phases may be considered active exposure phases. There was a fixation cross at the center of the computer screen to which participants should direct their attention while listening to these streams. Participants were not made aware of the transition between the learning and the implicit test phase. After the implicit test phase, behavioral measures of stream segmentation were collected. Here, participants had to perform three auditory LDTs (see Table1. APPENDIX 1. and Fig. 5). In each trial of these tasks, participants should direct their attention to a fixation cross that remained on the screen for 1 second. After that period, a word item (e.g., word, part-word or non-word) was randomly presented via earphones. Thus, participants heard each of the 15 different word-items in 15 different trials. After hearing each item, participants had unlimited time to decide whether the item presented was a word from the unknown language or was a foil by pressing a keyboard button. Even though most studies used a 2AFC procedure to evaluate speech segmentation (François et al., 2014; Francois & Schön, 2011), our procedure (i.e., LDTs) allows assessing not only if words were correctly recognized, but also if foils were correctly rejected. The beginning of the next trial was preceded by a silent period of 1500 ms.

Besides the prosodic test, participants performed two other LDTs (i.e., linguistic and musical tests) following the procedure of François and Schön (2010). These tests aimed to investigate if participants had learnt both the linguistic and musical structures embedded in the sung language to which they were exposed. The tests also allowed evaluating participants' ability to generalize knowledge of the statistical structure of the auditory streams to new items. In the linguistic test, participants were presented with the 15 words, part-words and non-words synthetized with a flat contour, i.e., test items that were devoid of the prosodic content that characterized the sung words. In the musical test, test items were midi-piano sequences derived from 'words', i.e., three-notes piano sequences that mimicked the melodic contour of the three categories of sung words. The musical items were created using Piano FxStudio Software. The order of the tests was counterbalanced across participants.

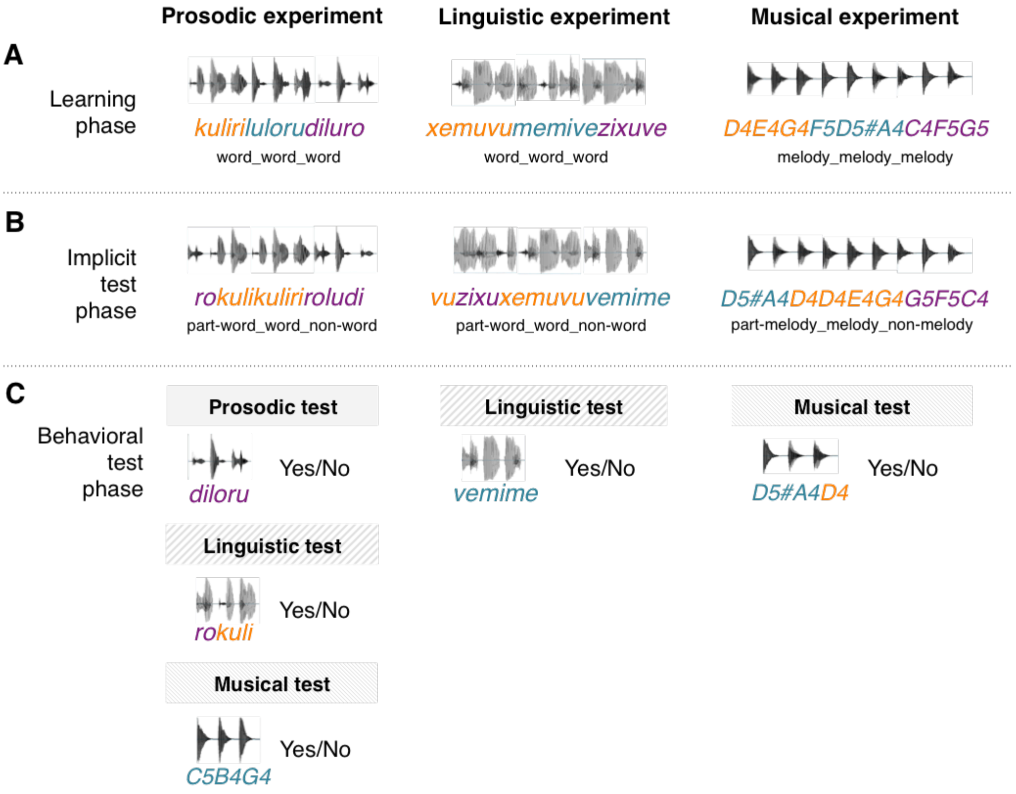

FIGURE 5. Illustration of the experimental procedure. **(A)** Illustration of the sound streams used in the learning phases of each experiment. The "_" represents the 50 ms ISI. Different words/melodies are displayed in different colors for illustrative purposes. **(B)** An implicit test phase followed the learning phase of each experiment. This phase consisted on the presentation of new legal/illegal word/melody items and previously presented words/melodies. **(C)** The implicit test phase was followed by behavioral tests. In the linguistic and musical experiments, the behavioral test phase consisted of a LDT in which words/melodies, part-words/melodies and non-words/melodies were presented and participants had to indicate whether they belonged to the sound stream or not. In the case of the prosodic experiment, there were three distinct tests: the prosodic test aimed to examine whether participants had learnt the prosodic words; the linguistic and musical tests, composed of non-prosodic words and melodies, respectively, aimed to evaluate the participants' ability to generalize the knowledge on the statistical structure of the sung language to new linguistic and musical items. The order of these latter tests was counterbalanced. EEG data was recorded during phases **(A)** and **(B)**.

As in the prosodic experiment (Experiment 1), the linguistic (Experiment 2) and musical experiments (Experiment 3) were each composed of three distinct phases. The instructions and the structure of the learning and of the implicit test phases were identical to those described above. Still, the explicit behavioral tests of these experiments (2 and 3) were composed of only one auditory LDT consisting on the presentation of the 15 non-prosodic words or musical items (see Table 7 and Fig. 5). EEG data was acquired during the first two phases of each experiment to obtain online measures of learning. The presentation of the three experiments was counterbalanced across participants. Stimulus presentation and timing were controlled with Presentation® Software (Neurobehavioral Systems, Albany, CA).

4.3.4 Data acquisition and processing

Data acquisition and processing underwent the same procedure for the three experiments. EEG data was recorded with a 64 channels BioSemi Active-Two system (BioSemi, Amsterdam, The Netherlands) according to the international 10-20 system (Jasper, 1958). The data was analyzed using Brain Vision Analyzer version 2.1.1 (Brain Products, Munich, Germany). Eye movements were monitored with 4 electrodes: horizontal movements were measured with two electrodes placed on the outer canthi of each eye, and vertical movements were monitored with one electrode placed at the infra-orbital ridge of the left eye and with FP1. Electrode offset was kept below 20 mV. The EEG was digitized at a rate of 512 Hz and filtered offline using 0.1-30 Hz Infinite Impulse Response filters. Ocular correction was performed using the method developed by Graton, Coles and Donchin (Gratton, Coles, & Donchin, 1983). Epochs containing eye blinks or movement artifacts exceeding ±100 μV were discarded.

For the ERP analysis, the signal was re-referenced offline to the algebraic average of the left and right mastoids. ERPs were time-locked to the onset of each item (e.g., words or melodies) and individual ERP epochs of 750 ms of duration were created. The baseline was defined as the 50 ms silent period preceding stimulus onset. For the analysis of the learning phase, segments considering the total phase duration were averaged for the five stimuli, for each participant at each electrode site. As we also aimed to investigate the time course of the electrophysiological responses during SL as a function of exposure time, a subsequent analysis was performed in which the learning phase was divided into five consecutive non-overlapping blocks, each representing 20% of the learning time: first block - trials 0-100 (20%); second block – trials 101-200 (40%); third block - trials 201-300 (60%); fourth block – trials 301-400 (80%); fifth block – 401-500 trials (100%).

For the analysis of the implicit test phase, segments were averaged separately for previously learned stimuli (e.g., word/melodies) and for each type of foils (e.g., part-words/melodies and non-words/melodies) and computed separately for each participant at each electrode site.

4.3.5 Statistical analysis

P-values were adjusted with the Greenhouse-Geisser correction for nonsphericity. Post-hoc tests for multiple comparisons were adjusted with Bonferroni correction.

4.3.5.1 Learning phase

After inspection of the grand averaged waveforms we did not find evidence of the canonical N1, P2 and N400-like components when considering their typical polarities, latencies and scalp topographies reported in previous SL studies (see, for example, de Diego-Balaguer et al., 2007; François et al., 2017, 2014; Francois & Schön, 2011; Vasuki et al., 2017a). Taking into consideration the methodological specificities of our paradigm, namely short ISIs (50 ms), a considerable overlap between target stimuli epochs was expected. As such, a stepwise analysis of evoked potentials was adopted, following previous studies (Kujala, Kallio, Tervaniemi, & Näätänen, 2001; Martín-Loeches, Sommer, & Hinojosa, 2005; Rellecke, Palazova, Sommer, & Schacht, 2011; Röder, Rösler, Hennighausen, & Näcker, 1996; Werheid, Alpay, Jentzsch, & Sommer, 2005). Specifically, mean amplitudes were computed for consecutive 50 ms intervals from 0-700 ms (e.g., the time range of a word/melody). This interval duration allowed the decomposition of the signal into several time windows whose analysis provided a dynamic perspective of the learning process. A significant advantage of this approach is that it does not rely on *a priori* defined, narrow time windows. Hence, it was expected to yield an adequate separation between overlapping components, allowing a more dynamic view on the neurocognitive processes taking place in response to the auditory streams' units (e.g., prosodic words, non-prosodic words and melodies) as a function of musical training, i.e. from stimulus onset until its offset.

Separate analyses of variance (ANOVAs) were performed to test the effects of the within- and between-subjects factors on ERP amplitude, and in order to avoid Type II errors. A repeated-measures ANOVA was carried out with ROI as within-subjects factor and Expertise (Musicians vs. Non-Musicians) as between-subjects factor. The ROI factor included five levels concerning frontal (Fz/3/4), fronto-central (FCz/3/4), central (Cz/3/4), centro-Parietal

(CPz/3/4) and parietal (Pz/3/4) topographical regions. Therefore, a total of 14 separate repeated-measures ANOVAs were performed for the different 50 ms intervals. To test for the effect of exposure time on the neural correlates of learning and for possible interaction effects between the timing of learning and the level of musical expertise, repeated-measures ANOVAs were computed with ROI (5 levels) and Block (5 levels: first, second, third, fourth and fifth blocks) as within-subject factors, and Expertise (2 levels) as between-subject factor for the 14 time windows. The same analyses were conducted for the three experiments separately. Only significant main effects or interactions involving the factors Expertise and Block ($p < .05$) are reported.

4.3.5.2 Implicit test phase

The statistical analyses of EEG activity in the implicit test phases were performed based on the mean amplitude extracted for consecutive 50 ms intervals from 0-700 ms for the three types of stimuli evaluated in this phase (e.g., items, part-items and non-items). Fourteen separate repeated-measures ANOVAs including the factors item type (3 levels: items, part-items and non-items) and ROI (5 levels) as within-subjects factor and Expertise (2 levels) as between-subjects factor were performed. This set of analyses was performed separately for the three experiments. Only main effects or interactions involving the factors Item type and Expertise ($p < .05$) are reported.

4.3.5.3 Behavioral test phase

Differences between behavioral performance and chance level (50%) were confirmed with *t*-tests. The performance of the two groups on the prosodic test (percentage of correct responses) was tested with an independent-samples *t*-test. Furthermore, a repeated-measures ANOVA was conducted including Expertise as between-subjects factor (2 levels) and Item type (3 levels) as within-subjects factor, separately for the linguistic and musical experiments. Additionally, in order to compare the results of the different behavioral tests of the prosodic experiment (e.g., prosodic, linguistic and musical), a three-way repeated-measures ANOVA with Expertise as between-subjects factor (2 levels), Item type (3 levels) and Stimulus type (3 levels: sung words, melodies and flat words) as within-subjects factors was computed. In order to compare the results of the three experiments, a three-way repeated-measures ANOVA with

Expertise as between-subjects factor (2 levels), and Item type (3 levels) and Stimulus type (3 levels) as within-subject factors was computed.

4.3.5.4 Brain-Behavior correlations

Spearman correlations (two-tailed) were performed in order to investigate if individual differences in the electrophysiological measures were associated with behavioral performance. We tested for possible associations between the measures obtained during the learning phase and the behavioral. These analyses were performed using the percentage of correct responses to words/melodies items in the LDTs and the mean amplitude values obtained for the time windows in which the ANOVAs showed significant differences between groups, across experiments.

4.4 Results

4.4.1 Experiment 1: Prosodic experiment

4.4.1.1 Electrophysiological Results

Fig. 6 shows grand averaged ERP waveforms. The repeated-measures ANOVAs revealed effects of musical expertise in specific time windows. Compared to non-musicians, musicians showed an enhanced negativity (i.e., more negative amplitude) in the 250-300 ms time window, $F(1, 38) = 10.599$, $p = .002$, $\eta_p^2 = .218$ (see Fig. 6A), and this effect was significant at all ROIs, $F(4, 152) = 4.202$, $p = .031$, $\eta_p^2 = .100$ (frontal: $p = .003$; fronto-central: $p = .001$; central: $p = .006$; centro-parietal: $p = .008$; parietal: $p = .032$) (see Fig. 6C and 7C).

When testing the effects of exposure time, the repeated-measures ANOVAs revealed several main effects of and/or interactions for different time windows (see Figs. 6B and 6C). Compared to non-musicians, musicians exhibited an enhanced negativity in the 200-250 ms time window in the first block of the learning phase (1-100 trials) at the parietal ROI ($p = .042$), $F(16, 608) = 2.647$, $p = .016$, $\eta_p^2 = .07$ (see Fig. 6C and 7B), and in the 250-300 ms time window during the first two blocks of the learning phase (1-200 trials) at several ROIs, $F(16, 608) = 2.183$, $p = .043$, $\eta_p^2 = .059$ (first block: fronto-central, $p = .038$, central, $p = .029$, centro-parietal, *p* = .010, parietal, *p* = .005; second block: frontal, *p* = .024, fronto-central, *p* = .012, central, *p* = .033, centro-parietal, *p* = .026) (see Figs. 6B, 6C, 7B and 7C).

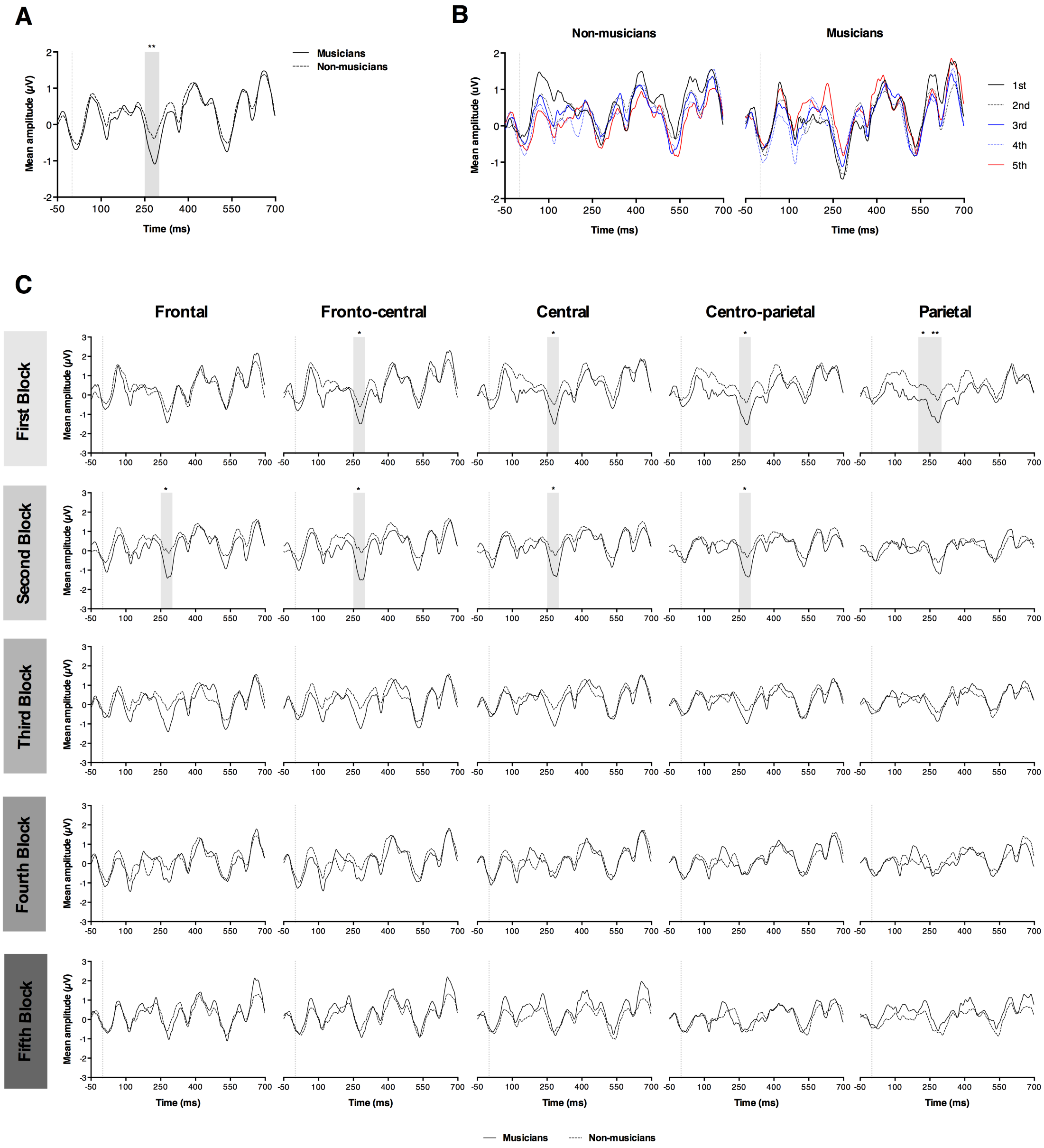

FIGURE 6. Grand-averaged ERP waveforms elicited by prosodic words during the learning phase of the prosodic experiment. **(A)** Grand-averaged waveforms obtained during the learning phase comparing musicians (solid line) and non-musicians (dashed line). **(B)** Each plot displays the averaged ERP waveforms for the five blocks of the learning phase (i.e., first [black line], second [dotted black line], third [blue line], fourth [dotted blue line] and fifth [red line] blocks), for each group separately. **(C)** The graphs show the evolution of group differences over the learning time (from the top to the bottom) for the different ROIs (i.e., frontal, fronto-central and, central, centro-parietal and parietal; from left to the right) for musicians [solid line] and non-musicians [dashed line]. Each ROI displays the averaged ERP signature of left (e.g., F3), right (e.g., F4) and midline (e.g., Fz) electrodes. The vertical grey lines indicate words' onset (0 ms). The grey shadowed areas indicate the time windows where significant group differences were found. * $p < .05$, ** $p < .01$, *** $p < .001$

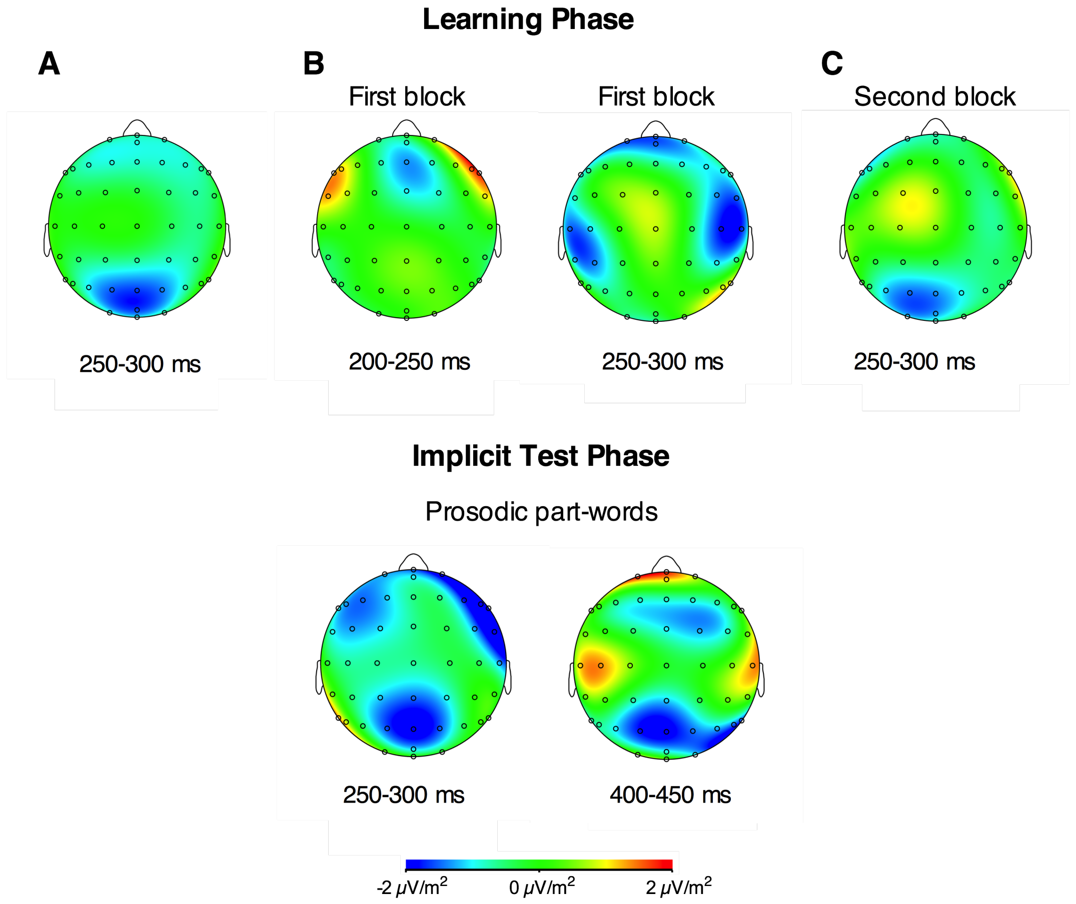

FIGURE 7. Topographic maps of the latency windows in which significant group effects were found for the prosodic experiment. *Note.* The maps (voltage mapping with spherical spline interpolation) illustrate the topographical distribution of the waveforms that were shown to significantly differ between groups, at distinct experimental phases and time windows. The topography of the significant difference found between musicians and non-musicians in the 250-300 ms time window during the learning phase is shown in **(A)**. The effects found during the first block for the 200-250 ms time window are presented in **B (left)** and for the subsequent time window also in **B (right)**; the topography of the difference found during the second block for the 250-300 ms time window is presented in **(C)**. Below, the maps display the spatial distribution of the effect found for part-words during the implicit test phase: on the left, the topography of the effect found at the 250-300 ms interval, and on the right the effect at the 400-450 ms interval.

In the implicit test phase, group differences emerged in the 250-300 ms and 400-450 ms time windows in response to part-words (see Fig. 8): amplitude was more positive in musicians compared to non-musicians (250-300 ms, $F(2, 76) = 3.580$, $p = .037$, $\eta_p^2 = .086$, $M = .566$ vs. $M = -.823$, $p = .028$; 400-450 ms, $F(2, 76) = 5.286$; $p = .010$, $\eta_p^2 = .122$, $M = 1.647$ vs. $M = -.082$, $p = .013$) (see the “Implicit test phase” panel of Fig. 7).

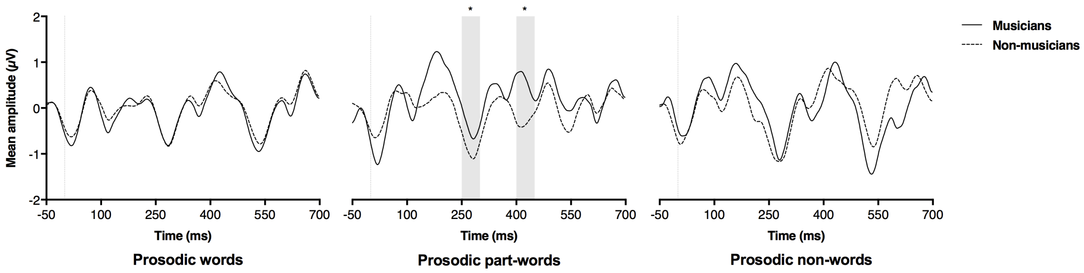

FIGURE 8. Illustration of the ERP results obtained for each type of prosodic item presented during the implicit test phase of Experiment 1. Each graph depicts the grand-averaged ERP waveforms in musicians (solid line) and non-musicians (dashed line) in response to prosodic words (e.g., previously learned items; left graph), prosodic part-words (e.g., new legal items; middle graph) and prosodic non-words (e.g., new illegal items; right graph) for the average of the five ROIs. The vertical grey lines indicate words' onset (0 ms). The grey shadowed areas indicate the time windows where significant differences between groups were found. * $p < .05$, ** $p < .01$, *** $p < .001$

4.4.1.2 Behavioral results

Comparing the performance in the prosodic test with chance level (50%) revealed that musicians performed significantly above chance (60.37 ± 18.36), $t(17) = 2.396$, $p = .028$, whereas non-musicians did not (56.33 ± 14.43), $t(21) = 1.964$, $p > .05$ (see Fig. 9A). Even though musicians showed a higher percentage of correct responses, the performance of the two groups did not differ significantly, $t(38) = -.758$, $p > .05$. Results of the two-way repeated-measures ANOVA showed no differences in the responses to the different item types (e.g., words, part-words and non-words) (see Fig. 9B).

In the musical and linguistic tests of the prosodic experiment, musicians' performance was not significantly above chance (Musical test: 51.48 ± 11.84, $t(17) = .531$, $p > .05$; Linguistic test: 51.11 ± 10.97, $t(17) = .430$, $p > .05$) as well as non-musicians' performance (Musical test: 49.67 ± 12.88, $t(21) = -.116$, $p > .05$; Linguistic test: 51.67 ± 13.83, $t(21) = .539$, $p > .05$) (see the middle and right bar graphs depicted in Fig. 9A). Moreover, the groups did not differ significantly (Musical test: $t(38) = -.450$, $p > .05$; Linguistic test: $t(38) = .136$, $p > .05$). The two-way ANOVAs performed on both tests revealed no differences in participants' responses as a function of item type ($p > .05$; see Fig. 9B).

The three-way ANOVA that compared the results of the three behavioral tests revealed a significant effect of stimulus type, $F(2, 76) = 4.157$, $p = .020$, $\eta_p^2 = .104$: Participants across groups responded more accurately to prosodic words ($M = 58.35$) than to melodies ($M = 50.57$, $p = .046$) and tended to respond more accurately to prosodic words than to words ($M =$

51.38, *p* = .097), although this last comparison was not statistically significant (see Fig. 9C). This analysis did not reveal additional significant main effects or interactions.

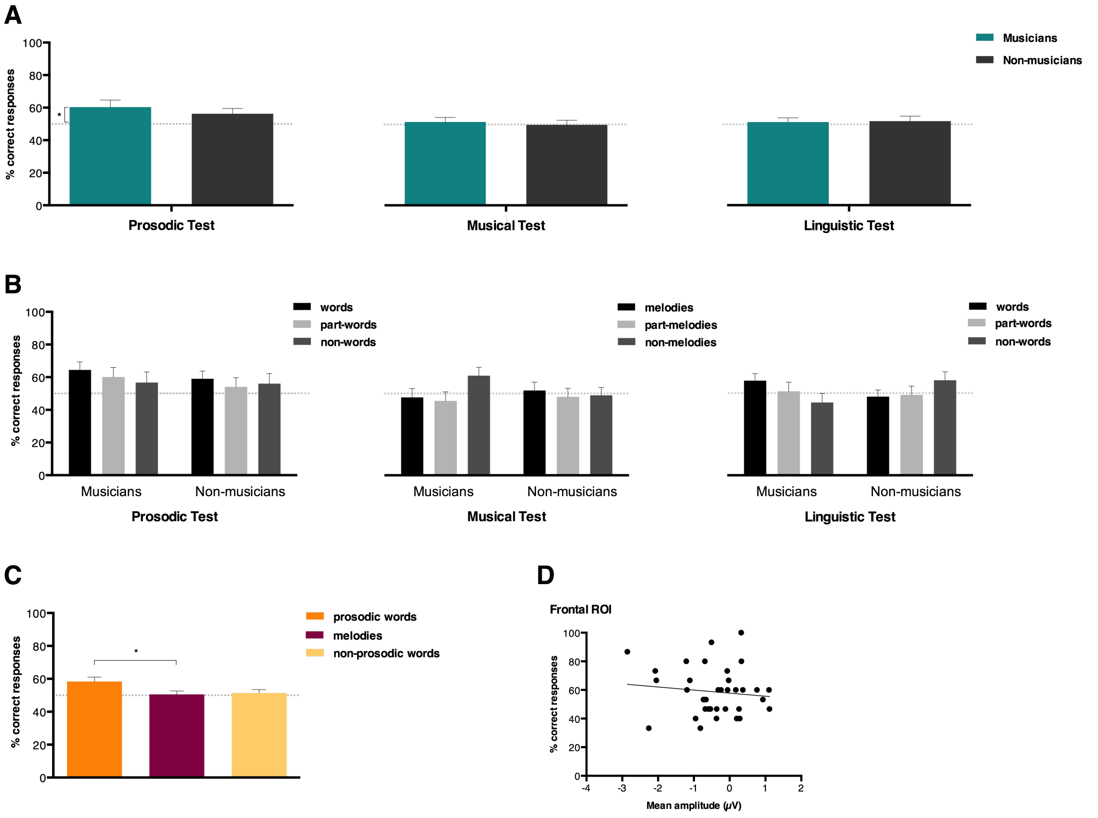

FIGURE 9. Participants' performance in the behavioral tests of the prosodic experiment. **(A)** The plots display the percentage of correct recognition of musicians (green) and non-musicians (dark grey) in the prosodic (left), musical (middle) and linguistic (right) tests. **(B)** The graphs compare the percentage of correct responses for each type of item (i.e., legal items, black; new legal items, soft grey; new illegal items, dark grey), in each test, in musicians vs. non-musicians. **(C)** Depiction of participants' percentage of correct responses for each type of stimuli: prosodic words (orange), melodies (purple) and non-prosodic words (yellow). **(D)** Illustration of the significant correlation between the mean amplitude (250-300 ms time window at frontal sites) of the averaged ERPs during the learning phase and the accuracy in the prosodic test. Bars indicate the standard error of the mean (SEM). The horizontal dotted lines refer to chance level. * *p* < .05, ** *p* < .01, *** *p* < .001

4.4.1.3 Brain-behavior correlations

We found a negative association between the accuracy in the prosodic test and the amplitude of the negative deflection occurring between 250-300 ms during the learning phase at the frontal ROI, *r* = -.384, *p* = .019: Performance was improved in participants showing increased negative amplitudes in the 250-300 ms latency range (see Fig. 9D). There were no other significant correlations (*p* > .05).

4.4.2 Experiment 2: Linguistic experiment

4.4.2.1 Electrophysiological Results

During the learning phase, musicians tended to exhibit an enhanced negativity relative to non-musicians (*M* = -.324 vs. *M* = -.024) in the 250-300 ms time window, *F*(1, 38) = 2.846, *p* = .099, η_p^2 = .066 (see Fig. 10A). When testing the effects of exposure time, the repeated-measures ANOVAs did not show group differences (*p* > .05; Fig. 10B and 10C).

The analysis of the ERPs during the implicit test phase revealed no differences between groups in the processing of words, part-words and non-words (see Fig. 11).

4.4.2.2 Behavioral results

The behavioral performance of musicians was significantly above chance (58.03 ± 12.42), *t*(17) = 2.669, *p* = .017, whereas the performance of non-musicians was not (53.33 ± 12.11), *t*(21) = 1.261, *p* > .05 (see Fig. 12A). Nevertheless, performance did not differ significantly between groups, *t(*38) = -1.178, *p* > .05. No significant main effects or interactions were found when comparing participants' responses to the different types of items (*p* > .05; see Fig. 12B and 12C).

4.4.2.3 Brain-behavior correlations

No significant correlations were found between brain and behavior measures (*p* > .05).

4.4.3 Experiment 3: Musical experiment

4.4.3.1 Electrophysiological Results

When considering the ERP responses for the total duration of the learning phase, no significant group differences were observed (*p* > .05; see Fig. 13A). However, the statistical analysis that tested the effects of exposure time on learning revealed that the effects of musical expertise interacted with the number of stimulus repetitions (i.e., Block; see Figs. 13A, 13C, 14A and 14B). Musicians showed an enhanced positivity in the first 50 ms after melodies' onset, during the first block of the learning phase, at frontal (*M* = 1.068 vs. *M* = .121, *p* = .007) and fronto-central (*M* = .848 vs. *M* = .166, *p* = .038) ROIs, *F*(16, 608) = 2.404, *p* =

.049, η_p^2 =.055 (see Figs. 13C and 14A).

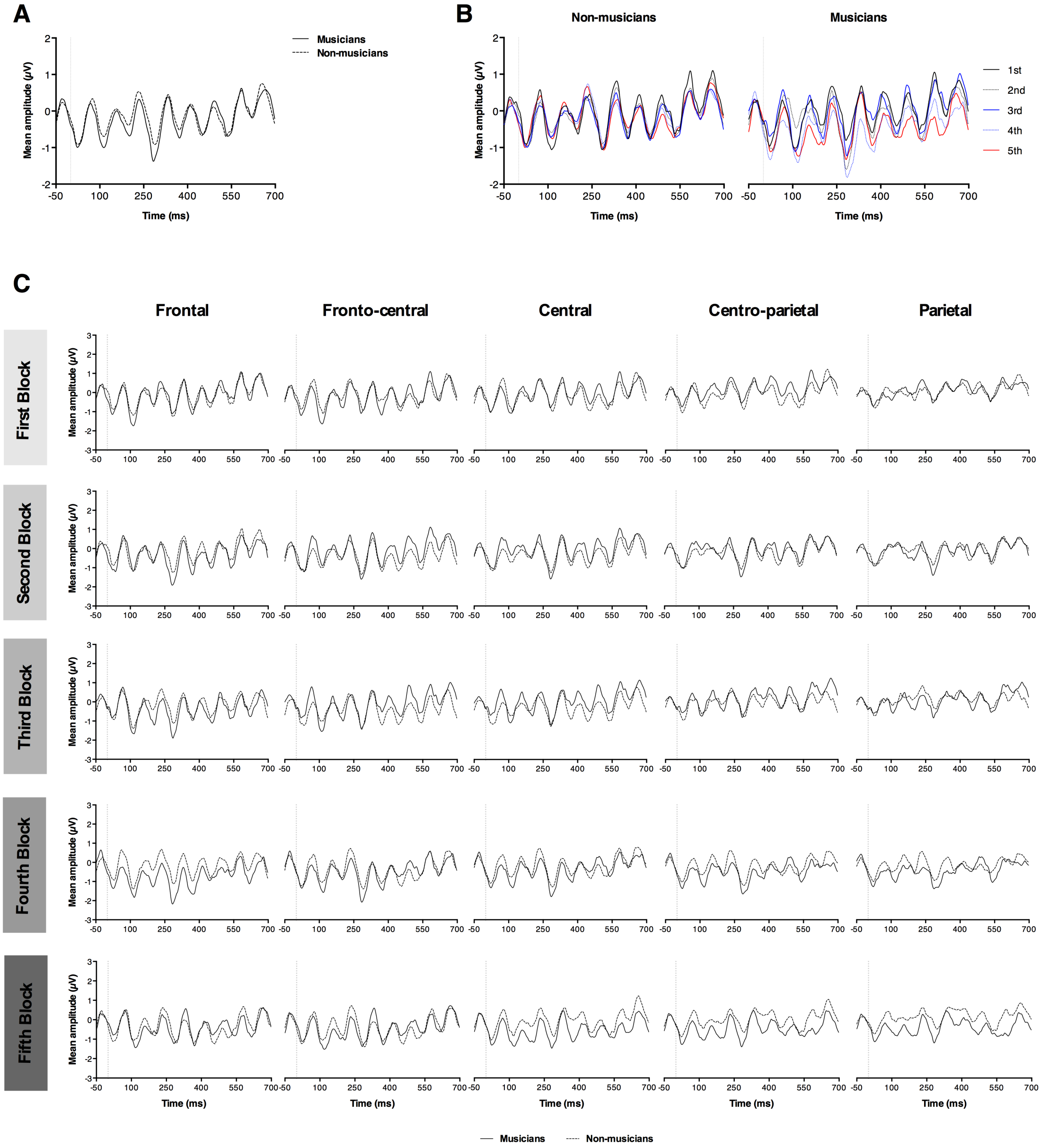

FIGURE 10. Grand-averaged ERP waveforms elicited by non-prosodic words in the learning phase of the linguistic experiment. **(A)** ERP waveforms averaged over the five ROIs, for the entire duration of the learning phase comparing musicians (solid line) and non-musicians (dashed line). **(B)** Each graph displays the ERP averages for the five blocks (i.e., first [black line], second [dotted black line], third [blue line], fourth [dotted blue line] and fifth [red line] blocks) of the learning phase for each group. **(C)** The graphs depict the evolution of brain responses over time for de different blocks (from the top to the bottom) and for the different ROIs (from the left to the right). Each ROI displays the averaged ERP waveforms over left (e.g., F3), right (e.g., F4) and midline (e.g., Fz) electrodes. The vertical grey lines indicate words' onset (0 ms).

The greater positivity observed in musicians in the first block (1-100 trials) of the learning phase at the frontal region was extended to the 50-100 ms time window (*M* = 1.110

vs. *M* = .329, *p* = .049), *F*(16, 608) = 2.751, *p* = .030, η_p^2 = .063 (see Figs. 13C and 14A).

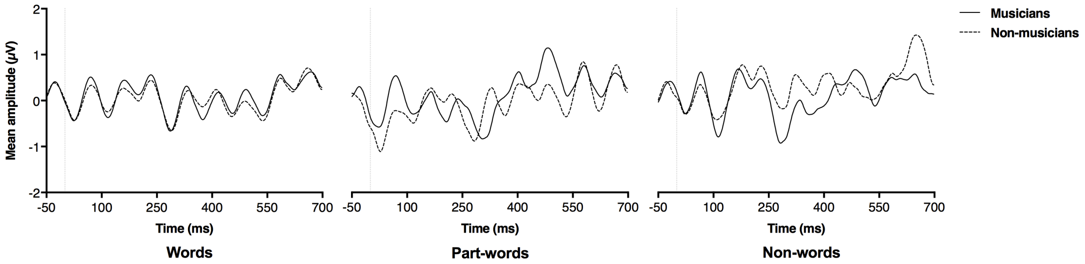

FIGURE 11. Grand-averaged ERP waveforms elicited by each type of non-prosodic words in the implicit test phase of the linguistic experiment. Each graph depicts the grand-averaged ERP waveforms of musicians (solid line) and non-musicians (dashed line) obtained for the entire duration of the implicit test phase in response to non-prosodic words (left graph), non-prosodic part-words (middle graph) and non-prosodic non-words (right graph) in the five ROIs. The vertical grey lines indicate words' onset (0 ms).

