Motor auditory interaction: a window towards associative and predictive processing

Vittoria Spinosa

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

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

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

MOTOR-AUDITORY INTERACTION: A WINDOW TOWARDS ASSOCIATIVE AND PREDICTIVE PROCESSING

Vittoria Spinosa

UNIVERSITAT DE
BARCELONA

Vittoria Spinosa | Motor-auditory interaction: a window towards associative and predictive processing

UNIVERSITAT DE BARCELONA

Department of Clinical Psychology and Psychobiology
Faculty of Psychology
University of Barcelona

Motor-auditory interaction: a window towards associative and predictive processing

Thesis presented by
Vittoria Spinosa

To obtain the degree of doctor
by the University of Barcelona

In accordance with the requirements for the
International PhD Diploma
Doctoral Programme in Biomedicine

Supervised by
Dr. Iria SanMiguel
Dr. Carles Escera

Barcelona, September 2019

Institut de Neurociències
UNIVERSITAT DE BARCELONA

Brainlab
cognitive neuroscience research group
grup de recerca en neurociència cognitiva

A mia sorella, Mariella,
per avermi trasmesso l’amore per il sapere.

Ed ai miei genitori,
per il supporto costante ed incondizionato

ACKNOWLEDGMENTS · AGRADECIMIENTOS · RINGRAZIAMENTI

The people supporting me during these years have been crucial to face most of the challenges of the PhD. This thesis starts acknowledging them, besides if they were close to me or faraway, if they were temporaneously crossing my life or constantly by my side. Probably the following words are not enough to express my acknowledgments to all of you, but I had to stop in order to save some trees.

My warmest acknowledgment goes to my family. Il piú sentito e immenso grazie è rivolto alla mia famiglia. Grazie per prendervi cura di me tanto a distanza, quanto in Italia. Un grazie in particolare ai miei genitori per aver sempre creduto in me, lasciandomi libera di allontanarmi dal mio porto sicuro, rispettando i miei desideri e sostenendo i miei sforzi. Un immenso grazie a mia sorella Mariella per alimentare la mia curiositá, fame di conoscenza, amore per la scienza e… tante altre cose (lo so, ti devo l'infinito!). Grazie a mia nipote Irene, dolce tempesta di pace, per farmi ricordare che non si vive di solo lavoro, ma soprattutto di relazioni umane, empatiche e curative. E grazie anche al mio nipotino Davide, ancora nel ventre materno: la voglia di conoscerti e coccolarti mi ha incentivato a terminare il dottorato. Grazie anche a Giulio, ai miei zii Tonia e Vincenzo, Franco e Agata ed ai pro zii Rosa e Pino, per ricaricarmi di affetto ai miei ritorni in patria. In particolare, grazie ai miei pro zii per le chiamate serali domenicali. Grazie di cuore a tutti voi. Il bene che vi voglio non ha confini, né unitá di misura.

My sincere acknowledgements go to the pillars of my PhD: Iria SanMiguel and Carles Escera… without special persons like you the PhD would not have become reality. I feel incredibly lucky for having found you: it allowed me to learn, as researcher and as person. Thank you very much Iria for allowing me to join the Brainlab, and being your first PhD candidate. Thanks for guiding me into the fascinating world of perception and neuroscience, providing me the tools to investigate (and to write...) always in a scientific, critic, and rigorous way. I am grateful to you not only for being a teacher willing to exchange ideas, but also for being a really humble person (despite having an immense scientific knowledge). Thanks for being always available to help me, for comprehending my difficulties and for recognizing when it was time to rest. Although the "final climbing" of the PhD is never simple, I feel really grateful also to my co-supervisor. Many thanks Carles for illuminating the pathway towards the deposit of the thesis. Thanks for providing me an external, neutral and expert point of view which helped me in putting black on white the immense work done during the PhD. A special thanks also to Dr. János Horváth and Bence Neszmélyi for allowing me to carry on the second experiment of this thesis in the Hungarian Academy of Science. Thank you both for the intense brainstorming preceding the actual experiments, and for your support in Budapest. Thanks to Bence also for all your efforts in programming the experiment, and for the nice time spent inside and outside the university.

A big thanks goes to all the members of the Brainlab: each one of you contributed in making the Lab such a special and enjoyable working place.

Marta Turró, thanks for guiding me in the infinite paperwork waiting for the foreigner candidates, for inviting me to play again the violin, to speak an initial Vit-Spanish and, especially, thanks for taking care of me. Jordi Costa-Faidella, thanks for nourishing my love for music and neuroscience with your passion and deep knowledge in auditory perception, thanks for reminding me when it was time to go home because night, for introducing me to the ICA correction, and especially for all the support, and help, you gave me in the most though moment of the PhD.

Thanks to the Lab colleagues whom accompanied me along the entire PhD: Kasia Żarnowiec, Natàlia Gorina, Fran Lopez and Teresa Ribas. Many thanks Kasia for having been the friend with whom to share not only the passion for Barcelona, but also the difficulties of being in a foreign country. I am grateful to you & Piotrek also for taking care of me, for reinforcing my achievements, and for allowing me to disconnect from work in your company. Natàlia, you have been an example of resolute personality for me. I thank you because, besides our divergent opinions, I could count on you in any moment in the Lab: you always offered me a suggestion for any doubt or drawback. Fran, I am really grateful to have lived with you each little step of this PhD, from getting the scholarship until finishing of writing up the thesis. Thanks for sharing with me not only the most joyful moments (i.e., during conferences and courses), but also the less funny moments (e.g., administration, etc...). And then, my dear Teresa: many thanks for being not only a colleague, but also a real friend, always supporting me and my dreams. Thanks for sharing and enjoying with me the little spare time left outside the university, and for the infinite energy and vitality you passed me in these years. You always brought me a pinch of lightheartedness, staying by my side, even in the darkest moments.

Paco Díaz, thanks for the patience you had with my technology and “Madlab”-related issues. Thanks to you, Marc Via and Roger (colleagues closed to my office) for your irony: it was a key ingredient to survive to the long days spent in Mundet. And many thanks also to the Lab colleagues more recently arrived Nadia Paraskevoudi, Marta Font, Giannina Puddu and Sonia Arenillas, for being great and super supportive office mates, and for encouraging me especially when the energy levels were getting lower and lower. Thanks also to all the people whom shared the office or worked temporally with me in the Lab: especially to Marta Ortega, Tin Senija, Miriam Cornella & Burcu Avci for sharing with me pleasant moments. I am grateful to Cristina Pulido and Marta Renato for the nice working-breaks spent together, and also to the members of the Psychology Department whom shared with me the lunch times both in Spain (Imma Clemente, MJ Corral, Raffaele Cacciaglia, Maria Mataró, Roser Pueyo, Mª Àngels Jurado, Montse Colell, Xavier Caldú, Noemi Lamonja, Jonatan Ottino, Julia Ballester, Olga Laporta y Xavier Prats) and in Hungary (Márta Volosin, Zsófia Szabó, Endre Takács, Petia Kojouharova, Orsolya Szalárdy, Anna Altbäcker and Irén Barkaszi). A special thanks also to the UB students whom helped me with the Template project (such as, Laura Casafont, Aina, Marta Blasco, Ignacio de Loyola Ruiz, Marta Colet), to Emese Várkonyi and Barbara Matulai for their help in collecting the data in Budapest, and to all the Spanish and Hungarian volunteers participating in the experiments of this thesis (without you I could not give any contribution to science). Thanks also to the researchers crossed in the conferences for their constructive comments.

This thesis would not have been possible without the support of the extraordinary people met outside the Lab, whom constantly encouraged my work and ideas, and whom taught me that joining the forces it is possible to achieve any goal.

Sin duda, un enorme agradecimiento va a Laura Díaz. Gracias con todo el corazón Laura por ser una amiga y un ejemplo de proactividad, por tu confianza en mis capacidades y en mi trabajo, por todo el apoyo incondicional que me has ofrecido, y por transformar nuestras dificultades en proyectos concretos. Muchísimas gracias también a Janire Gesto y Roger Ponce por saber como despertar mi buen humor, y junto a Ryan Armstrong, Laurent Ladépêche, Patricia Laiz, y Salva Guardiola por apoyarme y permitirme de desconectar del trabajo con actividades a favor de la ciencia.

Muchas gracias también a personas fuera del ámbito científico como Eva del Álamo y Susana Noguera por ofrecerme siempre vuestra ayuda y compañía, y a todas las personas que he conocido con baja o sin visión (e.g., Manel Martí, Mauro, Pep, Mari Pau, Graciela, Francesc, Dolors): gracias a todos vosotros por confiar en mí, por haberme permitido de explicaros de forma sencilla de que iba mi trabajo (y por vuestro interés en escucharme sin aburriros), por cuidaros de mi y en particular por el ejemplo que me habéis ofrecido con vuestra fuerza y determinación.

Gracias también a las personas con la que he tocado el violín en orquesta, en particular a Sànder Celma-Miralles, y en música de cámara, a José Manuel Soler y Paula Caserras, por las charlas pre- y post-ensayos musicales, y por permitir que mi mente no fuera ocupada sólo de ciencia, sino también de melodías.

Many thanks also to the landlord of my apartment, Michael Kennedy-Scanlon, y a mis compis, en particular a Monika, por cuidarse de mí y por haber hecho todo lo posible para respetar la tranquilidad en casa, sobre todo en los momentos más delicados de escritura de la tesis.

Among the people supporting me from faraway, a big big thanks goes to Marietta Forgó and to Alice Gozzi from Helsinki. Both of you were encouraging me from the phase of applying to this PhD until the writing up of the thesis. A special thanks to Marietta for helping me with the Hungarian language and for the visit in Budapest, ed un mega grazie ad Alice per l'immancabile e insostituibile supporto morale. Colgo l'occasione per ringraziare anche Andrea Ravignani per il suo supporto e le appassionanti chiacchierate su ritmo, musica e neuroscienze.