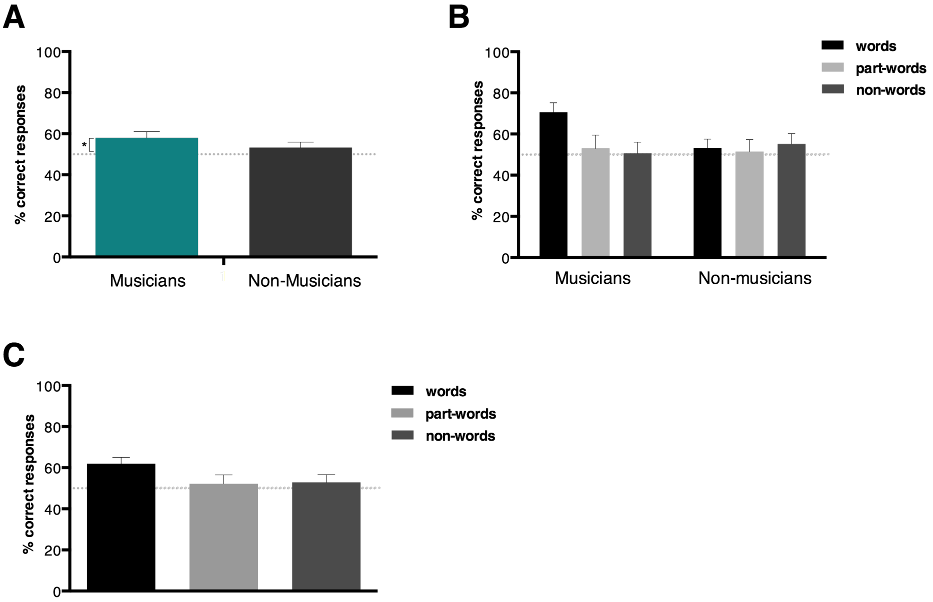

FIGURE 12. Participants' performance in the linguistic experiment. **(A)** Total percentage of correct recognition of musicians (green) and non-musicians (dark grey). **(B)** Musicians and non-musicians' percentage of correct responses for each type of item (i.e., words, black; part-words, soft grey; non-words, dark grey). **(C)** The graph depicts overall performance as a function of the type of item (i.e., words, black; part-words, soft grey; non-words, dark grey). Bars indicate SEM. The horizontal dotted lines refer to chance level. * $p < .05$, ** $p < .01$, *** $p < .001$

In the 300-350 ms period, musicians also exhibited an enhanced positivity relative to non-musicians, *F*(16, 608) = 2.898, *p* = .015, η_p^2 = .066: The group differences were observed during the first (1-100 trials) and the fourth block (300-400 trials) at specific ROIs (first block: frontal, *M* = .756 vs. *M* = -.639, *p* = .002, fronto-central, *M* = .617 vs. *M* = -.668, *p*

= .004, central, *M* = .624 vs. *M* = -.367, *p* = .032; fourth block: parietal, *M* = .496 vs. *M* = -.443, *p* = .015) (see Figs. 13C, 14A and 14B).

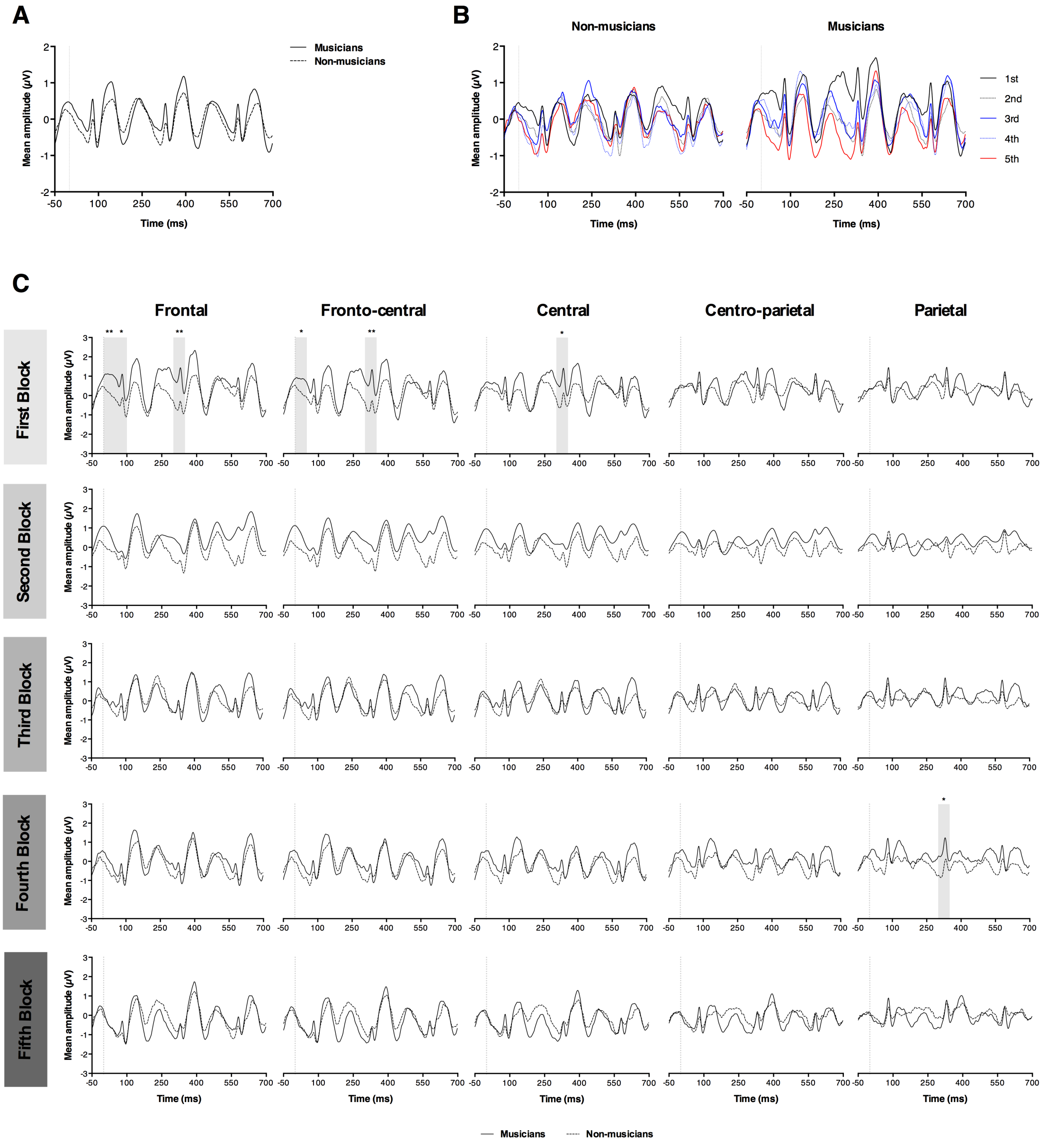

FIGURE 13. Grand-averaged ERP waveforms elicited by melodies during the learning phase of the musical experiment. **(A)** ERPs averaged over the five ROIs, for the entire duration of the learning phase comparing musicians (solid line) and non-musicians (dashed line). **(B)** Averaged ERP waveforms for the five blocks of the learning phase (i.e., first [black line], second [dotted black line], third [blue line], fourth [dotted blue line] and fifth [red line] blocks), for each group separately. **(C)** Evolution of group differences over the learning time (from the top to the bottom) for the different ROIs (i.e., frontal, fronto-central and, central, centro-parietal and parietal; from left to the right) for musicians [solid line] and non-musicians [dashed line]. Each ROI displays the averaged ERP signature of left (e.g., F3), right (e.g., F4) and midline (e.g., Fz) electrodes. The vertical grey lines indicate melodies' onset (0 ms). The grey shadowed areas indicate the time windows where significant differences between groups were found. * $p < .05$, ** $p < .01$, *** $p < .001$.

No group differences were observed in the processing of melodies, part-melodies and non-melodies during the implicit test phase (see Fig. 15).

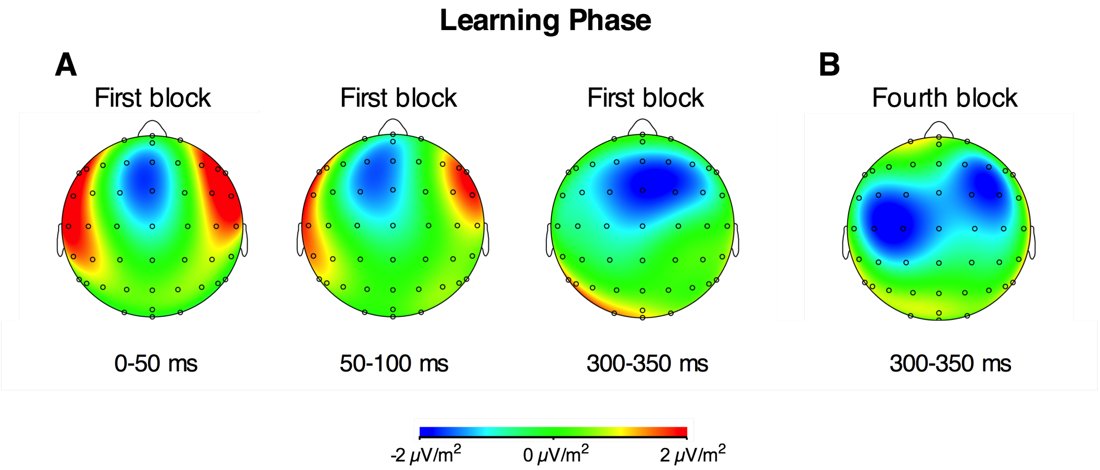

FIGURE 14. Topographic maps of the latency windows where significant group differences were found during the learning phase of the musical experiment. The topography of the differences found during the first block is displayed in panel **(A)**, whereas panel **(B)** shows the scalp distribution of the significant group difference found in the fourth block.

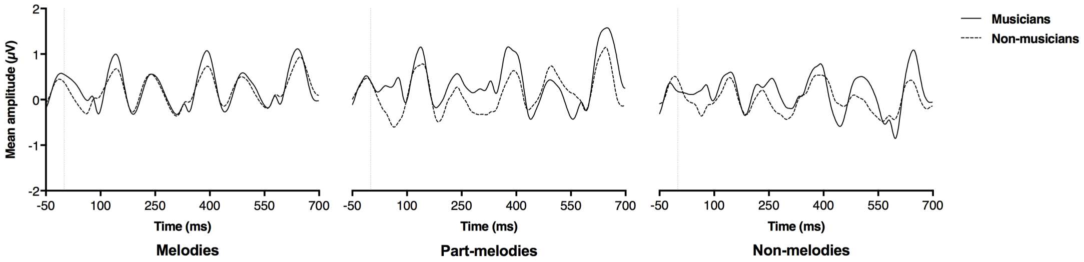

FIGURE 15. Grand-averaged ERP waveforms elicited by each melody item presented during the implicit test phase of the musical experiment. Each graph depicts the ERP waveforms of musicians (solid line) and non-musicians (dashed line) averaged over the entire duration of the implicit test phase in response to melodies (left graph), part-melodies (middle graph) and non-melodies (right graph) for the five ROIs. The vertical grey lines indicate melodies' onset.

4.4.3.2 Behavioral results

The comparison of performance with chance level showed that non-musicians performed below chance (54.33 ± 11.90), *t*(21) = 1.628, *p* > .05, while musicians significantly outperformed chance (59.26 ± 16.94), *t*(17) = 2.31, *p* = .033 (see Fig. 16A). Nevertheless, the performance of the two groups was not significantly different, *t*(38) = -1.046, *p* > .05. The results also showed that, irrespective of expertise, participants were significantly better at recognizing melodies than at correctly rejecting part-melodies (*M* = 63.56 vs. *M* = 49.28, *p* =

.012), $F(2, 76) = 3.931$, $p = .024$, $\eta_p^2 = .098$ (see Fig. 16B and 16C).

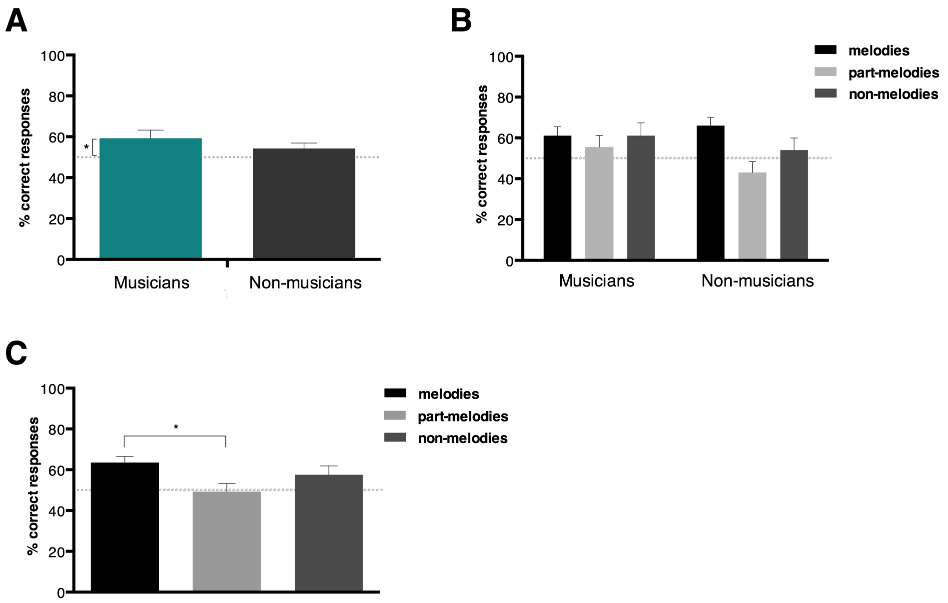

FIGURE 16. Participants' performance in the musical experiment. **(A)** Percentage of correct recognition of musicians (green) and non-musicians (dark grey). **(B)** Percentage of correct responses for each type of item (i.e., melodies, black; part-melodies, soft grey; non-melodies, dark grey) in musicians and non-musicians. **(C)** Depiction of participants' percentage of correct responses for each type of item (i.e., melodies, black; part-melodies, soft grey; non-melodies, dark grey). Bars indicate SEM. The horizontal dotted lines refer to chance level. * $p < .05$, ** $p < .01$, *** $p < .001$

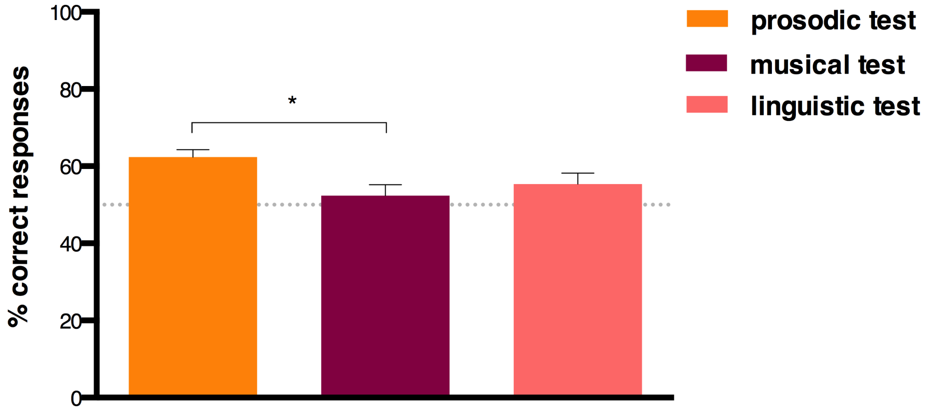

FIGURE 17. Behavioral performance across experiments. Performance in the different experiments as a function of stimulus type (i.e., prosodic words – prosodic test [orange], melodies – musical test [purple] and non-prosodic words – linguistic test [pink]). Bars indicate SEM. The horizontal dotted line refers to chance level. * $p < .05$, ** $p < .01$, *** $p < .001$

4.4.3.3 Brain-behavior correlations

ERP effects were not significantly associated with behavioral performance ($p > .05$).

4.4.4 Comparison of participants' performance across experiments

The number of correct responses was increased in the prosodic test of the prosodic experiment (M = 62.38) compared to the musical test of the musical experiment (M = 52.36, *p* = .023), *F*(2, 72) = 4.729, *p* =.012, η_p^2 = .122 (see Fig. 17). No other effects or interactions were found.

4.5 Discussion

The aim of the current study was to test whether musical training modifies the ASL of distinct types of auditory structures in order to clarify the extent of musical training's transfer effects. We examined brain responses to prosodic, non-prosodic and musical streams while learning unfolded over time. Further, the effects of musical training on ASL were examined through implicit (e.g., ERPs) and explicit (e.g., LDTs) measures. While an increased negativity was observed in musicians in the 250-300 ms latency range in response to prosodic words, differences between groups emerged earlier in time in response to melodies (i.e. within 100 ms after melodies' onset). Conversely, there were no differences between groups for non-prosodic (flat contour) words, despite a tendency for musicians to show an increased negativity in the 250-300 ms time window. Together, these results demonstrate that musicians and non-musicians use distinct neural mechanisms while tracking auditory regularities and segmenting auditory streams depending on the acoustic nature of the input. Violations of the statistical properties of the musical and non-prosodic streams did not elicit distinct ERP responses between groups. Conversely, a positivity emerged in response to prosodic part-words that had a larger amplitude in musicians in two distinct time windows (250-300 ms and 400-450 ms). This specific difference indicates that the effects of musical training on the detection of violations to the statistical structure of auditory streams depends on stimulus type, being more pronounced in response to speech sounds with melodic characteristics (i.e., sung words). The behavioral performance across experiments indicated that musicians learned the statistical structure of all auditory streams (indicated by above chance performance), whereas non-musicians did not. Furthermore, the recognition of prosodic words was facilitated when compared with the recognition of melodies for all participants. Importantly, performance in the prosodic test was associated with individual learning differences at the brain level.

4.5.1 ERP evidence for enhanced ASL in musicians

4.5.1.1 Speech segmentation

We showed that musical training facilitated the segmentation of prosodic words (Experiment 1). Notwithstanding, in the experiment in which non-prosodic words were tested (Experiment 2), no significant differences were observed between musicians and non-musicians in any of the time windows examined. This suggests that the presence of suprasegmental features in speech modulates the effects of musical training on ASL. Yet, we found a tendency for an increased negativity occurring in the 250-300 ms latency range in musicians during the linguistic experiment (Experiment 2). Despite being non-significant, this finding is relevant as it demonstrates consistency in the timing of the electrophysiological response to both types of artificial speech sounds.

The results of both experiments using speech sounds did not reveal a clear N400-like component in response to words' onset as previous studies have reported (i.e., in the 350-550 ms [François et al., 2014] and in the 200-350 ms [François et al., 2017] latency ranges). Instead, musicians exhibited an enhanced negativity in the 250-300 ms latency window. Typically, ERP modulations occurring around 200-300 ms after stimulus onset have been associated with processes related to stimulus discrimination (Patel & Azzam, 2005). For example, the N200 is a negativity elicited in oddball paradigms and is thought to index deviance processing elicited by rare stimuli in the context of STD or high-probability stimuli (Hoffman, 1990). Therefore, the negative deflection found in the current study (that differed between groups) may reflect a critical brain mechanism involved in speech segmentation. Along the same lines, a study from Vasuki and collaborators (Vasuki et al., 2017a) which investigated ASL abilities in children also found that the N250 ERP component was increased in musician relative to non-musician children. The authors conceptually compared the N250 response to tone-triplets in children with the N1 component typically elicited in response to word onsets in adult studies (Abla et al., 2008; Sanders, Ameral, & Sayles, 2009; Sanders et al., 2002). Moreover, they proposed that the N250 reflects a prediction process involving higher-order recruitment of attentional resources in musicians (Pallesen et al., 2010; Sanders et al., 2002). Furthermore, previous studies suggested that the N400-like component represents an index of lexical search and speech segmentation (Cunillera et al., 2009; de Diego-Balaguer et al., 2007; François et al., 2017, 2014; Vasuki et al., 2017b). Yet, the latency windows used to compute the mean amplitude of this component have diverged between studies. For example: 200 to 350 ms (François et al., 2017); 300 to 500 ms (Batterink & Paller, 2017; Cunillera et al., 2006; Vasuki et al., 2017b); 350 to 550 ms (Cunillera et al., 2009; de Diego-Balaguer et al., 2007; François et al., 2014); 350 to 550 ms, with stronger effects observed in the 400-550 ms time window (Cunillera, Laine, & Rodríguez-Fornells, 2016). Considering the variability in the time windows reported in previous studies and the fact that, in the current study, ERPs in musicians and non-musicians diverged precisely between 250 and 300 ms post-stimulus onset, it is plausible that the effects found here are related to the N400-like component reported in the SL literature.

The speech stimuli used here derived from a German diphone database provided by the Mbrola software whose phonotactic constraints are different from those characterizing the Portuguese language – the participants' mother tongue. Previous studies used speech stimuli based on the linguistic databases of their participants' mother languages (Cunillera et al., 2009; de Diego-Balaguer et al., 2007; François et al., 2014). Consequently, this may have compromised the researchers' ability to create word-items that were truly unfamiliar to their participants, which might have biased participants' responses in the SL tasks. Besides, here, we ensured that the word-items were completely unknown to participants by performing a stimulus validation, before EEG data collection, that confirmed the total unfamiliarity of the newly created sounds. These methodological differences between the present and early studies may have led to differences in the way participants perceived and processed the target stimuli. We hypothesize that these differences may have contributed to potential changes in the timing and length of the typical N400-like component observed in past experiments (Cunillera et al., 2009; de Diego-Balaguer et al., 2007; François et al., 2014; Vasuki et al., 2017b). However, future studies should address this issue by comparing ERP responses to speech stimuli created based on the participants' mother language vs. based on a foreign and unfamiliar language with different phonotactic constraints.

Considering that (1) previous studies found an enhanced N400-like component to be related to facilitated segmentation of speech streams (Cunillera et al., 2009; de Diego-Balaguer et al., 2007; François et al., 2014) and that (2) a significant association between a negativity in the 250-300 ms time window and participants' behavioral performance was observed in the current study, the 250-300 ms negativity might be an index of the sensitivity of musicians to the statistical structure of sung speech streams. Together with the observation of an increased accuracy in the prosodic test compared to both musical and linguistic tests (Experiment 1), this finding corroborates previous reports of facilitated segmentation of sung streams as opposed to spoken streams (i.e., composed of words with flat contour) (Schön et al., 2008). Besides, since this negative-going wave was found both for prosodic and non-prosodic speech sounds but not for melodies, it may reflect the participants' sensitivity to the statistical properties of the speech streams and not merely the sensitivity to differences between the acoustic profiles of the stimuli (Astheimer & Sanders, 2011).

In the implicit test of the prosodic experiment, the groups differed significantly in the 250-300 ms and 400-450 ms latency windows: Musicians exhibited an increased negativity in response to part-words in these time windows. Contrary to our hypothesis, there was no evidence of MMN/N200 components in response to structural auditory violations. The results did not reveal further group differences for words or non-words, which highlights the distinctiveness of part-words in the context of foils. Contrary to non-words, part-words were legal new items whose occurrence was five times less probable when compared with words. The online detection of this type of violations was demanding relative to non-words. Thus, it is worth noting that musical training influenced precisely the response to this type of foils.

4.5.1.2 Musical stream segmentation

The influence of musical training on the ERP responses to melodies' onset was observed very early in the time course of stimulus processing. The positive deflection observed in the 0-100 ms time window was increased in musicians, specifically during the first block (0-100 trials) of the learning phase. Hence, musicians' familiarity with musical stimuli, such as melodies like the ones used here, might make them more sensitive to musical sounds and more able to perceive these sounds distinctively, even at a very early processing stage. The timing of this group difference suggests an enhanced ability of musicians to process the sensory aspects of melodies since, typically, ERP modulations occurring within the first 200 ms after sound onset reflect sensory-based operations (Beres, 2017; Näätänen, 1992).

In previous studies, musical training was found to change the N1 and P2 responses to instrumental sounds and pure tones (Shahin et al., 2003). Likewise, and in good agreement with the results of the prosodic experiment, group differences were observed in the first trials of exposure to the auditory stream (exception made to the increased positivity observed in the 300-350 ms latency window during the fourth block [301-400 trials]). Therefore, the role played by musical training in ASL seems to be more prominent at the first stages of learning, not only when participants are presented with speech stimuli (as observed in Experiment 1) but also with musical stimuli. These findings are in accordance with previous reports (Abla et al., 2008; Sanders et al., 2002; Vasuki et al., 2017b) and suggest that musical training changes the computation of TPs in streams with melodic characteristics (i.e., both prosodic and musical streams contained items characterized by melodic variations).

At the behavioral level, evidence from the musical test showed that only musicians could segment the musical stream into its units. No effects of musical training were observed in the implicit test phase, which suggests that musical training did not modulate the processing of structural violations.

4.5.2 Faster sound stream segmentation in musicians

When the neural correlates of the prosodic stream were analyzed as a function of exposure time, we found that the enhanced negativity observed in musicians relative to non-musicians was present during the first two learning blocks (0-200 trials). However, this difference disappeared in the last blocks of the learning phase (201-500 trials). Moreover, differences between musicians and non-musicians in the processing of melodies emerged only during the first block (i.e., 0-100 trials). These findings confirm previous reports of the influence of musical training in the first stages of ASL. Vasuki and colleagues (Vasuki et al., 2017b) reported N1 and N400 latency modulations as a function of musical training in the first part of the familiarization phase of an embedded triplet paradigm with pure tones. In prior studies, increased N1 amplitude has been related to "expert" performance in behavioral SL tasks of speech segmentation (Sanders et al., 2009, 2002) and has also been reported in "high learners" during the first part of the learning phase (Abla et al., 2008). The current study also revealed that the group differences found in the earliest stages of ASL vanish as the exposure time increases. This indicates that musical training influences the speed at which SL occurs. It is plausible that musical experts recruit additional processing strategies when parsing out acoustic streams. For example, musicians may be more able to selectively focus attention on auditory regularities to perform more accurate computations of TPs. Selective attention is taken as an example of an effective learning strategy during the processing of auditory regularities (Daltrozzo & Conway, 2014) that some authors take as a requirement for SL to occur (Ahissar & Hochstein, 1993; Ding et al., 2018; Gilbert, Ito, Kapadia, &

Westheimer, 2000). Musicians show increased ERP responses (i.e., more negative MMN and N2b) during attentive listening when processing music and speech sounds (Tervaniemi et al., 2009). Therefore, musical training could lead to stronger representations of statistical regularities and, consequently, to advantages in auditory segmentation.

Another possibility is that faster SL in musicians is linked to better working memory capacities. The positive role of working memory capacities on SL (Palmer & Mattys, 2016). Batterink and Paller (2017) suggest that the speed at which listeners learn the statistical structure of a stream and the ability to create word-unit representations may be influenced by the auditory sensory memory span. Auditory sensory memory or echoic memory (Neisser, 1967) is responsible for maintaining very brief auditory memory traces available for further processing. Learners with good echoic memory abilities (e.g., increased memory span) might benefit from increased efficiency at retaining more auditory information in memory when recognizing correct or incorrect chunks of syllables. This ability, in turn, may increase the likelihood of high performance in offline tests of SL. Previous research has provided evidence of faster working memory updating processes in musicians (George & Coch, 2011). Hence, musicians might have an enhanced capacity to attend to auditory input and to retain and manage short-term statistical representations in memory.

4.5.3 The relationship between electrophysiological and behavioral measures of ASL

The results indicate that musical training induces changes in ASL abilities, which are reflected in an increased accuracy in behavioral ASL tasks and in alterations in the neurophysiological responses recorded during learning. It is important to highlight that only musicians performed above chance in the ASL tasks, which suggests that non-musicians did not learn the statistical properties of the streams. This finding agrees with previous reports (François et al., 2013; Vasuki et al., 2017a, 2017b) and raises the possibility that additional exposure time to each stream could have resulted in above-chance performance in both groups. Despite an overall poorer behavioral performance in non-musicians, the performance of both groups in the offline LDTs did not differ significantly. This is in agreement with some recent studies that did not find group differences at the behavioral level (Francois & Schön, 2011; Paraskevopoulos, Kuchenbuch, Herholz, & Pantev, 2012).

Importantly, the correlational analyses revealed an association between individual variability in the neural responses to prosodic words and behavioral performance: participants who showed an increased negativity in the 250-300 ms latency window also showed higher accuracy in the prosodic test. This evidence supports the notion that the capacity to decode word-like units from auditory input and to explicitly retrieve these units from long-term memory represents two interrelated processes. The 250-300 ms negativity observed during learning might represent an index of the perceptual encoding of the streams' regularities. The mental representations of the individual items composing the streams might be more effectively stored in the long-term memory of listeners who show an enhanced negativity at the neurophysiological level. Conversely, listeners exhibiting a less pronounced negativity might form more volatile representations of word-items, which compromises their ability to accurately recognize words and correctly reject foils in the explicit behavioral assessment of ASL. Nevertheless, caution is needed when interpreting these results since this ERP-behavior association was only found for prosodic words. Notwithstanding, this is a relevant finding in light of previous reports suggesting that SL is task-, modality- and stimulus-specific (Siegelman & Frost, 2015). Therefore, these results may be accounted for by the specific features of the sung language (e.g., the combination of linguistic and melodic information that prior studies have shown to facilitate auditory segmentation, see Schön et al., 2008) and/or the type of SL task chosen to evaluate stream segmentation. Of note, the participants of our study could have exhibited a different pattern of responses if tested with target detection or rating tasks instead of a recognition task (Batterink & Paller, 2017).

These results also provide evidence for the positive transfer effects from music to speech by showing that the impact of musical training on ASL is not circumscribed to musical stimuli. That is, musical training seems to bring advantages to the way musicians' brains process regularities in both musical and speech streams (cross-domain neuroplasticity effect; Bidelman & Alain, 2015; Bidelman, Hutka, & Moreno, 2013; Moreno & Bidelman, 2014; Ong, Burnham, Stevens, & Escudero, 2016).

4.6 Conclusion

The current study provided neural and behavioral evidence supporting the influence of musical expertise on the SL of distinct types of auditory stimuli, namely artificial speech (with or without prosodic characteristics) and melodies. Brain responses of musicians and non-musicians differed when exposed to prosodic and musical streams during the learning phase, at specific time windows: musicians showed an increased negativity 250-300 ms post-words onset and an enhanced positivity in two latency windows (i.e., 0-100 and 300-350 ms) in response to melodies. Nonetheless, musical training did not affect the segmentation of a non-prosodic speech stream. These findings suggest that the effects of musical training on the neural processing of auditory regularities occur selectively for streams with melodic features (e.g., melodies and sung words). Besides, whereas differences occur early in time during the processing of regularities in musical sounds, they emerge later for sung speech.

EEG measures also revealed that musical training impacts upon the processing of structural violations to the prosodic stream (viz., sung part-words, Experiment 1), with musicians showing increased positivities in 250-300 ms and 400-450 ms time windows. Hence, when learning is tested implicitly, there is only evidence of a training effect on prosodic words. Conversely, when learning was tested explicitly through behavioral tasks, musicians performed above chance in all ASL recognition tasks, irrespective of stimulus type, which is evidence of cross-domain plasticity. The divergent results in the implicit *vs.* explicit learning measures of learning indicate that the two measures might tap into distinct neurocognitive processes (which highlights the advantage of combining both in ASL research). Yet, this possibility needs to be clarified by further research. The relationship between brain and behavioral measures was confirmed in the prosodic experiment only: the capacity to behaviorally identify words and foils was associated with a more negative N250-300 during exposure to the prosodic stream (Experiment 1).

Together, these findings support the role of musical training in facilitating the computation of TPs between syllables/musical notes that ultimately lead to an enhanced sensitivity to auditory regularities and to an increased performance in SL recognition tasks. By showing that musical training modifies ASL, the present study supports the hypothesis that altered ASL mechanisms are at the core of the enhanced speech segmentation abilities observed in musicians. Thereby, these findings provide support to the claim of positive transfer effects from music to speech. The present work constitutes a relevant contribution to ASL research while shedding light on the neuroplasticity effects of musical training on auditory learning.

4.7 References

Abla, D., Katahira, K., & Okanoya, K. (2008). On-line assessment of statistical learning by event-related potentials. *Journal of Cognitive Neuroscience*, *20*(6), 952–964. http://doi.org/10.1162/jocn.2008.20058

Abla, D., & Okanoya, K. (2009). Visual statistical learning of shape sequences: An ERP study. *Neuroscience Research*, *64*(2), 185–190. http://doi.org/10.1016/j.neures.2009.02.013

Ahissar, M., & Hochstein, S. (1993). Attentional Control of Early Perceptual Learning. *Proceedings of the National Academy of Sciences*, *90*(12), 5718–5722.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8 month old infants. *Psychological Science*, *9*(4), 321–324. http://doi.org/http://doi.org/10.1111/1467-9280.00063

Astheimer, L. B., & Sanders, L. D. (2011). Predictability affects early perceptual processing of word onsets in continuous speech. *Neuropsychologia*, *49*(12), 3512–3516. http://doi.org/10.1016/j.neuropsychologia.2011.08.014

Batterink, L. J., & Paller, K. A. (2017). Online neural monitoring of statistical learning. *Cortex*, *90*, 31–45. http://doi.org/10.1016/j.cortex.2017.02.004

Beres, A. M. (2017). Time is of the Essence: A Review of Electroencephalography (EEG) and Event-Related Brain Potentials (ERPs) in Language Research. *Applied Psychophysiology Biofeedback*, *42*(4), 247–255. http://doi.org/10.1007/s10484-017-9371-3

Bermudez, P., & Zatorre, R. J. (2005). Differences in gray matter between musicians and nonmusicians. *Annals of the New York Academy of Sciences*, *1060*(2005), 395–399. http://doi.org/10.1196/annals.1360.057

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology*, *2*(MAY), 1–12. http://doi.org/10.3389/fpsyg.2011.00094

Bidelman, G. M., & Alain, C. (2015). Musical Training Orchestrates Coordinated Neuroplasticity in Auditory Brainstem and Cortex to Counteract Age-Related Declines in Categorical Vowel Perception. *Journal of Neuroscience*, *35*(3), 1240–1249. http://doi.org/10.1523/JNEUROSCI.3292-14.2015

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain and Cognition*, *77*(1), 1–10. http://doi.org/10.1016/j.bandc.2011.07.006

Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone Language Speakers and Musicians Share Enhanced Perceptual and Cognitive Abilities for Musical Pitch: Evidence for Bidirectionality between the Domains of Language and Music. *PLoS ONE*, *8*(4). http://doi.org/10.1371/journal.pone.0060676

Boh, B., Herholz, S. C., Lappe, C., & Pantev, C. (2011). Processing of complex auditory patterns in musicians and nonmusicians. *PLoS ONE*, *6*(7). http://doi.org/10.1371/journal.pone.0021458

Boutros, N. N., Korzyukov, O., Jansen, B., Feingold, A., & Bell, M. (2004). Sensory gating deficits during the mid-latency phase of information processing in medicated schizophrenia patients. *Psychiatry Research*, *126*(3), 203–215. http://doi.org/10.1016/j.psychres.2004.01.007

Carrión, R. E., & Bly, B. M. (2007). Event-related potential markers of expectation violation in an artificial grammar learning task. *NeuroReport*, *18*(2), 191–195. http://doi.org/10.1097/WNR.0b013e328011b8ae

Chobert, J., & Besson, M. (2013). Musical expertise and second language learning. *Brain Sciences*, *3*(2), 923–940. http://doi.org/10.3390/brainsci3020923

Chobert, J., François, C., Velay, J. L., & Besson, M. (2012). Twelve months of active musical training in 8-to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral Cortex*, *24*(4), 956–967. http://doi.org/10.1093/cercor/bhs377

Cunillera, T., Càmara, E., Toro, J. M., Marco-Pallares, J., Sebastián-Galles, N., Ortiz, H., ... Rodríguez-Fornells, A. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *NeuroImage*, *48*(3), 541–553. http://doi.org/10.1016/j.neuroimage.2009.06.069

Cunillera, T., Laine, M., & Rodríguez-Fornells, A. (2016). Headstart for speech segmentation: A neural signature for the anchor word effect. *Neuropsychologia*, *82*, 189–199. http://doi.org/10.1016/j.neuropsychologia.2016.01.011

Cunillera, T., Toro, J. M., Sebastián-Gallés, N., & Rodríguez-Fornells, A. (2006). The effects of stress and statistical cues on continuous speech segmentation: An event-related brain potential study. *Brain Research*, *1123*(1), 168–178.

http://doi.org/10.1016/j.brainres.2006.09.046

Daltrozzo, J., & Conway, C. M. (2014). Neurocognitive mechanisms of statistical-sequential learning: what do event-related potentials tell us? *Frontiers in Human Neuroscience*, *8*(June), 1–22. http://doi.org/10.3389/fnhum.2014.00437

de Diego-Balaguer, R., Rodríguez-Fornells, A., & Bachoud-Lévi, A.-C. (2015). Prosodic cues enhance rule learning by changing speech segmentation mechanisms. *Frontiers in Psychology*, *6*(September). http://doi.org/10.3389/fpsyg.2015.01478

de Diego-Balaguer, R., Toro, J. M., Rodriguez-Fornells, A., & Bachoud-Lévi, A. C. (2007). Different neurophysiological mechanisms underlying word and rule extraction from speech. *PLoS ONE*, *2*(11). http://doi.org/10.1371/journal.pone.0001175

Ding, N., Pan, X., Luo, C., Su, N., Zhang, W., & Zhang, J. (2018). Attention is required for knowledge-based sequential grouping: Insights from the integration of syllables into words. *The Journal of Neuroscience*, *38*(5), 1178–1188. http://doi.org/10.1523/JNEUROSCI.2606-17.2017

Dittinger, E., Barbaroux, M., D'Imperio, M., Jäncke, L., Elmer, S., & Besson, M. (2016). Professional music training and novel word learning: from faster semantic encoding to longer-lasting word representations. *Journal of Cognitive Neuroscience*, *28*(10), 1584–1602. http://doi.org/10.1162/jocn

Dittinger, E., Chobert, J., Ziegler, J. C., & Besson, M. (2017). Fast Brain Plasticity during Word Learning in Musically-Trained Children. *Frontiers in Human Neuroscience*, *11*(May), 1–16. http://doi.org/10.3389/fnhum.2017.00233

Dittinger, E., Valizadeh, S. A., Jäncke, L., Besson, M., & Elmer, S. (2018). Increased functional connectivity in the ventral and dorsal streams during retrieval of novel words in professional musicians. *Human Brain Mapping*, *39*(2), 722–734. http://doi.org/10.1002/hbm.23877

Ehrlé, N., & Samson, S. (2005). Auditory discrimination of anisochrony: Influence of the tempo and musical backgrounds of listeners. *Brain and Cognition*, *58*(1), 133–147. http://doi.org/10.1016/j.bandc.2004.09.014

Elmer, S., & Jäncke, L. (2018). Relationships between music training, speech processing, and word learning: a network perspective. *Annals of the New York Academy of Sciences*, 10–18. http://doi.org/10.1111/nyas.13581

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using

G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. http://doi.org/10.3758/BRM.41.4.1149

François, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex*, *23*(9), 2038–2043. http://doi.org/10.1093/cercor/bhs180

François, C., Cunillera, T., Garcia, E., Laine, M., & Rodriguez-Fornells, A. (2017). Neurophysiological evidence for the interplay of speech segmentation and word-referent mapping during novel word learning. *Neuropsychologia*, *98*(October), 56–67. http://doi.org/10.1016/j.neuropsychologia.2016.10.006

François, C., Jaillet, F., Takerkart, S., & Schön, D. (2014). Faster sound stream segmentation in musicians than in nonmusicians. *PLoS ONE*, *9*(7), 26–29. http://doi.org/10.1371/journal.pone.0101340

Francois, C., & Schön, D. (2010). Learning of musical and linguistic structures: Comparing event-related potentials and behavior. *NeuroReport*, *21*(14), 928–932. http://doi.org/10.1097/WNR.0b013e32833ddd5e

Francois, C., & Schön, D. (2011). Musical expertise boosts implicit learning of both musical and linguistic structures. *Cerebral Cortex*, *21*(10), 2357–2365. http://doi.org/10.1093/cercor/bhr022

François, C., & Schön, D. (2014). Neural sensitivity to statistical regularities as a fundamental biological process that underlies auditory learning: The role of musical practice. *Hearing Research*, *308*, 122–128. http://doi.org/10.1016/j.heares.2013.08.018

Gelfand, J. R., & Bookheimer, S. Y. (2003). Dissociating neural mechanisms of temporal sequencing and processing phonemes. *Neuron*, *38*(5), 831–842. http://doi.org/10.1016/S0896-6273(03)00285-X

George, E. M., & Coch, D. (2011). Music training and working memory: An ERP study. *Neuropsychologia*, *49*(5), 1083–1094. http://doi.org/10.1016/j.neuropsychologia.2011.02.001

Gilbert, C., Ito, M., Kapadia, M., & Westheimer, G. (2000). Interactions between attention, context and learning in primary visual cortex. *Vision Research*, *40*(10–12), 1217–1226. http://doi.org/10.1016/S0042-6989(99)00234-5

Gordon, E. E. (1989). *Advanced measures of music audiation.* Chicago, IL: G.I.A. Publications.

Graffar, M. (1956). Une méthode de classification d'échantillons de la population. *Courrier*, *6*, 455.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468–484.

Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., & Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology*, *42*(2), 180–190. http://doi.org/10.1111/j.1469-8986.2005.00272.x

Herholz, S. C., Lappe, C., & Pantev, C. (2009). Looking for a pattern: An MEG study on the abstract mismatch negativity in musicians and nonmusicians. *BMC Neuroscience*, *10*, 1–10. http://doi.org/10.1186/1471-2202-10-42

Hickok, G. (2012). The cortical organization of speech processing: Feedback control and predictive coding the context of a dual-stream model. *Journal of Communication Disorders*, *45*(6), 393–402. http://doi.org/10.1016/j.jcomdis.2012.06.004

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393–402. http://doi.org/10.1038/nrn2113

Hoffman, J. E. (1990). Event-related potentials and automatic and controlled processes. In J. J.W. Rohrbaugh, R. Parasuraman and R. Johnson (Ed.), *Event-Related Brain Potentials: Basic Issues and Applications* (Eds., pp. 145–157). New York: Oxford University Press.

Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). The effects of musical training on structural brain development: A longitudinal study. *Annals of the New York Academy of Sciences*, *1169*, 182–186. http://doi.org/10.1111/j.1749-6632.2009.04852.x

Jasper, H. H. (1958). The ten-twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 370–375.

Jonaitis, E. M. M., & Saffran, J. R. (2009). Learning harmony: The role of serial statistics. *Cognitive Science*, *33*(5), 951–968. http://doi.org/10.1111/j.1551-6709.2009.01036.x

Kishon-Rabin, L., Amir, O., Vexler, Y., & Zaltz, Y. (2001). Pitch discrimination: Are professional musicians better than non-musicians? *Journal of Basic and Clinical Physiology and Pharmacology*, *12*(2), 125–144. http://doi.org/10.1515/JBCPP.2001.12.2.125

Kramer, A. F., & Donchin, E. (1987). Brain Potentials as Indices of Orthographic and Phonological Interaction During Word Matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(1), 76–86. http://doi.org/10.1037/0278-7393.13.1.76

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Science and Society*, *11*(August), 599–606. http://doi.org/10.1038/nm2882

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*(11), 831–843. http://doi.org/10.1038/nrn1533

Kujala, T., Kallio, J., Tervaniemi, M., & Näätänen, R. (2001). The mismatch negativity as an index of temporal processing in audition. *Clinical Neurophysiology*, *112*(9), 1712–1719. http://doi.org/10.1016/S1388-2457(01)00625-3

Lijffijt, M., Lane, S. D., Meier, S. L., Boutros, N. N., Burroughs, S., Steinberg, J. L., ... Swann, A. C. (2009). P50, N100, and P200 sensory gating: Relationships with behavioral inhibition, attention, and working memory. *Psychophysiology*, *46*(5), 1059–1068. http://doi.org/10.1111/j.1469-8986.2009.00845.x

Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion*, *11*(5), 1021–1031. http://doi.org/10.1037/a0024521

Loui, P., Wessel, D. L., & Kam, C. L. H. (2010). Humans rapidly learn grammatical structure in a new musical scale. *Music Perception: An Interdisciplinary Journal*, *27*(5), 377–388. http://doi.org/10.1525/mp.2010.27.5.377.Humans

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique* (Second, Vol. 78). Cambridge, MA: MIT Press. http://doi.org/10.1118/1.4736938

Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, *18*(2), 199–211. http://doi.org/10.1162/jocn.2006.18.2.199

Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O., & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *Journal of Cognitive Neuroscience*, *23*(10), 2701–2715. http://doi.org/doi: 10.1162/jocn.2010.21585

Martín-Loeches, M., Sommer, W., & Hinojosa, J. A. (2005). ERP components reflecting stimulus identification: Contrasting the recognition potential and the early repetition effect (N250r). *International Journal of Psychophysiology*, *55*(1), 113–125. http://doi.org/10.1016/j.ijpsycho.2004.06.007

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience*, *7*(7), 703–704. http://doi.org/10.1038/nn1264

McNealy, K., Mazziotta, J. C., & Dapretto, M. (2010). The neural basis of speech parsing in children and adults. *Developmental Science*, *13*(2), 385–406. http://doi.org/10.1111/j.1467-7687.2009.00895.x

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, *219*(1–2), 36–47. http://doi.org/10.1016/j.heares.2006.05.004

Milovanov, R., Huotilainen, M., Välimäki, V., Esquef, P. A. A., & Tervaniemi, M. (2008). Musical aptitude and second language pronunciation skills in school-aged children: Neural and behavioral evidence. *Brain Research*, *1194*, 81–89. http://doi.org/10.1016/j.brainres.2007.11.042

Moreno, S., & Bidelman, G. M. (2014). Examining neural plasticity and cognitive benefit through the unique lens of musical training. *Hearing Research*, *308*, 84–97. http://doi.org/10.1016/j.heares.2013.09.012

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. http://doi.org/10.1093/cercor/bhn120

Münte, T. F., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience*, *3*(6), 473–478. http://doi.org/10.1038/nrn843

Näätänen, R. (1992). *Attention and Brain Function.* Lawrence Erlbaum Associates, Inc.

Näätänen, R. (2002). The perception of speech sounds by the human brain as reflected by the mismatch negativity brain response. *International Congress Series*, *1232*, 97–105. http://doi.org/10.1016/S0531-5131(01)00795-6

Näätänen, R., & Gaillard, A. W. K. (1983). 5 The Orienting Reflex and the N2 Deflection of the Event-Related Potential (ERP). In *Advances in psychology* (pp. 119–141). North-Holland. http://doi.org/0.1016/s0166-4115(08)62036-1

Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, *42*(4), 313–329. http://doi.org/10.1016/0001-6918(78)90006-9

Näätänen, R., Jacobsen, T., & Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): A review of the evidence. *Psychophysiology*, *42*(1), 25–32. http://doi.org/10.1111/j.1469-8986.2005.00256.x

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*, *118*(12), 2544–2590. http://doi.org/10.1016/j.clinph.2007.04.026

Nasreddine, Z., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., & Collin, I. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *American Geriatrics Society*, *53*(4), 695–699. http://doi.org/10.1111/j.1532-5415.2005.53221.x

Neisser, U. (1967). *Cognitive psychology: Classic edition.* Psychology Press.

Newport, E. L. (2016). Statistical language learning: computational, maturational, and linguistic constraints. *Language and Cognition*, *8*(03), 447–461. http://doi.org/10.1017/langcog.2016.20

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. http://doi.org/10.1016/0028-3932(71)90067-4

Ong, J. H., Burnham, D., Stevens, C. J., & Escudero, P. (2016). Naïve learners show cross-domain transfer after distributional learning: The case of lexical and musical pitch. *Frontiers in Psychology*, *7*(AUG), 1–10. http://doi.org/10.3389/fpsyg.2016.01189

Pallesen, K. J., Brattico, E., Bailey, C. J., Korvenoja, A., Koivisto, J., Gjedde, A., & Carlson, S. (2010). Cognitive control in auditory working memory is enhanced in musicians. *PLoS ONE*, *5*(6). http://doi.org/10.1371/journal.pone.0011120

Palmer, S. D., & Mattys, S. L. (2016). Speech segmentation by statistical learning is supported by domain-general processes within working memory. *Quarterly Journal of Experimental Psychology*, *69*(12), 2390–2401. http://doi.org/10.1080/17470218.2015.1112825

Paraskevopoulos, E., Kuchenbuch, A., Herholz, S. C., & Pantev, C. (2012). Statistical learning effects in musicians and non-musicians: An MEG study. *Neuropsychologia*, *50*(2), 341–349. http://doi.org/10.1016/j.neuropsychologia.2011.12.007

Patel, A. D. (1998). Syntactic Processing in Language and Music: Different Cognitive Operations, Similar Neural Resources? *Music Perception: An Interdisciplinary Journal*, *16*(1), 27–42. http://doi.org/10.2307/40285775

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, *6*(7), 674–681. http://doi.org/10.1038/nn1082

Patel, A. D. (2012). Language, music, and the brain: a resourche-sharing framework. *Language and Music as Cognitive Systems*, 204–223. http://doi.org/10.1111/j.1756-8765.2009.01080.x

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, *10*(6), 717–733.

Patel, A., & Morgan, E. (2017). Exploring Cognitive Relations Between Prediction in Language and Music. *Cognitive Science*, *41*, 303–320. http://doi.org/10.1111/cogs.12411

Patel, S. H., & Azzam, P. N. (2005). Characterization of N200 and P300: Selected studies of the Event-Related Potential. *International Journal of Medical Sciences*, *2*(4), 147–154. http://doi.org/10.7150/ijms.2.147

Paulmann, S., & Pell, M. D. (2010). Contextual influences of emotional speech prosody on face processing: How much is enough? *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 230–242. http://doi.org/10.3758/CABN.10.2.230

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233–238. http://doi.org/10.1016/j.tics.2006.03.006

Pinheiro, A. P., Vasconcelos, M., Dias, M., Arrais, N., & Gonçalves, Ó. F. (2015). The music of language: An ERP investigation of the effects of musical training on emotional prosody processing. *Brain and Language*, *140*, 24–34. http://doi.org/10.1016/j.bandl.2014.10.009

Reinke, K. S., He, Y., Wang, C., & Alain, C. (2003). Perceptual learning modulates sensory evoked response during vowel segregation. *Cognitive Brain Research*, *17*(3), 781–791. http://doi.org/10.1016/S0926-6410(03)00202-7

Rellecke, J., Palazova, M., Sommer, W., & Schacht, A. (2011). On the automaticity of emotion processing in words and faces: Event-related brain potentials evidence from a superficial task. *Brain and Cognition*, *77*(1), 23–32. http://doi.org/10.1016/j.bandc.2011.07.001

Ritter, W., Paavilainen, P., Lavikainen, J., Reinikainen, K., Alho, K., Sams, M., & Näätänen, R. (1992). Event-related potentials to repetition and change of auditory stimuli, *83*, 306–321.

Röder, B., Rösler, F., Hennighausen, E., & Näcker, F. (1996). Event-related potentials during auditory and somatosensory discrimination in sighted and blind human subjects.

Cognitive Brain Research, *4*(2), 77–93. http://doi.org/10.1016/0926-6410(96)00024-9

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. doi: 10.1126/science.274.5294.1926

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27–52. http://doi.org/10.1016/S0010-0277(98)00075-4

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*(4), 606–621. http://doi.org/10.1006/jmla.1996.0032

Saffran, J. R., Senghas, A., & Trueswell, J. C. (2001). The acquisition of language by children. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(23), 12874–12875. http://doi.org/10.1073/pnas.231498898

Sams, M., Alho, K., & Näätänen, R. (1983). Sequential effects on the ERP in discriminating two stimuli. *Biological Psychology*, *17*(1), 41–58. http://doi.org/10.1001/jamaoto.2016.3292

Sanders, L. D., Ameral, V., & Sayles, K. (2009). Event-related potentials index segmentation of nonsense sounds. *Neuropsychologia*, *47*(4), 1183–1186. http://doi.org/10.1016/j.neuropsychologia.2008.11.005

Sanders, L. D., Newport, E. L., & Neville, H. J. (2002). Segmenting nonsense: An event-related potential index of perceived onsets in continuous speech. *Nature Neuroscience*, *5*(7), 700–703. http://doi.org/10.1038/nn873

Schlaug, G., Norton, A., Overy, K., & Winner, E. (2005). Effects of music training on the child's brain and cognitive development. *Annals of the New York Academy of Sciences*, *1060*, 219–230. http://doi.org/10.1196/annals.1360.015

Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition*, *106*(2), 975–983. http://doi.org/10.1016/j.cognition.2007.03.005

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, *41*(3), 341–349. http://doi.org/10.1111/1469-8986.00172.x

Schultz, B. G., Stevens, C. J., Keller, P. E., & Tillmann, B. (2013). A Sequence Identification Measurement Model to Investigate the Implicit Learning of Metrical Temporal Patterns.

PLoS ONE, *8*(9), 1–13. http://doi.org/10.1371/journal.pone.0075163

Selchenkova, T., François, C., Schön, D., Corneyllie, A., Perrin, F., & Tillmann, B. (2014). Metrical presentation boosts implicit learning of artificial grammar. *PLoS ONE*, *9*(11). http://doi.org/10.1371/journal.pone.0112233

Shahin, A., Bosnyak, D. J., Trainor, L. J., & Roberts, L. E. (2003). Enhancement of Neuroplastic P2 and N1c Auditory Evoked Potentials in Musicians. *Journal of Neuroscience*, *23*(13), 5545–5552. http://doi.org/10.1523/JNEUROSCI.23-13-05545.2003

Sheehan, K. A., McArthur, G. M., & Bishop, D. V. M. (2005). Is discrimination training necessary to cause changes in the P2 auditory event-related brain potential to speech sounds? *Cognitive Brain Research*, *25*(2), 547–553. http://doi.org/10.1016/j.cogbrainres.2005.08.007

Shook, A., Marian, V., Bartolotti, J., & Schroeder, S. (2013). Musical experience influences novel language learning. *The American Journal of Psychology*, *126*(1), 95–104. http://doi.org/10.5406/amerjpsyc.126.1.0095

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, *81*, 105–120. http://doi.org/10.1016/j.jml.2015.02.001

Simões, M. R., Freitas, S., Santana, I., Firmino, H., Martins, C., Nasreddine, Z., & Vilar, M. (2008). Montreal cognitive assessment (MoCA): versão final portuguesa. *Coimbra: Serviço de Avaliação Psicológica, Faculdade de Psicologia e de Ciências Da Educação Da Universidade de Coimbra.*

Sluming, V., Barrick, T., Howard, M., Cezayirli, E., Mayes, A., & Roberts, N. (2002). Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. *NeuroImage*, *17*(3), 1613–1622. http://doi.org/10.1006/nimg.2002.1288

Snyder, J. S., Alain, C., & Picton, T. W. (2006). Effects of attention on neuroelectric correlates of auditory stream segregation. *Journal of Cognitive Neuroscience*, *18*(1), 1–13. http://doi.org/10.1162/089892906775250021

Strait, D. L., & Kraus, N. (2011). Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Frontiers in Psychology*, *2*(113), 1–10. http://doi.org/10.3389/fpsyg.2011.00113

Tervaniemi, M., Kruck, S., De Baene, W., Schröger, E., Alter, K., & Friederici, A. D. (2009). Top-down modulation of auditory processing: Effects of sound context, musical expertise and attentional focus. *European Journal of Neuroscience*, *30*(8), 1636–1642. http://doi.org/10.1111/j.1460-9568.2009.06955.x

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding Speech Prosody: Do Music Lessons Help? *Emotion*, *4*(1), 46–64. http://doi.org/10.1037/1528-3542.4.1.46

Tremblay, K., Kraus, N., & McGee, T. (1998). The time course of auditory perceptual learning: Neurophysiological changes during speech-sound training. *NeuroReport*, *9*(16), 3557–3560. http://doi.org/10.1097/00001756-199811160-00003

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*(4), 552–564. http://doi.org/10.1037/0096-3445.134.4.552

Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit Perceptual Anticipation Triggered by Statistical Learning. *Journal of Neuroscience*, *30*(33), 11177–11187. http://doi.org/10.1523/JNEUROSCI.0858-10.2010

Vaquero, L., Ramos-Escobar, N., François, C., Penhune, V., & Rodríguez-Fornells, A. (2018). White-matter structural connectivity predicts short-term melody and rhythm learning in non-musicians. *NeuroImage*, *181*, 252–262. http://doi.org/10.1016/j.neuroimage.2018.06.054

Vasuki, P. R. M., Sharma, M., Ibrahim, R., & Arciuli, J. (2017a). Statistical learning and auditory processing in children with music training: An ERP study. *Clinical Neurophysiology*, *128*(7), 1270–1281. http://doi.org/10.1016/j.clinph.2017.04.010

Vasuki, P. R. M., Sharma, M., Ibrahim, R. K., & Arciuli, J. (2017b). Musicians' Online Performance during Auditory and Visual Statistical Learning Tasks. *Frontiers in Human Neuroscience*, *11*(114), 1–12. http://doi.org/10.3389/fnhum.2017.00114

Werheid, K., Alpay, G., Jentzsch, I., & Sommer, W. (2005). Priming emotional facial expressions as evidenced by event-related brain potentials. *International Journal of Psychophysiology*, *55*(2), 209–219. http://doi.org/10.1016/j.ijpsycho.2004.07.006

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*(4), 420–422. http://doi.org/10.1038/nn1872

Chapter 5. Does attention affect ASL in musicians? EEG and behavioral insights.

5.1 Abstract

Previous ERPs and behavioral reports revealed that musicians have enhanced SL abilities, which are at the core of speech segmentation and language-learning skills. These studies tested SL abilities when subjects were actively paying attention to continuous sound streams. However, performance in SL tasks was found to decrease under high vs. low attentional demands, whereas musical training was found to enhance auditory selective attention. Hence, it is still not clear whether musicians' advantages in SL tasks persist under pre-attentive conditions (i.e., when sounds are task-irrelevant). The present study probed whether musical training affects how listeners learn different types of auditory structures (i.e., speech and musical streams) when attention is focused elsewhere. By examining ERPs of the EEG, the current study compared musicians and non-musicians during online SL of prosodic (i.e., sung words), non-prosodic (i.e., words with flat contour) and musical (i.e., tri-tone melodies) streams while participants were actively engaged in a non-demanding visual task. Additionally, learning outcomes were evaluated with implicit (i.e., ERP responses to structural violations of the auditory sequences) and explicit (i.e., behavioral performance in LDTs) measures. The ERP results revealed that musical training modulates ASL as a function of stimulus type: musicians showed larger positivities within 50 ms and 550-600 ms post-melodies onset during learning, but the two groups did not differ when processing regularities in artificial speech input (i.e., with or without prosodic modulations). Musical training also modulated the ERP responses to new melodies that did not match the structure of the musical stream. Conversely, musicians and non-musicians performed similarly in the SL recognition tests. Together, these findings indicate that, when attentional resources are not directed to the sounds, musical training modulates automatic neural responses to musical streams during SL, but does not bring behavioral advantages to explicit stream segmentation. These results shed light on the effects of music-related neuroplasticity and, specifically, on the conditions in which there is a positive transfer of training from music to auditory stream segmentation.

Keywords: ASL, musical training, ERPs, attention, stimulus-type

5.2 Introduction

5.2.1 Musical training as a model of neuroplasticity

A robust body of evidence indicates that the long-term multimodal training characterizing skilled musicians leads to changes in cortical (e.g., Bermudez & Zatorre, 2005; Foster & Zatorre, 2010; Hyde et al., 2009; Pantev et al., 1998; Schneider et al., 2002) and subcortical (Musacchia, Sams, Skoe, & Kraus, 2007; Parbery-Clark, Skoe, & Kraus, 2009; Wong, Skoe, Russo, Dees, & Kraus, 2007) brain regions involved in auditory processing. These changes have been related to behavioral benefits in auditory tasks such as an increased ability to process the acoustic features of sounds (Ehrlé & Samson, 2005; Micheyl, Delhommeau, Perrot, & Oxenham, 2006), namely enhanced pitch discrimination in speech and music stimuli (Magne, Schön, & Besson, 2006; Marques, Moreno, Castro, & Besson, 2007; Moreno et al., 2009; Schön, Magne, & Besson, 2004). At the neural level, musicians show larger amplitude of the N1 and P2 ERPs components of the EEG in response to synthetic and instrumental sounds (Shahin, Bosnyak, Trainor, & Roberts, 2003), larger N1m (the magnetic counterpart of the auditory N1) to piano notes than to pure tones (Pantev et al., 1998), enhanced subcortical encoding of pitch (i.e., higher FFRs, see Musacchia et al., 2007; Wong et al., 2007) and enhanced pre-attentive processing of chords (e.g., larger MMN – a neural correlate of pre-attentive deviance detection - and MMNm, see Brattico et al., 2009; Koelsch, Schröger, & Tervaniemi, 1999). The effects of musical training are not restricted to bottom-up changes in the processing of acoustic stimuli but also involve changes in higher-level cognitive functions such as selective attention (Strait & Kraus, 2011) or verbal memory (Ho, Cheung, & Chan, 2003).

5.2.2 The effects of musical training on attention

The regular and intensive practice of a musical instrument requires the constant integration of visual and auditory information with high temporal precision (Herholz & Zatorre, 2012). The fast integration of multimodal information relies on abilities related to cognitive control, namely the selective allocation of attention to target input (Strait & Kraus, 2011). Because of the high demands put on these functions during musical rehearsal and performance, some authors have claimed that musical training leads to improved executive functions (e.g., Bialystok & DePape, 2009; Hannon & Trainor, 2007), namely selective auditory attention (Fujioka, Ross, Kakigi, Pantev, & Trainor, 2006; Scott, 1992). Musicians show faster reaction times to target auditory cues in auditory tasks (Strait, Kraus, Parbery-Clark, & Ashley, 2010) and are more sensitive to differences in harmonic expectations (e.g., variations in chord progressions under attention demanding conditions) compared to non-musicians (Loui & Wessel, 2007). Children with musical training also demonstrate an increased ability to sustain focused attention on academic tasks (Scott, 1992) and musical sounds (Fujioka et al., 2006). At the neural level, musical training modulates the latency and amplitude of the P300 – an ERP component indexing deviance detection and the allocation of attentional resources to target stimuli (Polich, 2003) – during auditory processing: musicians showed earlier and larger P300 to DEV tones interspersed with STD tones (see George & Coch, 2011). Musical training is also related to enhanced cognitive control as indicated by larger blood oxygenation-level dependent activation in sustained attention and cognitive control networks (e.g., lateral prefrontal cortex, anterior cingulate gyrus, lateral parietal cortex, posterior dorsal prefrontal cortex) during working memory tasks (Pallesen et al., 2010). Together, these results suggest that selective attention might play an important role in musicians' capacity to process auditory stimuli (see Besson, Chobert, & Marie, 2011; Strait & Kraus, 2011).

Other studies suggest that differences in attention do not account for the enhanced auditory processing skills observed in musicians. For example, Marie and collaborators (Marie, Delogu, Lampis, Belardinelli, & Besson, 2011) found that the amplitude of the P200 ERP component to changes in the metric of syllables was larger in musicians, irrespective of the attentional focus (i.e., on metric vs. on semantic violations). Further, no differences were observed in automatic attention orienting as a function of musical training (Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005). Specifically, the P3a response (an ERP component indexing automatic attention orienting to salient stimuli; see Polich, 2007; Simons, Graham, Miles, & Chen, 2001) to pitch variations in spectrally complex tones did not differ between musicians and non-musicians. Along the same lines, Baumann and colleagues (Baumann, Meyer, & Jäncke, 2008) suggested that altered auditory ERPs in musical experts are due to musical training and not to selective attention: musicians exhibit larger N1 and P2 in response to sine-wave and instrumental tones irrespective of the attentional focus (towards or away from the sounds). Thus, it is still not clear whether the positive effects of musical training on auditory processing are mediated by attention or not (Moreno et al., 2009; Strait et al., 2010).

5.2.3 Positive transfer effects between music and speech

Besides the ongoing debate on the role of attention in the observed benefits of musical training, there is a growing body of evidence indicating common processing mechanisms between music and speech (Bidelman, Gandour, & Krishnan, 2011; Koelsch et al., 2002; Slevc, Rosenberg, & Patel, 2009). Transfer effects from music to speech were documented by studies showing that musicians excel in several speech-related skills: pitch perception (i.e., vowels generated with different levels of reverberation; Bidelman & Krishnan, 2010), phonemic processing (e.g., words rhymes; Anvari, Trainor, Woodside, & Levy, 2002), prosody processing (e.g., spoken sentences; Lima & Castro, 2011; Pinheiro, Vasconcelos, Dias, Arrais, & Gonçalves, 2015; Schön et al., 2004), speech-in-noise perception (i.e., spoken syklable played in background noise or quiet; Parbery-Clark et al., 2009), word learning (e.g., monosyllabic Thai words; Dittinger et al., 2016; Dittinger, Valizadeh, Jäncke, Besson, & Elmer, 2018), or second language proficiency (i.e., words, sentences and passages; Slevc & Miyake, 2006). Besides, when compared to non-musician children, children with musical training demonstrate better detection of pitch changes in speech (Kraus & Chandrasekaran, 2010; Magne et al., 2006; Moreno et al., 2009) and increased performance in language tasks such as reading (e.g., Moreno et al., 2009). This evidence supports the impact of short periods of training on brain plasticity, corroborating the notion of positive transfer effects from music to speech. These benefits are supported both by neuroimaging and electrophysiological findings showing that cortical (Chandrasekaran, Krishnan, & Gandour, 2009; Marie, Magne, & Besson, 2011; Pantev, Roberts, Schulz, Engelien, & Ross, 2001; Schön et al.,

2004) and subcortical (Bidelman et al., 2011; Wong et al., 2007) brain regions involved in auditory processing are modified by musical training, which may account for an enhanced speech processing capacity.

5.2.4 The role of ASL in speech segmentation: the case of musicians

The relationship between long-term musical training and speech in general has been in the spotlight in recent decades. Research in this field has emphasized speech segmentation – the capacity to extract meaningful words from continuous speech – as a key aspect of language learning and processing (Saffran, Newport, & Aslin, 1996; Saffran, Senghas, & Trueswell, 2001). Speech segmentation occurs in natural language environments based on the modulation of acoustic cues at word boundaries (e.g., pauses, stresses, etc.) and the statistical structure of the speech streams (Saffran et al., 1996). As often there are no cues at word boundaries, the segmentation of continuous speech greatly depends on SL - the implicit extraction of statistical regularities from the input (Kuhl, 2004; Saffran et al., 1996). SL allows the prediction of incoming sensory input (Saffran, Johnson, Aslin, & Newport, 1999; Siegelman & Frost, 2015).

Speech segmentation by SL operates based on the computation of TPs between syllable pairs (Saffran et al., 2001). In their seminal work, Saffran and collaborators (Aslin, Saffran, & Newport, 1998; Saffran et al., 1996) found that both adults and infants were able to segment new artificial language streams based only on TPs. The work of Saffran and collaborators was recently extended by François and colleagues (François, Chobert, Besson, & Schön, 2013; François, Jaillet, Takerkart, & Schön, 2014; François & Schön, 2011) using both behavioral and electrophysiological measures. The authors aimed to explore if musical expertise facilitates the SL of artificial sung languages. They found that adult musicians tend to outperform (i.e., show a higher percentage of correct responses) non-musicians in 2AFC tasks examining the learning of the prosodic and semantic aspects of the sung language. At the neural level, musicians showed alterations in the N1 and P2 components to words[6] and melodies during the linguistic and musical tests. Moreover, musicians showed an increased sensitivity to the statistical structure of the auditory stream – while being exposed to it -, as evidenced by alterations in the amplitude of the later N400-like component. The auditory N1 and the P2 are two early components observed within 200 ms post-

---

[6] In the context of the present work, ‘words’ refer to the constituents units of artificial speech streams.

sound onset that frequently emerge in the course of stimulus processing one after the other (Beres, 2017). These components are frequently found in artificial language learning paradigms not because they are exclusive signatures of speech processing, but because they index the processing of the sensory features of sounds (Tremblay, Kraus, & McGee, 1998). The N1 peaks at approximately 100 ms and is more prominent over fronto-central electrode sites (Abla, Katahira, & Okanoya, 2008; Sanders, Newport, & Neville, 2002; Vasuki, Sharma, Ibrahim, & Arciuli, 2017b). The N1 has been related to sensory processing mechanisms (Luck, 2005) and to predictive mechanisms implicated in speech decoding and comprehension (Heinks-Maldonado, Mathalon, Gray, & Ford, 2005). The P2 indexes sound categorization and discrimination processes (Ritter et al., 1992; Sheehan, McArthur, & Bishop, 2005), which are relevant for the perception of auditory regularities (de Diego-Balaguer, Toro, Rodriguez-Fornells, & Bachoud-Lévi, 2007; Snyder, Alain, & Picton, 2006). This component peaks at approximately 200 ms, and is also more prominent over fronto-central electrodes (Abla et al., 2008; Abla & Okanoya, 2009; Cunillera et al., 2009; Cunillera, Toro, Sebastián-Gallés, & Rodríguez-Fornells, 2006; de Diego-Balaguer, Rodríguez-Fornells, & Bachoud-Lévi, 2015; Reinke, He, Wang, & Alain, 2003; Vasuki et al., 2017b). The N400-like component (which is distinct from the classical N400 component; Marta Kutas & Hillyard, 1980) is thought to represent an ERP signature of lexical search and auditory stream segmentation (Abla et al., 2008; Cunillera et al., 2009, 2006; de Diego-Balaguer et al., 2007; François, Cunillera, Garcia, Laine, & Rodriguez-Fornells, 2017; François et al., 2014; Sanders et al., 2002; Vasuki et al., 2017b). In comparison with the classical N400, the N400-like component presents a fronto-central topographical distribution and typically peaks between 200 and 500 ms post-stimulus onset in auditory learning tasks (Cunillera et al., 2009; Dittinger, Chobert, Ziegler, & Besson, 2017; François et al., 2017, 2014; Sanders et al., 2002; Vasuki et al., 2017b).

François and colleagues extended these investigations to young children by examining speech segmentation skills before and after training programs consisting of either two years of musical training or painting lessons. They found that only musician children showed behavioral improvements in speech segmentation. Further, after two years of musical training, these children evidenced alterations in a fronto-central N400-like component in response to words - a result that parallels previous findings with adults (François & Schön, 2014). These findings were interpreted as evidence that musical training may increase the strength of representations of the statistical regularities characterizing speech input, which in turn may explain the advantage of musicians in second language and word learning tasks (Dittinger et al., 2016; François et al., 2014; François & Schön, 2011). These results are in accordance with behavioral and electrophysiological evidence (N1 and N400-like latencies modulation) demonstrating that musicians show increased ASL skills (Vasuki, Sharma, Ibrahim, & Arciuli, 2017). Nonetheless, it is worth mentioning that the studies reported above relied on experimental designs in which participants' attention was actively focused on the auditory input. Thereby, these studies do not clarify whether and how attention modulates the positive impact of musical training on ASL.

5.2.5 The interplay between attention and ASL

Studies of SL that have actively manipulated attention are scarce and the extent to which attention is required for SL to occur remains controversial. The existing evidence suggests that SL is automatic (Fiser & Aslin, 2002a, 2002b; Saffran et al., 1996; Turk-Browne, Jungé, & Scholl, 2005) and incidental (Fiser & Aslin, 2005), occurring without conscious awareness (Perruchet & Pacton, 2006). However, this is not to say that SL outcomes do not vary with manipulations in attentional load or selective attention. Early studies claimed that SL is only possible if attention is actively oriented towards the stimuli to be learned (Ahissar & Hochstein, 1993; Gilbert, Ito, Kapadia, & Westheimer, 2000). When the performance on dual task experimental settings (e.g., high attentional load) is compared with the results of active, single-task settings, for example, SL is increased in the latter than in the former conditions, irrespective of the sensory modality of the stimulation (Toro, Sinnett & Soto-Faraco, 2005; Turk-Browne et al., 2005). Toro, Sinnett and Soto-Faraco (Toro, Sinnett, & Soto-Faraco, 2005) attempted to clarify this issue by testing the listeners' ability to segment an artificial speech stream during passive listening while a concurrent task was being performed. They examined the impact of high attentional load on speech segmentation using three types of concurrent tasks: a) detecting changes in auditory noise coming from a concurrent stream, b) detecting pitch changes in the actual stream, and c) detecting picture repetitions on a concurrent visual stream. The results of recognition tests showed that, independently of the nature of the task competing for attention, participants' performance was at chance level as a result of diverted attention. Conversely, the performance was significantly above chance in the group of participants who only passively[7] listened to the speech stream. Similarly, a recent EEG study from Ding and colleagues (Ding et al., 2017) revealed that attention is a necessary condition for the brain to chunk sequences of syllables into words. Therefore, while the encoding of syllable units seems to be automatic, the chunking of syllables into more complex and meaningful sequences (e.g., words) seems to be dependent on attention.

5.2.6 SL in musicians under pre-attentive task requirements

So far, the existing studies that compared musicians and non-musicians have examined ASL abilities by explicitly asking listeners to focus their attention on the speech input. One exception is the magnetoencephalography (MEG) study of Paraskevopoulos and colleagues (Paraskevopoulos, Kuchenbuch, Herholz, & Pantev, 2012), who investigated neural responses to auditory DEVs embedded in an oddball paradigm with tone-sequences when participants' attention was focused on concurrent visual stimuli. This study revealed no differences between musicians and non-musicians in the MMN elicited in response to DEV tone-triplets. The MMN component emerges in response to perceptual changes (e.g., physical or abstract, simple or complex, see Boh, Herholz, Lappe, & Pantev, 2011; Chobert, François, Velay, & Besson, 2012; Herholz, Lappe, Knief, & Pantev, 2009; Näätänen, Gaillard, & Mäntysalo, 1978) in a regular continuum of sounds (Näätänen, 2002; Näätänen, Paavilainen, Rinne, & Alho, 2007). This early component - that peaks between 150-250 ms and has fronto-central distribution (Duncan et al., 2009) – indexes automatic deviance detection mechanisms (Näätänen, Jacobsen, & Winkler, 2005), shown to be enhanced in musicians when processing changes in complex sounds (e.g., melodies, Tervaniemi, Rytkönen, Schröger, Ilmoniemi, & Näätänen, 2001; chords, Koelsch, Schröger, & Tervaniemi, 1999b). Despite the absence of effects of training on the MMN amplitude, the abovementioned study (Paraskevopoulos et al., 2012) reported increased P50 amplitude to DEV compared to STD triplets in musicians. Behaviorally, neither musicians nor non-musicians were able to learn the statistical structure of the tonal stream. Overall, these findings reveal several important aspects: first, ASL seems to be affected when listeners' attention is not actively focused on the auditory stream (behavioral results); second, even when sounds are task-irrelevant, the brain seems to encode TPs between tones providing a template for the forthcoming auditory input (ERP

---

[7] It is worth noting that, despite being called "passive", the passive listening condition was a condition in which participants were instructed to pay attention to the auditory stream without having any particular information regarding its structure.

modulations). Therefore, in pre-attentive conditions, musical training does not bring advantages to learning varying pitch-sequences, but it might alter the way in which the brain processes TPs among tone pairs.

Research combining neural and behavioral measures of ASL under preattentive task requirements is almost non-existent. To the best of our knowledge, the study of Paraskevopoulos et al. (2012) was the first using ERP measures to examine the effects of musical training on the pre-attentive SL of tone-sequences. There are no similar studies examining the role of training in the SL of other types of auditory structures, namely, artificial speech, under low attentional conditions. Besides, most studies investigating the role of musical training in speech segmentation tasks used sung materials that have specific features that should be considered. First, words with melodic contours facilitate the segmentation of speech streams since the parallel mapping of linguistic and prosodic cues promotes the computation of TPs between syllables pairs and phonological discrimination (Schön et al., 2008). Second, musical training facilitates the processing of sounds with melodic variations (François & Schön, 2010). It remains to be clarified whether musical training affects speech segmentation when stimuli are devoid of melodic contours and when participants are not actively attending to this input. A systematic investigation of whether the positive impact of musical training on ASL (François et al., 2014; François & Schön, 2011; Vasuki, Sharma, Ibrahim, & Arciuli, 2017a; Vasuki et al., 2017b) depends on stimulus characteristics is warranted. This is particularly relevant to clarify whether the effects of musical training on SL are due to within-domain (i.e., musical sounds) and/or cross-domain (i.e., other types of sounds) neuroplasticity.