Tra le persone lontane, ringrazio anche gli amici italiani che mi hanno sempre supportato da Bari e dintorni, in particolare a Rossella Tridente, Concetta Civita, Giacomo & Giusy Stano, Manu Amato, Lucia Benevento, Rosa Raspatella e Marinella Simone. Ed un ultimo, ma speciale, ringraziamento va a Rossella De Giorgio ed a Stella della Queva per darmi coraggio ed aiutare a credere in me stessa nei momenti piú ardui.

Grazie · Thanks · Gracias · Gràcies · Kiitos · Köszönöm

ABSTRACT

Perception can be considered as the bridge necessary to interact with the surroundings. Indeed, through our senses we are continuously exchanging information between the physical world around us and our internal models, made of prior beliefs and expectations, which allow to constantly create meanings and interpretations about the events caused by external sources or by us. Illusions and cross-modal effects are proofs of the incessant and constructive activity of the human brain.

By experience, we learn also to make action-effect associations and to predict the sensory consequences of each motor act. For instance, we learn that by touching a warm object we can get burnt, thus we can predict that this action can be dangerous and avoid making it in the future.

Current research is attempting to investigate the mechanisms underlying associative and predictive processing, and how these mechanisms are modulated by the sensory experience. A great deal of this research is carried out with the electroencephalogram (EEG), and some intriguing results have suggested that a dysregulation of associative and predictive processing may underlie some clinical phenomena, such as psychosis.

The present PhD thesis focuses on expectation effects arising from motor-auditory interactions, in order to investigate action-effect associative and predictive processing. Two independent studies were carried out in which behavioral (action timing and action force) and electrophysiological (event-related brain potentials, ERPs) measurements were collected

while participants were instructed to press different buttons that were associated with either the presentation or the omission of a sound, and with violations of these associations.

Study I focused on predictive processing and aimed to investigate the matching- and mismatching-related effects between an event and an established expectation. The findings suggest that these effects are manifestations of a unitary underlying process of prediction. Study II focused on associative processing and aimed to investigate the buildup of action-effect associations, in a context in which there were no stable regularities. The findings suggest that the repetition of an action-effect contingency can foster an expectation, and that the consequences of an action can modify the action itself.

Overall, the results from this thesis suggest that associative and predictive processing arising from a motor act modulate behavior and neuronal activity in a close-loop manner. That is, the motor act causes specific consequences which modulate behavior and neuronal processing, and internal action-effect associations and predictions seem to affect, in turn, the neuronal processing and the action itself.

RESUMEN

La percepción puede considerarse como el puente necesario para interactuar con el entorno. De hecho, a través de nuestros sentidos, estamos continuamente intercambiando información entre el mundo físico que nos rodea y nuestros modelos internos, formados a partir de creencias anteriores y expectativas, que permiten dar constantemente significado e interpretación a los sucesos causados por fuentes externas o por nosotros mismos. Las ilusiones y los efectos intersensoriales son pruebas de la actividad incesante y constructiva del cerebro humano.

Por experiencia, aprendemos también a hacer asociaciones de acción-efecto y a predecir las consecuencias sensoriales de cada acto motor. Por ejemplo, aprendemos que al tocar un objeto caliente nos quemamos, y por tanto podemos predecir que esta acción puede ser peligrosa y así evitarla en el futuro.

Actualmente, muchos investigadores están estudiando los mecanismos subyacentes del procesamiento asociativo y predictivo, y la modulación de este procesamiento en la experiencia sensorial. El electroencefalograma (EEG) es una de las técnicas que permiten esta investigación. Algunas teorías proponen, además, que una desregulación del procesamiento asociativo y predictivo podría explicar fenómenos clínicos, tales como la psicosis.

En la presente tesis doctoral, nos centramos en los efectos de la expectativa derivados de la interacción motora-auditiva, con el fin de estudiar los procesos asociativos y predictivos entre acciones - efecto. Para alcanzar este objetivo, realizamos dos estudios independientes en los que registramos medidas de comportamiento (tiempo de acción y fuerza de la acción) y electrofisiológicas (potenciales evocados) mientras los participantes realizaron pulsaciones de

botones que estaban asociadas con la presentación o la omisión de un sonido, y con las violaciones de estas asociaciones.

El estudio I se centró en procesos predictivos consolidados y tuvo como objetivo investigar los efectos relacionados con la coincidencia y la falta de coincidencia entre un evento y una expectativa. Los resultados sugieren que estos efectos son manifestaciones del mismo proceso subyacente de predicción. El estudio II se centró en el procesamiento asociativo y tuvo como objetivo investigar la construcción de asociaciones de acción-efecto, en un contexto en el que no había regularidades establecidas. Los resultados sugieren que la repetición de una contingencia acción-efecto puede fomentar una expectativa, y que la consecuencia de una acción puede modificar la acción en sí.

Los resultados sugieren que el procesamiento asociativo y predictivo que surge de un acto motor modula circularmente el comportamiento y la actividad neuronal. Es decir, el acto motor causa consecuencias específicas que modulan el comportamiento y el procesamiento neuronal, pero también las asociaciones de acción-efecto y las predicciones relacionadas parecen afectar el procesamiento neuronal y la acción en sí.

List of studies

This thesis includes two completed studies that are in the process of being submitted for publication to international peer-reviewed scientific journals.

STUDY I.

Spinosa, V., & SanMiguel, I. Expecting sounds and silences: Disentangling match- and mismatch-related effects on auditory ERPs.

STUDY II.

Spinosa, V., Neszmélyi, B., Horváth, J., SanMiguel, I. An auditory ERP study tracking the buildup of action-effect associations.

This thesis is submitted to obtain the Degree of Doctor by the University of Barcelona (International Doctor Mention). Study I is the result of the work carried out under the direct supervision of Dr. Iria SanMiguel and Prof. Carles Escera at the *Brainlab – Cognitive Neuroscience Research Group* (Excellence Research Group established by the Generalitat de Catalunya SGR2014-177 and SGR2017-974), at the Department of Clinical Psychology and Psychobiology of the Faculty of Psychology, University of Barcelona (Barcelona, Spain); and Study II is the result of the work performed in a three-month research visit at the Experimental Psychology Research Group, led by Dr. János Horváth, located in the Institute of Cognitive Neuroscience and Psychology, Hungarian Academy of Sciences (Budapest, Hungary).

This work was funded by the Spanish Ministry of Economy and Knowledge through funds awarded to Dr. Iria SanMiguel (PSI20145273P, RYC2013312577), and by a doctoral grant from the University of Barcelona (Ajuts de Personal Investigador Predoctoral en Formació-APIF) awarded to Vittoria Spinosa. Part of the research was also supported by the Hungarian Scientific Research Fund (K 108783) awarded to Dr. János Horváth.

GLOSSARY OF ABBREVIATIONS

ANOVA Analysis of Variance

CS Control Sound

CO Control Omission

DSM Diagnostic and Statistical Manual

EEG Electroencephalography

ES Expected Sound

EO Expected Omission

ERP Event-Related Potential

IPI Inter-Press Interval

M Match

MM Mismatch

MMN Mismatch Negativity

RP Repetition Positivity

US Unexpected Sound

UO Unexpected Omission

INDEX

1. INTRODUCTION

PERCEPTION: A BRIDGE BETWEEN EXTERNAL AND INTERNAL WORLD

There is a general agreement in conceiving human beings as not passive receptors of an external reality, but as constant creators of meaning and active interpreters of reality (Purves & Lotto, 2011; Frith, 2013). It goes beyond the goals of this thesis to define the concept of reality, thus we assume that reality corresponds to the physical world, and that the human interaction with the surrounding world is made possible by discovering, experiencing and interpreting the physical world around us.

We could imagine perception as a bridge, which connects the external physical world with our internal models made up of thoughts, beliefs and expectations. The activity along this bridge never stops, because we are continuously exchanging signals between the world outside and inside our body. Is there a window toward this incessant activity?

Illusions reveal the ongoing interaction between the external sensory input and our internal knowledge, and they have often been subject of study from experimental psychology (Pressnitzer, Suied, & Shamma, 2011).

The most well-known illusions are the optical ones (Purves & Lotto, 2011), maybe thanks also to their application in graphic art, such as in Escher's figures. Among the visual illusions, we can mention the Kanizsa's triangle (1955), in which we perceive a white bright triangle, pointing upwards, in the middle of the figure. The illusion is given by an incomplete inverted triangle with black outlines, pointing downwards, and three black disks with cutouts (also named as "pacmen"). Different developmental studies reported that the ability to complete a Kanizsa illusory figure starts early during infancy (Condry, Smith, & Spelke, 2001; Csibra, 2001), but an eye-tracking study reported that different attentional operations can occur for 6-months-old infants compared to adults when they were subjected to figures with Kanizsa's illusion (Bulf, Valenza, & Simion, 2009), suggesting the impact of development on the perception of the illusion.