5.2.7 The current study and hypotheses

In an attempt to systematically investigate the influence of stimulus type and musical training on ASL under pre-attentive conditions, we compared the EEG responses to linguistic (e.g., prosodic and non-prosodic words) and musical (e.g., melodies) SL tasks of musicians and non-musicians. The ERP technique was used as it provides an on-line neurophysiological measure of learning with ms precision that is advantageous for two reasons: (1) it is sensitive to implicit mechanisms that cannot be evaluated with behavioral measures, and (2) it provides a more dynamic view on the neurofunctional mechanisms involved in ASL and how they unfold over time

(François et al., 2017, 2014). Participants were presented with distinct experiments composed by three experimental phases: a learning phase, an implicit test phase and an explicit test phase. During the learning phase, participants heard streams of artificial sung words (prosodic experiment [Experiment 1]), non-prosodic words (linguistic experiment [Experiment 2]) and tri-tone melodies (musical experiment [Experiment 3]). The analysis of the EEG data aimed to characterize the electrophysiological responses to words/melodies while learning took place and, specifically, to examine changes in the neural responses as a function of exposure time or stimulus repetitions. In the implicit test phase, legal items (e.g., items presented during the learning phase) where pseudo-randomly intermixed with new legal and illegal items that violated the statistical structure of the auditory streams. The aim of this experimental phase was to overcome methodological limitations imposed by post-learning measures, namely the impossibility to test participants' capacity to implicitly detect statistical mismatches between previously learned and new information. Moreover, it allowed testing participants' SL abilities implicitly and to study if musical training can benefit the detection of deviations in the statistical structure of the input. Since the neurophysiological correlates of SL were studied under pre-attentive conditions and the data was collected in a unique experimental session, it was not possible to acquire behavioral data for all the experiments. An explicit, post-learning SL task after the first streams to be tested would have made participants aware of the purpose of the auditory stimulation. Thus, we tested participants' recognition of the prosodic words and their ability to learn the linguistic and musical dimensions of the prosodic stream with LDTs.

To the best of our knowledge, there is no previous report systematically examining the neural correlates of ASL for different types of auditory stimuli under pre-attentive conditions. Therefore, our hypotheses regarding expected ERP responses during learning are exploratory. However, taking into account that the N400-like component has been consistently found in response to streams' items while ASL unfolds (François et al., 2017, 2014; Vasuki et al., 2017b), we expected this ERP component to be elicited in response to stimulus onset and to be modulated by musical training (e.g., larger amplitude in musicians). Specifically, we expected to observe cross-domain effects, with ERP differences between groups emerging for all auditory streams tested (i.e., independently of their acoustic properties). This evidence would confirm the idea that musical training modulates ASL in general. Also, we expected to observe ERP differences between groups as a function of exposure time (see, for example, François et al., 2014), namely a linear increased in N400-like amplitude in non-musicians as opposed to an increase, followed by a decrease in N400's amplitude (inverted U-shape learning curve) in musicians.

Regarding the implicit test phase, MMN/N200 components were expected in response to structural violations (François et al., 2016). The N200, along with the MMN, is an electrophysiological index of mismatch detection (Näätänen & Gaillard, 1983; Sams, Alho, & Näätänen, 1983), and its amplitude increases (i.e., it becomes more negative) to violations in the grammaticality of sequences (Carrión & Bly, 2007; Selchenkova et al., 2014). Specifically, musicians were expected to show larger amplitudes in response to these violations compared to non-musicians. Behaviorally, if musical training benefits speech segmentation independently of attention, we predicted that musicians would learn more items and would more accurately recognize them in the LDT. Moreover, we hypothesized that individual learning-related differences in the ERP components would be correlated with the performance in the LDTs, i.e., enhanced ERP responses would be associated with a higher number of correct responses in the recognition tasks.

5.3 Material and Methods

5.3.1 Participants

In order to ensure that the statistical analyses had sufficient power, we performed an a priori power analysis using the G*Power software (Faul, Erdfelder, Buchner, & Lang, 2009) and selecting the following criteria: F tests family - repeated-measures ANOVA; within-between participants design with 5 measures; non-sphericity correction (ε) of 0.25; effect size = 0.5 at a significance level of α= 0.05; power of 1 - β = 0.95. The analysis revealed that a minimum of 12 participants per group were needed to achieve the desired statistical power level in the data analyses.

Twenty musicians (13 women; 20.75 ± 2.95 years [*M* ± *SD*]) and 20 non-musicians (12 women; 22.8 ± 4.49 years [M ± *SD*]) participated in this study. Groups did not differ in age, *t*(38) = 1.71, *p* = .096 (see Table 8). The group of musicians was composed of participants who had a minimum of 8 years of musical experience (years of training: 12.35 ± 3.31; training onset: 8.45 ±

2.89 years [M ± *SD*]) that involved both formal musical classes and an almost daily practice of at least one musical instrument (see Table 9). The non-musicians were participants who never had musical classes or played a musical instrument. All participants met the following inclusion criteria: were right-handed (Edinburgh Handedness Inventory; Oldfield, 1971), had normal hearing, had normal to corrected-to-normal vision, had no history of psychopathological/neurological disorders or substance abuse, were native speakers of European Portuguese, were monolingual and never learned German. Participants were assessed with the Advanced Measures of Music Audiation (AMMA; Gordon, 1989) in order to ensure that the two groups differed in their musical aptitude abilities (see Table 8). Additionally, in order to ensure that the groups were similar regarding socioeconomic status and cognitive functioning, participants were evaluated using the Graffar Scale (Graffar, 1956) and The Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005; Simões et al., 2008), respectively. The Ethics Committee for Life and Health Sciences of the University of Minho approved the present study.

TABLE 8

Socio-demographic and Other Relevant Characteristics of the Sample

	Musicians (*N* = 20)	Non-Musicians (*N* = 20)	Statistics	
	M (SE)	*M (SE)*	*t*	*p*
Age	20.75 (0.66)	22.80 (1.00)	1.706	.096
Graffar Scale	11.75 (0.66)	12.35 (0.65)	.647	.521
MoCA	29.00 (0.24)	29.05 (0.19)	-.165	.870
AMMA Total	71.65 (3.22)	51.45 (4.57)	**-3.612**	**.001**
AMMA Tonal	70.40 (3.58)	52.75 (4.68)	**-2.993**	**.005**
AMMA Rhythm	70.35 (3.24)	51.55 (4.18)	**-3.552**	**.001**
Musical training (years)	12.35 (0.74)	N.A.	N.A.	
Onset of training (years)	8.45 (0.63)	N.A.	N.A.	

Note. The MoCA is from Nasreddine et al. (2005) and the MoCA adaptation to the Portuguese population is from Simões et al. (2008); The AMMA is from Gordon (1989); SE – Standard Error; N.A. – Non-Applicable. Significant differences are highlighted in bold.

All participants filled in a written informed consent and agreed on voluntary participation according to the Declaration of Helsinki. Participation was rewarded with course credit or gift vouchers.

Table 9

Characterization of Musicians' Musical Background and Training Routines

Participant Code	Musical training[1]	Onset of Training[1]	Instrument[2]	Practice days/week	Practice hours/day
M1	13	6	piano	7	4
M2	8	12	piano	6	2
M3	11	10	clarinet	7	4
M4	15	6	violin	4	2
M5	8	11	clarinet	7	2.5
M6	9	11	saxophone	7	6
M7	10	9	clarinet	7	3
M8	13	6	piano	7	5.5
M9	13	6	violoncello	7	6.5
M10	10	15	trumpet	7	4
M11	11	12	percussion	7	2
M12	13	6	percussion	5	3
M13	15	11	piano	5	3
M14	14	5	piano	6	4
M15	20	7	violin	7	2
M16	17	7	piano	7	3.5
M17	17	6	transverse flute	4	2
M18	10	6	violin	4	1.5
M19	8	9	violin	5	3
M20	12	8	piano	7	2

Note. [1] Measured in years; [2] The column displays the participant's main instrument despite several participants reporting expertise in more than one musical instrument.

5.3.2 Stimuli and Tasks

The study was composed of three distinct experiments – prosodic (Experiment 1), linguistic (Experiment 2) and musical (Experiment 3). These experiments investigated the neural and behavioral correlates of the SL of prosodic, non-prosodic and musical streams, respectively. In order to create an artificial language stream with prosodic content, five different words with specific intonations were used in the prosodic experiment (Experiment 1; see Table 10). These words were selected considering the combination of ten different consonant-vowel syllables according to François and colleagues (François et al., 2014). The syllables were created with the Mbrola speech synthesizer (http:// tcts.fpms.ac.be/synthesis/mbrola.html) using a German database (ge1). Syllables were created based on German phonotactic constraints since none of the participants was fluent in German: this ensured stimuli had different phonotactic constraints relative to the native language of the participants.

The syllables were digitally spoken with a specific pitch (e.g., C3, D3, F3, E4, etc., according to musical letter notation) and had 200 ms of duration, with 10 ms of rise and decay time. Syllables were concatenated using the Audacity® software (version 2.1.1) to create five trissillabic words (e.g., gifaca [C3 D3 F3], fifeci [E4 Db4 G3], begaci [D4 C4 G3], cibaga [G3 B3 C4], bafice [B3 E4 F4]) with specific and distinct prosodic contours between each other. The artificial language stream was the result of the concatenation of the five words, each repeated 100 times, in a pseudo-random order (i.e., the same word could not be repeated consecutively). The words lasted 700 ms and were separated by 50 ms ISI. The ISI and the pause periods functioned as 50 ms silent baselines for each syllable, ideal for ERP analyses. TPs within words ranged from 0.5 to 1 (*M* = 0.7) and TPs across word boundaries ranged from 0.07 to 0.31 (*M* = 0.19). The pitch contours of words occurred 50% of the time at word boundaries. This language stream was presented during the learning phase of the prosodic experiment.

Another speech stream was presented during the implicit test phase. The stream contained the words that were previously presented as well as foils. There were two categories of foils: i) legal part-words – words composed by the two last syllables of a word and the first syllable of another, or vice-versa; ii) illegal non-words – items that corresponded to the reversed (backwards) syllable sequence of each legal word of the artificial language (e.g., 'gifaca' originated the non-word 'cafagi'; see Table 10). Therefore, legal items had ABC forms (e.g., A stands for first syllable, B for second, and C for third), new legal items had BCA or CAB forms, and illegal items had CBA forms.

Table 10

Characterization of the Sounds Used in the Learning and Implicit Test Phases of Each Experiment

Experimental Phase	Category	Experiment		
		Prosodic	Linguistic	Musical
Learning phase	Words/melodies	gifaca (C3 D3 F3)	B2 C2 E2	tanisi
		fifeci (E4 Db4 G3)	D3 A2# F2	nanosa
		begaci (D4 C4 G3)	C3 B3 F2	pobisa
		bafice (B3 E4 F4)	A2 D3 E3	pinaso
		cibaga (G3 B3 C4)	F2 A2 B3	sapiti
Implicit Test Phase	Words/melodies	gifaca (C3 D3 F3)	B2 C2 E2	tanisi
		fifeci (E4 Db4 G3)	D3 A2# F2	nanosa
		begaci (D4 C4 G3)	C3 B3 F2	pobisa
		bafice (B3 E4 F4)	A2 D3 E3	pinaso
		cibaga (G3 B3 C4)	F2 A2 B3	sapiti
	Part-words/melodies (new legal items)	bagafi (B3 C4 E4)	A#2 F2 B2	bisana
		cabega (F3 D4 C4)	A2 B3 F3	nosata
		cegifa (F4 C3 D3)	B3 F2 D3	patina
		fecigi (Db4 G3 C3)	E2 C3 B3	sipobi
		gacifi (C4 G3 E4)	E3 B2 C2	sotani
	Non-words/melodies (new illegal items)	cafagi (F3 D3 C3)	E2 C2 B2	Sinita
		cifefi (G3 Db4 E4)	F2 A#2 D3	Sanona
		cigabe (G3 C4 D4)	F2 B3 C3	Sabipo
		cefiba (F4 E4 B3)	E3 D3 A2	Sonapi
		gabaci (C4 B3 G3)	B3 A2 F2	Tipisa

Note. The pitch (Hz) of the notes that compose each melody is reported in musical letter notation according to the English convention. Each note is named by a scientific pitch notation with a letter-name and a number identifying pitch's octave; additionally, "#" stands for *sharp*, "b" stands for *flat*.

In the stream, words were repeated 90 times, and part and non-words were repeated 8 times, each making up a total of 530 stimuli. Before data collection, 31 first-year college students (29 women; 18.6 ± 1.85 years) evaluated the stimuli to ensure complete unfamiliarity. All participants were native speakers of European Portuguese and had no linguistic fluency in more than two languages (German excluded). Participants were asked to rate the familiarity of the words using a scale from 1 to 7 (1 – absolutely unfamiliar; 7 – highly familiar). Participants were also instructed to rate the degree of confidence in their previous response using a scale from 1 to 7 (1 – not confident; 7 – absolutely confident). Participants classified the words as highly non-familiar (1.51 ± 0.17 [M ± *SD*]) and with a high degree of confidence (6.43 ± 0.17 [M ± *SD*]).

For the linguistic (Experiment 2) and musical (Experiment 3) experiments, new words and melodies were created. The characteristics of these stimuli are reported in Table 10. The procedure adopted to create the new items was the same used in the prosodic experiment. The number of stimuli, duration, rise and decay times, pauses, ISI and TPs for these streams followed the same procedure used in the prosodic experiment. The artificial words of the linguistic experiment (Experiment 2) were non-prosodic words: syllables were synthesized at a rate of 1000 Hz so that the words had no variations in musical contour. In the musical experiment (Experiment 3), 5 distinct tri-tone melodies were presented. Syllables were synthesized with the Mbrola software and tones were created using the Piano Fx Studio software; both were combined into words and melodies using the Audacity® software package (version 2.1.1). The only cues to words' or melodies' boundaries were the TPs between syllables or tones, respectively. The two new streams of sounds corresponded to the learning phases of the linguistic and musical experiments.

Similarly to the prosodic experiment (Experiment 1), the stimuli that composed the implicit test phase of the linguistic (Experiment 2) and musical (Experiment 3) experiments were derived from the stimuli used in the learning phases, following the procedure that was previously described (see Table 10). Prior to the SL task, the linguistic stimuli were also validated with the same 31 participants who did not participate in the EEG experiment to ensure they were unfamiliar to the participants. Participants were confident (6.26 ± 0.18 [M ± *SD*]) that the words were completely unfamiliar to them (1.57 ± 0.24 [M ± *SD*]).

5.3.3 Procedure

A schematic representation of the procedure is illustrated in Fig. 18. Stimuli were presented using Presentation® software (Neurobehavioral Systems, Albany, CA). The prosodic experiment (Experiment 1) was composed of three distinct phases in the following order: learning phase, implicit test phase and explicit behavioral test phase (see Fig. 18). The experiment started with the learning phase during which participants listened to the artificial stream of sounds via Sennheiser earphones. During stimulus presentation, participants watched a movie (a documentary without sound and neutral in content) presented on a computer screen at a distance of approximately1 m. Before the implicit test phase, there was a break that participants were not aware of. During the learning and implicit test phases, participants were instructed to ignore the sounds and to pay attention to the movie so that they would be able to correctly answer a small questionnaire at the end of the experiment. The instructions intended to ensure that participants did not focus their attention on the auditory stimuli, providing the conditions needed to study ASL in the absence of overt attention to the acoustic stream. EEG data was collected during both phases to obtain online measures of learning.

After the EEG experiments, participants performed three LDTs (see Table 2. APPENDIX 1. for details on the stimuli) in a counterbalanced order. The prosodic test was designed to behaviorally evaluate the learning of the artificial language. If learning occurred, participants were expected to accurately recognize each word item and to reject the foil items; if not, participants were expected to respond indiscriminately to words and foils. The second and third LDTs were named ‘linguistic’ and ‘musical’ tests as they assessed how regularities were derived from the sung language from a linguistic and musical point of view (following François & Schön, 2010). The items of the linguistic test were words, part-words and non-words without prosodic intonation (e.g., the syllables’ pitch was 1000 Hz). The items of the musical test were midi-piano sequences that mimicked the ‘melody’ of the words, i.e., three-notes sequences with the same musical contours of prosodic words, prosodic part-words and prosodic non-words.

As in the prosodic experiment (Experiment 1), the linguistic (Experiment 2) and musical (Experiment 3) experiments were composed of a learning phase and an implicit test phase. These experiments (2 and 3) were presented before the prosodic experiment. These experiments did not have explicit test phases since they could change the attention focus of the participants to the auditory stimuli, despite opposite instructions. The order in which these two experiments were presented was counterbalanced across participants.

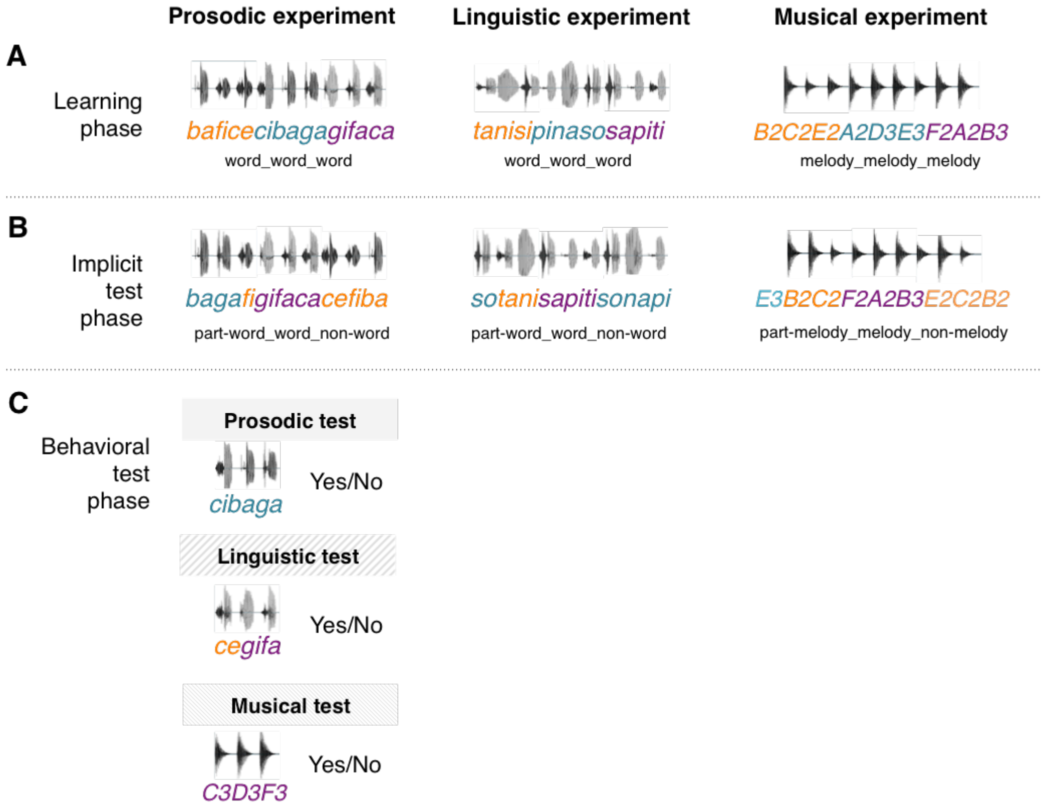

FIGURE 18. Illustration of the experimental procedure. **(A)** Illustration of the sound streams used in the learning phase of each experiment. The ''_'' represents the 50 ms ISI. Different items are displayed in different colors. **(B)** The learning phase of each experiment was followed by an implicit test phase in which new legal/illegal word/melody items were presented among previously learned words/melodies. **(C)** In the prosodic experiment, a behavioral test phase was administered after the implicit test phase. This offline set of tests evaluated if participants had learnt the words/melodies through LDTs. In the prosodic test, participants listened to a word item, i.e., a word, a part-word or a non-word (700 ms), and judged whether the item was presented at the beginning of the experiment or not (e.g., Yes/No answer). Five words, part-words and non-words were evaluated in 15 trials. This test was followed by musical and linguistic tests that had the same structure of the prosodic test but presenting melodies and non-prosodic words. These latter tests aimed to evaluate participants' ability to generalize the knowledge on the statistical structure of the sung language to new musical and word items. The order of the latter tests was counterbalanced. *Note.* EEG data were recorded during phases **(A)** and **(B)**.

5.3.4 Data acquisition and processing

EEG was recorded from the scalp at 64 standard locations with the Active-Two system (Biosemi, Amsterdam, The Netherlands) and according to the extended International 10-20 system. BrainVision Analyzer software (version 2.1.1; Brain Products, Inc., 2000) was used for

EEG analysis. The horizontal and vertical eye movements were monitored with 4 electrodes placed in the outer canthi of the eyes and at the infra and superior (FP1) orbital ridge of the left eye. The offset of all electrodes was kept below 20 mV and the EEG was digitized at a rate of 512 Hz. The EEG signal was filtered offline with an Infinite Impulse Response filter of 0.1-30 Hz and was referenced to the algebraic average of the left and right mastoids. Trials were discarded offline based on an eye blink detection method (Gratton, Coles, & Donchin, 1983) and on movement artifacts or eye blinks exceeding ±100 μV. For the analysis of the learning and implicit test phases, ERPs were time-locked to the onset of each stimulus and averaged over epochs of 750 ms (i.e., 50 ms pre-stimulus and 700 ms post-stimulus). The baseline was defined as the 50 ms silent period preceding the stimulus onset.

The following procedure was adopted in the analysis of EEG data in each experiment. For the analysis of the learning phase, segments were averaged over the entire duration of the phase, for each stimulus, participant and electrode site. An additional analysis investigated changes in the electrophysiogical data as a function of exposure time (i.e., number of stimulus repetitions): here, the learning phase was divided in five consecutive non-overlapping blocks (first block: trials 0-100 [20%]; second block: trials 101-200 [40%]; third block: trials 201-300 [60%]; fourth block: trials 301-400 [80%]; fifth block: 401-500 trials [100%]). Each block represented 20% of the learning time and segments were averaged for each stimulus, participant and electrode site. For the analysis of the implicit test phase, segments were averaged separately for previously learned stimuli (e.g., words/melodies) and each type of foils (e.g., part-words/melodies and non-words/melodies) and computed separately for each participant and electrode site.

After the visual inspection of the grand averaged waveforms we computed mean amplitudes for consecutive 50 ms intervals from 0 to 700 ms (e.g., the time range of a word/melody). The motivation for this approach was twofold. First, the visual inspection of the data did not show evidence of the expected ERP components when taking into account their typical latencies, polarities or scalp topographies (see, for example, de Diego-Balaguer et al., 2007; François et al., 2017, 2014; François & Schön, 2011; Vasuki et al., 2017a). Second, the short ISIs (i.e., 50 ms) presented in our experiments could lead to substantial overlap between components. Therefore, we designed a step-wise analysis (for a similar analysis procedure, see Kujala, Kallio, Tervaniemi, & Näätänen, 2001; Martín-Loeches, Sommer, & Hinojosa, 2005; Rellecke, Palazova, Sommer, & Schacht, 2011; Röder, Rösler, Hennighausen, & Näcker, 1996; Werheid, Alpay, Jentzsch, & Sommer, 2005) in which mean amplitudes were extracted from 50 ms intervals that encompassed the entire duration of stimulus epochs. By implementing this approach, we overcame the limitations imposed by *a priori* defined time windows (to compute mean amplitudes) and benefited from a more dynamic view of changes in the evoked potentials elicited by the processes involved in the segmentation of the auditory streams.

5.3.5 Statistical analyses

5.3.5.1 Learning phase

Separate analyses of variance (ANOVAs) were performed. Repeated-measures ANOVAs were carried out with Region of Interest (ROI) as within-subjects factor and Expertise (2 levels: Musicians vs. Non-Musicians) as between-subjects factor. The ROI factor included Frontal (Fz/3/4), Fronto-Central (FCz/3/4), Central (Cz/3/4), Centro-Parietal (CPz/3/4) and Parietal (Pz/3/4) topographical regions. A total of 14 separate repeated measures ANOVAs were performed for the 14 consecutive 50 ms time windows of interest. *P*-values were adjusted with the Greenhouse-Geisser correction for non-sphericity. The significance values of Post-hoc comparisons were adjusted with Bonferroni correction. Only main effects or interactions involving the factor Expertise are reported.

To investigate changes in the ERP responses along the learning phase, stimulus-locked EEG data were divided in five blocks. To test for the effect of exposure time on learning and for possible interaction effects between the timing of learning and musical training, repeated-measures ANOVAs were computed with ROI (5 levels) and block (5 levels: 20%, 40%, 60%, 80%, 100%) as within-subject factors, and expertise (2 levels) as between-subject factor for the 14 time windows. The results concerning this last ANOVA only include significant main effects or interactions ($p < .05$) involving the block factor. These analyses were conducted for the three experiments, separately.

5.3.5.2 Implicit test phase

Statistical analyses were performed based on the mean amplitude extracted for consecutive 50 ms intervals from 0-700 ms for the three types of items evaluated in this phase (e.g., items, part-items and non-items). Fourteen separate repeated-measures ANOVAs including the factors item type (3 levels: words, part-words, non-words) and ROI (5 levels) as within-subjects factor and expertise (2 levels) as between-subjects factor were conducted. Only main effects or interactions involving the factors item type and expertise are reported. These analyses were conducted for the three experiments, separately.

5.3.5.3 Behavioral test phase

t-tests examined differences between performance (percentage of correct responses) and chance level (50%) in the three LDTs. The performance of the two groups of participants on the tasks was tested by means of *t*-tests. Also, two-way repeated-measures ANOVAs were separately computed for the three tests including expertise (2 levels) as between-subjects factor and item type (3 levels: item, part-item and non-item) as within-subjects factor. Additionally, to compare participants' performance as a function of the type of stimuli tested in the three test experiments, a three-way repeated-measures ANOVA including expertise as between-subjects factor (2 levels), and stimulus type (3 levels: prosodic-words, melodies and non-prosodic words) and item type (3 levels: item, part-item and non-item) as within-subject factors was computed.

5.3.5.4 Brain-Behavior correlations

Spearman correlations (two-tailed) tested whether significant group differences in the electrophysiological measures were associated with differences in behavioral performance. With this analysis, we examined the association between the measures obtained during the learning phase and the LDTs in the prosodic experiment. These analyses were performed using the percentage of correct responses in the LDTs and the mean amplitude values obtained for the time windows in which the ANOVAs showed significant differences between groups.

5.4 Results

5.4.1 Experiment 1: Prosodic experiment

5.4.*1.1 Electrophysiological Results*

Fig. 19 shows the grand average waveforms illustrating the time course of neural responses during the learning phase of the prosodic experiment. There were no significant differences between musicians and non-musicians in the grand averaged ERPs computed for the total duration of the learning phase (see Fig. 19A).

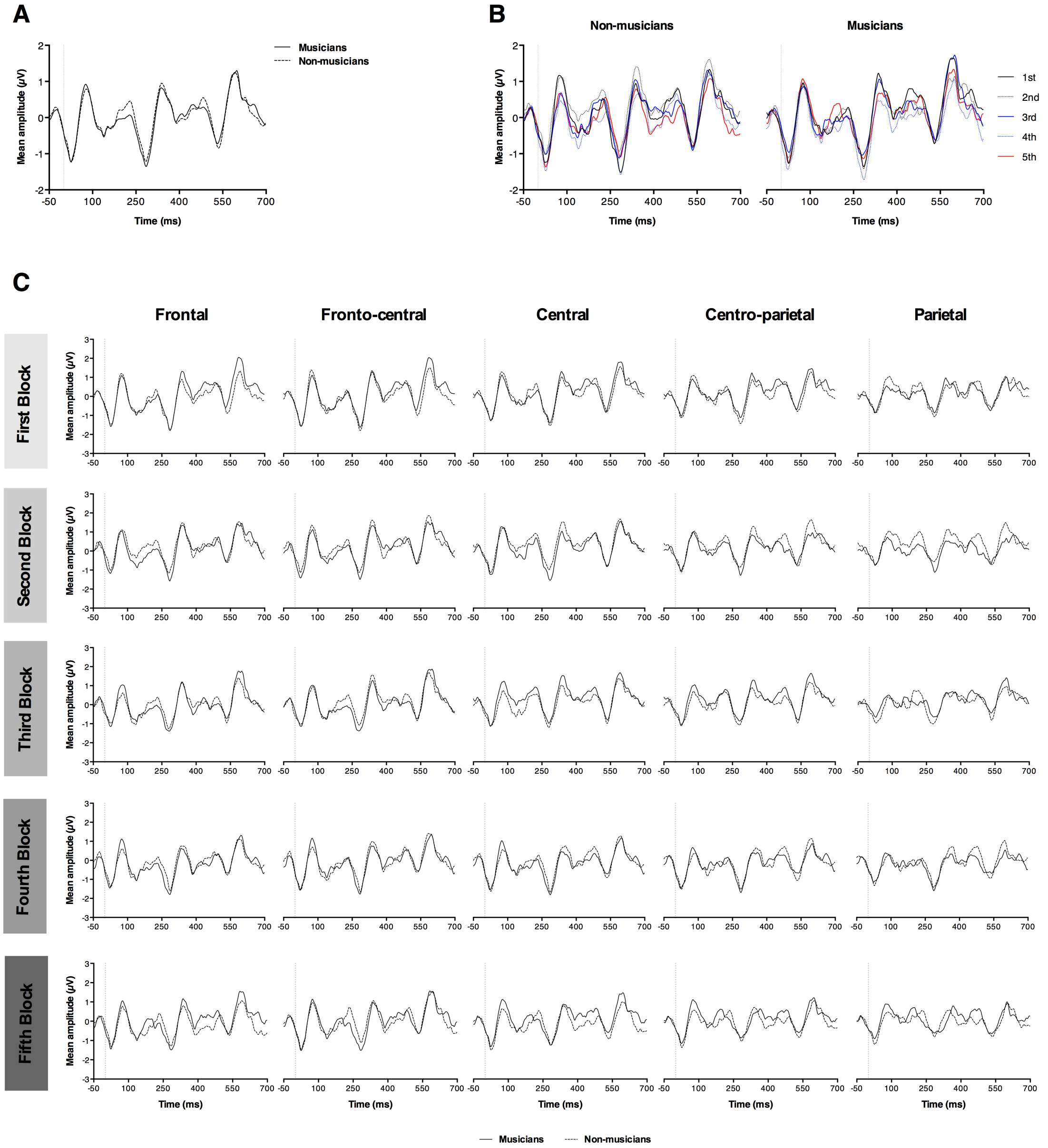

FIGURE 19. Grand-averaged ERP waveforms elicited by prosodic words during the learning phase of the prosodic experiment. **(A)** ERPs averaged for the five ROIs, for the entire duration of the learning phase comparing musicians (solid line) and non-musicians (dashed line). **(B)** Each graph displays the averaged ERP waveforms for the five blocks of the learning phase (i.e., first [black line], second [dotted black line], third [blue line], fourth [dotted blue line] and fifth [red line] blocks), for each group separately. **(C)** The graphs show the evolution of group differences over learning time (from the top to the bottom) at different ROIs (i.e., frontal, fronto-central and, central, centro-parietal and parietal; from the left to the right; from left to the right), in musicians [solid line] and non-musicians [dashed line]. Each ROI displays the averaged ERP signature of left (e.g., F3), right (e.g., F4) and midline (e.g., Fz) electrodes. The vertical grey lines indicate words' onset (0 ms).

However, the analysis of ERP amplitude in the learning phase as a function of exposure time (see Fig. 19B and 19C) revealed a tendency for musicians ($M = .054$, $p = .065$) to exhibit a decreased positivity in the 450-500 ms time window when compared to non-musicians ($M = .750$) in the second block at the parietal ROI, $F(16, 608) = 2.283$; $p = .036$, $\eta_p^2 = .054$. There were no other significant differences between groups during the learning phase, as well as during the implicit test phase (see Fig. 20).

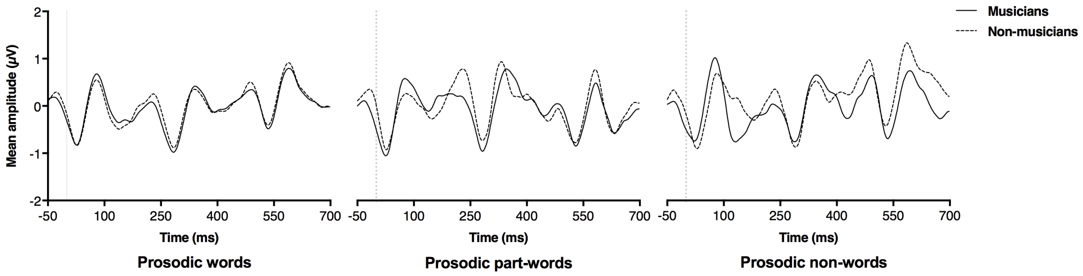

FIGURE 20. Illustration of the ERP results obtained for each type of prosodic item presented during the implicit test phase of the prosodic experiment. Each graph depicts the grand averaged ERP waveforms in musicians (solid line) and non-musicians (dashed line) in response to prosodic words (e.g., previously learned items; left graph), prosodic part-words (e.g., new legal items; middle graph) and prosodic non-words (e.g., new illegal items; right graph) for the average of the five ROIs. The vertical grey lines indicate words' onset (0 ms).

5.4.1.2 Behavioral results

Behavioral results are illustrated in Fig. 21. Comparison of the performance in the prosodic test with chance level (50%) showed that both musicians and non-musicians learned the artificial language (Musicians: 61.67% of correct responses, $t(19) = 2.950$, $p = .008$; Non-musicians: 61.25% of correct responses, $t(19) = 4.217$, $p = .001$) (Fig. 21A). The groups did not differ in the total number of correct responses, $t(38) = -.083$, $p > .05$. In the musical test, musicians' performance tended to be significantly above chance, whereas non-musicians' performance was at the chance level (Musicians: 54.67% of correct responses, $t(19) = 1.837$, $p = .082$; Non-musicians: 46.25% of correct responses, $t(19) = -1.065$, $p > .05$). The comparative analysis of the groups' performance revealed that musicians tended to outperform non-musicians, $t(38) = -1.986$, $p = .055$ (Fig. 21A). On the contrary, in the linguistic test, non-musicians performed above chance while musicians did not (Musicians: 52.00% of correct responses, $t(19) = .587$, $p > .05$.; Non-musicians: 58.04% of correct responses, $t(19) = 3.086$, $p = .007$), even though the performance of both groups did not differ significantly, $t(38) = 1.369$, $p > .05$ (Fig. 21A).

The results of the two-way ANOVAs are illustrated in Fig. 21B. The two-way ANOVAs that tested the effects of item type and expertise on performance in the prosodic and linguistic tests revealed no significant main effects or interactions ($p > .05$). On the contrary, the two-way ANOVA testing performance in the musical test revealed a higher percentage of correct responses for non-melodies ($M = 57.75\%$) relative to melodies ($M = 44.75\%$, $p = .008$) but not part-melodies ($M = 48.88\%$, $p > .05$), $F(2, 76) = 4.719$; $p = .013$, $\eta_p^2 = .122$ (see Fig. 21C). Additionally, participants responded more correctly to the prosodic test than to the musical test ($p = .007$), $F(2, 76) = 6.442$, $p = .003$, $\eta_p^2 = .159$ (see Fig. 21D).

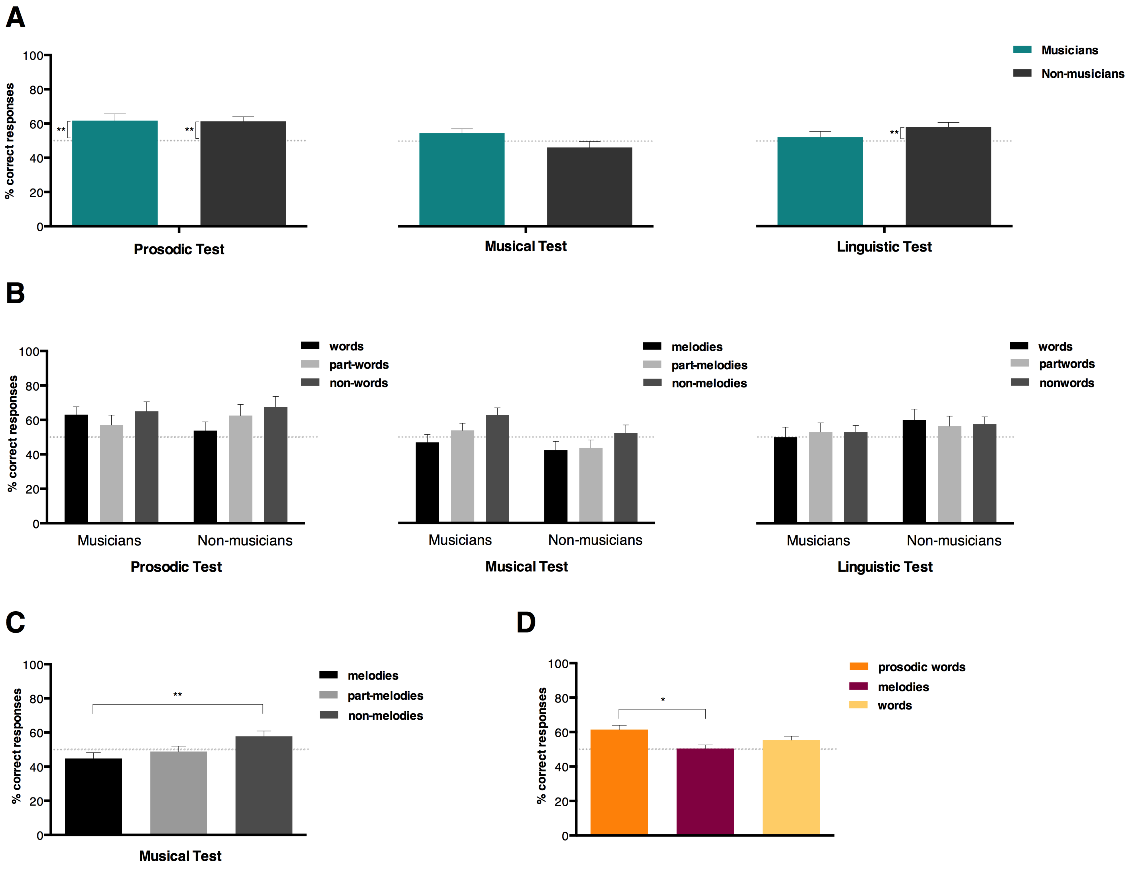

FIGURE 21. Illustration of the behavioral results of the prosodic experiment. **(A)** The plots display the percentage of correct recognition of musicians (green) and non-musicians (dark grey) in the prosodic (left), musical (middle) and linguistic (right) tests. **(B)** The graphs compare musicians and non-musicians percentage of correct responses for each type of item (i.e., legal items, black; new legal items, soft grey; new illegal items, dark grey) at each test. **(C)** The graph shows participants' performance as a function of item type in the musical test. **(D)** Depiction of participants' percentage of correct responses for each type of stimuli (across tests): prosodic words (orange), melodies (purple) and

non-prosodic words (yellow). Bars indicate SEM. The horizontal dotted lines refer to chance level. * $p < .05$, ** $p < .01$, *** $p < .001$.