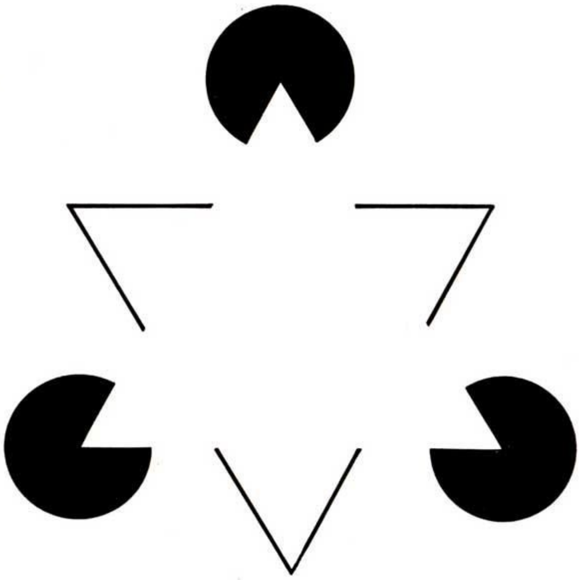

Figure 1. Kanizsa's triangle (1955)

The development of magnetic tape recorders allowed to isolate sounds and to study auditory illusions. One of these auditory illusions is the phonemic restoration discovered by Warren & Obusek (1971), who replaced a phoneme of a sentence by extraneous non-speech sounds (e.g., a cough or a tone), or by a silent gap, of the same duration as the missing phoneme. This study showed that in the case of non-speech sounds, participants can easily restore the missing phoneme, failing in recognizing the exact moment of the interference. However, when the phoneme was replaced by a silence, participants were able to detect the exact location of the missing phoneme and phonemic restoration occurred less easily. Thus, this study leaves questions about the brain's ability to restore information and how silent interruptions are processed in the brain.

Although there are illusions for each sensory modality (e.g., visual, auditory, and tactile illusions), showing that it is possible to trick each one of our senses, there are also particular types of illusions, such as multisensory illusions, which show that our perception of the physical reality occurs simultaneously through different senses, although often we are not aware of it. One of these cross-modal effects is the McGurk Effect, discovered by McGurk and MacDonald (1976), in which conflicting information coming from the sight (lips movement mimicking "ba") and the hearing (auditory input "ga") can affect what is actually perceived (auditory sound "da"). Another cross-modal effect is the illusory flashing which confirms the constant interaction between different modalities (Shams et al., 2000). Indeed, although the experimenters presented a single visual flash on the screen, the number of visual flashes experienced by the participants depended on the number of auditory beeps produced concomitantly to the single visual flash.

Along the entire lifespan, we develop the ability of perceiving the external inputs simultaneously from different senses through cross-modal experiences, until reaching a multisensory integration which is vital for our survival (Stein, 2014). Illusions and cross-modal effects suggest that the brain is a dynamic system and not only a passive receptor of the external reality.

Gestalt psychology has been the most prominent theory describing the perceptive system and offered a fundamental contribution to the understanding of human perception. The core defended by the Gestaltists was that the whole is more than the sum of the single parts. That means that it is not possible to decompose perception by the single elements that elicit it, since these elements interact with each other and with the person. In the same way, one's personal perception of a concert does not depend only on the single notes that were played, since two people who assist to the same concert can perceive them differently. This idea can be particularly evident also when we look at visual illusions, since the illusion effect emerges by the combination of all the elements together and not only observing each of the single elements separately. Indeed, each person can focus on a different element, perceiving the figures differently.

Although Gestaltists could not take advantage of modern technology that allows us to explore the nervous system, they raised interesting questions which have inspired many current experimental psychologists and neuroscientists (Rock & Palmer, 1990). For instance, among Gestalt theories, Köhler's theory (1938) has been particularly relevant for modern theories on perception, since he describes the brain system as a physical structure that tends towards a balanced state, necessary to waste no energy.

THEORIES ON ASSOCIATIVE AND PREDICTIVE PROCESSING

The illusions and the cross-modal effects described in the previous paragraph have shown that perception is an active system, which is affected by previous experiences of interaction with the environment. During these interactions we learn to recognize the repetitive patterns to which we are exposed along the entire life span, and to build-up associations between phenomena. These associations are fundamental for the creation of internal models with which we are able to make predictions about upcoming events.

Every day we make use of associative and predictive processing, although without being aware of that. For instance, let’s imagine we are watching TV and in a given moment we take the remote control to change the TV channel. How do we know that pressing a button will change the channel? We first learn to establish associations between an action and its possible effects. Learning these action-effect associations we are then able to build-up internal models which allow us to finally make predictions about the possible consequences of our actions. For instance, once one learns the association between pressing a button (action) and its sensory consequence (effect), one can decide how to make use of the TV remote control. Indeed, by experience, one can learn the specific sensory consequences that could be caused by the actions (Stock & Stock, 2004), then can retrieve this learning from memory (Bar, 2007), and based on this knowledge one can predict the consequences and select the action which fits better with the goals (Baum, Wolfensteller, & Ruge, 2017; Horváth, Bíró, & Neszmélyi, 2018). In this section we reported some of the most influential theories about associative and predictive processing.

Associative processing

A big contribution to the understanding of the causal relationship between action and its effects derives from instrumental learning theories. For instance, Thorndike (1931) demonstrated that a cat was able to reach a goal (i.e., to escape from a puzzle box) using an instrument (i.e., pulling some strings). Specifically, this discovery revealed that learning was possible not only by trial but also by error, suggesting the importance of this latter (Frith, 2013). In humans, the understanding of the ability of making use of the body to reach specific goals has nourished the debates of philosophers, psychologists and physiologists. In this regard, one of the most prominent frameworks is the ideomotor theory, according to which actions and effects are characterized by a bidirectional connection. That is, the knowledge (*ideo*) about the effect of an action can trigger an action (*motor*) (Carpenter, 1852). Therefore, not only an action elicits a sensory input, but it can also be modified by the elicited input (Greenwald, 1970; Prinz, 1987; Elsner & Hommel, 2001; Hommel, Müsseler, Aschersleben, & Prinz, 2001; Kunde, Koch, & Hoffmann, 2004). The ideo-motor theory has inspired other theoretical models, such as the event coding theory which proposed that the action and its sensory consequence share the same representational code (Prinz, 1990; Hommel et al., 2001), and the two stage model according to which there is a first stage in which an action-effect regularity is detected and a second one in which the repetition of the action-effect contingency allow the emergence of goal-directed actions (Elsner & Hommel, 2001, 2004). However, the ideomotor theory leaves further open questions regarding the buildup of action-effect associations (Stock & Stock, 2004).

Through the experiments on classical conditioning carried out by Pavlov (1927) we can observe the impact of repetitions on associative processing and on behavior (Shanks, 2007). Indeed, this physiologist demonstrated that the repetitive co-occurrence between two stimuli (i.e., metronome sound and food) could cause a change in behavior, demonstrating that not only an association was created between a sound and a rewarding experience, but also that the learning of this association could cause a prediction observable through the dog's salivation, even in the absence of the stimulus typically causing this effect (i.e., the food).

However, as suggested by Shank (2007), it is possible to distinguish between associative and cognitive representations. That is, both associative and cognitive representations are based on prior knowledge, but their distinction consists in the degree of awareness which underlie the action, since our behavior can be elicited by associative (i.e., automatic) and/or cognitive (i.e., rational) thoughts. Overall, the theoretical models about action-effect causality can be roughly grouped into two types: statistical and associative models (Shank, 1993). What distinguishes the two types of models is that the statistical models imply the idea that there are cognitive mechanisms to infer the frequencies of events, whilst for the associative models these processes are exempt from cognitive mechanisms. The degree of awareness that dominate the relation between action and effect still leaves an open question toward a deeper understanding of causal learning. Highly interesting would be to investigate the existence of a continuum in which our thinking is dominated by associative and cognitive processes.

Predictive processing

The establishment of an association between an action and an effect allows to generate predictions (Bar, 2007). Indeed, going back to the previous example, once an association has been established between pressing a button and changing a TV channel, one can also predict what will happen by pressing a different button, based on the action-effect associations previously created.

Predictive coding is one of the most influential theories suggesting that the brain is able to capture the regularities in the surrounding environment, creating internal models which can produce predictions about upcoming events (Friston, 2005). This theory describes the brain as a hierarchical system, in which predictions are formulated as top-down regulation. According to this theory, the main aim of the brain is to minimize "free energy", that is coming from uncertainty, entropy, and surprise (Friston, 2010). Thus, the brain keeps constantly updated its internal models, made up of prior beliefs, through the comparison between the predictions about the causes of the sensory stimuli and the actual flux of information received from the senses. When there is a match between the prediction and the current sensory inputs there is no information to be sent upwards to higher cerebral areas. But, when there is a mismatch between the prediction and the current sensory inputs, the error in the prediction, named "prediction error", is sent up to the higher areas in order to update the internal models. Overall, this model suggests that also the entropy coming from the error is fundamental in order to successfully perceive sensory information, since it allows to update our prior beliefs (Frith, 2013). One of the strengths of predictive coding is the possibility of explaining both brain functioning and malfunctioning in computational terms, allowing to test in quantitative terms this brain function theory (Sterzer et al., 2018). The idea of a hierarchical brain system which modulates the sensory processing through different stages has been proposed also by others (Winkler, Takegata, & Sussman, 2005; Bekinschtein et al., 2009; Diekhof, Biedermann, Ruebsamen, & Gruber, 2009; Wacongne et al., 2011). However, many experiments are still attempting to disentangle the behavioral and neuronal mechanisms underlying the processes described by the predictive coding theory.

The predictive coding principles can also be applied to the specific association that is the focus of this thesis: the pairing of an action (a button press) with its sensory consequence (a sound). In this domain, it has been shown that the processing of the sensory consequences of our actions (i.e., self-generated stimuli) differs from the processing of other (externally-generated) stimulation, and this difference has been attributed to predictive processing of the self-generated stimulation.

For example, Blakemore, Wolpert, and Frith (2000) aimed to test the following dilemma: is it possible to tickle ourselves? The results from their experiment showed that, when the actions are self-generated, the tickling effects are attenuated and the associated brain responses are suppressed (Blakemore, Wolpert, & Frith, 1998). One possible explanation is that the brain predicts the sensory consequences of self-generated actions, cancelling out their effects. Indeed, current theories propose that for each self-generated action there is a copy of this motor-command (i.e., an efference-copy) which is sent to the sensory cortices allowing the estimation of the expected sensory consequences (Wolpert, Ghahramani, & Jordan, 1995). The match between the expected sensory consequences and the current inputs can cause a sensory attenuation (i.e., attenuated neuronal response) (Hughes, Desantis, & Waszak, 2013) and an intentional binding effect (i.e., perceiving a shorter timing between action and effect) (Haggard, Clark, & Kalogeras, 2002). Nevertheless, there is some debate as to whether these effects always reflect predictive processing, since most of the studies contrast sensory consequences elicited by self-generated and externally-generated actions, without specifically differentiating between predictable versus unpredictable stimuli (Hughes et al., 2013).

Associative and predictive processing: fundamental for cognitive function

All in all, different neuroscientific frameworks propose that human beings need to make use of associations and predictions in order to survive (Bar, 2007; Albright, 2012). Some theories even propose that memory and predictive processes could be strongly related, since both of them make use of regularity representations in order to maintain a stable perception of the environment (Bar, 2007; Winkler, Denham, & Nelken, 2009; Albright, 2012). Indeed, going back to the previous example, after having created an action-effect association (e.g., button-channel association), we need to use our memory to recall this association, in order to generate a prediction about the sensory consequences of a button-press, and thus to change TV programme.

Associations and predictions can be very helpful, since they allow to actively interact with the surrounding environment, as long as this processing is flexible enough to be updated based on the different cues coming from multiple sources (Griffin & Fletcher, 2017). Indeed, there are researchers suggesting that dysregulation of these associative and predictive processing could be responsible of clinical phenomena (Ford & Mathalon, 2005), whose roots could be related to different stages of these processing, such as to their generation, verification and updating (Bar, 2007). Indeed, many psychotic symptoms are related with an incorrect representation of the reality (Corlett, Honey, Krystal, & Fletcher, 2011), which brings difficulties in distinguishing reality from imagination, and internal from external events (Griffin & Fletcher, 2017; Sterzer et al., 2018). In line with this proposal, it could be possible to explain negative psychotic symptoms (DSM-V), such as social withdrawal and lack of motivation, as an error related with giving too much weight on memories; and positive psychotic symptoms, such as hallucinations and delusions, as an error related with giving too much weight on a robust, but wrong prediction (Corlett et al., 2011). Nevertheless, it is necessary to examine the complexity of psychosis from multiple levels, such as neurobiological, cognitive, social, individual in order to bridge the gaps and to be able to unravel the mechanisms underlying its functioning (Griffin & Fletcher, 2017; Maia & Frank, 2017).

AUDITORY PREDICTIVE PROCESSING: PARADIGMS AND FINDINGS

The present PhD thesis is a modest attempt to contribute with empirical data to test associative and predictive models, in particular regarding how the associations between an action (i.e., a button press) and an event (i.e., a sound) are created and how they can modify behavioral and brain responses. The present section provides an overview of the relevant empirical studies investigating auditory predictive processing, focusing on the expectation-related effects originated by the human brain's ability in capturing auditory regularities from the environment. An example could help to better illustrate some of the topics investigated in this literature. Imagine staying at a concert hall, where musicians are playing Ravel's Bolero (1928), characterized by several repetitions of the same sequence of sounds from the percussion player, dominating the whole duration of the musical piece (i.e., around 15 minutes). We can hear the sequence of sounds played by the percussion player, or we can be the player. This simple example can help to understand open experimental questions: is it possible to extract an auditory regularity and thus to generate an expectation just by hearing sound sequences? What would differentiate the neuronal response between a person listening to externally-generated sounds (i.e., the attender) and another person self-generating the sounds (i.e., the musician)? In case of differences, are they due to the origin of the sound or to the degree of predictability characterizing self-generated and externally-generated sounds? Numerous empirical studies have attempted to disentangle the answers arising from similar questions.

Two main lines of research have investigated these issues making use of event-related potentials: a line focusing on regularities in the auditory-alone domain (e.g., regularities that can be extracted by the people listening to the concert); and a line focusing on motor-auditory regularities (e.g., those affecting the sound processing for the concert players). In both lines, the studies have described modulations of exogenous ERPs, that is, those generated mainly by the physical characteristics of a sensory inputs (i.e., sounds). Moreover, these studies have also shown modulations in motor ERPs and specific endogenous components related to associative and predictive processing. Indeed, endogenous components are generated by neural processes which are not necessarily associated with the physical input per se, providing information about cognitive processes (e.g., associative and predictive processing). In the following, we describe the main paradigms and findings stemming from this research.

Passive listening paradigms

In order to study the predictive processes that take place when auditory regularities are extracted from the auditory environment, a prolific line of studies has investigated the brain activity of people listening passively to sound sequences. These studies have analyzed the neuronal responses elicited by sounds matching (expected) and by sounds mismatching (unexpected) internal models, and have shown the existence of specific elements and indices characterizing the confirmation and the violation of internal models about auditory regularities.

The studies focusing on the auditory domain and investigating the cases of violation of a regular auditory pattern have been possible through the so-called oddball paradigm. This paradigm employes a rare deviant sound (i.e., mismatching the expectation) which differs from the frequent standard one in at least one physical feature (e.g., pitch, duration, intensity, spatial location) (Näätänen, Gaillard, & Mäntysalo, 1978; Deouell & Bentin, 1998; Näätänen, Paavilainen, Rinne, & Alho, 2007). Typically, these ERP studies examine difference waves, resulting from the subtraction between the auditory ERPs elicited to standard and deviant sounds, to reveal that stimuli mismatching the expectation elicit two endogenous components named, respectively, mismatch negativity (MMN) and P3. The MMN occurs at around 200 ms and indicates the detection of a stimulus deviating from an expectation (Näätänen et al., 2007; Escera & Malmierca, 2014). The P3 occurs at around 300 ms (Escera, Alho, Winkler, & Näätänen, 1998), and indicates the updating of the internal models about auditory regularities (Donchin & Coles, 1988; Polich, 2007). In parallel, there are studies investigating the confirmation of an auditory regularity through a further paradigm named roving standard, in which different sequences of repeated sounds are alternated and characterized by a physical feature. The name of the paradigm is due to the fact that the standard and the deviant sound are variable along the sequences, that is the first sound of each sequence is considered a deviant sound, and the last sound is considered a standard one. These studies reveal that the repetition allows the extraction of a regularity (Cowan, Winkler, Teder, & Näätänen, 1993; Grill-Spector, Henson, & Martin, 2006). Indeed, it has been shown that the repetition of an auditory stimulus elicits a specific endogenous response, termed repetition positivity, RP (Haenschel, Vernon, Dwivedi, Gruzelier, & Baldeweg, 2005; Baldeweg, 2007; Costa-Faidella, Baldeweg, Grimm, & Escera, 2011) which characterizes the neuronal responses to stimuli matching the expectation. A further confirmation about the importance of repetition in fostering an expectation has been reported in neuroimaging studies, showing that the detection of a sound mismatching an expectation activates the superior temporal gyrus, and that this activation was larger for longer sound sequences (Cacciaglia, Costa-Faidella, Żarnowiec, Grimm, & Escera, 2019).

Contingent self-generation paradigms

The effects of predictive processing stemming from motor-auditory regularities have been mostly investigated using the so-called contingent self-generation paradigm. In this paradigm, responses to self-generated (e.g., generating an auditory regularity) auditory stimuli are compared to responses to externally-generated auditory stimuli (e.g., passive listening to an auditory regularity), in order to examine any differences due to the origin of the sound (self-generated vs. externally-generated). Specifically, this paradigm is typically composed of three conditions. First, a motor-auditory condition in which each sensory event is consequent and hence caused by a motor act (i.e., self-generated stimulus). Second, a motor condition in which the action produces no sound, allowing to subtract the motor activity from the response recorded in the motor-auditory condition. Finally, an auditory-alone condition in which the participant only listens to a sound sequence allowing the comparison between the purely auditory self-generated (isolated from motor-auditory condition) and the externally-generated response (recorded in the auditory-alone condition) (Horváth, 2015). The most critical finding from these studies on motor-auditory interactions is that the exogenous neuronal response (e.g., P1, N1, P2 components of the stimulus-evoked response) to self-generated stimuli is attenuated, compared to the response to externally-generated stimuli (Schäfer & Marcus, 1973; Martikainen, Kaneko, & Hari, 2005; Baess, Jacobsen, & Schröger, 2008; Aliu, Houde, & Nagarajan, 2009; Baess, Widmann, Roye, Schröger, & Jacobsen, 2009; Hesse, Nishitani, Fink, Jousmaki, and Hari, 2010; Baess, Horváth, Jacobsen, & Schröger, 2011; Knolle, Schröger, & Kotz, 2012; Sowman, Kuusik, & Johnson, 2012; Horváth, 2013a,b; SanMiguel, Todd, & Schröger, 2013 b; Saupe, Widmann, Trujillo-Barreto, & Schröger, 2013; Timm, SanMiguel, Saupe, & Schröger, 2013; van Elk, Salomon, Kannape, & Blanke, 2014; Horváth, 2015; Mifsud & Whitford, 2017; Neszmélyi & Horváth, 2017). Currently the most accepted explanation is that this sensory attenuation effect reflects predictive processing (Friston, 2005; Loehr, 2013; Kaiser & Schütz-Bosbach, 2018; Bendixen, SanMiguel, & Schröger, 2012). Interestingly, several studies have shown that this sensory attenuation is impaired in schizophrenic patients, suggesting impaired predictive processing in this population (Ford et al., 2001, 2007; Ford & Mathalon, 2005).