5.4.1.3 Brain-behavior correlations

No significant correlations were found between brain and behavior measures ($p > .05$).

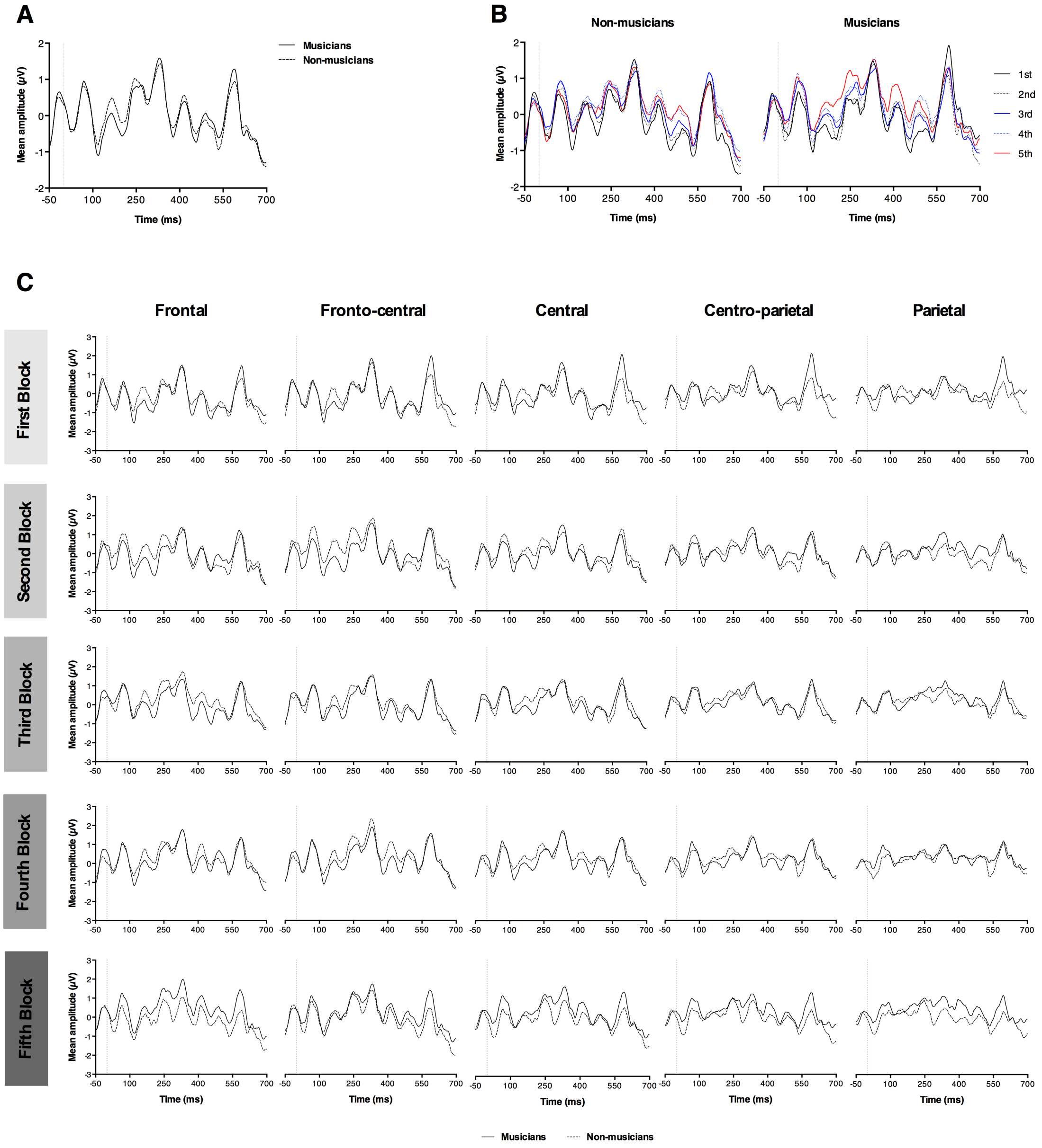

FIGURE 22. Grand-averaged ERPs of the non-prosodic words obtained during the learning phase of the linguistic experiment. **(A)** ERPs averaged over the five ROIs, for the entire duration of the learning phase comparing musicians (solid line) and non-musicians (dashed line). **(B)** Grand-averaged ERP waveforms for the five blocks of the learning phase (i.e., first [black line], second [dotted black line], third [blue line], fourth [dotted blue line] and fifth [red line] blocks), for each group separately. **(C)** Evolution of the group differences over learning time (from the top to the bottom) at different ROIs (from the left to the right) in musicians [solid line] and non-musicians [dashed line]. Each ROI displays the averaged ERP signature of left (e.g., F3), right (e.g., F4) and midline (e.g., Fz) electrodes. The vertical grey

lines indicate words' onset (0 ms).

5.4.2 Experiment 2: Linguistic experiment

5.4.2.1 Electrophysiological results

In the 150-200 ms window, the interaction between ROI and Expertise was significant, $F(4,\ 152) = 4.717$, $p = .009$, $\eta_p^2 = .10$. Nevertheless, the pairwise comparisons were non-significant, revealing only a tendency for more negative amplitude at the frontal ($M = -.181$, $p = .086$) and fronto-central ($M = -.239$, $p = .095$) ROIs in musicians compared to non-musicians (frontal: $M = .220$; fronto-central: $M = .223$) (see Fig. 22A). There were no significant differences between groups in the learning phase as a function of exposure time (see Fig. 22B and 22C).

Furthermore, there were no significant group differences during the implicit test phase (see Fig. 23).

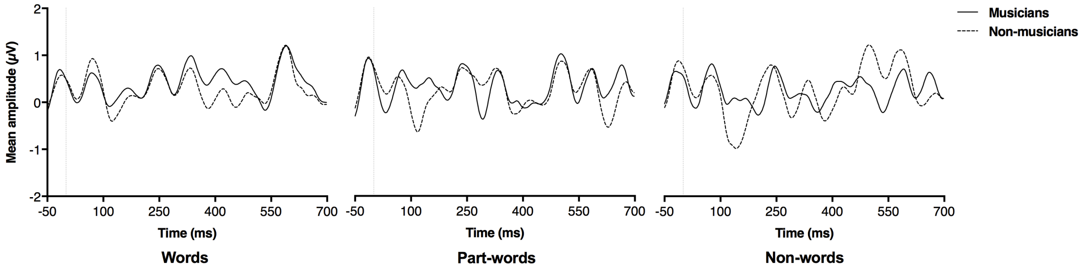

FIGURE 23. Illustration of the ERP results obtained for each type of word item presented during the implicit test phase of the linguistic experiment. Each graph depicts the grand-averaged ERP waveforms in musicians (solid line) and non-musicians (dashed line) in response to non-prosodic words (left), non-prosodic part-words (middle) and non-prosodic non-words (right) for the average of the five ROIs. The vertical grey lines indicate words' onset (0 ms).

5.4.3 Experiment 3: Musical experiment

5.4.3.1 Electrophysiological results

The grand average waveforms obtained during the learning phase of the musical experiment can be visualized in Fig. 24. In the 0-50 ms time window, there was a main effect of Expertise, $F(1,\ 38) = 4.485$; $p = .041$, $\eta_p^2 = .106$: The pairwise comparisons revealed an increased positivity in this time window in musicians ($M = .460$) when compared with non-musicians ($M = -.018$, $p = .041$) (see Figs. 24A and 25A).

The analysis of the ERP responses in the learning phase as a function of exposure time (see Fig. 24B, 24C and 25B) revealed that, in the 550-600 ms window, the interaction between

Block, ROI and Expertise was significant, *F*(16, 608) = 2.885; *p* = .008, η_p^2 = .067: Musicians showed a larger positivity in the second block at both frontal (*M* = .497 vs. *M* = -.628, *p* = .004) and fronto-central (*M* = .415 vs. *M* = -.596, *p* = .008) ROIs, whereas the same tendency was observed during the first block at the frontal ROI (*M* = 1.136 vs. *M* = .224, *p* = .080).

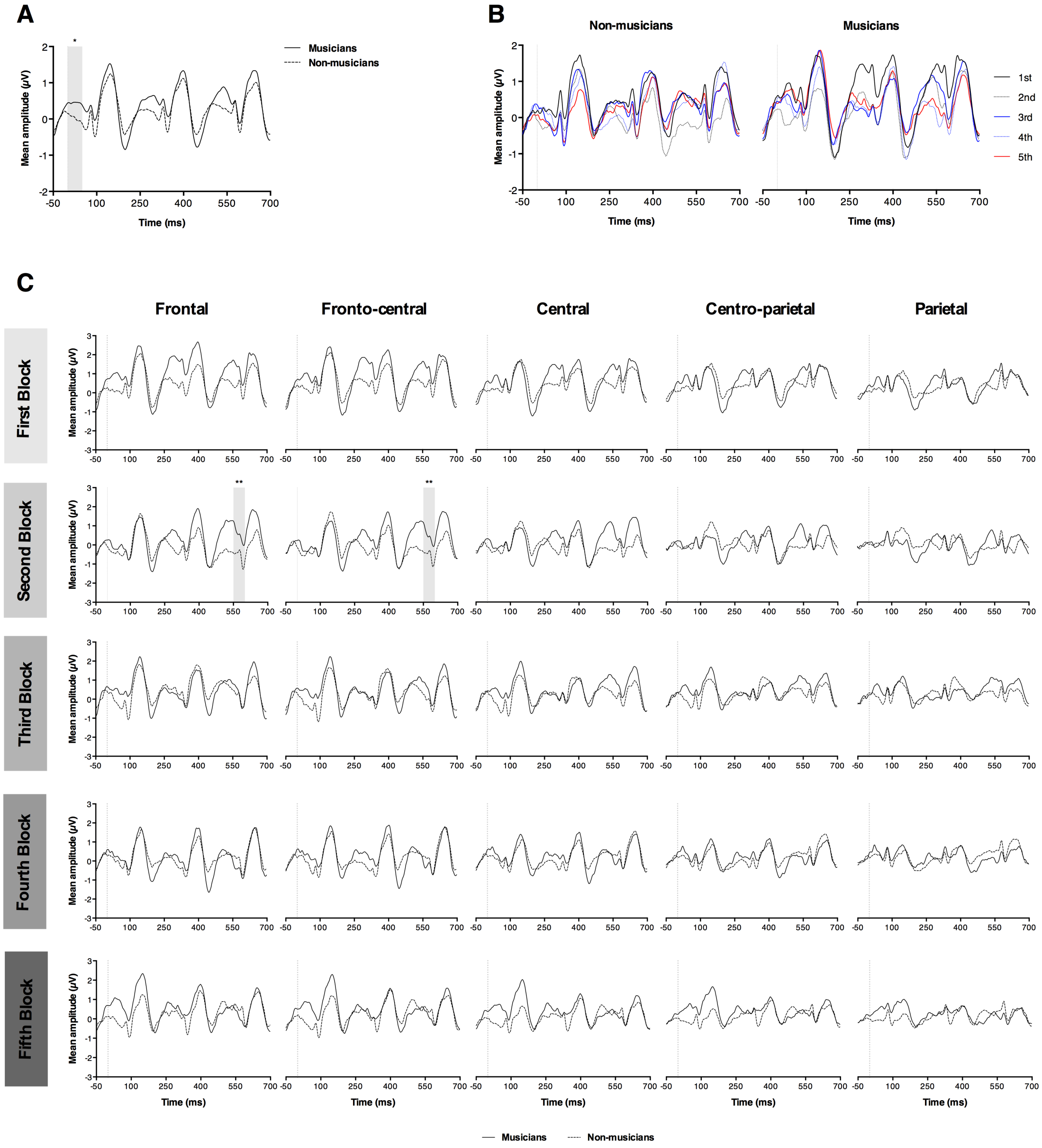

FIGURE 24. Grand average ERPs of the melodies obtained during the learning phase of the musical experiment. **(A)** ERPs averaged over the five ROIs, for the entire duration of the learning phase comparing musicians (solid line) and non-musicians (dashed line). **(B)** Each graph displays the averaged ERP waveforms for the five blocks of the learning phase (i.e., first [black line], second [dotted black line], third [blue line], fourth [dotted blue line] and fifth [red line] blocks), for each group separately. **(C)** The graphs show the evolution of group differences over learning time (from the top to the bottom) at different ROIs (from left to the right) in musicians [solid line] and non-musicians [dashed line]. Each ROI displays the averaged ERP signature of left (e.g., F3), right (e.g., F4) and midline (e.g., Fz) electrodes. The vertical grey lines indicate melodies' onset (0 ms). The grey shadowed areas indicate the time windows where

significant differences between groups were found. * $p < .05$, ** $p < .01$, *** $p < .001$

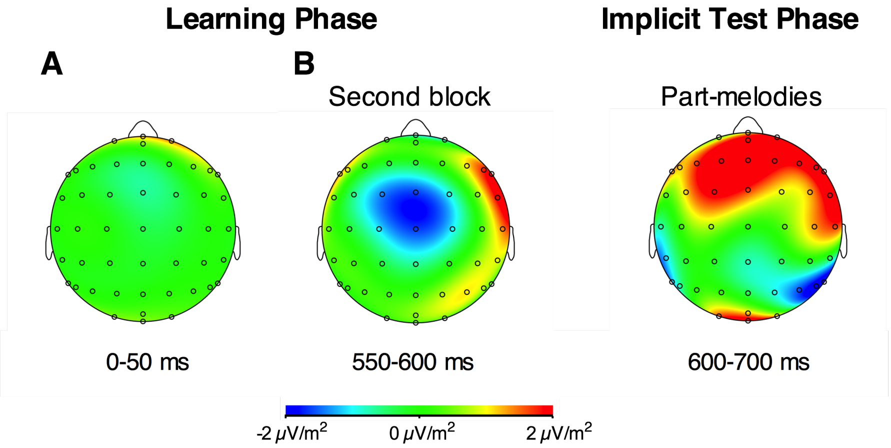

FIGURE 25. The maps depict the scalp distribution of the latency windows in which the significant group differences were found during the learning phase of the musical experiment (panels A and B). Panel **(A)** shows the topography of the effect observed for the 0-50 ms time window calculated for all trials of the learning phase. Panel **(B)** shows the difference emerging 550-600 ms after melodies' onset during the second block (101-200 trials) of exposure. The map on the right depicts the scalp topography of the waveform emerging at the 600-700 ms latency range in response to part-melodies during the implicit test phase.

The results of the implicit test phase (see the "implicit test phase" panel of Fig. 25 and Fig. 26) showed that part-melodies elicited a more positive amplitude in musicians ($M = 2.101$) compared to non-musicians ($M = -.318$, $p = .011$) in the 600-650 window, $F(2, 76) = 4.616$, $p = .021$, $\eta_p^2 = .108$. In the 650-700 ms time window, the amplitude was more positive in musicians ($M = 2.637$) relative to non-musicians ($M = .382$) in response to part-melodies ($p = .007$), $F(2, 76) = 3.997$, $p = .038$, $\eta_p^2 = .095$.

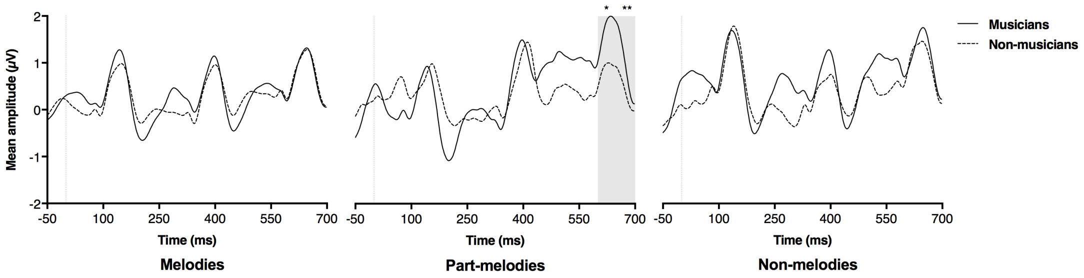

FIGURE 26. Illustration of the ERP results obtained for each type of melodies presented during the implicit test phase of the musical experiment. Each graph depicts the grand averaged ERPs in musicians (solid line) and non-musicians (dashed line) obtained during the entire duration of the implicit test phase in response to melodies (left graph), part-melodies (middle graph) and non-melodies (right graph) for the five ROIs. The vertical grey lines indicate

melodies' onset (0 ms). The grey shadowed areas indicate the time windows where significant differences between groups were found. * $p < .05$, ** $p < .01$, *** $p < .001$

5.5 Discussion

We examined whether musical training modulates the SL of different types of auditory structures when sounds are task-irrelevant. EEG provided ideal temporal resolution to probe the brain dynamics of SL as a function of training and stimulus type (i.e., prosodic words vs. non-prosodic words vs. melodies). Our results support the implicit nature of SL mechanisms: Even though participants' attention was not focused on the sounds, behavioral data indicated that participants learned the sung language (Experiment 1). Moreover, musical training affected how participants learned the statistical regularities of the different types of auditory streams. Learning distinct auditory structures based on their statistical properties resulted in unique ERP patterns. In the case of the prosodic words (Experiment 1), non-musicians tended to exhibit a larger positivity 450-500 ms post-stimulus onset during the second block (101-200 trials) of the learning phase. When learning words that were devoid of prosodic contour (Experiment 2), musicians evidenced a tendency for an enhanced negativity in the 150-200 ms latency range. During the SL of musical streams (Experiment 3), musicians showed an overall enhanced positivity 50 ms after melodies' onset and an increased positivity in the 550-600 ms period during the early part of the learning phase. Despite no evidence for the modulation of typical ERP markers of online auditory segmentation (e.g., the N400-like component), the effects of musical training were reflected in specific time windows of auditory processing. Critically, these effects were modulated by stimulus type and occurred irrespective of attention. The absence of group differences in the behavioral test challenges the interpretation of previous reports of enhanced speech segmentation skills in musicians and highlights the role of selective attention as a mediator variable in the association between musical training and ASL.

5.5.1 The influence of attention on ASL

The online neural measures of ASL revealed that, irrespective of stimulus type (e.g., linguistic or musical) and of musical expertise, the typical SL-related components (e.g., N400-like) were not elicited. The existing studies that investigated the impact of attention on SL indicate that attentional load affects SL performance at the behavioral level (Shanks, Rowland, & Ranger, 2005; Turk-Browne et al., 2005). However, as mentioned before, research probing the neural mechanisms of SL under pre-attentional requirements is scarce. Until now, only the study of Paraskevopoulos and collaborators (Paraskevopoulos et al., 2012) probed the effects of musical training on SL of tone sequences when the input was task-irrelevant. They observed P50 and MMN effects in response to the sequences' onsets, even though behavioral performance in a post-learning task indicated that participants did not learn the sequences. No other studies examined how the brain processes regularities in auditory input presented as background stimulation. Our findings indicate that when the learning of auditory sequences occurs under pre-attentive conditions, the brain uses distinct mechanisms to process statistical regularities. Even though the learning conditions were challenging (i.e., under low attentional focus), the responses to the offline recognition task (prosodic test) confirmed that participants accurately recognized the words embedded in the stream.

5.5.2 The effects of musical training on ASL

5.5.2.1 Pre-attentive segmentation of speech sounds

Non-musicians showed a tendency for an enhanced positivity 450-500 ms after prosodic words' onset (Experiment 1) during the second block of trials (101-200 trials, i.e. 40% of word repetitions). The 450-500 ms time window is the period in which the transition between the second and the third syllable of a word occurs. Syllable transitions that happen within and between word boundaries may represent a critical time for a memory update based on the information conveyed by the latest syllable heard. Based on this update, the brain might generate an expectation regarding the upcoming syllable. As long as SL is taking place, the second-to-third syllable transition within words (450-500 ms) is more predictable than the first-to-second one (200-250 ms). At the second syllable transition, there is more information available to compare the incoming input with the existent mental representation for words. The timing of the musical training effect might be related to the relevance of this syllable transition for the learning process. However, this is an explanatory hypothesis that needs to be experimentally addressed

In recent decades, a considerable number of studies have associated the emergence of late posterior ERP components to the processing of anomalies in structured sequences (Besson &

Faïta, 1995; Donchin, 1981; Hagoort, Brown, & Groothusen, 1993; Neville, Nicol, Barss, Forster, & Garrett, 1991; Nittono, Bito, Hayashi, Sakata, & Hori, 2000; Osterhout & Holcomb, 1992; Paller, McCarthy, & Wood, 1992). The late positive component (LPC), for example, typically occurs 300 ms following stimulus onset at central and parietal sites (Donchin, 1981). The LPC has been related to violations of expectations regarding the occurrence of a given stimulus, such as notes (Besson & Macar, 1987) or melodies (Paller et al., 1992). Further, it is modulated by factors such as musical expertise, the degree of unexpectedness of a stimulus and its familiarity (Besson & Faïta, 1995; Nittono et al., 2000). Also, late positive potentials (LPPs), such as the P300, are related to the evaluation of stimulus salience and increases in their amplitude have been related to decreased stimulus probability (see, for example, Polich, 2007). The existing evidence indicates that these late components emerge both for linguistic and non-linguistic stimuli (Lelekov-Boissard & Dominey, 2002). Hence, irrespective of stimulus type, late components might be related to the processing of incongruences in rule-governed sequences (Tabullo et al., 2011). In our study, both the latency and the scalp distribution of the effect of training during prosodic words' processing refer to a late positive modulation of the EEG signal that has commonalities with the late positive potentials. Based on evidence that LPPs amplitude reflects participants' expectations and also that the amplitude of this component is inversely proportional to the expectation of its eliciting event, we speculate that the enhanced positivity observed in non-musicians might index increased difficulty in predicting these words compared to musicians.

In the linguistic experiment (Experiment 2), differences between groups tended to emerge in the form of a negative-going wave, in the 150-200 ms time window, that was enhanced in amplitude for musicians. This finding suggests that the computation of regularities of speech streams is affected by its suprasegmental features (e.g., prosodic information) and by training. While musical experts exhibited a late modulation of the ERP response when presented with a sung language, they showed an earlier modulation of the ERP response to a speech stream devoid of pitch contours. Of note, neither musicians nor non-musicians could detect mismatches between words, part-words and non-words. It is plausible that this occurred both due to the short time of exposure to the speech stream and to the lack of attention to the sounds.

5.5.2.2 Pre-attentive segmentation of musical sounds

The enhanced positivity exhibited by musicians in the first 50 ms after the melodies' onset indicates that musical training impacts upon the very early stages of melodies processing (Experiment 3). The enhanced positivity shown by musicians was observed from the beginning of the learning phase. The early modulation may be associated with stimulus familiarity: musicians might be more familiar with musical sounds than non-musicians due to extensive training and practice of a musical instrument. The timing of this group difference may also indicate differences in the sensory processing of musical sounds. For example, Paraskevopoulos and colleagues (Paraskevopoulos et al., 2012) found evidence of an enhanced P50 response in musicians in response to an oddball paradigm composed of tone sequences.

Moreover, musicians exhibited an enhanced frontal positivity in the 550-600 ms latency window, overlapping in time with the presentation of the third tone of each melody. This group difference emerged in the early part of the learning phase (1-200 trials). A previous study (Besson & Faïta, 1995) has reported LPC modulations by musical training: LPC amplitude was enhanced in musicians compared to non-musicians. Also, late ERP positivities have been proposed to reflect effects of musical expectancy – the ability to predict the emergence of musical sounds (Besson & Macar, 1987; Paller et al., 1992). The enhanced P550/600 observed in musicians may indicate a better capacity to compute TPs between tones and to detect musical regularities. This enhancement was observed in the first 200 trials of exposure to the musical stream, suggesting that the effects of musical expertise are manifested right at the beginning of the learning phase, but vanish with increased stimulus repetitions. Consistently, other studies have reported that musicians (Vasuki et al., 2017b) and participants with high learning profiles (Abla, Katahira, & Okanoya, 2008; Sanders, Newport, & Neville, 2002) show enhanced ERP changes at the beginning of the exposure to the streams. Thereby, even when task-irrelevant, musical sounds might represent 'special' stimuli that musicians and non-musicians process differently. This assumption is also corroborated by the observation of an enhanced positivity in musicians approximately at 600 ms post-stimulus (i.e., part-melodies) onset in the implicit test phase.

5.5.3 Implicit and explicit outcomes of pre-attentive ASL

5.5.3.1 Implicit evidence of online detection of statistical violations

Contrary to our hypotheses, MMN/N200 components were not elicited by structural violations of the auditory streams. These hypotheses were based on evidence from a speech segmentation study of François and collaborators (François, Cunillera, Garcia, Laine, & Rodriguez-Fornells, 2017). In this report, participants were presented with learning and implicit test (e.g., words vs. non-words) phases and were instructed to pay attention to the streams in order to identify the words of and "alien" language. In the current study, participants were instructed to ignore the sounds and to focus their attention on a video. Such early ERP components (MMN/N200) are sensitive to attention. A plausible explanation for the current results is that participants have not developed a strong representation of the streams' structure. Therefore, there were no differences in the MMN response to streams' items and new part/non-items. This is not to say that the segmentation process was not achieved: explicit measures of learning revealed that participants learned the sung language (Experiment 1). Moreover, musicians showed a larger positivity 600-700 ms in response to part-melodies during the implicit test phase (Experiment 3). Part-melodies had BCA or CAB forms that represented sequences of tones that could have occurred in the learning phase but to a much lesser extent than the actual melodies. It was harder to discriminate between melodies and part-melodies than between melodies and non-melodies since the latter were novel, illegal items. Based on evidence from studies probing late positive components (see Osterhout & Holcomb, 1992; Osterhout & Mobley, 1995; Patel, Gibson, Ratner, Besson, & Holcomb, 1998), we may consider that the positivity elicited by part-melodies represent one of such components both due to its morphology (e.g., polarity, latency) and functional significance. Late components are associated with memory-related context update processes (Besson & Faïta, 1995; Loui, Grent-'t-Jong, Torpey, & Woldorff, 2005). The N400 is known to index cognitive processes related to semantic expectancy (Kutas & Federmeier, 2011; Kutas & Hillyard, 1980; Osterhout & Mobley, 1995) whereas LPC components were found to index musical expectancy (Besson & Macar, 1987; Besson & Faïta, 1995). Considering this evidence, the 600-700 ms positivity found in response to part-melodies might represent a response to the unexpectedness of these items based on the preceding statistical context. The increased positivity found in musicians may reflect their enhanced ability to compute TPs and to detect violations in the statistical structure of the continuous musical stream. This interpretation is supported by prior studies (Paraskevopoulos et al., 2012) demonstrating that the segmentation of tone sequences occurs at a pre-attentive level. The analysis of the implicit test phase also revealed no group differences in the online detection of statistical mismatches in the artificial speech streams (Experiments 1 and 2).

5.5.3.2 Explicit measures: behavioral performance in LDTs

Due to experimental constraints, only the prosodic experiment (Experiment 1) included an explicit SL test phase. The current behavioral data do not support an advantage of musical training in speech segmentation when attention is focused elsewhere. Despite electrophysiological group differences during speech processing, the behavioral performance of participants in the prosodic LDT did not differ as a function of musical training. Based on the divergent findings of the current study and previous reports (François et al., 2013, 2014), it is plausible that the benefits of musical training observed in behavioral indices of speech segmentation represent a consequence of the enhanced attentional skills of musicians. Thereby, these benefits are observed only when participants were actively trying to parse out speech streams. However, the results of the musical test showed a tendency of musicians to outperform non-musicians. The prosodic and the musical tests examined distinct learning outcomes: the first investigated word recognition while the later tested the ability to transfer the knowledge of the auditory regularities to new musical sequences. Therefore, whereas there was no effect of training on the behavioral recognition of the streams' items, training might still be advantageous in the generalization of knowledge of the statistical structure of a stream to new sequences of sounds with melodic properties. These results challenge prior claims for cross-domain neuroplasticity in extensive musical training. However, they highlight two relevant aspects that SL research should take into account. First, the relevance of properly defining the cognitive processes under study when specific behavioral tasks are used (and consequently when comparing the results of the SL literature). Second, the fact that training may only benefit specific cognitive processes related to pre-attentive ASL (i.e., generalization of knowledge). Moreover, the performance of participants was increased in the prosodic test compared to the musical test, which is not surprising giving that the musical test was challenging.

5.5.4 Limitations and future directions

Due to experimental constraints, the learning of the non-prosodic words and melodies was not evaluated behaviorally. This aspect prevented the investigation of the representations of the auditory streams that were effectively stored in long-term memory and that would, in the future, be available for subsequent processing (e.g., word-object mapping). This limitation also constrained our ability to discuss the implications of the effects of musical training on ASL. Besides performing post-learning SL tasks aiming to overcome this limitation, future studies should investigate if the electrophysiological responses observed during the encoding of linguistic and musical regularities in the absence of attention are associated with specific EEG response patterns during behavioral test phases.

5.6 Conclusion

To the best of our knowledge, the present study is the first to provide a systematic investigation of the electrophysiological signatures of ASL under pre-attentive task requirements, probing the effects of musical training and stimulus type. The results show that, even when attention is not focused on the auditory streams, musical training modulates the neurofunctional mechanisms underlying SL distinctly as a function of stimulus type. Specifically, the electrophysiological responses characterizing the learning of musical sequences, and also the detection of mismatches in these sequences, were distinct in musicians. Our findings also challenge previous studies that examined the impact of musical training on speech segmentation: in the absence of an active attentional effort, musicians exhibit similar explicit speech segmentation abilities compared to musically-naïve individuals. Nevertheless, musicians tend to be better than non-musicians at generalizing the learned statistical regularities to new musical input. These findings do not provide evidence of cross-domain neuroplasticity. Together with previous studies, the current work indicates that the impact of musical training upon auditory learning mechanisms is modulated by selective attention. These findings contribute to the clarification of the role of attention on auditory stream segmentation and add to the broad scientific discussion on the automatic/implicit nature of SL.

5.7 References

Abla, D., Katahira, K., & Okanoya, K. (2008). On-line assessment of statistical learning by event-related potentials. *Journal of Cognitive Neuroscience*, *20*(6), 952–964. http://doi.org/10.1162/jocn.2008.20058

Abla, D., & Okanoya, K. (2009). Visual statistical learning of shape sequences: An ERP study. *Neuroscience Research*, *64*(2), 185–190. http://doi.org/10.1016/j.neures.2009.02.013

Ahissar, M., & Hochstein, S. (1993). Attentional Control of Early Perceptual Learning. *Proceedings of the National Academy of Sciences*, *90*(12), 5718–5722.

Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology*, *83*(2), 111–130. http://doi.org/10.1016/S0022-0965(02)00124-8

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8 month old infants. *Psychological Science*, *9*(4), 321–324. http://doi.org/http://doi.org/10.1111/1467-9280.00063

Baumann, S., Meyer, M., & Jäncke, L. (2008). Enhancement of auditory-evoked potentials in musicians reflects an influence of expertise but not selective attention. *Journal of Cognitive Neuroscience*, *20*(12), 2238–2249. http://doi.org/10.1162/jocn.2008.20157

Beres, A. M. (2017). Time is of the Essence: A Review of Electroencephalography (EEG) and Event-Related Brain Potentials (ERPs) in Language Research. *Applied Psychophysiology Biofeedback*, *42*(4), 247–255. http://doi.org/10.1007/s10484-017-9371-3

Bermudez, P., & Zatorre, R. J. (2005). Differences in gray matter between musicians and nonmusicians. *Annals of the New York Academy of Sciences*, *1060*(2005), 395–399. http://doi.org/10.1196/annals.1360.057

Besson, M., & Faïta, F. (1995). An Event-Related Potential (ERP) Study of Musical Expectancy: Comparison of Musicians With Nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(6), 1278–1296. http://doi.org/10.1037/0096-1523.21.6.1278

Besson, M., & Macar, F. (1987). An event-related potential analysis of incongruity in music and other non-linguistic contexts. *Psychophysiology*, *24*(1), 14–25.

Bialystok, E., & DePape, A. M. (2009). Musical Expertise, Bilingualism, and Executive Functioning. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(2), 565–574. http://doi.org/10.1037/a0012735

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain and Cognition*, *77*(1), 1–10. http://doi.org/10.1016/j.bandc.2011.07.006

Bidelman, G. M., & Krishnan, A. (2010). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Research*, *1355*, 112–125. http://doi.org/10.1016/j.brainres.2010.07.100

Boh, B., Herholz, S. C., Lappe, C., & Pantev, C. (2011). Processing of complex auditory patterns in musicians and nonmusicians. *PLoS ONE*, *6*(7). http://doi.org/10.1371/journal.pone.0021458

Boutros, N. N., & Belger, A. (1999). Midlatency evoked potentials attenuation and augmentation reflect different aspects of sensory gating. *Biological Psychiatry*, *45*(7), 917–922. http://doi.org/10.1016/S0006-3223(98)00253-4

Brattico, E., Pallesen, K. J., Varyagina, O., Bailey, C., Anourova, I., Järvenpää, M., … Tervaniemi, M. (2009). Neural discrimination of nonprototypical chords in music experts and laymen: An MEG study. *Journal of Cognitive Neuroscience*, *21*(11), 2230–2244. http://doi.org/10.1162/jocn.2008.21144

Carrión, R. E., & Bly, B. M. (2007). Event-related potential markers of expectation violation in an artificial grammar learning task. *NeuroReport*, *18*(2), 191–195. http://doi.org/10.1097/WNR.0b013e328011b8ae

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2009). Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain and Language*, *108*(1), 1–9. http://doi.org/10.1016/j.bandl.2008.02.001

Chobert, J., François, C., Velay, J. L., & Besson, M. (2012). Twelve months of active musical training in 8-to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral Cortex*, *24*(4), 956–967. http://doi.org/10.1093/cercor/bhs377

Cunillera, T., Càmara, E., Toro, J. M., Marco-Pallares, J., Sebastián-Galles, N., Ortiz, H., ... Rodríguez-Fornells, A. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *NeuroImage*, *48*(3), 541–553. http://doi.org/10.1016/j.neuroimage.2009.06.069

Cunillera, T., Toro, J. M., Sebastián-Gallés, N., & Rodríguez-Fornells, A. (2006). The effects of stress and statistical cues on continuous speech segmentation: An event-related brain potential study. *Brain Research*, *1123*(1), 168–178. http://doi.org/10.1016/j.brainres.2006.09.046

de Diego-Balaguer, R., Rodríguez-Fornells, A., & Bachoud-Lévi, A.-C. (2015). Prosodic cues enhance rule learning by changing speech segmentation mechanisms. *Frontiers in Psychology*, *6*(September). http://doi.org/10.3389/fpsyg.2015.01478

de Diego-Balaguer, R., Toro, J. M., Rodriguez-Fornells, A., & Bachoud-Lévi, A. C. (2007). Different neurophysiological mechanisms underlying word and rule extraction from speech. *PLoS ONE*, *2*(11). http://doi.org/10.1371/journal.pone.0001175

Ding, N., Pan, X., Luo, C., Su, N., Zhang, W., & Zhang, J. (2017). Attention is required for knowledge-based sequential grouping: Insights from the integration of syllables into words. *The Journal of Neuroscience*, 2606–17. http://doi.org/10.1523/JNEUROSCI.2606-17.2017

Dittinger, E., Barbaroux, M., D'Imperio, M., Jäncke, L., Elmer, S., & Besson, M. (2016). Professional music training and novel word learning: from faster semantic encoding to longer-lasting word representations. *Journal of Cognitive Neuroscience*, *28*(10), 1584–1602. http://doi.org/10.1162/jocn

Dittinger, E., Chobert, J., Ziegler, J. C., & Besson, M. (2017). Fast Brain Plasticity during Word Learning in Musically-Trained Children. *Frontiers in Human Neuroscience*, *11*(May), 1–16. http://doi.org/10.3389/fnhum.2017.00233

Dittinger, E., Valizadeh, S. A., Jäncke, L., Besson, M., & Elmer, S. (2018). Increased functional connectivity in the ventral and dorsal streams during retrieval of novel words in professional musicians. *Human Brain Mapping*, *39*(2), 722–734. http://doi.org/10.1002/hbm.23877

Donchin, E. (1981). Surprise!...Surprise? *Psychophysiology*. http://doi.org/10.1111/j.1469-8986.1981.tb01815.x

Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., ... Van Petten,

C. (2009). Event-related potentials in clinical research: Guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clinical Neurophysiology*, *120*(11), 1883–1908. http://doi.org/10.1016/j.clinph.2009.07.045

Ehrlé, N., & Samson, S. (2005). Auditory discrimination of anisochrony: Influence of the tempo and musical backgrounds of listeners. *Brain and Cognition*, *58*(1), 133–147. http://doi.org/10.1016/j.bandc.2004.09.014

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. http://doi.org/10.3758/BRM.41.4.1149

Fiser, J., & Aslin, R. N. (2002a). Statistical Learning of Higher-Order Temporal Structure from Visual Shape Sequences. *Journal of Experimental Psychology: Learning Memory and Cognition*, *28*(3), 458–467. http://doi.org/10.1037/0278-7393.28.3.458

Fiser, J., & Aslin, R. N. (2002b). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(24), 15822–6. http://doi.org/10.1073/pnas.232472899

Fiser, J., & Aslin, R. N. (2005). Encoding multielement scenes: Statistical learning of visual feature hierarchies. *Journal of Experimental Psychology: General*, *134*(4), 521–537. http://doi.org/10.1037/0096-3445.134.4.521

Foster, N. E. V, & Zatorre, R. J. (2010). Cortical structure predicts success in performing musical transformation judgments. *NeuroImage*, *53*(1), 26–36. http://doi.org/10.1016/j.neuroimage.2010.06.042

François, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex*, *23*(9), 2038–2043. http://doi.org/10.1093/cercor/bhs180

François, C., Cunillera, T., Garcia, E., Laine, M., & Rodriguez-Fornells, A. (2017). Neurophysiological evidence for the interplay of speech segmentation and word-referent mapping during novel word learning. *Neuropsychologia*, *98*(October), 56–67. http://doi.org/10.1016/j.neuropsychologia.2016.10.006

François, C., Jaillet, F., Takerkart, S., & Schön, D. (2014). Faster sound stream segmentation in musicians than in nonmusicians. *PLoS ONE*, *9*(7), 26–29.

http://doi.org/10.1371/journal.pone.0101340

François, C., & Schön, D. (2010). Learning of musical and linguistic structures: Comparing event-related potentials and behavior. *NeuroReport*, *21*(14), 928–932. http://doi.org/10.1097/WNR.0b013e32833ddd5e

François, C., & Schön, D. (2011). Musical expertise boosts implicit learning of both musical and linguistic structures. *Cerebral Cortex*, *21*(10), 2357–2365. http://doi.org/10.1093/cercor/bhr022

François, C., & Schön, D. (2014). Neural sensitivity to statistical regularities as a fundamental biological process that underlies auditory learning: The role of musical practice. *Hearing Research*, *308*, 122–128. http://doi.org/10.1016/j.heares.2013.08.018

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*, *129*(10), 2593–2608. http://doi.org/10.1093/brain/awl247

George, E. M., & Coch, D. (2011). Music training and working memory: An ERP study. *Neuropsychologia*, *49*(5), 1083–1094. http://doi.org/10.1016/j.neuropsychologia.2011.02.001

Gilbert, C., Ito, M., Kapadia, M., & Westheimer, G. (2000). Interactions between attention, context and learning in primary visual cortex. *Vision Research*, *40*(10–12), 1217–1226. http://doi.org/10.1016/S0042-6989(99)00234-5

Gordon, E. E. (1989). *Advanced measures of music audiation.* Chicago, IL: G.I.A. Publications.