Furthermore, when examining responses to self-generated sounds, negative and positive components have been observed in responses to errors and deviants in time-windows similar to the MMN and P3 found using passive listening paradigms (Falkenstein et al., 1990; Gehring et al., 1993; Falkenstein, Hohnsbein, Hoormann, Blanke, 2000; Nittono & Ullsperger, 2000; Nittono, 2004; Overbeek, Nieuwenhuis, & Ridderinkhof, 2005; Nittono, 2006; Waszak & Herwing, 2007; Katahira, Abla, Masuda, & Okanoya, 2008; Band, van Steenbergen, Ridderinkhof, Falkenstein, & Hommel, 2009; Iwanaga & Nittono, 2010), suggesting that similar neural operations take place when auditory-alone and motor-auditory regularities are violated. Moreover, several studies have shown that when the stimulus elicited by an action violated an established expectation and/or elicited an error, this resulted in slowing effects on the subsequent action (Barcelo, Escera, Corral, & Periáñez, 2006; Fogelson et al., 2009; Notebaert et al., 2009; Iwanaga & Nittono, 2010) providing behavioral evidence that predictive processing is taking place, and specifically that the violation of predictions triggers mental operations that delay subsequent actions.

In parallel, several studies employing self-generated sounds have focused rather on investigating associative processing, and aimed to uncover the validity of the ideomotor theory and thus the idea of bidirectional relation between an action and an effect (Elsner & Hommel, 2001; Greenwald, 1970; Hommel et al., 2001; Kunde et al., 2004; Prinz, 1987), investigating how the sensory consequence of an action can modify the action itself. In this regard, it has been shown that participants used less force when the action (e.g., a button press) was associated with a sound than when the same action was associated with no sensory consequences, and when the button press elicited a louder sound compared to a softer sound (Kunde et al., 2004). These differences were detectable even through online adjustments on the action force while pressing the button (Neszmélyi & Horváth, 2017; Horváth et al., 2018).

Action-sound coincidence studies

The studies in the previous section focused on consolidated action-effect associations that lead to predictive processing of the sounds. However a separate line of studies, focusing on motor-auditory interactions, has employed a different type of paradigm, namely the coincidence paradigm (Horváth, Maess, Baess, & Tóth, 2012), in which the series of button presses and the series of sounds are uncorrelated. That is, participants are instructed to repeatedly press a button, which has no particular consequences, and in parallel a series of sounds is played with a pre-established schedule. Despite the lack of contingent associations between the presses and the sounds, random temporal coincidences can occur between button presses and sounds. This type of paradigm was designed to critically evaluate the relation between the phenomenon of sensory attenuation (e.g., N1 attenuation) and predictive processing. Surprisingly, these studies have shown that even when the sound was only coinciding temporally with the button press, in a situation characterized by the absence of predictability, there is an N1 and P2 attenuation, similar to the one reported in studies using contingent self-generation paradigms. These studies opened interesting questions about the nature (e.g., by contingency vs. coincidence) of the action-effect associations and prediction, and how these processes affect the neuronal responses to the auditory stimuli. Indeed, it has been suggested that even the solely action-effect coincidence can lead to a neuronal attenuation (Makeig, Müller, & Rockstroh, 1996; Horváth, 2013 a b, 2014). If this was the case, there would be an open question about the source of the sensory attenuation: is it due to associative and predictive processing or it is unspecifically gated during movement?

Omission studies

In addition to the EEG studies investigating the neuronal responses elicited by the presence of auditory stimuli, a few interesting studies have investigated the endogenous signals elicited in response to the omission of an expected auditory stimulus. This line of research has been carried out focusing both on the auditory domain, and on motor-auditory interactions, and provides specific information about the endogenous response strictly related to the prediction (Bendixen et al., 2009; Wacongne et al., 2011). Indeed, the modulation of neuronal responses observed in the absence of sensory inputs reveals that endogenous processes are taking place. Specifically, the studies focusing on the auditory-alone domain, which employed a passive listening paradigm, have shown that omitting an expected sound elicits an MMN-like response, although only when the interval between successive stimuli is shorter than ca. 150 ms (Tervaniemi, Saarinen, Paavilainen, Danilova, & Näätänen, 1994; Yabe et al., 1998, 2001). In contrast, the studies focusing on motor-auditory interactions, and employing a variation of the self-generation paradigm, reveal that the omission of the self-generated sound elicits very similar responses to the sound-evoked response (e.g., omission N1, N2 and P3) (Raij, McEvoy, Mäkelä, & Hari, 1997; Hughes et al., 2001; Bendixen et al., 2009; SanMiguel, Saupe, & Schröger, 2013 a; SanMiguel, Widmann, Bendixen, Trujillo-Barreto, & Schröger, 2013 c).

SUMMARY

Perception bridges our senses with the surrounding world, and our body is one of the learning sources that we have available to interact with the physical environment. By experience, the brain constructs internal models, made of prior beliefs, which allow us to interpret and to give sense to the flux of information coming simultaneously from different senses. Illusions and cross-modal effects are two of the many windows showing the incessant brain activity in interpreting the inputs from the surroundings. Through goal-oriented actions we can learn to make action-effect associations, but also to predict the sensory consequences of our motor actions. Empirical studies support the impact of associative and predictive processing in modulating behavioral and neuronal responses. Specifically, behavioral studies have shown action-effects associations characterized by bidirectional effects (e.g., modulations of both the sound processing and the action eliciting the sounds) and expectation-mismatch effects on the action timing (e.g., slowing effects). EEG studies have shown the modulation of ERPs both in the case of stimuli matching expectations (e.g., N1, P2 and RP), and in the case of stimuli mismatching expectations (e.g., MMN and P3). However, the specific mechanisms underlying these modulations remains still under debate.

2. OBJECTIVES & HYPOTHESIS

OVERVIEW

Overall, several studies investigating the auditory regularities have been carried out through different research lines and paradigms. These studies have reported the modulation of behavioral and electrophysiological responses, nevertheless the actual phenomena underlying these modulations remain still under debate. Predictive coding (Friston, 2005) attempts to find a unified explanation for the modulations observed at neuronal level (e.g., MMN, P3), suggesting that match- and mismatch-related effects are manifestations of the same phenomenon, since they reflect the amount of prediction error, and hence the degree of match between sensory inputs and expectations. However, there are no studies examining expectation-match and -mismatch effects within the same study. Thus, a direct comparison between expectation-related effects, which would be necessary to validate the explanations provided by predictive coding, has not been reported yet.

Furthermore, most of the studies reporting the modulatory effects of expectation on electrophysiological responses, have focused on stimulus-evoked responses (driven by the presence of an auditory stimulus), but not on endogenous responses (driven by the absence of a sensory stimulus) modulating the sensonsory responses. Thus, the exact processes underlying these modulations are still unclear. A possible solution to investigate these endogenous processes (strictly related to the expectation effects) is given by examining the modulations of the electrophysiological signals in the absence of any sensory input. Some studies have already investigated the cases in which a regular standard sound was omitted (mismatching a sound expectation). However, those studies have not examined the cases in which the regularity was established by an omission (mismatching an omission expectation), thus the exact electrophysiological processes behind the cases of matching an expectation have not been revealed yet.

Finally, to date a large body of studies has investigated expectation-related effects in contexts characterized by stable regularities, focusing only on consolidated effects given by established and regular action-effect associations. However, among the studies focusing on motor-auditory interactions, particularly relevant are the results reporting an attenuation of the electrophysiological responses also in case of absence of predictability and presence of temporal coincidence between action and sound. Thus, it is still necessary to clarify whether the modulation of the electrophysiological indices is due to consolidated predictive processes, and how the motor action contributes in this sensory attenuation. Moreover, there is little knowledge about the different stages characterizing the actual buildup of the action-effect associations. That is, from the initial detection of a novel association to its consolidation, and which role has the repetition in this consolidation, especially in case of absence of a stable regularity.

The general aim of this thesis was to investigate the mechanisms behind associative and predictive processing, focusing specifically on those involved in the processing of action-sound pairings. We were particularly interested in how these processes modulate the behavioral and neuronal responses. We attempted to unify findings and ideas from different paradigms and lines of research, with a special interest in merging ideas stemming from passive and self-generation studies. Indeed, we attempted to relate the different effects to the theoretical framework of predictive coding, and to provide a more unified perspective on the processes of forming associations, discovering regularities, and building and maintaining internal models. Moreover we were interested in how these processes affect our actions and the processing of their sensory consequences. These general aims were explored in two studies, the first one focusing on consolidated predictive processing, and the second one focusing on the buildup of action-effect associations.

In both studies we employed a variation of the contingent self-generation paradigm. In Study I, the design variation allowed us to examine within the same study the cases in which the expectation match- and mismatch-related effects were driven by the stimulus-evoked responses (by the presence of the sound) and by the endogenous responses (by the omission of the sound). Then, we compared a predictable condition, characterized by the presence of consolidated expectation effects, to an unpredictable condition, characterized by their absence. In Study II the design variation allowed us to examine a context characterized by the impossibility of predicting the beginning of a novel regularity. Then, we compared a motor-auditory condition to an auditory-alone condition, in order to isolate the motor contribution in fostering an expectation, and to analyze the entire underlying process and its evolution.

SPECIFIC AIMS & HYPOTHESIS

Study I

In the first study, we employed a variation of the contingent self-generation paradigm in which the events (i.e., sound, omission) generated by a button press in a predictable condition were compared to the corresponding events recorded in an unpredictable control condition. The aim of this experimental design was to be able to investigate, within the same study, all the indices which have been reported so far in the literature as modulated by expectation-related effects (match: RP, N1 & P2 attenuation; mismatch: MMN- and P3-like responses), in order to analyze how they relate to one another. In particular, we aimed to compare match- and mismatch-related effects driven by the presence and by the omission of a sound in order to examine the relationship between these expectation-related effects, under the predictive coding perspective. Additionally, we aimed to isolate the modulations strictly related to the expectation effects, contrasting the endogenous responses (driven by the omission of a sound) elicited in the predictable and unpredictable conditions.

We hypothesized that differences between a predictable and an unpredictable situation should reflect the presence of consolidated expectation-related effects due to the establishment of action-effect associations. These expectation-related effects should be observable through changes both in behavioral (action timing) and electrophysiological (ERPs) measures. Assuming that expectation-match and -mismatch effects are the result of the same process of comparison between sensory inputs and expectations, we hypothesized to find similar temporal dynamics, but opposite polarities when analyzing the relation between match and mismatch expectation-related effects.

Study II

In the second study, we employed a variation of the contingent self-generation paradigm simulating a “broken device”, in which most of the times the button press “did not work” (it generated an omission), but occasionally “started working”, generating a micro-sequence of sounds time-locked to several consecutive button presses. The main aim of this design was to track the buildup of associations between an action and a tone. Thus, we created a dynamic context characterized by the absence of a stable motor-auditory regularity, and we focused on the responses obtained during the micro-sequences, in which a new press-tone association was established. Specifically, we aimed to investigate the evolution of the modulation of the electrophysiological responses (e.g., N1 attenuation), and to uncover the impact of contingency- and coincidence-related effects on them. In parallel, we also aimed to investigate the evolution of bidirectional effects (i.e., how the action itself is modified by its effect) during the buildup of the action-effect association. Finally, we aimed to investigate the specific contribution of the motor action in fostering the auditory expectation. To this aim, we also compared the “broken device” condition to a passive replay of the same sound sequences.

We hypothesized that by studying a dynamic context, we will be able to observe both the buildup and the decay of the internal models related to the action-effect regularities. Assuming that press-omission associations have been established during the long sequence simulating a “broken device”, we hypothesized that the beginning of a new sound micro-sequence should produce coincidence-like effects (i.e., detection of a new press-tone association), and that the repetition of the same press-tone contingency should allow the consolidation of the new internal model. The process of evolution of the action-effect association should be visible both on the modulation of the electrophysiological responses (e.g., N1-attenuation), and on the modulation of the behavioral responses (i.e., action force and action timing). Finally, we hypothesized that the establishment of an action-effect association (as in a motor-auditory condition) can strengthen a prediction of a sound more than by passive listening to an auditory sequence containing sounds (as in a auditory-alone condition), thus changes in the ERPs should reflect the specific contribution of the motor action.

3. METHODOLOGY

Participants

The participants (69% women) in the studies of this thesis were healthy human volunteers ranging from 18-32 years. All participants had normal hearing, and reported no history of neurological or psychiatric disease. Participants gave written informed consent to participate after the nature of the experiments was explained to them. The participation in both studies was compensated by a monetary payment. Both experiments were conducted in accordance to the Declaration of Helsinki and approved by the corresponding ethical local committees.

Stimuli and procedure

In both studies, we used a contingent paradigm in which either the presence or absence of an auditory stimulus was associated with a motor action. Participants were instructed to press a button on a silent device (i.e., not producing any mechanical sound) (Study I) or to pinch a force sensitive resistor (FSR) (Study II). The auditory stimulation was presented through over-ear headphones. In Study I the auditory stimulus was a phonetic syllable “Ta” of 101 ms in duration, and in Study II the tone was a sinusoidal wave of 50 ms in duration. The manipulation of the contingency effects elicited by the action varied based on the conditions of each study. Specifically, in Study I, there was a predictable experimental condition and a non-predictable control condition. In both conditions, participants were instructed to press one of two buttons randomly. In the experimental condition, each button was associated with a high probability of eliciting a sound or an omission, and with occasional violation of these associations. In the control condition, each button was associated with a 50% probability of eliciting a sound or an omission. In Study II, we used a motor-auditory condition (MA), auditory-alone condition (A) and motor-alone condition (M). In the motor-auditory, each motor action had in most of the cases no sensory consequence (i.e., silence/omission of the tone), but occasionally consecutive pinches could trigger a sound micro-sequence of at least five consecutive tones.

Behavioral measurements

The action timing was measured in both studies as the interval between subsequent button presses/pinches (inter-press interval, IPI). Furthermore, in Study II we used a force-sensitive resistor (FSR) in order to collect data regarding the actual force applied to the device. Specifically, the FSR-signal was recorded by using the high level input of a SynAmps2 EEG amplifier (Compumedics NeuroScan), with a sampling rate of 1000 Hz, and online low-pass filtered at 200 Hz. The force applied to each motor act (i.e., a pinch) was determined by the maximal peak in the corresponding FSR signal.

Behavioral data analysis

In order to assess the expectation-related effects on the action timing we computed ANOVAs and post-hoc t-tests. In order to assess the repetition-related effects on the force maxima peaks we computed ANOVAs and Wilcoxon signed-rank tests. Significance was defined for $p \leq 0.05$. The Greenhouse-Geisser correction was applied when the assumption of sphericity was violated.

Electrophysiological measurements

The EEG was continuously acquired at a sampling rate of 500 Hz, from Ag/AgCl scalp electrodes mounted on an elastic nylon cap (Quick-Cap, Compumedics NeuroScan, Charlotte, NC, USA). The signal was amplified by SynAmps (Compumedics NeuroScan) amplifiers and recorded with an online 0.05-100 Hz bandpass filter. The reference electrode was placed on the tip of the nose. In Study I the ground electrode was placed between Fz and FPz, and in Study II in the middle of the forehead. In addition, an electrode was placed on each mastoid (M1, M2), in both studies. The horizontal eye movements were monitored with electrodes placed at the outer canthi of each eye, and the vertical eye movement with electrodes placed above and below the left eye.

The EEG analysis was performed with EEGlab (Delorme & Makeig, 2004), and ERPs were visualized with Eeprobe (ANT). We run an independent component analysis (ICA) (Delorme & Makeig, 2004) aimed to remove eye and muscle artifacts. Thus, the raw EEG data was first filtered with a 0.5 Hz high pass, then we manually rejected the sections containing non-stereotypical artefacts from the continuous file. Subsequently, we applied a binary compiled version of runica (binica), that uses the logistic infomax ICA algorithm (Onton & Makeig, 2006). The extracted ICA weights were then applied to the original unfiltered raw datasets. Components related to eye movements and muscle activity were pruned. On average, in Study I we rejected eight components (range: 5 - 11) per participant, and in Study II we rejected nine components (range: 6 - 13) per participant. After the ICA correction, the EEG was bandpass filtered from 0.5 to 30 Hz (windowed sinc FIR filter, Kaiser window, Kaiser beta 4.53351, filter order 734). Subsequently, in Study I we defined epochs of 700 ms around each button press (-200 to +500 ms), and in Study II we defined epochs of 600 ms

(-200 to 400 ms), time-locked to the FSR pinch-threshold. Only in Study II we applied a baseline correction of 200 ms (-200 to 0 ms).

Electrophysiological data analysis

In both studies, we extracted event-related potentials (ERPs), time-locked to each action, separately for sound and omission stimuli. In Study I we were mainly interested in the modulation of consolidated predictive processing on exogenous (i.e., P1, N1 and P2) and on endogenous components (i.e., RP, MMN and P3), both in the case of sounds and omissions. In Study II we were mainly interested in the buildup of the action-effect associative processing in the N1, P2/MMN and P3 time windows, and we focused on the tone responses. Subsequently, we computed ANOVAs and t-tests per each comparison of interest for each study. Significance was defined for $p \leq 0.05$. The Greenhouse-Geisser correction was applied when the assumption of sphericity was violated.

4. RESULTS

In this chapter, we present the two studies which are in the process of being submitted for publication.

STUDY I

The first study of this thesis is entitled “Expecting sounds and silences: Disentangling match- and mismatch-related effects on auditory ERPs”, by Vittoria Spinosa and Iria SanMiguel.

Introduction

As humans, we are steadily surrounded by enormous amounts of sensory information. In order to deal with this plethora of information while keeping energy expenditure at a minimum, it is necessary to extract regularities from repetitive patterns coming from the surroundings and to create internal models about the potential causes of our sensory experiences. For instance, when attending a concert, we have learnt that at the end of the musical performance, the audience will applaud the musicians, so we can decide whether to join this common action or not. The particular sound elicited by clapping the hands is an example of auditory regularity that we can extract from the environment. Based on this regularity, we create an internal model of the auditory consequence elicited by the clapping of the hands (auditory regularity extracted from motor-auditory interactions), or by hearing the clapping (auditory regularity extracted from the passive listening). Based on these internal models, humans are also able to detect when these regularities are confirmed, or violated.

Currently, a large body of research has studied how the expectation arising from internal models about auditory regularities can modulate the electrophysiological responses to the sounds, particularly through the recording of event-related brain potentials (ERPs), both when the expectation is confirmed (match), and when it is violated (mismatch). The auditory regularities have been examined through different experimental paradigms, which we can mainly group in paradigms focusing on the auditory-alone domain, and paradigms focusing on motor-auditory interactions.

In studies examining the expectation-related effects stemmed from the auditory-alone domain using passive listening paradigms, several endogenous components related to the violation of an established expectation (i.e., mismatch due to unexpected sound) have been identified. Of these, the most well-known are the so-called mismatch-negativity (MMN/N2) and the P3-family responses (including the P3a, the novelty P3 and the P3b). The MMN occurs at around 150 ms after the onset of a sound deviating from an established regularity (i.e., "deviant" sound) and it is elicited both when the subject is attending and not attending to the auditory stimulation (Näätänen et al., 2007; Escera & Malmierca, 2014). The MMN component has been suggested to reflect the need of adjusting the confidence in our internal model of regularities (Winkler et al., 2009). The P3a, in turn, occurs at around 300 ms after the presentation of a deviant stimulus and it is characterized by a frontocentral positivity (Escera et al., 1998, Escera et al., 2000; Schröger et al., 2000; Friedman et al., 2001; Polich, 2007; Joos et al., 2014). Since the deviant sounds (i.e., the unexpected stimuli) are mismatching the internal models (Parmentier et al., 2011), some researchers suggested that P3 responses indicate the need of finding a solution to the uncertainty of unexpected events (Sutton et al., 1965) and of updating the context and internal models (Donchin & Coles, 1988; Polich, 2007).