Graffar, M. (1956). Une méthode de classification d'échantillons de la population. *Courrier*, *6*, 455.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468–484.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (sps) as an erp measure of syntactic processing. *Language and Cognitive Processes*, *8*(4), 439–483. http://doi.org/10.1080/01690969308407585

Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: effects of enculturation and formal training on development. *Trends in Cognitive Sciences*, *11*(11), 466–472. http://doi.org/10.1016/j.tics.2007.08.008

Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., & Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology*, *42*(2), 180–190. http://doi.org/10.1111/j.1469-8986.2005.00272.x

Herholz, S. C., Lappe, C., Knief, A., & Pantev, C. (2009). Imagery mismatch negativity in musicians. *Annals of the New York Academy of Sciences*, *1169*, 173–177. http://doi.org/10.1111/j.1749-6632.2009.04782.x

Herholz, S. C., & Zatorre, R. J. (2012). Musical Training as a Framework for Brain Plasticity: Behavior, Function, and Structure. *Neuron*, *76*(3), 486–502. http://doi.org/10.1016/j.neuron.2012.10.011

Ho, Y.-C., Cheung, M.-C., & Chan, A. S. (2003). Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. *Neuropsychology*, *17*(3), 439–450. http://doi.org/10.1037/0894-4105.17.3.439

Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). The effects of musical training on structural brain development: A longitudinal study. *Annals of the New York Academy of Sciences*, *1169*, 182–186. http://doi.org/10.1111/j.1749-6632.2009.04852.x

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*(2), 159–201. http://doi.org/10.1080/016909600386084

Koelsch, S., Gunter, T. C., V. Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: A cortical "language-network" serves the processing of music. *NeuroImage*, *17*(2), 956–966. http://doi.org/10.1016/S1053-8119(02)91154-7

Koelsch, S., Schröger, C. A. E., & Tervaniemi, M. (1999a). Superior pre-attentive auditory processing in musicians, *10*(6), 1309–1313.

Koelsch, S., Schröger, E., & Tervaniemi, M. (1999b). Superior pre-attentive auditory processing in musicians. *NeuroReport*, *10*(6), 1309–1313. http://doi.org/10.1097/00001756-199904260-00029

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Science and Society*, *11*(August), 599–606. http://doi.org/10.1038/nrn2882

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews*

Neuroscience, *5*(11), 831–843. http://doi.org/10.1038/nrn1533

Kujala, T., Kallio, J., Tervaniemi, M., & Näätänen, R. (2001). The mismatch negativity as an index of temporal processing in audition. *Clinical Neurophysiology*, *112*(9), 1712–1719. http://doi.org/10.1016/S1388-2457(01)00625-3

Kutas, M., & Federmeier, K. D. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology*. http://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*. http://doi.org/10.1126/science.7350657

Lelekov-Boissard, T., & Dominey, P. F. (2002). Human brain potentials reveal similar processing of non-linguistic abstract structure and linguistic syntactic structure. *Neurophysiologie Clinique*, *32*(1), 72–84. http://doi.org/10.1016/S0987-7053(01)00291-X

Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion*, *11*(5), 1021–1031. http://doi.org/10.1037/a0024521

Loui, P., Grent-'t-Jong, T., Torpey, D., & Woldorff, M. (2005). Effects of attention on the neural processing of harmonic syntax in Western music. *Cognitive Brain Research*, *25*(3), 678–687. http://doi.org/10.1016/j.cogbrainres.2005.08.019

Loui, P., & Wessel, D. (2007). Harmonic expectation and affect in Western music: Effects of attention and training. *Perception and Psychophysics*, *69*(7), 1084–1092. http://doi.org/10.3758/BF03193946

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique* (Second, Vol. 78). Cambridge, MA: MIT Press. http://doi.org/10.1118/1.4736938

Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, *18*(2), 199–211. http://doi.org/10.1162/jocn.2006.18.2.199

Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O., & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *Journal of Cognitive Neuroscience*, *23*(10), 2701–2715. http://doi.org/doi: 10.1162/jocn.2010.21585

Marie, C., Magne, C., & Besson, M. (2011). Musicians and the metric structure of words. *Journal of Cognitive Neuroscience*, *23*(2), 294–305. http://doi.org/10.1162/jocn.2010.21413

Marques, C., Moreno, S., Castro, S. L., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience*, *19*(9), 1453–1463. http://doi.org/10.1162/jocn.2007.19.9.1453

Martín-Loeches, M., Sommer, W., & Hinojosa, J. A. (2005). ERP components reflecting stimulus identification: Contrasting the recognition potential and the early repetition effect (N250r). *International Journal of Psychophysiology*, *55*(1), 113–125. http://doi.org/10.1016/j.ijpsycho.2004.06.007

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, *219*(1–2), 36–47. http://doi.org/10.1016/j.heares.2006.05.004

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. http://doi.org/10.1093/cercor/bhn120

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. http://doi.org/10.1073/pnas.0701498104

Näätänen, R. (2002). The perception of speech sounds by the human brain as reflected by the mismatch negativity brain response. *International Congress Series*, *1232*, 97–105. http://doi.org/10.1016/S0531-5131(01)00795-6

Näätänen, R., & Gaillard, A. W. K. (1983). 5 The Orienting Reflex and the N2 Deflection of the Event-Related Potential (ERP). In *Advances in psychology* (pp. 119–141). North-Holland. http://doi.org/0.1016/s0166-4115(08)62036-1

Näätänen, R., Gaillard, A. W. K., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, *42*(4), 313–329. http://doi.org/10.1016/0001-6918(78)90006-9

Näätänen, R., Jacobsen, T., & Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): A review of the evidence. *Psychophysiology*, *42*(1), 25–32.

http://doi.org/10.1111/j.1469-8986.2005.00256.x

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*, *118*(12), 2544–2590. http://doi.org/10.1016/j.clinph.2007.04.026

Nasreddine, Z., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., & Collin, I. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *American Geriatrics Society*, *53*(4), 695–699. http://doi.org/10.1111/j.1532-5415.2005.53221.x

Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically Based Sentence Processing Classes: Evidence from Event-Related Brain Potentials. *Journal of Cognitive Neuroscience*, *3*(2), 151–165. http://doi.org/10.1162/jocn.1991.3.2.151

Nittono, H., Bito, T., Hayashi, M., Sakata, S., & Hori, T. (2000). Event-related potentials elicited by wrong terminal notes: Effects of temporal disruption. *Biological Psychology*, *52*(1), 1–16. http://doi.org/10.1016/S0301-0511(99)00042-3

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. http://doi.org/10.1016/0028-3932(71)90067-4

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785–806. http://doi.org/10.1016/0749-596X(92)90039-Z

Osterhout, L., & Mobley, L. A. (1995). Event-Related Brain Potentials Elicited by Failure to Agree. *Journal of Memory and Language*. http://doi.org/10.1006/jmla.1995.1033

Paller, K. A., McCarthy, G., & Wood, C. C. (1992). Event-Related Potentials Elicited by Deviant Endings to Melodies. *Psychophysiology*. http://doi.org/10.1111/j.1469-8986.1992.tb01686.x

Pallesen, K. J., Brattico, E., Bailey, C. J., Korvenoja, A., Koivisto, J., Gjedde, A., & Carlson, S. (2010). Cognitive control in auditory working memory is enhanced in musicians. *PLoS ONE*, *5*(6). http://doi.org/10.1371/journal.pone.0011120

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, *392*(6678), 811–814. http://doi.org/10.1038/33918

Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., & Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *NeuroReport*, *12*(1), 169–174. http://doi.org/10.1097/00001756-200101220-00041

Paraskevopoulos, E., Kuchenbuch, A., Herholz, S. C., & Pantev, C. (2012). Statistical learning effects in musicians and non-musicians: An MEG study. *Neuropsychologia*, *50*(2), 341–349. http://doi.org/10.1016/j.neuropsychologia.2011.12.007

Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical Experience Limits the Degradative Effects of Background Noise on the Neural Processing of Sound. *Journal of Neuroscience*, *29*(45), 14100–14107. http://doi.org/10.1523/JNEUROSCI.3256-09.2009

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations inlanguage and music: an event-related potential study. *J Cog Neurosci*, *10*, 717–733. http://doi.org/10.1162/089892998563121

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233–238. http://doi.org/10.1016/j.tics.2006.03.006

Pinheiro, A. P., Vasconcelos, M., Dias, M., Arrais, N., & Gonçalves, Ó. F. (2015). The music of language: An ERP investigation of the effects of musical training on emotional prosody processing. *Brain and Language*, *140*, 24–34. http://doi.org/10.1016/j.bandl.2014.10.009

Polich, J. (2003). *Detection of Change: event-related potential and fMRI findings.* (J. Polich, Ed.). Springer Science & Business Media. http://doi.org/10.1007/978-1-4615-0294-4

Polich, J. (2007). Updating P300: An Integrative Theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. http://doi.org/10.1016/j.clinph.2007.04.019.Updating

Reinke, K. S., He, Y., Wang, C., & Alain, C. (2003). Perceptual learning modulates sensory evoked response during vowel segregation. *Cognitive Brain Research*, *17*(3), 781–791. http://doi.org/10.1016/S0926-6410(03)00202-7

Rellecke, J., Palazova, M., Sommer, W., & Schacht, A. (2011). On the automaticity of emotion processing in words and faces: Event-related brain potentials evidence from a superficial task. *Brain and Cognition*, *77*(1), 23–32. http://doi.org/10.1016/j.bandc.2011.07.001

Ritter, W., Paavilainen, P., Lavikainen, J., Reinikainen, K., Alho, K., Sams, M., & Näätänen, R. (1992). Event-related potentials to repetition and change of auditory stimuli, *83*, 306–321.

Röder, B., Rösler, F., Hennighausen, E., & Näcker, F. (1996). Event-related potentials during auditory and somatosensory discrimination in sighted and blind human subjects. *Cognitive Brain Research*, *4*(2), 77–93. http://doi.org/10.1016/0926-6410(96)00024-9

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27–52. http://doi.org/10.1016/S0010-0277(98)00075-4

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*(4), 606–621. http://doi.org/10.1006/jmla.1996.0032

Saffran, J. R., Senghas, A., & Trueswell, J. C. (2001). The acquisition of language by children. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(23), 12874–12875. http://doi.org/10.1073/pnas.231498898

Sams, M., Alho, K., & Näätänen, R. (1983). Sequential effects on the ERP in discriminating two stimuli. *Biological Psychology*, *17*(1), 41–58. http://doi.org/10.1001/jamaoto.2016.3292

Sanders, L. D., Newport, E. L., & Neville, H. J. (2002). Segmenting nonsense: An event-related potential index of perceived onsets in continuous speech. *Nature Neuroscience*, *5*(7), 700–703. http://doi.org/10.1038/nn873

Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience*, *5*(7), 688–694. http://doi.org/10.1038/nn871

Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition*, *106*(2), 975–983. http://doi.org/10.1016/j.cognition.2007.03.005

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, *41*(3), 341–349. http://doi.org/10.1111/1469-8986.00172.x

Scott, L. (1992). Attention and Perseverance Behaviors of Preschool Children Enrolled in Suzuki Violin Lessons and Other Activities. *Journal of Research in Music Education*, *40*(3), 225–235. http://doi.org/10.2307/3345684

Selchenkova, T., François, C., Schön, D., Corneyllie, A., Perrin, F., & Tillmann, B. (2014). Metrical presentation boosts implicit learning of artificial grammar. *PLoS ONE*, *9*(11). http://doi.org/10.1371/journal.pone.0112233

Shahin, A., Bosnyak, D. J., Trainor, L. J., & Roberts, L. E. (2003). Enhancement of Neuroplastic P2 and N1c Auditory Evoked Potentials in Musicians. *Journal of Neuroscience*, *23*(13), 5545–5552. http://doi.org/10.1523/JNEUROSCI.23-13-05545.2003

Shanks, D. R., Rowland, L. A., & Ranger, M. S. (2005). Attentional load and implicit sequence learning. *Psychological Research*, *69*(5–6), 369–382. http://doi.org/10.1007/s00426-004-0211-8

Sheehan, K. A., McArthur, G. M., & Bishop, D. V. M. (2005). Is discrimination training necessary to cause changes in the P2 auditory event-related brain potential to speech sounds? *Cognitive Brain Research*, *25*(2), 547–553. http://doi.org/10.1016/j.cogbrainres.2005.08.007

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, *81*, 105–120. http://doi.org/10.1016/j.jml.2015.02.001

Simões, M. R., Freitas, S., Santana, I., Firmino, H., Martins, C., Nasreddine, Z., & Vilar, M. (2008). Montreal cognitive assessment (MoCA): versão final portuguesa. *Coimbra: Serviço de Avaliação Psicológica, Faculdade de Psicologia e de Ciências Da Educação Da Universidade de Coimbra.*

Simons, R. F., Graham, F. K., Miles, M. A., & Chen, X. (2001). On the relationship of P3a and the Novelty-P3. *Biological Psychology*, *56*(3), 207–218. Retrieved from papers3://publication/uuid/32F1F356-4F45-4D03-A28F-1F0F872E0679

Slevc, L. R., & Miyake, A. (2006). Individual Differences in Second- Language Proficiency, *17*(8), 675–681. http://doi.org/10.1111/j.1467-9280.2006.01765.x

Slevc, L. R., Rosenberg, J. C., & Patel, A. D. (2009). Making psycholinguistics musical: Self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychonomic Bulletin and Review*, *16*(2), 374–381. http://doi.org/10.3758/16.2.374

Snyder, J. S., Alain, C., & Picton, T. W. (2006). Effects of attention on neuroelectric correlates of auditory stream segregation. *Journal of Cognitive Neuroscience*, *18*(1), 1–13. http://doi.org/10.1162/089892906775250021

Strait, D. L., & Kraus, N. (2011). Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Frontiers in Psychology*, *2*(113), 1–10. http://doi.org/10.3389/fpsyg.2011.00113

Strait, D. L., Kraus, N., Parbery-Clark, A., & Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: Evidence from masking and auditory attention performance. *Hearing Research*, *261*(1–2), 22–29. http://doi.org/10.1016/j.heares.2009.12.021

Tabullo, Á., Sevilla, Y., Pasqualetti, G., Vernis, S., Segura, E., Yorio, A., ... Wainselboim, A. (2011). Expectancy modulates a late positive ERP in an artificial grammar task. *Brain Research*, *1373*(1428), 131–143. http://doi.org/10.1016/j.brainres.2010.12.023

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: An event-related potential and behavioral study. *Experimental Brain Research*, *161*(1), 1–10. http://doi.org/10.1007/s00221-004-2044-5

Tervaniemi, M., Rytkönen, M., Schröger, E., Ilmoniemi, R. J., & Näätänen, R. (2001). Superior formation of cortical memory traces for melodic patterns in musicians. *Learning and Memory*, *8*(5), 295–300. http://doi.org/10.1101/lm.39501

Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, *97*(2), 25–34. http://doi.org/10.1016/j.cognition.2005.01.006

Tremblay, K., Kraus, N., & McGee, T. (1998). The time course of auditory perceptual learning: Neurophysiological changes during speech-sound training. *NeuroReport*, *9*(16), 3557–3560. http://doi.org/10.1097/00001756-199811160-00003

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*(4), 552–564. http://doi.org/10.1037/0096-3445.134.4.552

Vasuki, P. R. M., Sharma, M., Ibrahim, R., & Arciuli, J. (2017a). Statistical learning and auditory processing in children with music training: An ERP study. *Clinical Neurophysiology*, *128*(7), 1270–1281. http://doi.org/10.1016/j.clinph.2017.04.010

Vasuki, P. R. M., Sharma, M., Ibrahim, R. K., & Arciuli, J. (2017b). Musicians' Online Performance during Auditory and Visual Statistical Learning Tasks. *Frontiers in Human Neuroscience*, *11*(114), 1–12. http://doi.org/10.3389/fnhum.2017.00114

Werheid, K., Alpay, G., Jentzsch, I., & Sommer, W. (2005). Priming emotional facial expressions as evidenced by event-related brain potentials. *International Journal of Psychophysiology*, *55*(2), 209–219. http://doi.org/10.1016/j.ijpsycho.2004.07.006

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*(4), 420–422. http://doi.org/10.1038/nn1872

PART III

Concluding remarks

Chapter 6. Final Considerations

6.1 Introduction

The present Dissertation was inspired by musical training as a model of neuroplasticity that allows the investigation of the effects of long-term instrumental training on auditory abilities. Simultaneously, it builds upon an intellectual curiosity for the efficient and complex cognitive capacity to extract patterns from external, intricate auditory inputs. The existing literature reveals positive transfer effects between musical training and speech processing (Besson, Chobert, & Marie, 2011; Chobert, François, Velay, & Besson, 2012; Gordon, Fehd, & McCandliss, 2015; Lima & Castro, 2011; Magne, Schön, & Besson, 2006; Moreno et al., 2009; Swaminathan & Schellenberg, 2017) and instigated us to examine possible neural and behavioral mechanisms accounting for musicians' enhanced speech processing skills, such as ASL.

To shed light on this question, we designed a set of three empirical studies aiming to investigate if ASL skills are at the core of the abovementioned transfer effects between domains. More specifically, the studies reported in this Dissertation aimed to examine ASL as a function of musical expertise, stimulus type and attention. In particular, we sought to clarify how musicians and non-musicians process statistical regularities among musical and linguistic sequences under attentive and pre-attentive conditions. Two experimental studies, each composed of three distinct experiments, were performed to shed light on the processes involved in the SL of auditory regularities. One of the major strengths of these studies was the collection of EEG data while participants were exposed to different types of auditory structures (i.e., in the learning stage). Thus, besides evaluating the behavioral outcomes of SL during post-learning tasks, we investigated the neural correlates while the acoustic streams were being processed. This approach allowed us to better understand the neurofunctional mechanisms underlying online ASL as a function of learning time.

Additionally, we also aimed to clarify the role of musical training in the SL of simple – pure tones - and complex – linguistic and musical sequences - auditory contingencies when the auditory input was task-irrelevant.

Together, the results of the present Dissertation support the notion that the SL of complex auditory regularities is constrained by musical expertise, stimulus type and attention. The results demonstrate that, irrespective of attentional focus, musical expertise modulates the brain mechanisms recruited during the extraction of regularities present in streams of words and melodies - and that this effect interacts with the acoustic nature of the stimuli. Musicians showed behavioral advantages in the learning of both musical and linguistic streams when participants' attention was actively and explicitly focused on these materials. Nonetheless, when prosodic words were presented as background stimulation, the behavioral advantage of musicians in the learning of these sounds disappeared. Therefore, the behavioral evidence from different studies indicates that the parsing of sung speech only benefits from musical training if individuals are actively allocating attention to the auditory input. Interpreted from a different perspective, our findings also reveal that individuals, in general, are able to compute the regularities of sung speech streams even when not paying attention to them. We also found that musical training is not associated with enhanced pre-attentive neural responses to simple pitch deviances but only with a tendency for faster deviance detection. This finding suggests that, when the auditory input has a simple statistical structure (i.e., a structure dictated by changes in only one acoustic parameter), musicians and non-musicians process acoustic regularities similarly.

Below, we first discuss specific aspects of these results, namely: the effects of musical expertise both on neural and behavioral ASL outcomes, the effects of attention on ASL and its interaction with expertise, the distinct outcomes of SL based on stimulus type, along with the innovative aspect of studying how the brain implicitly detects structural violations in the input's structure. The limitations of these studies are highlighted. We further outline possible strategies to overcome these limitations in future studies and refer relevant opportunities for additional research into these questions. We conclude by evaluating the theoretical implications of the present findings for the understanding of musical training as a neuroplasticity model and for a broad comprehension of the transfer effects from music to speech.

6.1.1 ASL skills in musicians and non-musicians

During the last decades, a growing body of literature has dedicated efforts to understand the cognitive alterations stemming from long-term and intense contact with a musical instrument (Kraus & Chandrasekaran, 2010; Patel, 2003; Patel & Iversen, 2007; Strait & Kraus, 2011; Strait, Kraus, Skoe, & Ashley, 2009; Tierney & Kraus, 2013). The cognitive effects of musical training have emerged as a model of auditory learning (Herholz & Zatorre, 2012; Strait & Kraus, 2011). Research converges to demonstrate that musical training is associated with structural and functional changes in the auditory cortex, namely increased grey matter volume (right-lateralized) in the posterolateral aspect of Heschl's gyrus (Bermudez, Lerch, Evans, & Zatorre, 2009), enhanced functional connectivity between auditory and motor cortices (Du & Zatorre, 2017), larger (higher) amplitude of evoked responses to auditory stimuli (Baumann, Meyer, & Jäncke, 2008; Brown et al., 2017; Geiser, Sandmann, Jäncke, & Meyer, 2010; Shahin, Bosnyak, Trainor, & Roberts, 2003) and greater FFRs to sounds (Bidelman & Alain, 2015; Bidelman, Gandour, & Krishnan, 2011; Bidelman & Krishnan, 2010). One of the most intriguing alterations associated to musical training is the distinctive way in which musicians process speech stimuli which has been observed both at neural and behavioral levels (Flaugnacco et al., 2015; François, Jaillet, Takerkart, & Schön, 2014; Intartaglia, White-Schwoch, Kraus, & Schön, 2017; Magne et al., 2006; Marie, Delogu, Lampis, Belardinelli, & Besson, 2011; Schön, Magne, & Besson, 2004; Wong, Skoe, Russo, Dees, & Kraus, 2007). Based on these results and on evidence of neural overlap between music and speech, a growing body of literature probed the positive transfer effects between the two domains. A bulk of studies are now available confirming these positive effects (Besson et al., 2011; Kraus & Chandrasekaran, 2010; Schellenberg & Peretz, 2008). However, a comprehensive explanation of the cognitive processes and neural mechanisms accounting for the enhanced ability of musicians to process speech is lacking. Some research suggests that the great ease of musicians to process such complex stimuli as speech derives from an increased sensitivity to process acoustic information in general (Magne et al., 2006; Nikjeh, Lister, & Frisch, 2008; Schön et al., 2004; Wong et al., 2007), more specifically a fine-grained sensitivity to process pitch. Other studies indicate that the explanation for these effects goes beyond common acoustic processing across domains: musical training might change the way the brain builds up percepts and cognitive representations of linguistic sounds (for a review, see Besson et al., 2011). From our perspective, the long-term interaction that musicians have with music and with the practice of musical instruments develops their internal cognitive models of the statistical regularities governing the musical pieces to which they are exposed. This experience enhances the probabilistic prediction mechanisms that enable listeners to process and categorize the mental representations of auditory input. These learned skills might be modality-general and apply to other types of auditory information that have a statistical structure, such as speech. The present Dissertation was built up from the need to gather further evidence for this theoretical rationale.

Before further discussing our findings, it should be mentioned that the studies that integrate this work were inspired by the recent research developed by Clément François and colleagues (François, Cunillera, Garcia, Laine, & Rodriguez-Fornells, 2017; François et al., 2014; Francois & Schön, 2010, 2011). In a sense, the prosodic experiment of the second study of this Dissertation (see Chapter 4) can be considered a conceptual replication of the study of François and colleagues (2014). Yet, the experiment presented here – along with the other experiments composing Chapters 4 and 5 - benefited from methodological changes comparatively with the experimental design of François and collaborators (2014). We would like to highlight these changes in the lines ahead before proceeding to the discussion of the results. First, the words used in our studies were composed of new syllables, distinct from the ones used before. Second, while the words were generated using a German phonetic database provided by the Mbrola software, the words used in François and colleagues' study (François et al., 2014) were created using a Spanish database. The option for creating stimuli based on different phonotactic constraints relative to the subjects' native language is explained by the need of generating word items that are not semantically related to anything existent in subjects' lexicon and also completely unfamiliar regarding its intonation properties. Otherwise, we would be compromising an essential criterion of SL studies that is related to stimulus unfamiliarity. Recent evidence points to an effect of prior knowledge about languages' structures on the SL of new artificial languages (Noam Siegelman, Bogaerts, Elazar, Arciuli, & Frost, 2018). We find this a crucial aspect of our experimental manipulation since, in our opinion, studies testing artificial languages with word items that resemble the subjects' linguistic background undermine the researchers' ability to generalize their results and to argue for the investigation of pure SL processes. Additionally, we controlled for stimulus duration, making all the words' length equal to 700 ms. This aspect is relevant both due to the effort of normalizing stimulus features and to comply with the specificities of the ERP technique. Related to this methodological specificity towards the ERPs is also the introduction of pauses in between stimuli: we added a 50 ms ISI (e.g., in between words) along the entire duration of the stream to create silent baseline periods, essential for the ERP analysis. Resulting from these baseline periods, 50 ms intervals in between syllables were also added. If the words were only separated by silent periods participants would need just a few seconds to learn them. Besides, no SL processes could be investigated since participants would very easily retrieve the representations of the words stored in the long-term memory. The insertion of gaps of the same length between syllables and words ensured equal spacing between the streams' most basic entities. This allowed that only TPs between syllable pairs could account for stream segmentation. We consider that the adaptation of the François and collaborators' (2014) paradigm presented here is a positive upgrade on the paradigm's structure, stimuli and, importantly, adequateness to the ERP technique.

The findings reported in Chapter 4 revealed that musical training modulates the way in which the brain derives statistical information pertaining to the acoustic streams' structure, independently of the stimulus nature (e.g., linguistic or musical). Musicians and non-musicians' EEG responses to the speech streams revealed differences 250-300 ms after words' onset. Since these responses were followed by higher accuracy rates in musicians (during the post-learning LDT) the N250-300 might index mechanisms involved in lexical search or semantic integration. It is worth highlighting that, in contrast to our hypotheses, we failed to find clear N400-like effects in this study. However, the lack of these effects needs to be distinguished from the absence of N400-like components. The non-emergence of a N400-like component or the absence of N400-like differences between conditions or groups cannot be taken as evidence that the cognitive processes underlying the N400-like have not occurred (Bornkessel-Schlesewsky & Schlesewsky, 2008; Kolk & Chwilla, 2007). The divergent latency windows reported in the ASL literature (see the Discussion section of Chapter 4; Batterink & Paller, 2017; Cunillera et al., 2009; Cunillera, Laine, & Rodríguez-Fornells, 2016; Cunillera, Toro, Sebastián-Gallés, & Rodríguez-Fornells, 2006; de Diego-Balaguer, Toro, Rodriguez-Fornells, & Bachoud-Lévi, 2007; François et al., 2017, 2014; Vasuki, Sharma, Ibrahim, & Arciuli, 2017) along with specificities of our paradigm that distinguish it from previous studies, make us confident that the negativity we found in response to words (N250-300) indexes predictive processes linked to auditory segmentation. Moreover, since the behavioral evidence shows that only the group of musicians learned the speech streams (prosodic and non-prosodic) and additionally, that the amplitude of the negativity was correlated with the performance on the prosodic test, the differences in this ERP deflection might well index regularity processing mechanisms accounting for participants' ability to learn the statistical structure of the artificial languages. Nevertheless, further studies are needed to investigate if this negativity is indeed a neural marker of stream segmentation.

The findings of the musical experiment of Chapter 4 also revealed differences between groups but in two distinct processing windows (e.g., positivities at 0-100 ms and 300-350 ms). These results corroborate the findings of the linguistic experiments: The online tracking of acoustic regularities is changed by training irrespective of stimulus type. The behavioral performance in the musical LDT (Chapter 4) also showed musicians to outperform non-musicians. This suggests that the amplitude differences found during the learning phase may index relevant time windows for the processing of regularities among musical streams. From a broad perspective, these findings indicate three important aspects about musical trainings' impact on ASL during attentive conditions: (1) musical training benefits the learning of auditory structures independently of the acoustic nature of the stimuli and, (2) despite leading to similar behavioral responses, the SL of linguistic and musical input evokes specific electrophysiological responses, that (3) are distinct between musicians and non-musicians.

Long-term training may enhance the computation of TPs between sounds and contribute to enhance the musicians' ability to parse continuous streams of input into meaningful segments. A recent longitudinal study with children supports this view: Two years of training led to increments in ASL abilities reflected in behavioral responses to a LDT (François, Chobert, Besson, & Schön, 2013). In this context, our findings are critical. We showed for the first time that the improved ASL skills of musicians are not circumscribed to sounds with melodic characteristics, which is evidence of cross-domain neuroplasticity. Furthermore, we also showed that, at the neural level, musicians and non-musicians respond differently irrespective of the presence of modulations in pitch cues. Thus, when subjects allocated attention to the stimuli, musical training facilitated regularities' processing and sequence learning in the auditory modality in general. This reveals that the greater sensitivity of musicians to process auditory regularities (François et al., 2014; Francois & Schön, 2011) is not due to enhanced perceptual/memory processes for pitch. In fact, our MMN study (see Chapter 3) also supports this idea since it demonstrated no significant ERP differences between groups in the processing of regularities based on pitch variations. Additionally, we found that participants showing enhanced N250-300 amplitude while being exposed to the sung speech stream had improved performance on the post-learning prosodic task. This finding is important for several reasons. First, it demonstrates that the online tracking of statistical regularities in sung speech as measured by ERPs predicts offline measures of SL, namely the correct recognition of learned words. Second, our data supports the idea that ERP responses are valid measures of the online processing of regularities, and furthermore, they are associated with classical behavioral, post-learning measures. Third, it supports the notion that the negativity found to diverge between groups in response to sung words might constitute an index of speech segmentation, very much like the N400-like component reported in the Batterink and Paller's study (2017), for example. Lastly, it also suggests that the two components of SL – identification and memory storage (Batterink & Paller, 2017) – interact with each other, with the long-term memory storage of integrated representations possibly depending on the perceptual processes engaged in streams' segmentation.

From our perspective, there are two complementary explanations for the enhanced ASL abilities observed in musicians. The most plausible explanation resides in musicians' intense contact with music that is characterized by high demands on acoustic analysis and auditory feedback integration (Jäncke, 2009; Klein, Liem, Hänggi, Elmer, & Jäncke, 2016; Münte, Altenmüller, & Jäncke, 2002; Schlaug, Norton, Overy, & Winner, 2005). Musical pieces are characterized by the combination of hundreds or thousands of acoustic elements that have to be decoded at various timescales, and that musicians need to memorize and evoke. The long-term experience with music provides a great sense of how musical time is structured and might develop efficient input grouping strategies (Chen, Penhune, & Zatorre, 2008). Some studies suggest that neural oscillators are entrained by the rhythmic aspects of musical pieces, which account for an enhanced synchronization of perceptual processes towards the musical beat (Large, 2000; Large & Palmer, 2002). For musicians, rapidly learning sequential regularities in the auditory input is an important component of musical skills. The learning of musical sequences might be generalized to the learning of other sequences governed by TPs within the auditory modality, such as speech.

On the other hand, we should also consider that the intense use of a musical instrument develops motor sequence learning skills (Bergstrom, Howard, & Howard, 2012; Chen et al., 2008; Haslinger et al., 2004; Hund-Georgiadis & Yves Von Cramon, 1999; Landau & D'Esposito, 2006) that might generalize to other forms of learning, namely auditory learning. Besides timing and spatial organization of movement, sequencing is an important motor control function that allows the correct orchestration of movements when playing individual notes (Zatorre, Chen, & Penhune, 2007). Motor sequence learning depends on predictive processes that allow the anticipation of events and the planning of sequences of actions (Gonzalez & Burke, 2018). Studies that investigated the impact of musical expertise on motor learning revealed greater implicit sequence learning in musicians (e.g., alternating serial reaction time task, see Bergstrom et al., 2012) and greater performance in rhythmic tapping tasks (Chen et al., 2008; Hund-Georgiadis & Yves Von Cramon, 1999). Pianists perform better than non-musicians in complex finger-tapping tasks, showing a differential activation of primary and secondary motor areas (i.e., increased activation in the contralateral primary motor cortex, and progressive decrease in the activation of secondary areas, such as bilateral supplementary, premotor, and cerebellar areas, when compared to non-musicians, see Hund-Georgiadis & Yves Von Cramon, 1999). Pianists also show enhanced sequence acquisition and faster reaction times in comparison with non-pianists in sequenced key presses that are accompanied by increased activations in a lateralized (right) network of motor, prefrontal and cingulate brain areas (Landau & D'Esposito, 2006). Musicians and musically-naïve individuals show differences in the patterns of activation of a network of regions (e.g., anterior cingulate cortex, right dorsal premotor cortex, both cerebellar hemispheres, and right basal ganglia) while performing over-practiced bimanual parallel movements (Haslinger et al., 2004). When the task is to tap in synchrony with auditory rhythms that progressively change in complexity, musicians and non-musicians activate secondary motor areas to the same extent but musicians evidence higher recruitment of the prefrontal cortex (Chen et al., 2008). These results suggest that non-musicians show lower efficiency in cortical and subcortical regions recruited during motor tasks. They also suggest that long-term motor training changes the functional activity of brain regions involved in sequence learning, leading to the recruitment of an extended network of areas. High-level motor skills are essential to allow musicians to focus on more artistic aspects of their performance and to promote excellence. Since, in music, the execution of sequences of motor commands is paired with sound, we wonder if greater ASL skills might emerge as an epiphenomenon of rich motor training. Motor abilities might transfer to other sensory modalities, namely the auditory modality, contributing to explain the enhanced capacity of musicians to learn auditory sequences. This is a hypothesis that deserves to be examined by future empirical studies.

6.1.2 The role of attention in mediating the effects of musical training on ASL

While the effect of attention on auditory perception is well established in general terms (Alho, Rinne, Herron, & Woods, 2014; Fritz, Elhilali, David, & Shamma, 2007a, 2007b; Rinne et al., 2005), there is a gap in the literature concerning its effects on ASL. Additionally, no previous research has examined the differential role of musical training in ASL as a function of attention. In the study reported in Chapter 4, the attentional mechanisms used during the attentive listening of the streams might have differed between groups. Musicians and non-musicians may use attentional networks in a distinctive way, employing distinct attentional strategies and using distinct neuronal resources while processing acoustic regularities. Enhanced selective attention abilities were documented in musicians (Besson et al., 2011; Strait, Slater, O'Connell, & Kraus, 2015) and some studies report the impact of selective attention on SL (Jiang & Chun, 2001; Shanks, Rowland, & Ranger, 2005; Turk-Browne, Jungé, & Scholl, 2005). Thus, the distinct use of attention during the online processing of regularities may affect the electrophysiological and behavioral correlates of ASL beyond direct effects related to musical training. This is a relevant aspect that undermines one's ability to infer the impact of musical training on the SL of musical and linguistic stimuli and, more specifically, to further generalize the conclusions from Chapter 4.

Research examining ASL in musicians under distinct attentional conditions is almost inexistent. To the best of our knowledge, there is only one study on this topic, led by Paraskevopoulos and colleagues (Paraskevopoulos, Kuchenbuch, Herholz, & Pantev, 2012). This study demonstrated (1) P50 amplitude modulations in response to the statistical structure of varying pitch sequences, (2) evidence of increased P50 as a function of musical expertise, (3) no modulation of the MMN to tonal violations by expertise, and (4) no learning of tone-sequences under pre-attentive conditions. This study also indicated that the segmentation of simple tone sequences occurs at a pre-attentive level: a representation of regular (STD) sequences allows the brain to make predictions about incoming input and to respond differently to sound transitions with varying probabilities. Moreover, it also showed that neither musicians nor non-musicians explicitly learned the tone-sequence sets, which suggested that the implicit knowledge of the statistical structure of the tonal stream is only explicitly available if attention is allocated to the task. Nevertheless, as the authors addressed pre-attentive SL using an oddball paradigm with tone sequences, there was still no evidence of how SL proceeds pre-attentively for other types of auditory input (e.g., sequences of words/melodies). For that reason, in Chapter 5 we overcame this gap in the literature by studying how musicians and non-musicians process acoustic regularities in linguistic and musical input when attention is not directed to the processing of these stimuli. Additionally, this work was motivated by specific concerns, namely: 1) the need of understanding if the ERP effects reported in previous SL studies with musicians are markers of SL per se or if they constitute indirect effects of attention, and 2) the need to investigate if the behavioral advantage brought by musical training to the learning of speech structures (François et al., 2013, 2014) would be maintained when speech is ignored.