Although deviance processing effects have been originally investigated with passive listening paradigms, similar effects have been observed also when the expectation stemmed from a motor-auditory interaction and not from the auditory domain exclusively. These studies have used mainly the so-called contingent self-generation paradigm (Horváth, 2015), in which each motor act (e.g., a button press) is contingent with and causes a specific sensory consequence (e.g., a sound or a feedback). Specifically, the studies investigating motor-auditory interactions have found a negative component occurring at around 80 ms (Falkenstein et al., 1990; Gehring et al., 1993; Falkenstein et al., 2000), and a positive component occurring at around 200-500 ms (Falkenstein et al., 2000; Nittono & Ullsperger, 2000; Nittono, 2004; Overbeek et al., 2005; Nittono, 2006; Waszak & Herwing, 2007; Katahira et al., 2008; Band et al., 2009; Iwanaga & Nittono, 2010) in response to negative feedback which is akin to an unexpected effect due to a deviant stimulus. It has been proposed that the negative component is related to the MMN/N2 family and the positive component to the P3 response, typically found in passive listening studies (Folstein & Van Petten, 2008), and further that these components are associated with the degree of expectation and not only with the valence of the feedback (Holroyd & Coles, 2002). At behavioral level, it has been shown that both unexpected events and errors tend to slow down a subsequent motor action (Barcelo et al., 2006; Fogelson et al., 2009; Notebaert et al., 2009; Iwanaga & Nittono, 2010), and it has been suggested that this is probably due to the infrequency of their presentation which captures the attention of participants for these events (Notebaert et al., 2009).

In a further approach to investigate the violation of an expectation with ERPs, several studies have also shown that when an expected sound is omitted (i.e., mismatch due to an unexpected omission), despite the absence of any physical stimulation, an ERP response is obtained time-locked to the omission, which mimics in its morphology, time-course, and brain sources the response evoked by the actual expected sound (Raij et al., 1997; Hughes at al., 2001; Bendixen et al., 2009; Wacongne et al., 2011; SanMiguel et al., 2013 a, c). Furthermore, using a self-generation paradigm, it has been shown that omitted sounds result in an N1-like, N2/MMN-like and P3-like omission response (SanMiguel et al., 2013 b). However, in studies using passive listening paradigms the N2/MMN-like response to an omitted stimulus is elicited only when the interval between stimuli is shorter than around 150 ms, suggesting that also the timing between stimuli plays an important role in generating this endogenous response (Tervaniemi et al., 1994; Yabe et al., 1998, 2001).

On the other hand, the ERP literature investigating the cases in which the occurrence of a particular stimulus confirms an established expectation (i.e., match due to an expected sound) has shown the modulation of several exogenous auditory ERP components, such as P1/P50, N1 and P2 (Schäfer & Marcus, 1973; Lü et al., 1992; Boutros et al., 1995; Martikainen et al., 2005; Baess et al., 2008; Aliu et al., 2009; Baess et al., 2009; Lijffijt et al., 2009; Hesse et al., 2010; Baess et al., 2011; Lasaponara et al., 2011; Knolle et al., 2013; Sowman et al., 2012; Horváth, 2013 a, b; SanMiguel et al., 2013 b; Saupe et al., 2013; Timm et al., 2013, 2014; van Elk et al., 2014; Horváth, 2015; Mifsud & Whitford, 2017; Neszmélyi & Horváth, 2017), and the attenuation of the auditory N1 subcomponent, Tb (SanMiguel et al., 2013 b; Saupe et al., 2013). Of these studies, those focusing on motor-auditory interactions have shown an attenuation of sensory responses elicited by the self-generated sensory consequences of the motor act (Blakemore et al., 2000; Hughes et al., 2001). Yet, the actual nature of these effects is disputed, as some identified the N1 attenuation as a reduction not only due to the self-generation per se, but also to the context, to the predictability and to attentional levels (Loehr, 2013; Kaiser & Schütz-Bosbach, 2018).

A further research line has focused on the effects of stimulus repetition as confirmation of the internal model about regularities, mainly in passive listening paradigms. In these cases, the occurrence of a stimulus that matches the expected sound results in a particular endogenous

ERP component, termed as the repetition-positivity (RP), which increases with the number of repetitions of the stimulus. Specifically, the RP is seen as the modulation of different ERPs generated along a wide temporal span, indeed it tends to develop in the P1 time-window and to extend until the N1 and P2 time windows (Haenschel et al., 2005; Baldeweg, 2007; Costa-Faidella et al., 2011; Recasens et al., 2015). These studies examined only the cases in which the presence of a sensory stimulus was expected, but there is a lack of studies examining the cases in which the absence of the sensory stimulus is expected (i.e., match due to an expected omission). Specifically, the study of the purely endogenous signals in response to the omissions could help to unravel the expectation-related effects which are supposedly causing the modulation of the ERPs (Bendixen et a., 2009; Wacongne et al., 2011).

In summary, across different studies and experimental paradigms there is a general trend to find increased negativities in the 100-250 ms time-window (e.g., N1, MMN, N2, action-effect negativity), and increased positivities (e.g., P3a) in subsequent time-windows (ca. 300 ms) for events mismatching expectation, whereas there is a general trend to find increased positivities for input matching expectation (e.g., attenuated N1, RP). Overall, these findings suggest that the electrophysiological responses are modulated by the expectation-match and -mismatch effects (Bendixen et al., 2012), however to the best of our knowledge, these ERPs have always been studied in separated time-windows and there is a lack of studies examining, within the same experimental design, the expectation-match and -mismatch effects, both driven by the presence of the sensory input (i.e., stimulus-evoked response) and by the omission of the same sensory input (i.e., endogenous response). Furthermore, since most of the study designs used so far did not allow to directly compare the expectation-match and

-mismatch effects to each other, the relation between these modulations on the ERPs is still unclear.

The “predictive coding theory” is one of the major influential perspectives describing how the brain makes use of internal models (Friston, 2010). Predictive coding proposes that the brain uses the prior beliefs to generate predictions about the causes of our sensory experiences, in order to better select our actions and to minimize the level of uncertainty and surprise coming from unexpected events. Specifically, this theory views the brain as a hierarchical system and proposes that the prediction is instantiated as a top-down modulation. The difference between the prediction and the actual sensory input generates a prediction error (PE), which plays a central role in keeping updated the internal models. Indeed, only the brain signals related to the prediction error are transmitted upwards in the sensory hierarchy.

Predictive coding provides also a unifying explanation to the modulation of the expectation-related effects on the ERPs (e.g., MMN and P3) (Friston, 2005). Indeed, according to predictive coding, this modulation makes the neuronal response less excitable and then suppressed in the case of expected stimuli (match-related effect), and more excitable and then enhanced in the case of unexpected stimuli (mismatch-related effect). Thus, this theory accounts match- and mismatch-related effects as two sides of the same coin: the neuronal response depends on the degree of match (match vs. mismatch) between the current sensory input and the prediction.

The present study focused on the motor-auditory interaction, in order to uncover whether expectation-match and -mismatch effects arising from a motor act are the manifestations of the same underlying process, as proposed by predictive coding (Friston, 2005). Thus, we used a contingent paradigm, in which each motor action (i.e., a button press) was associated with a specific probability of eliciting an event (i.e., sound or silence). Specifically, in the experimental condition, the action was associated either with a sound or with an omission of the sound (expected events), and occasionally with violations of these associations (unexpected events). Then, we compared these predictable responses recorded in the experimental condition, to non-predictable responses recorded in the control condition. We first examined the ERPs typically reported in literature as affected by expectation-match and -mismatch effects (hypothesis-driven analysis), and then we investigated how the different expectation effects relate to one other (data-driven analysis). Specifically, we studied both the case of presence of a sound, and of omission of a sound, in order to understand whether the ERPs only reflect the prediction error signals, or whether the top-down modulatory activity can also be observed by itself on the ERPs, in the absence of the input (i.e., endogenous signals). Based on previous studies, at behavioral level, we hypothesized to observe slowing effects after an event mismatching with the expectation. At electrophysiological level, from the analysis on the ERP components previously reported in literature as affected by expectation-related effects (hypothesis-driven analysis), we hypothesized to find the attenuation of the exogenous components (P1, N1, P2) and the presence of the RP in case of expectation-match, and the presence of the N2/MMN-like and P3 responses in case of expectation-mismatch, both in the case of sounds and omissions. Furthermore, from the study on the endogenous signals in response to the omission (i.e., absence of inputs) of the sound events, we hypothesized that the presence of differences in the ERPs recorded in the experimental (predictable) and control (non-predictable) condition could reflect the top-down activity and/or the prediction error triggered by expectation-match and -mismatch effects on the omission responses. Assuming that the modulations on the ERPs reflect the same underlying process (Friston, 2005), studying the relationships between expectation-match and -mismatch effects (data-driven analysis), we hypothesized that they present similar time-courses and scalp distributions, albeit with different polarities, since the subtraction between the presence or absence of the sensory input (sound vs. omission) minus the expectation should result in less prediction error (in the case of expectation-match), or more prediction error (in the case of expectation-mismatch), compared to a non-predictable control condition.