The combined results of the fourth and fifth chapters confirmed the role of attention on ASL and its interaction with musical training, revealed by differences between musicians and non-musicians in the neural and behavioral mechanisms underlying streams' segmentation. Previous research manipulating attentional load demonstrated that SL performance decreases when individuals are attending to multiple targets or when there is an overload of attentional resources (Toro, Sinnett, & Soto-Faraco, 2005; Turk-Browne et al., 2005). Yet, these experimental conditions are just two examples of challenging environments in which SL may occur. Another challenging environment for the computation of statistical regularities is the situation in which the auditory input is not being attended to, despite being captured and processed by the brain, which is the situation Chapter 5 mimic. The behavioral results of that Chapter revealed that participants were able to learn the statistical structure of the sung speech stream (i.e., prosodic experiment). This finding goes against that of Paraskevopoulos and colleagues (2012), who did not found evidence of tonal-sequence learning using an identical experimental set-up. The most plausible reason to account for this discrepancy deals with stimulus nature: individuals might have a greater ease in the processing of speech materials comparatively with tonal sequences, despite both sharing a melodic structure. Speech sounds are among the most relevant auditory stimuli to which humans are exposed to given the important cues they carry for communicative purposes (e.g., the affect, the intents, the personality of an interlocutor, see Green, Kuhl, Meltzoff, & Stevens, 1991; Schuller, Batliner, Steidl, & Seppi, 2011). These aspects may render speech signals more salient in comparison to other sounds and might explain why individuals, even when not actively attending to words' sequences, can still learn from them.

This finding is also surprising in light of the results of the prosodic experiment reported in Chapter 4. In that experiment, only musicians performed significantly above chance. In the next section (see section 6.2.4), we point a methodological limitation of Chapter 5 that might explain these discrepant results. However, from our perspective, there is a better explanation for the lower performance of participants in Chapter 4 as opposed to Chapter 5. When explicitly trying to parse the prosodic stream into words, individuals might develop several possible representations of the words-to-be that remain in a labile state until individuals are asked to perform 'yes/no' decisions. When the time for the LDT comes, participants are faced with words and part-words that, for them, are both considered word candidates given that both occurred in the learning phase. This aspect, along with possible biases stemming from decisional judgments, might lead to more false alarms and decreased performance rates in attentive conditions.

The behavioral results of Chapter 5 also revealed the null effect of training on speech segmentation when individuals are not actively attending to the linguistic sounds. When attention is focused on a cross-modal, low demanding task, musicians do not show evidence of having a greater ability to segment speech. These results suggest that attention is a necessary condition for musicians to show higher learning of sung speech or, from another point of view, that the enhanced selective attention skills of musicians play a role in their capacity to accurately recognize learned words.

Chapter 5 was important to clarify how musicians and non-musicians learn acoustic regularities implicitly, without intent or explicit awareness of what they learn. At the neural level, individuals showed distinct patterns of electrophysiological activity according to the type of stimulus being processed (e.g., words/melodies). This indicates that the way the brain processes regularities in acoustic streams is very different depending on the active use of attentional resources and on the acoustics of the input. When tracking regularities pre-attentively, the groups differed at specific latency windows considering the onset of the streams' items: in the case of speech streams, they tended to differ between 450-500 ms after prosodic words' onset, whereas the differences tended to emerge earlier for flat contour words (N150-200 ms); in the case of the musical stream, groups differed at the very beginning of melodies onset (e.g., P0-50 ms) and at the end (e.g., P550-600 ms). If analyzed carefully, the EEG results of Chapters 4 and 5 reveal that (1) differences between groups emerge very early when participants are processing musical notes, which might be a consequence of musicians' long-term experience with this type of sounds, and that (2) the presence/absence of attentional resources allocated to the processing of speech sounds (independently of having or not melodic characteristics) strongly modifies the way the brain processes TPs between syllables-pairs.

The results of Chapter 5 are also interesting from another point of view: they show that two samples of participants can have similar success rates in a behavioral SL task having different online learning trajectories (i.e., EEG responses during learning phases). This finding points to the advantage of choosing a neural approach to the study of SL by showing that post-learning outcomes do not fully characterize the processing of acoustic regularities. Experimental paradigms that congregate distinct operational measures of ASL and that additionally shed light on neural changes accompanying learning trajectories provide invaluable insights for a comprehensive understanding of SL (Siegelman, Bogaerts, Kronenfeld, & Frost, 2017).

Furthermore, the results of Chapters 4 and 5 call our attention to the emergence of neural differences between groups right at the beginning of the learning phases. For example, the higher negativity found in musicians in response to the attentive processing of prosodic words was only observed for the first 200 trials of the phase, vanishing as exposure to the words increased. Similarly, when attentively processing the melodic stream, the groups differed significantly during the first 100 trials; when these sounds were presented as background stimulation, musicians and non-musicians showed differences in a P600-like component only during the first and second blocks of the learning phase. These results indicate that musical training changes musicians' brains by providing them with an enhanced sensitivity to sensory aspects of musical sounds that is observed at early stages of sound processing (e.g., 0-200 ms after sounds' onset). Secondly, they show that specific changes in the electrophysiological activity as a function of training occur at the very beginning of the exposure to the auditory streams. This finding might indicate that musicians engage distinct neural strategies to assimilate the structure of auditory input as soon as they hear it. It also suggests that, as the exposure to this input increases, the differences between groups tend to disappear. This aspect might indicate that non-musicians catch up with musicians at a certain point of the learning phase and, probably, at that point, both groups entail the same type of strategies to process the acoustic regularities. Previous research has also found that musical training modifies the earliest stages of SL (Vasuki et al., 2017) and another report with musically naïve individuals showed that expert learners (i.e. the participants showing the highest accuracy in behavioral performance) typically demonstrate increased amplitude of the N1 component in the first part of the exposure phase (Abla, Katahira, & Okanoya, 2008).

Overall, these findings support the importance of attention for many cognitive processes and further boost the investigation of musical training in relation to attention. Given our findings and previous research revealing altered selective attention abilities in musicians (and also its relation with working memory which is allegedly improved in musical experts; see George & Coch, 2011; Talamini et al., 2017; for a review, see Besson et al., 2011) we suggest that attention might be one of the mechanisms accounting for the many cognitive advantages related to musical training.

6.1.3 ERP differences as a function of stimulus type

Most studies investigating the effect of stimulus properties on SL have focused on modality-based differences (Conway & Christiansen, 2005; Milne, Petkov, & Wilson, 2017; Raviv & Arnon, 2018; Silva, Folia, Inácio, Castro, & Petersson, 2018; Vasuki et al., 2017). Only a few studies have gone further to explore the role of sensory modality in SL (Raviv & Arnon, 2018; Shufaniya & Arnon, 2018). Yet, none explored ASL outcomes in musicians and non-musicians depending on stimulus type within a sensory modality. Chapters 4 and 5 were dedicated to this question, specifically testing musicians and non-musicians' processing of statistical regularities with distinct types of auditory sequences: words with prosodic content, words with flat contour (e.g., non-prosodic) and piano melodies. The findings from these studies corroborated previous evidence while offering a systematic investigation of the neural processes underlying the SL of both music and artificial speech sounds. The use of the same experimental paradigm to study the EEG and behavioral responses for all the distinct types of stimuli represents an advantage of the present studies in face of previous research. Yet, before elaborating more on the results, it is important to consider additional discrepancies between the sounds used in our studies. The pitch of the linguistic stimuli varied across experiments because prosodic and non-prosodic words differed precisely in its intonation properties that result from pitch manipulations. Linguistic and musical sounds were also different: whereas non-prosodic words had flat contour, prosodic words and melodies were characterized by varying pitches and melodic patterns. Apart from that, all sounds had the exact same duration and intensity, and rise and decay times. Thus, excluding the pitch, the acoustic properties and presentation rates of the stimuli were identical. This aspect ensures that there were no additional physical variables accounting for the differences found in the neural and behavioral correlates of ASL and, therefore, rules out potential challenges to our interpretations.

The electrophysiological results showed that words elicited negativities between 250 and 300 ms post-onset that differed in amplitude between groups (see Chapter 4) when participants were actively trying to segment the streams. Despite the presence of only a tendency for a difference in the case of flat words, the grand-averaged waveforms revealed the same pattern of EEG responses for both types of linguistic sounds. This pattern is consistently different from the one that emerged in response to melodies: amplitude differences between groups occurred very early during the course of processing, specifically between 50-100 ms post-stimulus onset. These findings indicate that the computation of statistical regularities in musical and speech streams elicits distinct electrophysiological responses, suggesting that there are distinct neural mechanisms involved in the learning of the statistical structure of continuous auditory input depending on stimulus features. At the behavioral level, musicians outperformed non-musicians in all experiments and there were no significant differences between participants' performances as a function of stimulus type.

The fact that neural measures revealed differences in the online processing of words and melodies but behavioral performances did not also corroborates the idea that SL has dissociable components (Batterink & Paller, 2017). Additionally, considering that (1) the ERP results demonstrated that ASL is characterized by stimulus-specificity (which is consistent with claims from previous research, see Siegelman & Frost, 2015), that (2) post-learning behavioral responses did not vary as a function of stimulus type, and that (3) some studies have shown SL' outcomes to be highly task-specific, we wonder if the administration of an alternative SL task (e.g., target detection task, rating task) would have revealed distinct learning outcomes.

The modulatory effect of stimulus type on ASL was observed both while subjects were attending or not attending to the streams. The results of Chapter 5 revealed distinct EEG patterns in musicians and non-musicians in response to prosodic words, non-prosodic words and melodies. Unfortunately, due to methodological constraints, it was impossible to collect post-learning measures of all experiments, which prevented us from examining if the participants' long-term representations of learned items exhibited stimulus specificity in pre-attentive conditions.

A study on congenital amusia has proposed SL to rely on different processing systems based on evidence that amusic subjects could learn a language based on SL but they could not do the same for musical sequences (Peretz, Saffran, Schön, & Gosselin, 2012). This is revealing of distinct SL processing systems depending on the input nature (e.g., syllables or tones). The evidence also suggests that SL proceeds differently within the same sensory modality, in this case, the auditory one. For example, while there is a positive effect of age on SL performance in visual and non-linguistic auditory (environmental sounds) tasks (i.e., higher performance as children get older, see Shufaniya & Arnon, 2018), there is age invariance in the learning of linguistic structures (e.g., syllables, see Raviv & Arnon, 2018). This is clear evidence of stimulus-specificity regarding the effect of age on SL abilities.

6.1.4 Implicit measures of the detection of online structural violations

ERPs have brought significant contributions to our understanding of how speech and music are encoded in the brain. The EEG represents a powerful tool to deepen our knowledge on complex aspects of cognition that would otherwise be opaque to behavioral inquisition. Besides, it allows tracking changes in brain dynamics supporting cognitive operations and to distinguish between specific mechanisms engaged in complex cognitive processes. Adding to these facts, the possibility of inquiring cognitive skills without having individuals actively performing a task makes it ideal to study SL phenomena. One of the more innovative aspects of the present Dissertation was the acquisition of implicit measures related to the detection of online violations to the auditory streams' structure before explicitly assessing learning outcomes. The items of the streams were intermixed with part- and non-items that violated the statistical regularities of the streams and, therefore, allowed the examination of brain responses involved in the online tracking of mismatches between expected and actual input.

Chapter 4 revealed that groups distinctively processed prosodic part-words, with musicians showing enhanced positivities in two time windows (250-300 ms and 400-450 ms). However, there were no differences between groups in the EEG responses to violations present in linguistic (i.e., flat words) or musical streams. These results are worth of several considerations. First, they demonstrate that training only seems to modify the neural responses to violations of specific auditory stimuli. Sung speech facilitates phonological discrimination, and therefore, speech segmentation, by promoting the parallel mapping of linguistic and melodic cues (Francois & Schön, 2010; Schön et al., 2008). If segmentation is facilitated in this context, then the cognitive representations of the stream's items are well defined and violations of the statistical structure of the input might be more easily detected. Secondly, the differences between groups were found only for part-words, which compared to non-words are legal items and, for that reason, more difficult to disentangle from actual words. The distinct neural responses of musicians to part-words corroborate the idea that training modifies the cognitive strategies used to process acoustic regularities, which in consequence, might strengthen their ability to detect violations in sung speech. This would explain, on the one hand, the similar EEG response of musicians and non-musicians to non-words that could be more easily identified as foils, and, on the other hand, the amplitude differences found for part-words whose identification was cognitively demanding.

The role of musical training in the implicit detection of online mismatches was also examined under pre-attentive conditions (Chapter 5). Under these conditions, there were no effects of musical training on the processing of violations contained in sung speech. The fact that attentional resources were allocated to a cross-modal task might explain the divergent findings reported in the prosodic experiments of Chapters 4 and 5. Additionally, these divergent results called our attention to the similar divergent findings in the explicit assessment of learning: after actively attending to the stream, musicians outperformed non-musicians in the recognition of the prosodic words; when allocating attention to a visual task while listening to the stream, musicians ended up being as good as non-musicians. The evidence points to a decisive role of attention in modulating the impact of training upon the computation of TPs in speech. It also suggests that the ERP responses characterizing learning time might be associated with the learning outcomes. The negative association between the magnitude of the N250-300 and the performance on the prosodic test (see Section 3.1.3 of Chapter 4) supported this idea. Future experiments might also look for correlations between implicit and explicit learning assessments.

Additionally, the finding of neural differences between groups in response to part-melodies (i.e., enhanced P600-like in musicians) was surprising given that no effects of training on the implicit test phase were expected when the streams were presented as background auditory stimulation. Given that no training effects were found for the other types of sounds, this finding suggests that melodies may be more distinctive sounds for musicians. Possibly due to their intense contact with music, musicians might process instrumental notes in a special way compared to musically naïve people, even when these sounds are not in the focus of their attention.

6.1.5 Musical training and simple vs. complex ASL

The experimental studies reported here collectively examined the role of musical training in the SL of simple and complex auditory contingencies. The results of Chapter 3 revealed that individuals are able to detect changes in simple pitch sequences independently of having long-term musical experience, as demonstrated by the absence of amplitude differences in the MMN as a function of training. Other studies investigating MMN responses to pitch deviations in musicians corroborate our findings (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2004; Nikjeh et al., 2008; Nikjeh, Lister, & Frisch, 2009; Paraskevopoulos et al., 2012; Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005). These studies raised the possibility that the absence of training effects on the MMN was not due to musicians and non-musicians recruiting similar sequence learning skills; on the contrary, it was suggested that individuals were being tested with pitch sequences that were so simply structured that average pitch discrimination skills were enough to process them. In this sense, musicians' enhanced SL skills would only be demonstrated when the processing of more complex sequential auditory input was examined. Chapters 4 and 5 addressed this question by examining the impact of musical training on standard ASL tasks composed of trisyllabic words or tri-tone melodies. The only cue to learn the statistical structure of these tasks' streams was the computation of TPs between syllable/tone pairs. In fact, the study reported in Chapter 4 revealed that musical training benefits the processing of statistical regularities, independently of stimulus type. Moreover, it demonstrated that the effects of training on ASL are manifested during the online tracking of regularities and, in the case of sung speech, also during the implicit testing for the detection of statistical violations. Conversely, when these same paradigms were tested under pre-attentive conditions, there was no behavioral evidence of an advantage of musicians in the segmentation of speech, despite the distinct electrophysiological responses of musicians and non-musicians to the different auditory streams.

Comparing the findings of the three experimental reports described in the present Dissertation it is possible to conclude that musical training plays a role in ASL under specific task conditions. When the task of parsing out the acoustic input is demanding, namely when it depends on the computation of TPs and when it occurs pre-attentively, musicians show distinct online tracking of acoustic regularities; when it is relatively easy to figure out a pattern in the acoustic stimulation, musicians demonstrate similar responses to non-musicians at the neural level. However, if the input is more complex, musical training brings advantages to ASL but only when individuals are actively engaged in segmenting the auditory streams, not when these streams are task-irrelevant. These findings suggest that attention plays an active role in mediating the effects of training on ASL. Additionally, they suggest that training distinctively impacts the learning of auditory structures depending on its complexity.

6.2 Limitations

Despite the relevant contributions of this Dissertation, specific shortcomings are worth of attention. Below, we highlight several aspects we consider that would have benefited from a different theoretical/methodological approach.

6.2.1 The silent intervals in between streams' units

The acoustic streams used in the experiments depicted in this Dissertation were characterized by silent 50 ms gaps in between streams' units: pauses were inserted in between syllables/tones and simultaneously in between words/melodies. The motivation for the introduction of these pauses is explained before (see section 6.1.1) and essentially relates to ERP analysis requirements. Nevertheless, this feature of our experiments has no parallel in previous studies. From our perspective, the absence of ISIs in previous SL studies using the ERP technique is a shortcoming that interferes with researchers' pre-processing analyses, and decisively influences the EEG signal characterizing epochs of interest (e.g., the epoch starting from words' onset towards its end, for example). The continuous presentation of stimuli, without baseline periods, restricts the possibility of studying these epochs without the interference of pre-stimulus activity that will confound the analysis of ERP components. Thus, the decision to include ISIs in the current Dissertation aimed to surpass this gap, ensuring the best quality of the EEG data. However, it also limited our ability to firmly compare the results of the current studies with previous ones. In that sense, we may consider this methodological option simultaneously a gain and a limitation of the current Dissertation.

6.2.2 The duration of the learning phase

The behavioral results of both Chapters 4 and 5 revealed low word/melody recognition rates among all participants. Despite the differences found between groups in some conditions, participants' performance was barely above chance, which indicates that participants struggled to learn the streams' structure. Possibly, the duration of the learning phase of our experiments was too short in order for subjects to be able to extract the patterns (i.e., words/melodies instead of syllables/tones) embedded in the streams. In the case of the current experiments, the length of the learning phases, which is related to stimulus repetitions, was also determined by previous research on speech segmentation, namely the studies that inspired our work (François et al., 2013; François & Schön, 2014). Some studies ensured that participants learned the acoustic streams by increasing learning time, often by presenting streams repetitions in different blocks separated by silent intervals of a few sec or min (see Abla et al., 2008; Vasuki et al., 2017, for example). Nevertheless, other studies suggest that SL can occur after brief learning times (Aslin, Saffran, & Newport, 1998; Kirkham, Slemmer, & Johnson, 2002; Saffran, Aslin, & Newport, 1996; Singh, Daltrozzo, & Conway, 2017). Therefore, in face of discrepancies regarding the effect of the length of exposure time on SL outcomes, increasing the duration of the learning phases might have represented a better methodological choice to ensure that learning would have indeed occurred. However, a too long learning phase could also have overshadowed any possible positive effects of musical training over ASL. Future studies should address this issue by comparing the outcomes of learning phases with different durations.

6.2.3 The implications of the implicit test stream on the test phase

For theoretical reasons that were addressed before (see section 6.1.4), all the experiments reported in Chapters 4 and 5 included an implicit test phase. In our perspective, the introduction of this intermediate phase added important new information to the understanding of the brain processes engaged in SL, namely related to structural violations' detection. However, the potential biases introduced by this phase on the learning outcomes deserve some attention. First, since we were not sure that participants have learned the items during the learning phase, it is possible that the presentation of new stimuli (e.g., part-words/non-words) might have had a confounding effect on participants' knowledge of the stream presented during the learning phases. This confounding effect might have been driven by the similar duration of the learning and implicit test phases. Therefore, even if the items were distinct between phases, the insertion of this new stream might have led to more unsecure judgments of participants during the subsequent phase (e.g., behavioral test). At last, we should consider also a potential familiarity effect on the behavioral performance: the part-words and non-words used in the implicit test and the behavioral tests were the same. The repetition of the items could have biased participants' decision towards the acceptance of these items as belonging to the original stream. The aspects listed here might have accounted for the participants' low accuracy in the recognition of words/melodies across experiments. To overcome these specific limitations, the creation of new part-words/non-words to use in the behavioral test phase and the lengthening of the learning phase comparatively to the implicit test phase would be desirable.

6.2.4 The order of the prosodic experiment (Chapter 5)

According to recent research suggesting that attention is a necessary condition for the SL of sound sequences to occur (Ding et al., 2017; Paraskevopoulos et al., 2012), it would be expected that participants listening to the prosodic stream as background stimulation (Chapter 5) had more difficulty in accurately differentiating words from part-words and non-words during the behavioral test. However, musicians and non-musicians significantly outperformed chance in the behavioral test. This is a surprising finding for two reasons: first, it does not conform to the findings of previous research; and, second, it differs from the results obtained in the study of Chapter 4 (in which participants were actively trying to parse out the auditory materials) that revealed that one of the groups (non-musicians) performed at chance level. Thus, why do participants recognize better prosodic words when they are not attending to them comparatively to when they are actively focusing their attention on them? A possible explanation for this result resides in the fact that the prosodic experiment was always the last experiment to which participants were exposed to in the pre-attentive study (Chapter 5), while in the study of Chapter 4 the order of experiments was counterbalanced across participants. Since musical and prosodic streams have the same melodic structure, participants might have benefited from the previous exposition to the musical stream. A recent study found that explicit pre-training before passive exposure to a language stream benefited participants' explicit memory representations of the words (Batterink, Reber, & Paller, 2015). The strengthening of the memory representations led to high overall accuracy when participants were faced with the discrimination of words against foils. Although, here, participants were not explicitly trained on the statistical structure of the stream or presented with the streams' items before the learning phase, the presentation of the musical stream could have worked as an implicit pre-training for the prosodic experiment. To the best of our knowledge, there is no study explicitly addressing the impact of previous implicit knowledge on the SL of language streams. Thus, assuming that participants generalized the implicit knowledge of the musical stream to the prosodic one, the results of the LDT may echo the benefits of a long exposure to the streams' melodic structure.

6.2.5 The absence of behavioral test phases in the musical and linguistic experiments (Chapter 5)

In Chapter 5 we investigated ASL under pre-attentive conditions. For that purpose, participants were instructed to focus attention on a visual task while ignoring the sounds presented via earphones. If, during the tasks, participants had been asked about the auditory streams, they would logically ignore the researcher's instruction and allocate attentional resources to the auditory stimulation. For this reason, we chose to present only a behavioral test phase at the end of the experiments. Also due to this we chose to test all participants' SL abilities regarding one specific stimulus: prosodic words. If, on the one hand, this experimental set up allowed us to collect behavioral measures of the learning of prosodic structures, on the other hand it made it impossible to test what participants had learnt from musical and linguistic (i.e., flat words) streams. This methodological option does not overshadow the potential and informative value of EEG data to elucidate on the regularity processing mechanisms recruited during the learning of the musical and linguistic streams. However, it limits our capacity to discuss the effects of training on pre-attentive ASL, as a whole. Ideally, to answer our research question in full, three distinct samples of participants should be recruited and each one would participate in one of the three auditory experiments. Future research aiming to replicate Chapter 5 should take these aspects into consideration.

6.3 Future Studies

Future research should address the mechanisms underlying the processing of statistical regularities using more ecological paradigms that could closely mimic real environmental settings. For that purpose, researchers should use stimuli with more ecological validity, such as real recorded excerpts of unknown languages (i.e., languages that participants do not know). These examples may convey human speech prosody in a more authentic way and enhance our comprehension of how intonation cues modulate the neural responses to speech. Additionally, it would be interesting to explore the impact of stimulus duration on the SL outcomes. Stimulus presentation rate has an impact on implicit SL: whereas fast rates typically benefit auditory learning, slow rates have a detrimental effect on it (Emberson, Conway, & Christiansen, 2011). That is not surprising since auditory processing is highly sensitive to the temporal aspects of information especially because sounds are temporally variable signals (Chen, Repp, & Patel, 2002). Future studies could extend the present work by presenting participants with similar SL paradigms under distinct timing conditions. Maintaining the statistical information the same, researchers could investigate the differential role of the presentation rates on learning different auditory stimuli and examine if and how musical training interact with those variables.

It would also be relevant to address these issues while investigating the rules or memory-based judgments that listeners used during SL tasks. The use of distinct cognitive strategies could likely impact the neural responses to the streams during the learning time. Thus, researchers could investigate these cognitive strategies by means of self-reports or brief questionnaires and relate these data to the subjects' ERP profiles and post-learning performance.

It is hard to recruit Musical-college students or professional musicians matching our criteria (e.g., minimum number of training years, number of languages spoken, etc.) who agree to dedicate several hours of their time to participate in experimental research. We trust in the reliability of our findings since the between-conditions/groups differences were large and significant. Despite the relatively homogeneous samples characterizing the different studies, increasing the sample sizes would increase the confidence in the results and expand the possibilities for statistical analyses in future studies. Also, researchers should creatively undertake efforts to recruit bigger samples of musicians.

Another aspect worth of exploration in future research is the mediating role of task instructions in SL outcomes. In Chapter 4, participants were told to actively try to identify the 'alien' words and melodies embedded in the acoustic streams. Previous studies have compared individuals' performance on visual SL tasks under intentional (i.e., active effort to segment the input) or incidental (i.e., without instructions regarding the input content or structure) instructions and distinct presentation rates (e.g., slow vs. fast) (Arciuli, von Koss Torkildsen, Stevens, & Simpson, 2014; Bertels, Destrebecqz, & Franco, 2015). They revealed that intentional instructions improved participants' explicit knowledge of the input sequences but only when visual sequences were presented at slow presentation rates. Word learning studies have also tested the effect of intentionality suggesting that word-referent mapping is increased under intentional conditions, despite also occurring when there is no statistical information on the items (Hamrick & Rebuschat, 2012; Kachergis, Yu, & Shiffrin, 2014). This evidence indicates that, irrespective of stimulus modality, explicit efforts lead to a better performance, which might be a consequence of the use of strategic processes during learning. Nevertheless, the existing studies show confusion in the terminology used to illustrate different task instructions: whereas some authors use the term "incidental" to refer to conditions in which participants are attending to the streams but have no further information about it, others use the same term to refer to conditions in which participants, aside from not being told explicit instructions, are focused on a concurrent task. In the last scenario, there is an additional manipulation of attentional resources and possibly cognitive load such, which makes interpretations based purely on the presence/absence of explicit instructions unfeasible. We consider that future studies addressing the role of musical training in ASL should take these aspects into account. By replicating the study reported in Chapter 4 while adding a manipulation in which participants learn incidentally (e.g., focusing attention on the streams without further knowledge of their structure), it would be possible to examine if explicit learning strategies play a mediating role in the effect of musical training on auditory learning and how it changes learning (e.g., electrophysiological and behavioral outcomes).

The analysis of evoked responses to different types of stimuli is the most classical way to analyze EEG data. Nevertheless, the classical ERP approach raises some methodological problems when it is applied to paradigms in which there is a rapid and continuous presentation of stimuli (such as those included in our studies). First, there are baseline issues arising from absent or too short silent pre-stimulus periods; second, these periods can be contaminated by previous stimulus processing (Cohen, 2014; Luck, 2005). In addition to ERPs, frequency analysis through Fourier transformations can inform on oscillatory effects and how they are modulated over time. Previous research has shown that oscillatory phase coherence and the synchronization of EEG at different frequencies of the spectrum, and over time, can enlighten the brain mechanisms underlying auditory processing, particularly speech processing (Luo & Poeppel, 2007). Neural entrainment to target frequencies has been related to the decomposition of streams into meaningful units (Riecke, Sack, & Schroeder, 2015; Teng, Tian, Doelling, & Poeppel, 2018). In this domain, cortical entrainment – that is, cortical activity that is entrained to the temporal envelope of specific stimuli (e.g., speech) - to the theta frequency band has been related to the perceptual grouping of auditory stimuli (Ghitza, 2012; Ghitza & Greenberg, 2009; Riecke et al., 2015). There is evidence that oscillatory synchronization enhances speech processing since speech intelligibility is predicted by neural phase locking to the speech envelope (Ahissar et al., 2001; Doelling, Arnal, Ghitza, & Poeppel, 2014). The existing studies indicate that theta-band oscillations (i.e., increased phase coherence across trials) promote speech segmentation by synchronizing with syllable onsets (Doelling et al., 2014; Luo & Poeppel, 2007; Peelle et al., 2013). The EEG time-frequency analysis can, therefore, complement the typical ERP analysis by providing additional information on the mechanisms underlying the detection of statistical regularities both in linguistic and musical input. ITPC - phase-locked consistency in neural activity to an event’s onset, across trials, at individual electrodes (Cohen, 2014; Roach & Mathalon, 2008) - can be used as a measure of neural coding because it indexes the coherence of phase values at specific EEG time-frequency points or bands related to the temporal processing of stimuli (Ding & Simon, 2014; Doelling et al., 2014; Kayser, Wilson, Safaai, Sakata, & Panzeri, 2015; Luo, Tian, Song, Zhou, & Poeppel, 2013). Research suggests that ITPC is a measure that reliably tracks speech comprehension (Ahissar et al., 2001; Ding et al., 2017; Kerlin, Shahin, & Miller, 2010; Luo & Poeppel, 2007, 2012) and simultaneously overcomes some limitations of power spectrum estimations, such as the vulnerability to low frequency fluctuations (Forget, Buiatti, & Dehaene, 2010). Based on this evidence, some interesting questions arise: is cortical entrainment to the theta band a neural marker of the SL of auditory sequences in general and not only of speech?; does musical training modify the cortical entrainment to specific frequency bands (e.g., theta) during ASL?; is there an association between neural oscillatory activity during auditory learning, independently of stimulus type, and behavioral performance on post-learning tasks? Future studies need to address these questions in order to identify crucial aspects of auditory learning that may relate to entrainment of neural oscillations to temporal features of these types of auditory input, across frequency bands and over time.

Ultimately, there is an important issue - common to the vast majority of studies produced in the field of musical training research - that the present Dissertation also could not address: the question of causation (Fujioka, Ross, Kakigi, Pantev, & Trainor, 2006; Wong et al., 2007). The correlational nature of our studies does not allow determining whether the neural and behavioral results in musicians are a consequence of long-term musical practice or the consequence of a specific predisposition for music. Despite the relevance of the current findings, missing a causal link between training and cognitive outcomes leaves aside the possibility of safely claiming that musical training produces positive alterations in ASL mechanisms and skills. It might be that the musicians recruited here are simply a group of people with enhanced ASL skills who were attracted to a musical career instead of being individuals whose ASL skills benefited from training. Recently, a study with monozygotic twins revealed increased cortical thickness in the left auditory-motor network only in twins who were actively engaged in musical training, which suggests that part of the anatomical differences between brothers was not explained by genetics (de Manzano & Ullén, 2018). In fact, most brain alterations that have been associated with musical training, namely structural and functional responses, have been corroborated by longitudinal research conducted with children integrating training programs, which provides convincing evidence in favor of developmental plasticity (Fujioka et al., 2006; Habibi, Cahn, Damasio, & Damasio, 2016; Hyde et al., 2009; Tierney & Kraus, 2013; Tierney, Krizman, & Kraus, 2015).

Therefore, this field would strongly benefit from studies with longitudinal designs that randomly assign participants to different groups in which musical training is provided or not. It would be of significant interest to clarify whether facilitated ASL competences and distinct ASL correlates (e.g., electrophysiological activity) are innate or the consequence of training. Additionally, researchers should also be aware of the need to contrast different forms of auditory training in order to rule out the specific effects of instrumental music practice (i.e., instrumental training) in comparison with other types of auditory stimulation (e.g., singing). Of course, these manipulations would be even more valuable if the homogeneity of groups is guaranteed in terms of relevant variables, such as socio-economic background, gender and/or neuropsychological functioning. If future research is able to clearly link the practice of a musical instrument to improved ASL skills and to changes in the brain processes accompanying them, then training could be more strongly advised in the remediation of language disorders stemming from deficient speech segmentation, for example (see next section).

6.4 Implications

The results described in the current Dissertation have theoretical implications to our understanding of ASL skills in musical experts and non-experts. First, these data enrich our understanding of experience-driven cortical plasticity in general. As such, the outcomes of this Dissertation are an important tool for the examination of musical training as a model of neuroplasticity since they provide evidence regarding specific cognitive processes that seem to be modulated by long-term training. These findings highlight the benefit of conducting SL research with quantitative approaches, congregating both the behavioral performance on specific tasks and the electrophysiological signatures underlying the processes that lead to that performance. Moreover, they provide complementary – behavioral and neural - information on a broad phenomenon, such as SL.

The experimental work reported in the present Dissertation provided important information on processes involved in ASL in musically–trained and control samples of individuals. The data on the neural and behavioral processes reflecting the segmentation of continuous streams of sounds, both melodic and linguistic, allow us to better understand the extent of musically-induced changes on auditory cognition. Additionally, they provided relevant information on the processing of artificial speech input that can inform research on speech perception and comprehension. By examining the SL of linguistic sounds, we unraveled how the brain decomposes sensory information reaching the auditory system for further integration, along with informing on the translation of perceptual input into long-term memory representations. Understanding the brain mechanisms involved in ASL in the healthy adult brain provides a common ground for the investigation of normal and atypical language development. In line with what is advocated by recent neuroscientific studies on musical training and language processes (Dittinger et al., 2016; Dittinger, Chobert, Ziegler, & Besson, 2017; François et al., 2013; François & Schön, 2014), we consider that the present findings can have important educational and societal impacts.

From a therapeutic perspective, the information resulting from our studies can provide important insights to the understanding of the biomarkers of speech processing and related disorders. The combination of EEG and other techniques (e.g., fMRI, transcranial direct current stimulation) with neurofeedback protocols now offers a new avenue for the understanding and remediation of abnormal speech processing and auditory processing in general. With these neuroimaging techniques and the insights brought up by studies such as the ones reported here, clinicians can tackle speech disorders which much more detail and efficiency than before. The results of this Dissertation also offers new insights into the processes involved in second/new language learning and that might be disrupted in children presenting speech disorders, such as dyslexia (e.g., poor reading abilities due to impaired implicit learning skills, see Arciuli & Simpson, 2012; Gabay, Thiessen, & Holt, 2015) or other deficits. Rhythm processing abilities have been linked to reading-related cognitive skills (Anvari, Trainor, Woodside, & Levy, 2002; Ozernov-Palchik & Patel, 2018; Strait & Kraus, 2011; Thomson & Goswami, 2008) due to the sharing of temporal processing mechanisms. Also, Hung and collaborators (Hund-Georgiadis & Yves Von Cramon, 1999) identified a network of brain regions involved both in SL (e.g., SRT task) and word reading. Importantly, the authors found that more skilled readers show a higher discrepancy between the reaction times to ordered vs. random sequence conditions. Our results demonstrate that musicians are more able to use the statistical information contained in the auditory input to form predictions and to segment continuous streams of speech. This ability is highly relevant for the development of reading skills (Arciuli & Simpson, 2012). Thus, our findings can help to design new interventions for children with reading disabilities and advocate for musical training as a strategy to promote early reading skills. A greater understanding of the basic neural mechanisms that underlie speech perception will allow psychologists and clinicians in general to design new and individually tailored strategies to treat the abovementioned conditions, improving the efficiency and the timing of the therapeutics.

The evidence gathered here might also inform the design of new strategies and tools for developing word learning in children and/or adults with learning disorders (e.g., developmental dyslexia, specific language impairment) or acquired clinical disorders (e.g., aphasia). Recent research has revealed the positive impact of music-based interventions in patients with Broca's aphasia (see Van Der Meulen, Van De Sandt-Koenderman, Heijenbrok, Visch-Brink, & Ribbers, 2016, on the beneficial effects of melodic intonation therapy for the improvement of naming, repetition and connected speech) and in patients with Parkinson's Disease (see Di Benedetto et al., 2009; and Tanner, Rammage, & Liu, 2016, on the effects of singing-based interventions on expressive prosody and articulation in Parkinson's patients).

From an educational point of view, the data described in the current Dissertation (along with evidence from previous studies, see Dittinger et al., 2016; Dittinger et al., 2017) supports the use of musical training as a pedagogical and instrumental strategy to improve word learning in foreign languages. Besides, musical training promotes the contact with a form of art – music - that further develops subjects' executive functioning (Bialystok & DePape, 2009), working memory (Besson et al., 2011; D'Souza, Moradzadeh, & Wiseheart, 2018; Suárez, Elangovan, & Au, 2016) and emotional intelligence (Schellenberg, 2011).

6.5 Conclusion

Our knowledge of the information contained in millions of neurons, travelling through uncountable neural connections and reflected in electrical brain signals still has a long way to go. We find it critical to have a deeper knowledge of the neural changes underlying perceptual and cognitive processes in musicians, namely ASL mechanisms, in order to gain a broad understanding of musical training as an experience-dependent model of neuroplasticity. In our perspective, one of the ways to unravel the brain dynamics of ASL is to bring together the knowledge from cognitive psychology and neuroscience and to put it in service of the design of well-adapted cognitive paradigms complemented by the best neuroimaging techniques combined with suitable analysis approaches.

The present Dissertation aimed to provide that information. Two of the core studies that integrate this work (Chapters 4 and 5) adopted an innovative approach to the examination of the processing of auditory regularities: besides characterizing the behavioral responses to ASL tasks, these studies investigated the online learning of sequential sounds. While post-learning measures, such as those collected in the LDT tasks, inform on the representations of repetitive patterns that individuals memorized and how they were retrieved, online measures (i.e., EEG responses) of learning provide information on how those representations were encoded, updated and integrated. These measures are essential for characterizing the predictive processes that the auditory system utilizes to learn sequential patterns. Our approach to investigate ASL provided reliable signatures of SL processes and performance, clearly disambiguating between what characterizes the learning processes *per se* vs. the learning outcomes. With this approach, we showed how different operational measures of ASL contribute to a broader understanding of how regularities are perceived and retrieved in the auditory modality and depending on musical expertise, attention and stimulus type.

Our studies demonstrated that the extended practice of a musical instrument - that implicates long-term contact with auditory sequencing - facilitates the computation of sequential regularities among new auditory input, independently of its acoustic nature (e.g., linguistic vs. melodic) but only when individuals are actively engaged with this input. Overall, our findings suggest that the ASL skills of musical experts and naïve individuals, and the nature of the accompanying electrophysiological changes, are highly dependent on the allocation of attentional resources and on the complexity and the acoustics of the sounds. Because of that and of the absence of a clear causal link between neural changes and higher accuracy rates in behavioral ASL tasks, care must be taken in the generalization of the present results. Though our methodology does not allow causal deductions to be made, our findings strengthen the idea that the online EEG patterns characterizing the tracking of auditory regularities provide important information on the musicians and non-musicians' capacity to segment acoustic streams. Moreover, the results of this Dissertation clearly demonstrate that musicians' advantages in the segmentation of artificial speech are not a product of enhanced pitch sensitivity but instead the result of a fine-grained ability to decode statistical information from continuous auditory input when individuals are actively attending to the auditory environment. It further provides evidence of positive transfer effects between music and speech, shedding more light onto the extension of (possible) musically-induced cognitive changes.

We hope that the studies encompassing this Dissertation contributed to enrich our knowledge on the thrilling human capacity to decode patterns from the auditory environment, while shedding light on the transformations that musical training entails over our brain machinery. We wish the present work might instigate novel scientific endeavors both on auditory cognition and on musical training as a unique model of neuroplasticity.

6.6 References

Abla, D., Katahira, K., & Okanoya, K. (2008). On-line assessment of statistical learning by event-related potentials. *Journal of Cognitive Neuroscience*, *20*(6), 952–964. http://doi.org/10.1162/jocn.2008.20058

Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., & Merzenich, M. M. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proceedings of the National Academy of Sciences*, *98*(23), 13367–13372. http://doi.org/10.1073/pnas.201400998

Alho, K., Rinne, T., Herron, T. J., & Woods, D. L. (2014). Stimulus-dependent activations and attention-related modulations inthe auditory cortex: A meta-analysis of fMRI studies. *Hearing Research*, *307*, 29–41. http://doi.org/10.1016/j.heares.2013.08.001

Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology*, *83*(2), 111–130. http://doi.org/10.1016/S0022-0965(02)00124-8

Arciuli, J., & Simpson, I. C. (2012). Statistical Learning Is Related to Reading Ability in Children and Adults. *Cognitive Science*, *36*(2), 286–304. http://doi.org/10.1111/j.1551-6709.2011.01200.x

Arciuli, J., von Koss Torkildsen, J., Stevens, D. J., & Simpson, I. C. (2014). Statistical learning under incidental versus intentional conditions. *Frontiers in Psychology*, *5*(JUL), 1–8. http://doi.org/10.3389/fpsyg.2014.00747

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8 month old infants. *Psychological Science*, *9*(4), 321–324.

Batterink, L. J., & Paller, K. A. (2017). Online neural monitoring of statistical learning. *Cortex*, *90*, 31–45. http://doi.org/10.1016/j.cortex.2017.02.004

Batterink, L. J., Reber, P. J., & Paller, K. A. (2015). Functional differences between statistical learning with and without explicit training. *Learning and Memory*, *22*(11), 544–556. http://doi.org/10.1101/lm.037986.114

Baumann, S., Meyer, M., & Jäncke, L. (2008). Enhancement of auditory-evoked potentials in musicians reflects an influence of expertise but not selective attention. *Journal of Cognitive Neuroscience*, *20*(12), 2238–2249. http://doi.org/10.1162/jocn.2008.20157

Bergstrom, J. C., Howard, J. H., & Howard, D. V. (2012). Enhanced Implicit Sequence Learning in College-age Video Game Players and Musicians. *Applied Cognitive Psychology*, *26*(1), 91–96. http://doi.org/10.1002/acp.1800

Bermudez, P., Lerch, J. P., Evans, A. C., & Zatorre, R. J. (2009). Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. *Cerebral Cortex*, *19*(7), 1583–1596. http://doi.org/10.1093/cercor/bhn196

Bertels, J., Destrebecqz, A., & Franco, A. (2015). Interacting effects of instructions and presentation rate on visual statistical learning. *Frontiers in Psychology*, *6*(NOV). http://doi.org/10.3389/fpsyg.2015.01806

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: Common processing, attention, and memory. *Frontiers in Psychology*, *2*(MAY), 1–12. http://doi.org/10.3389/fpsyg.2011.00094

Bialystok, E., & DePape, A. M. (2009). Musical Expertise, Bilingualism, and Executive Functioning. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(2), 565–574. http://doi.org/10.1037/a0012735

Bidelman, G. M., & Alain, C. (2015). Musical Training Orchestrates Coordinated Neuroplasticity in Auditory Brainstem and Cortex to Counteract Age-Related Declines in Categorical Vowel Perception. *Journal of Neuroscience*, *35*(3), 1240–1249. http://doi.org/10.1523/JNEUROSCI.3292-14.2015

Bidelman, G. M., Gandour, J. T., & Krishnan, A. (2011). Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain and Cognition*, *77*(1), 1–10. http://doi.org/10.1016/j.bandc.2011.07.006

Bidelman, G. M., & Krishnan, A. (2010). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Research*, *1355*, 112–125. http://doi.org/10.1016/j.brainres.2010.07.100

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "semantic P600" effects in language comprehension. *Brain Research Reviews*, *59*(1), 55–73. http://doi.org/10.1016/j.brainresrev.2008.05.003

Brown, C. J., Jeon, E.-K., Driscoll, V., Mussoi, B., Deshpande, S. B., Gfeller, K., & Abbas, P. J. (2017). Effects of Long-Term Musical Training on Cortical Auditory Evoked Potentials. *Ear and Hearing*, *38*(2), e74–e84. http://doi.org/10.1097/AUD.0000000000000375

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Moving on time: Brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience*, *20*(2), 226–239. http://doi.org/10.1162/jocn.2008.20018

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Moving on time: Brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience*, *20*(2), 226–239. http://doi.org/10.1162/jocn.2008.20018

Chen, Y., Repp, B. H., & Patel, A. (2002). Spectral decomposition of variability in synchronization and continuation tapping: Comparisons between auditory and visual pacing and feedback conditions. *Human Movement Science*, *21*(4), 515–532. http://doi.org/10.1016/S0167-9457(02)00138-0

Chobert, J., François, C., Velay, J. L., & Besson, M. (2012). Twelve months of active musical training in 8-to 10-year-old children enhances the preattentive processing of syllabic duration and voice onset time. *Cerebral Cortex*, *24*(4), 956–967. http://doi.org/10.1093/cercor/bhs377

Cohen, X. M. (2014). *Analyzing Neural Time Series Data*. *MIT Press*. http://doi.org/10.1007/s13398-014-0173-7.2

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning*

Memory and Cognition, *31*(1), 24–39. http://doi.org/10.1037/0278-7393.31.1.24

Cunillera, T., Càmara, E., Toro, J. M., Marco-Pallares, J., Sebastián-Galles, N., Ortiz, H., ... Rodríguez-Fornells, A. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *NeuroImage*, *48*(3), 541–553. http://doi.org/10.1016/j.neuroimage.2009.06.069

Cunillera, T., Laine, M., & Rodríguez-Fornells, A. (2016). Headstart for speech segmentation: A neural signature for the anchor word effect. *Neuropsychologia*, *82*, 189–199. http://doi.org/10.1016/j.neuropsychologia.2016.01.011

Cunillera, T., Toro, J. M., Sebastián-Gallés, N., & Rodríguez-Fornells, A. (2006). The effects of stress and statistical cues on continuous speech segmentation: An event-related brain potential study. *Brain Research*, *1123*(1), 168–178. http://doi.org/10.1016/j.brainres.2006.09.046

D'Souza, A. A., Moradzadeh, L., & Wiseheart, M. (2018). Musical training, bilingualism, and executive function: working memory and inhibitory control. *Cognitive Research: Principles and Implications*, *3*(1), 11. http://doi.org/10.1186/s41235-018-0095-6

de Diego-Balaguer, R., Toro, J. M., Rodriguez-Fornells, A., & Bachoud-Lévi, A. C. (2007). Different neurophysiological mechanisms underlying word and rule extraction from speech. *PLoS ONE*, *2*(11). http://doi.org/10.1371/journal.pone.0001175

de Manzano, Ö., & Ullén, F. (2018). Genetic and environmental influences on the phenotypic associations between intelligence, personality, and creative achievement in the arts and sciences. *Intelligence*, *69*(February), 123–133. http://doi.org/10.1016/j.intell.2018.05.004

Di Benedetto, P., Cavazzon, M., Mondolo, F., Rugiu, G., Peratoner, A., & Biasutti, E. (2009). Voice and choral singing treatment: A new approach for speech and voice disorders in Parkinson's disease. *European Journal of Physical and Rehabilitation Medicine*, *45*(1), 13–19. http://doi.org/R49081917 [pii]

Ding, N., Pan, X., Luo, C., Su, N., Zhang, W., & Zhang, J. (2017). Attention is required for knowledge-based sequential grouping: Insights from the integration of syllables into words. *The Journal of Neuroscience*, 2606–17. http://doi.org/10.1523/JNEUROSCI.2606-17.2017

Ding, N., & Simon, J. Z. (2014). Cortical entrainment to continuous speech: functional roles and interpretations. *Frontiers in Human Neuroscience*, *8*(May), 1–7.

http://doi.org/10.3389/fnhum.2014.00311

Dittinger, E., Barbaroux, M., D'Imperio, M., Jäncke, L., Elmer, S., & Besson, M. (2016). Professional music training and novel word learning: from faster semantic encoding to longer-lasting word representations. *Journal of Cognitive Neuroscience*, *28*(10), 1584–1602. http://doi.org/10.1162/jocn

Dittinger, E., Chobert, J., Ziegler, J. C., & Besson, M. (2017). Fast Brain Plasticity during Word Learning in Musically-Trained Children. *Frontiers in Human Neuroscience*, *11*(May), 1–16. http://doi.org/10.3389/fnhum.2017.00233

Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage*, *85*, 761–768. http://doi.org/10.1016/j.neuroimage.2013.06.035

Donchin, E. (1981). Surprise!...Surprise? *Psychophysiology*. http://doi.org/10.1111/j.1469-8986.1981.tb01815.x

Du, Y., & Zatorre, R. J. (2017). Musical training sharpens and bonds ears and tongue to hear speech better. *Proceedings of the National Academy of Sciences*, 201712223. http://doi.org/10.1073/pnas.1712223114

Emberson, L. L., Conway, C. M., & Christiansen, M. H. (2011). Timing is everything: Changes in presentation rate have opposite effects on auditory and visual implicit statistical learning. *Quarterly Journal of Experimental Psychology*, *64*(5), 1021–1040. http://doi.org/10.1080/17470218.2010.538972

Flaugnacco, E., Lopez, L., Terribili, C., Montico, M., Zoia, S., & Schön, D. (2015). Music training increases phonological awareness and reading skills in developmental dyslexia: A randomized control trial. *PLoS ONE*, *10*(9). http://doi.org/10.1371/journal.pone.0138715

Forget, J., Buiatti, M., & Dehaene, S. (2010). Temporal integration in visual word recognition. *Journal of Cognitive Neuroscience*, *22*(5), 1054–1068. http://doi.org/10.1162/jocn.2009.21300

François, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex*, *23*(9), 2038–2043. http://doi.org/10.1093/cercor/bhs180

François, C., Cunillera, T., Garcia, E., Laine, M., & Rodriguez-Fornells, A. (2017).

Neurophysiological evidence for the interplay of speech segmentation and word-referent mapping during novel word learning. *Neuropsychologia*, *98*(October), 56–67. http://doi.org/10.1016/j.neuropsychologia.2016.10.006

François, C., Jaillet, F., Takerkart, S., & Schön, D. (2014). Faster sound stream segmentation in musicians than in nonmusicians. *PLoS ONE*, *9*(7), 26–29. http://doi.org/10.1371/journal.pone.0101340

Francois, C., & Schön, D. (2010). Learning of musical and linguistic structures: Comparing event-related potentials and behavior. *NeuroReport*, *21*(14), 928–932. http://doi.org/10.1097/WNR.0b013e32833ddd5e

Francois, C., & Schön, D. (2011). Musical expertise boosts implicit learning of both musical and linguistic structures. *Cerebral Cortex*, *21*(10), 2357–2365. http://doi.org/10.1093/cercor/bhr022

François, C., & Schön, D. (2014). Neural sensitivity to statistical regularities as a fundamental biological process that underlies auditory learning: The role of musical practice. *Hearing Research*, *308*, 122–128. http://doi.org/10.1016/j.heares.2013.08.018

Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007a). Auditory attention - focusing the searchlight on sound. *Current Opinion in Neurobiology*, *17*(4), 437–455. http://doi.org/10.1016/j.conb.2007.07.011

Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007b). Does attention play a role in dynamic receptive field adaptation to changing acoustic salience in A1? *Hearing Research*, *229*(1–2), 186–203. http://doi.org/10.1016/j.heares.2007.01.009

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*, *129*(10), 2593–2608. http://doi.org/10.1093/brain/awl247

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive Neuroscience*, *16*(6), 1010–1021. http://doi.org/10.1162/0898929041502706

Gabay, Y., Thiessen, E. D., & Holt, L. L. (2015). Impaired statistical learning in developmental dyslexia. *Journal of Speech, Language, and Hearing Research*, *58*(3), 934–945. http://doi.org/10.1044/2015

Geiser, E., Sandmann, P., Jäncke, L., & Meyer, M. (2010). Refinement of metre perception - training increases hierarchical metre processing. *European Journal of Neuroscience*, *32*(11), 1979–1985. http://doi.org/10.1111/j.1460-9568.2010.07462.x

George, E. M., & Coch, D. (2011). Music training and working memory: An ERP study. *Neuropsychologia*, *49*(5), 1083–1094. http://doi.org/10.1016/j.neuropsychologia.2011.02.001

Ghitza, O. (2012). On the role of theta-driven syllabic parsing in decoding speech: Intelligibility of speech with a manipulated modulation spectrum. *Frontiers in Psychology*, *3*(JUL), 1–12. http://doi.org/10.3389/fpsyg.2012.00238

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. http://doi.org/10.1159/000208934

Gonzalez, C. C., & Burke, M. R. (2018). Motor Sequence Learning in the Brain: The Long and Short of It. *Neuroscience*, *389* (February), 85–98. http://doi.org/10.1016/j.neuroscience.2018.01.061

Gordon, R. L., Fehd, H. M., & McCandliss, B. D. (2015). Does music training enhance literacy skills? A meta-analysis. *Frontiers in Psychology*, *6*(DEC). http://doi.org/10.3389/fpsyg.2015.01777

Green, K. P., Kuhl, P. K., Meltzoff, A. N., & Stevens, E. B. (1991). Integrating speech information across talkers, gender, and sensory modality: Female faces and male voices in the McGurk effect. *Perception & Psychophysics*, *50*(6), 524–536. http://doi.org/10.3758/BF03207536

Habibi, A., Cahn, B. R., Damasio, A., & Damasio, H. (2016). Neural correlates of accelerated auditory processing in children engaged in music training. *Developmental Cognitive Neuroscience*, *21*, 1–14. http://doi.org/10.1016/j.dcn.2016.04.003

Hamrick, P., & Rebuschat, P. (2012). How implicit is statistical learning. In *Statistical learning and language acquisition* (pp. 365–382).

Haslinger, B., Erhard, P., Altenmüller, E., Hennenlotter, A., Schwaiger, M., Von Einsiedel, H. G., ... Ceballos-Baumann, A. O. (2004). Reduced recruitment of motor association areas during bimanual coordination in concert pianists. *Human Brain Mapping*, *22*(3), 206–215. http://doi.org/10.1002/hbm.20028

Herholz, S. C., & Zatorre, R. J. (2012). Musical Training as a Framework for Brain Plasticity: Behavior, Function, and Structure. *Neuron*, *76*(3), 486–502. http://doi.org/10.1016/j.neuron.2012.10.011

Hund-Georgiadis, M., & Yves Von Cramon, D. (1999). Motor-learning-related changes in piano players and non-musicians revealed by functional magnetic-resonance signals. *Experimental Brain Research*, *125*(4), 417–425. http://doi.org/10.1007/s002210050698

Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). The effects of musical training on structural brain development: A longitudinal study. *Annals of the New York Academy of Sciences*, *1169*, 182–186. http://doi.org/10.1111/j.1749-6632.2009.04852.x

Intartaglia, B., White-Schwoch, T., Kraus, N., & Schön, D. (2017). Music training enhances the automatic neural processing of foreign speech sounds. *Scientific Reports*, *7*(1), 1–7. http://doi.org/10.1038/s41598-017-12575-1

Jäncke, L. (2009). Music drives brain plasticity Major recent advances Structural brain plasticity, (October). http://doi.org/10.3410/B1-78

Jiang, Y., & Chun, M. M. (2001). Selective attention modulates implicit learning. *The Quarterly Journal of Experimental Psychology Section A*, *54*(4), 1105–1124. http://doi.org/10.1080/713756001

Kachergis, G., Yu, C., & Shiffrin, R. M. (2014). Cross-situational word learning is both implicit and strategic. *Frontiers in Psychology*, *5*(JUN). http://doi.org/10.3389/fpsyg.2014.00588

Kayser, C., Wilson, C., Safaai, H., Sakata, S., & Panzeri, S. (2015). Rhythmic Auditory Cortex Activity at Multiple Timescales Shapes Stimulus-Response Gain and Background Firing. *Journal of Neuroscience*, *35*(20), 7750–7762. http://doi.org/10.1523/JNEUROSCI.0268-15.2015

Kerlin, J. R., Shahin, A. J., & Miller, L. M. (2010). Attentional Gain Control of Ongoing Cortical Speech Representations in a “Cocktail Party.” *Journal of Neuroscience*, *30*(2), 620–628. http://doi.org/10.1523/JNEUROSCI.3631-09.2010

Kirkham, N. Z., Slemmer, J. a, & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain-general learning mechanism. *Cognition*, *83*, 35–42. http://doi.org/10.1016/S0010-0277(02)00004-5

Klein, C., Liem, F., Hänggi, J., Elmer, S., & Jäncke, L. (2016). The "silent" imprint of musical training. *Human Brain Mapping*, *37*(2), 536–546. http://doi.org/10.1002/hbm.23045

Kolk, H., & Chwilla, D. (2007). Late positivities in unusual situations. *Brain and Language*, *100*(3), 257–261. http://doi.org/10.1016/j.bandl.2006.07.006

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Science and Society*, *11*(August), 599–606. http://doi.org/10.1038/nrn2882

Landau, S. M., & D'Esposito, M. (2006). Sequence learning in pianists and nonpianists: An fMRI study of motor expertise. *Cognitive, Affective and Behavioral Neuroscience*, *6*(3), 246–259. http://doi.org/10.3758/CABN.6.3.246

Large, E. W. (2000). On synchronizing movements to music. *Human Movement Science*, *19*(4), 527–566. http://doi.org/10.1016/S0167-9457(00)00026-9

Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music., *26*, 1–37.

Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion*, *11*(5), 1021–1031. http://doi.org/10.1037/a0024521

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique* (Second, Vol. 78). Cambridge, MA: MIT Press. http://doi.org/10.1118/1.4736938

Luo, H., & Poeppel, D. (2007). Phase Patterns of Neuronal Responses Reliably Discriminate Speech in Human Auditory Cortex. *Neuron*, *54*(6), 1001–1010. http://doi.org/10.1016/j.neuron.2007.06.004

Luo, H., & Poeppel, D. (2012). Cortical oscillations in auditory perception and speech: Evidence for two temporal windows in human auditory cortex. *Frontiers in Psychology*, *3*(MAY), 1–10. http://doi.org/10.3389/fpsyg.2012.00170

Luo, H., Tian, X., Song, K., Zhou, K., & Poeppel, D. (2013). Neural response phase tracks how listeners learn new acoustic representations. *Current Biology*, *23*(11), 968–974. http://doi.org/10.1016/j.cub.2013.04.031

Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, *18*(2), 199–211. http://doi.org/10.1162/jocn.2006.18.2.199

Marie, C., Delogu, F., Lampis, G., Belardinelli, M. O., & Besson, M. (2011). Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *Journal of Cognitive Neuroscience*, *23*(10), 2701–2715. http://doi.org/doi: 10.1162/jocn.2010.21585

Milne, A. E., Petkov, C. I., & Wilson, B. (2017). *Auditory and visual sequence learning in humans and monkeys using an artificial grammar learning paradigm*. *Neuroscience*. The Author(s). http://doi.org/10.1016/j.neuroscience.2017.06.059

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. http://doi.org/10.1093/cercor/bhn120

Münte, T. F., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience*, *3*(6), 473–478. http://doi.org/10.1038/nrn843

Nikjeh, D. A., Lister, J. J., & Frisch, S. A. (2008). Hearing of note: An electrophysiologic and psychoacoustic comparison of pitch discrimination between vocal and instrumental musicians. *Psychophysiology*, *45*(6), 994–1007. http://doi.org/10.1111/j.1469-8986.2008.00689.x

Nikjeh, D. A., Lister, J. J., & Frisch, S. A. (2009). Preattentive cortical-evoked responses to pure tones, harmonic tones, and speech: Influence of music training. *Ear and Hearing*, *30*(4), 432–446. http://doi.org/10.1097/AUD.0b013e3181a61bf2

Ozernov-Palchik, O., & Patel, A. D. (2018). Musical rhythm and reading development: Does beat processing matter? *Annals of the New York Academy of Sciences*, (June). http://doi.org/10.1111/nyas.13853

Paraskevopoulos, E., Kuchenbuch, A., Herholz, S. C., & Pantev, C. (2012). Statistical learning effects in musicians and non-musicians: An MEG study. *Neuropsychologia*, *50*(2), 341–349. http://doi.org/10.1016/j.neuropsychologia.2011.12.007

Patel, A. D. (2003). Music, language, syntax and the brain. *Nature Neuroscience*, *6*(7), 674–681.

Patel, A. D., & Iversen, J. R. (2007). The linguistic benefits of musical abilities. *Trends in Cognitive Sciences*, *11*(9), 369–372. http://doi.org/10.1016/j.tics.2007.06.007

Patel, S. H., & Azzam, P. N. (2005). Characterization of N200 and P300: Selected studies of the Event-Related Potential. *International Journal of Medical Sciences*, *2*(4), 147–154.

http://doi.org/10.7150/ijms.2.147

Pavlidou, E. V., Louise Kelly, M., & Williams, J. M. (2010). Do children with developmental dyslexia have impairments in implicit learning? *Dyslexia*, *16*(2), 143–161. http://doi.org/10.1002/dys.400

Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cerebral Cortex*, *23*(6), 1378–1387. http://doi.org/10.1093/cercor/bhs118

Peretz, I., Saffran, J., Schön, D., & Gosselin, N. (2012). Statistical learning of speech, not music, in congenital amusia. *Annals of the New York Academy of Sciences*, *1252*(1), 361–366. http://doi.org/10.1111/j.1749-6632.2011.06429.x

Polich, J. (2007). Updating P300: An Integrative Theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. http://doi.org/10.1016/j.clinph.2007.04.019.Updating

Raviv, L., & Arnon, I. (2018). The developmental trajectory of children's auditory and visual statistical learning abilities: modality-based differences in the effect of age. *Developmental Science*, *21*(4), 1–13. http://doi.org/10.1111/desc.12593

Riecke, L., Sack, A. T., & Schroeder, C. E. (2015). Endogenous Delta/Theta Sound-Brain Phase Entrainment Accelerates the Buildup of Auditory Streaming. *Current Biology*, *25*(24), 3196–3201. http://doi.org/10.1016/j.cub.2015.10.045

Rinne, T., Pekkola, J., Degerman, A., Autti, T., Jääskeläinen, I. P., Sams, M., & Alho, K. (2005). Modulation of auditory cortex activation by sound presentation rate and attention. *Human Brain Mapping*, *26*(2), 94–99. http://doi.org/10.1002/hbm.20123

Roach, B. J., & Mathalon, D. H. (2008). Event-related EEG time-frequency analysis: An overview of measures and an analysis of early gamma band phase locking in schizophrenia. *Schizophrenia Bulletin*, *34*(5), 907–926. http://doi.org/10.1093/schbul/sbn093

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Schellenberg, E. G. (2011). Music lessons, emotional intelligence, and IQ. *Music Perception: An Interdisciplinary Journal*, *29*(2), 185–194. http://doi.org/10.3390/min8070278

Schellenberg, E. G., & Peretz, I. (2008). Music, language and cognition: unresolved issues. *Trends in Cognitive Sciences*, *12*(2), 45–46.

http://doi.org/10.1016/j.tics.2007.11.002

Schlaug, G., Norton, A., Overy, K., & Winner, E. (2005). Effects of music training on the child's brain and cognitive development. *Annals of the New York Academy of Sciences*, *1060*, 219–230. http://doi.org/10.1196/annals.1360.015

Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition*, *106*(2), 975–983. http://doi.org/10.1016/j.cognition.2007.03.005

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, *41*(3), 341–349. http://doi.org/10.1111/1469-8986.00172.x

Schuller, B., Batliner, A., Steidl, S., & Seppi, D. (2011). Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge. *Speech Communication*, *53*(9–10), 1062–1087. http://doi.org/10.1016/j.specom.2011.01.011

Shahin, A., Bosnyak, D. J., Trainor, L. J., & Roberts, L. E. (2003). Enhancement of Neuroplastic P2 and N1c Auditory Evoked Potentials in Musicians. *Journal of Neuroscience*, *23*(13), 5545–5552. http://doi.org/10.1523/JNEUROSCI.23-13-05545.2003

Shanks, D. R., Rowland, L. A., & Ranger, M. S. (2005). Attentional load and implicit sequence learning. *Psychological Research*, *69*(5–6), 369–382. http://doi.org/10.1007/s00426-004-0211-8

Shufaniya, A., & Arnon, I. (2018). Statistical Learning Is Not Age-Invariant During Childhood: Performance Improves With Age Across Modality. *Cognitive Science*, 1–16. http://doi.org/10.1111/COGS.12692

Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., & Frost, R. (2018). Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition*, *177*(April), 198–213. http://doi.org/10.1016/j.cognition.2018.04.011

Siegelman, N., Bogaerts, L., Kronenfeld, O., & Frost, R. (2017). Redefining "Learning" in Statistical Learning: What Does an Online Measure Reveal About the Assimilation of Visual Regularities? *Cognitive Science*, *42*, 692–727. http://doi.org/10.1111/cogs.12556

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, *81*, 105–120. http://doi.org/10.1016/j.jml.2015.02.001

Silva, S., Folia, V., Inácio, F., Castro, S. L., & Petersson, K. M. (2018). Modality effects in implicit artificial grammar learning: An EEG study. *Brain Research*, *1687*, 50–59. http://doi.org/10.1016/j.brainres.2018.02.020

Singh, S., Daltrozzo, J., & Conway, C. M. (2017). Effect of pattern awareness on the behavioral and neurophysiological correlates of visual statistical learning. *Neuroscience of Consciousness*, *3*(1), 1–14. http://doi.org/10.1093/nc/nix020

Strait, D. L., & Kraus, N. (2011). Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise. *Frontiers in Psychology*, *2*(113), 1–10. http://doi.org/10.3389/fpsyg.2011.00113

Strait, D. L., Kraus, N., Skoe, E., & Ashley, R. (2009). Musical experience and neural efficiency - Effects of training on subcortical processing of vocal expressions of emotion. *European Journal of Neuroscience*, *29*(3), 661–668. http://doi.org/10.1111/j.1460-9568.2009.06617.x

Strait, D. L., Slater, J., O'Connell, S., & Kraus, N. (2015). Music training relates to the development of neural mechanisms of selective auditory attention. *Developmental Cognitive Neuroscience*, *12*, 94–104. http://doi.org/10.1016/j.dcn.2015.01.001

Suárez, L., Elangovan, S., & Au, A. (2016). Cross-sectional study on the relationship between music training and working memory in adults. *Australian Journal of Psychology*, *68*(1), 38–46. http://doi.org/10.1111/ajpy.12087

Swaminathan, S., & Schellenberg, E. G. (2017). Musical competence and phoneme perception in a foreign language. *Psychonomic Bulletin and Review*, *24*(6), 1929–1934. http://doi.org/10.3758/s13423-017-1244-5

Talamini, F., Altoè, G., Carretti, B., & Grassi, M. (2017). Musicians have better memory than nonmusicians: A meta-analysis. *PLoS ONE*, *12*(10), 1–21. http://doi.org/10.1371/journal.pone.0186773

Tanner, M., Rammage, L., & Liu, L. (2016). Does singing and vocal strengthening improve vocal ability in people with Parkinson's disease? *Arts and Health*, *8*(3), 199–212. http://doi.org/10.1080/17533015.2015.1088047

Teng, X., Tian, X., Doelling, K., & Poeppel, D. (2018). Theta band oscillations reflect more than entrainment: Behavioral and neural evidence demonstrates an active chunking process. *European Journal of Neuroscience*, *48(*(8), 2770–2782. http://doi.org/10.1111/ejn.13742

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: An event-related potential and behavioral study. *Experimental Brain Research*, *161*(1), 1–10. http://doi.org/10.1007/s00221-004-2044-5

Thomson, J. M., & Goswami, U. (2008). Rhythmic processing in children with developmental dyslexia: auditory and motor rhythms link to reading and spelling. *Journal of Physiology-Paris*, *102*(1–3), 120–129. http://doi.org/10.1016/j.jphysparis.2008.03.007

Tierney, A., & Kraus, N. (2013). *Music training for the development of reading skills*. *Progress in Brain Research* (1st ed., Vol. 207). Elsevier B.V. http://doi.org/10.1016/B978-0-444-63327-9.00008-4

Tierney, A. T., Krizman, J., & Kraus, N. (2015). Music training alters the course of adolescent auditory development. *Proceedings of the National Academy of Sciences*, *112*(32), 10062–10067. http://doi.org/10.1073/pnas.1505114112

Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition*, *97*(2), 25–34. http://doi.org/10.1016/j.cognition.2005.01.006

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*(4), 552–564. http://doi.org/10.1037/0096-3445.134.4.552

Van Der Meulen, I., Van De Sandt-Koenderman, M. W. M. E., Heijenbrok, M. H., Visch-Brink, E., & Ribbers, G. M. (2016). Melodic Intonation Therapy in Chronic Aphasia: Evidence from a Pilot Randomized Controlled Trial. *Frontiers in Human Neuroscience*, *10*(November), 1–9. http://doi.org/10.3389/fnhum.2016.00533

Vasuki, P. R. M., Sharma, M., Ibrahim, R. K., & Arciuli, J. (2017). Musicians' Online Performance during Auditory and Visual Statistical Learning Tasks. *Frontiers in Human Neuroscience*, *11*(114), 1–12. http://doi.org/10.3389/fnhum.2017.00114

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, *10*(4), 420–422. http://doi.org/10.1038/nn1872

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*(7), 547–558. http://doi.org/10.1038/nrn2152

APPENDIX 1

TABLE 1. APPENDIX 1. *Characterization of the Sounds Used in the Behavioral Test Phases of Each Experiment described in Chapter 4.*

TABLE 2. APPENDIX 1. *Characterization of the Sounds Used in the Test Phase of the Prosodic Experiment described in Chapter 5.*

TABLE 1. APPENDIX 1

Characterization of the Sounds Used in the Behavioral Test Phases of Each Experiment Described in Chapter 4

Experiment	Test	Words/melodies	Part-words/melodies	Non-words/melodies
Prosodic experiment	Prosodic	kuliri (C4 D4 F4) luloru (E5 Db5 G4) dokiru (D5 C5 G4) diluro (B4 E5 F5) rudiki (G4 B4 C5)	kirulu (C5 G4 E5) loruku (Db5 G4 C4) dikilu (B4 C5 E5) rokuli (F5 C4 D4) ridoki (F4 D5 C5)	kidiru (C5 B4 G4) riliku (F4 D4 C4) roludi (F5 E5 B4) rukido (G4 C5 D5) rulolu (G4 Db5 E5)
	Linguistic (flat contour words)	kuliri luloru dokiru diluro rudiki	kirulu loruku dikilu rokuli ridoki	kidiru riliku roludi rukido rulolu
	Musical (piano melodies)	C4 D4 F4 E5 Db5 G4 D5 C5 G4 B4 E5 F5 G4 B4 C5	C5 G4 E5 Db5 G4 C4 B4 C5 E5 F5 C4 D4 F4 D5 C5	C5 B4 G4 F4 D4 C4 F5 E5 B4 G4 C5 D5 G4 Db5 E5
Linguistic experiment		xemuvu memive zixuve zumevi vezuxu	mivevu vixemu vuzixu xuveme zuxume	vumuxe vemime vexuzi vimezu xuzuve
Musical experiment		D4 E4 G4 F5 D5# A4 E5 D5 A4 C4 F5 G5 A5 C4 D5	C4 D5 F5 D#5 A4 D4 D5 A4 F5 G4 A5 D5 G5 D4 E4	G4 E4 D4 A4 D5# F5 A4 D5 E5 G5 F5 C4 D5 C4 A5

Note. The pitch (Hz) of the notes that compose each melody is reported in musical letter notation according to the English convention. Each note is named by scientific pitch notation with a letter-name and a number identifying pitch's octave; additionally, "#" stands for *sharp*, "b" stands for *flat*.

TABLE 2. APPENDIX 1

Characterization of the Sounds Used in the Test Phase of the Prosodic Experiment Described in Chapter 5

Test	Words/melodies	Part-words/melodies	Non-words/melodies
Prosodic	gifaca C3 D3 F3 fifeci E4 Db4 G3 begaci D4 C4 G3 bafice B3 E4 F4 cibaga G3 B3 C4	bagafi (B3 C4 E4) cabega (F3 D4 C4) cegifa (F4 C3 D3) fecigi (Db4 G3 C3) gacifi (C4 G3 E4)	cafagi (F3 D3 C3) cifefi (G3 Db4 E4) cigabe (G3 C4 D4) cefiba (F4 E4 B3) gabaci (C4 B3 G3)
Linguistic (flat contour words)	gifaca fifeci begaci bafice cibaga	bagafi cabega cegifa fecigi gacifi	cafagi cifefi cigabe cefiba gabaci
Musical (piano melodies)	C3 D3 F3 E4 Db4 G3 D4 C4 G3 B3 E4 F4 G3 B3 C4	B3 C4 E4 F3 D4 C4 F4 C3 D3 Db4 G3 C3 C4 G3 E4	F3 D3 C3 G3 Db4 E4 G3 C4 D4 F4 E4 B3 C4 B3 G3

Note. The pitch (Hz) of the notes that compose each melody is reported in musical letter notation according to the English convention. Each note is named by scientific pitch notation with a letter-name and a number identifying pitch's octave; additionally, "#" stands for *sharp*, "b" stands for *flat*.

APPENDIX 2

Verbatim copy of the informed consent provided to the participants.

CONSENTIMENTO INFORMADO

(De acordo com a “Declaração de Helsínquia” da Associação Médica Mundial (Helsínquia 1964; Tóquio 1975; Veneza 1983; Hong Kong 1989; Somerset West 1996 e Edimburgo 2000))

1. Denominação do Projeto de Investigação: Efeitos de transferência entre a música e a linguagem: o impacto do treino musical sobre os correlatos comportamentais e neuronais da aprendizagem de estruturas linguísticas e musicais.

2. Informação sobre o Projeto de Investigação

Ex.$^{mo(a)}$ Senhor(a),

Convidámo-lo(a) a participar num projeto de investigação que visa examinar a influência do treino musical sobre o processamento de estruturas linguísticas e musicais. O projeto é desenvolvido pela doutoranda Margarida Vasconcelos e coordenado pela Investigadora Doutora Ana Pinheiro da Escola de Psicologia da Universidade do Minho. O objectivo deste documento é prestar-lhe informação sobre o projeto. Deverá ler o documento e decidir se pretende participar do projeto de forma voluntária, devendo para tal assinar o consentimento que lhe é apresentado na página seguinte.

2.1 Descrição do projeto e procedimentos: O objetivo deste projeto é perceber se o treino musical tem impacto sobre a linguagem e, em particular, compreender como esse treino influencia os processos neuronais subjacentes ao processamento de regularidades sonoras em estruturas linguísticas e musicais. O projeto envolve a realização de tarefas que visam a aquisição de dados clínicos, comportamentais e electrofisiológicos. As tarefas durante as quais serão recolhidos dados comportamentais serão tarefas em que é pedido ao participante para distinguir sons. Com vista à obtenção dos dados electrofisiológicos, que permitirão examinar o processamento neuronal, serão realizados electroencefalogramas. Um electroencefalograma é um exame que regista graficamente mudanças na voltagem das correntes eléctricas do cérebro. Este registo não é invasivo ou doloroso, não induz quaisquer efeitos secundários e consiste na colocação de alguns eléctrodos no couro cabeludo, juntamente com um gel electrolítico. Durante os electroencefalogramas, o participante estará a realizar tarefas simples, tais como, ouvir um sequência de palavras ou melodias. Também poderá ser solicitado ao participante que responda a questões simples sobre as tarefas que se encontra a realizar. Durante cada tarefa, pode existir um ou mais intervalos determinados pelo investigador responsável e/ou o participante. Outros procedimentos do projeto incluem a recolha de dados sócio-demográficos e cognitivos. A duração total da participação no projeto é variável e não inferior a 6 horas. A participação neste projeto implica, assim, a realização de um conjunto de tarefas a ter lugar na Universidade do Minho (Braga).

2.2 Questões associadas à participação na investigação: Não há quaisquer riscos associados à participação nesta investigação.

2.3 Uso dos dados do participante e confidencialidade: Os dados resultantes deste estudo serão mantidos confidenciais e anónimos. Apenas um número limitado de pessoas (membros da equipa de investigação) terão acesso aos dados. Esta informação não será usada para nenhum outro fim que não o dos objetivos aqui descritos. Apenas serão divulgados os resultados globais por grupos de indivíduos sem qualquer informação que leve à identificação dos respectivos participantes. Os resultados desta investigação poderão ser publicados para objetivos científicos, mas a sua identidade ou dados pessoais não serão revelados.

2.4 Direitos do(a) participante da investigação: Questões acerca da investigação ou acerca dos direitos do participante poderão ser colocadas aos investigadores responsáveis (Dr. Ana P. Pinheiro, 253 601 398; Margarida Vasconcelos, 911 542 955). A participação no projeto é voluntária, podendo o participante desistir de participar em qualquer momento.

3. Identificação

Nome:____Data: ______________________

Caso aceite fazer parte desta investigação, assine por favor a secção referente ao consentimento.

CONSENTIMENTO

Eu, abaixo-assinado,____ (nome completo), fui informado(a) e compreendi os objectivos e procedimentos do projeto no qual participarei de forma voluntária. Tomei conhecimento de que, de acordo com as recomendações da Declaração de Helsínquia, a informação ou explicação que me foi prestada versou os objetivos, os métodos, os benefícios previstos, os riscos potenciais e o eventual desconforto do projeto. Todas as minhas questão foram devidamente esclarecidas. Fui também informado(a) de que em qualquer momento poderei recusar a minha participação se assim o entender. Assim, autorizo que os meus dados integrem esta investigação e que sejam apresentados anónima e confidencialmente em apresentações públicas, congressos científicos e publicações.

____________________, _________ /_________/_________

(Local e Data)

____

Assinatura do(a) participante voluntário(a)

____

Assinatura da investigadora responsável