Materials and Methods

Participants

The present study was conducted in accordance with the Declaration of Helsinki and approved by the Bioethics Committee of the University of Barcelona (IRB00003099). All participants gave written informed consent for their participation after the nature of the study was explained to them. Thirty-two healthy adult volunteers were enrolled. Data from seven participants were excluded both from the behavioral and EEG analysis, due to excessive artefacts and/or button-press timing errors resulting in high rejection rates (> 40% of trials). Thus, the final sample consisted of 25 participants (15 women, 10 men, 3 left-handed) ranging in age 18–32 years (mean = 22.9 years). All participants had normal or corrected-to-normal vision, and reported no hearing impairment or history of psychiatric or neurological disease. They were compensated by a monetary payment of 25€.

Stimuli and procedure

The experimental task was delivered with Psychtoolbox-3 (Kleiner et al., 2007) running on Matlab. Participants sat comfortably inside an electrically shielded chamber and were instructed to fixate their gaze on a fixation cross displayed on a screen placed at a distance of approximately 100 cm from their eyes. They were also instructed to use two buttons (A-B) of a wireless numeric touchpad (Glassedpad – Typhoon, Schalksmühle, Germany) with the index and the medium finger of their dominant hand. The participant's task was to randomly select and press either one of the two buttons on every trial. Participants were asked to balance the presses of the two buttons across trials while avoiding to follow any obvious repeating sequence (e.g., A-B-A-B-A-B or A-A-B-B-A-A), and to allow approximately 800 ms between two consecutive presses.

The experimental design (Figure 1) consisted of an experimental condition (E) and a control condition (C), performed in different blocks. In the experimental condition, each button was associated with a different pattern of sensory consequences. Specifically, pressing button A produced a sound in 88% of the presses (expected sound, ES) and no sound in the remaining 12% of the presses (unexpected omission, UO). Hence, button A was associated with sound expectation. Conversely, pressing button B produced no sound in 88% of the presses (expected omission, EO) and the sound in 12% of the presses (unexpected sound, US), and this button B was associated with the omission expectation. The assignment of buttons A and B to the middle and index fingers was counterbalanced across participants. In the control condition, the auditory stimulation was not determined by the button pressed, but it was an exact replay of the auditory sequence generated in the experimental block. That is, each button press in the control condition triggered the next event (either sound or omission) of the sequence generated in the experimental block. Thus, the events generated in the control condition were not necessarily generated by the same button used in the experimental condition. Furthermore, in the experimental condition participants produced an equal distribution of presses of each button, and the probabilities of sound generation were crossed for the two buttons. Therefore, in the control condition, the button presses randomly produced the sound in 50% of the cases (control sound, CS), and the omission of the sound in the remaining 50% (control omission, CO). Consequently, the experimental condition was a predictable situation and the control condition was an unpredictable situation. In total, there were 1056 expected events and 144 unexpected events per each button in the experimental condition, and an equal number of corresponding events in the control condition. Expected and unexpected events were pseudorandomly placed with the only restrictions that there were always at least two expected event types before each unexpected event type, and in every block the first five presses of each button only included expected event types.

The auditory stimulus was the phonetic syllable ‘Ta’ presented with a duration of 101 ms (with voice onset time at 20 ms) and a pitch of 167 Hz. The stimulus was created with Klatt speech synthesizer (Klatt, 1980) and Matlab software (Matlab R2007a, MathWorks) based on parameters used in previous studies (Steinschneider et al., 2003; Partenen et al., 2011) and adjusted for Spanish language speakers. Sounds were presented through over-ear noise-cancelling headphones (Sennheiser KD 380 PRO, Marlow, UK) in order to silence any external noise. Sounds were presented at a fixed intensity of 40 dB above the individual threshold that was measured for each participant through an audiometric test.

The start of the experiment was preceded by a one minute training block in order to adjust to the requested timing between button presses (inter-press interval, IPI). Presses separated by less than 600 ms or more than 1200 ms were considered timing errors. During training, visual feedback on the timing between button presses was presented on each trial. When the optimal performance was reached and the participant felt able of keeping the required pace without feedback, the first experimental block could start. Each experimental block was followed by a control block. Every block lasted approximately 4 minutes. Participants were allowed short resting breaks between blocks. In total, 8 blocks per condition were presented. Blocks were repeated if more than 15 timing errors were committed, or the unbalance of presses between the two buttons exceeded 40-60%. Total experimental time excluding breaks and preparation was around one hour.

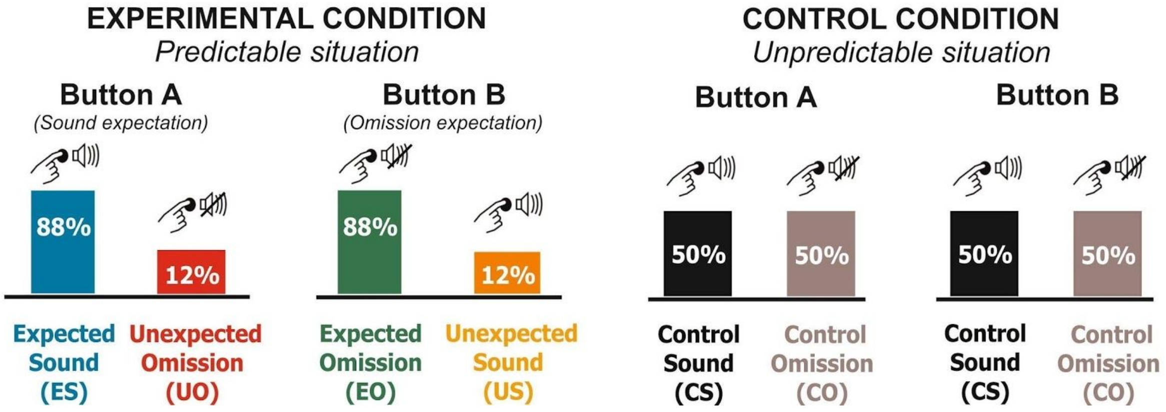

Figure 1*. Schematic representation of the experimental design. In the experimental condition, button A produced most of the times a sound and hence it was associated with a sound expectation, and button B produced most of the times an omission (no sound), thus being associated with an omission expectation. In the control condition, there were no specific expectations associated to any button.*

Behavioral analysis

At behavioral level, we mainly aimed to investigate the expectation-related effects on button selection and on inter-press intervals (IPI). Specifically, these behavioral analyses were necessary to exclude the existence of confounding effects, which could impede the correct comparisons between the experimental and the control condition.

Thus, we intended to assess the distribution of button presses (button A vs. button B), in order to make sure that participants were balancing the presses among the two buttons, as requested in the instructions. To this aim, we tested for any difference in the percentage of presses of each button in the two conditions computing a repeated measures ANOVA with the factors Condition (E, C) and Button (A, B). Subsequently, we also intended to assess whether participants were more prompt into particular button press sequences, in order to avoid any bias on the results due to button repetitions (R; e.g., A-A or B-B in consecutive trials) and button alternations (A; e.g., A-B or B-A in consecutive trials). Thus, we computed a repeated measures ANOVA with the factors Condition (E, C) and Button Sequence (R, A).

Then, we intended to assess whether the expectation-related effects affected the button press timing (inter-press interval, IPI), in order to exclude any difference in the action timing between the two conditions. That is, we assessed whether the expectation of producing a sound or an omission affected the time taken to press the buttons, considering that only in the experimental condition each button was associated with a particular expectation: sound (button A) or omission (button B). To this aim, we computed a repeated measures ANOVA with the factors Condition (E, C) and Button (A, B). Additionally, we wanted to analyze whether the nature of the preceding event (sound vs. omission; match vs. mismatch) had an impact on the timing of the subsequent press, in order to avoid confounding effects on the

IPI. Furthermore, we wanted to replicate the presence of slowing effects after stimuli mismatching the expectation (Fogelson et al., 2009; Notebaert et al., 2009; Iwanaga & Nittono, 2010), as proof that expectation effects were formed and affected the behavior only in the experimental condition. To do so, we took into account on the one hand whether the preceding button press produced a sound or an omission, and on the other hand whether the preceding button press was matching or mismatching with the button expectation in the experimental condition. Note that in the control condition the buttons are not associated with any strong expectation, thus there we simply took the corresponding events following those that matched or mismatched the expectation in the experimental condition. We computed two separate repeated measure ANOVAs: 1) Condition (E, C) times Preceding Event (sound, omission); 2) Condition (E, C) times Preceding Event (match, mismatch). Finally, we wanted to test whether the press sequence influenced the timing between button presses, in order to avoid any bias due to repetitions and alternations on IPI. To this aim we considered whether the current press was a button repetition or a button alternation, and analyzed the timing between the two presses as dependent measure. Thus, we ran a repeated measures ANOVA Condition (E, C) times Button Sequence (repetition, alternation). All the interactions were followed up by post-hoc t-tests.

EEG acquisition

The EEG was continuously acquired at a sampling rate of 500 Hz from 64 Ag/AgCl scalp electrodes mounted on an elastic nylon cap (Quick-Cap, Compumedics NeuroScan, Charlotte, NC, USA) according to the 5% electrode placement system (Oostenveld & Praamstra, 2001). The signal was amplified by SynAmps RT amplifiers and recorded with

Scan 4.4 (Compumedics), applying an online 0.05-100 Hz bandpass filter. An electrode placed on the tip of the nose served as reference, and the ground electrode was placed between Fz and FPz. In addition, an electrode was placed on each mastoid (M1, M2). Eye movements were monitored with electrodes placed above and below the left eye (vertical electrooculogram) and at the outer canthi of each eye (horizontal electrooculogram). During the EEG recording, all electrodes impedances were kept below 10 k**Ω**.

EEG processing

EEG analysis was performed with EEGlab (Delorme & Makeig, 2004), and ERPs were visualized with Eeprobe (ANT). In order to run an independent component analysis (ICA) (Delorme & Makeig, 2004), the raw EEG data was first filtered with a 0.5 Hz high pass (windowed sinc FIR filter, Kaiser window, Kaiser beta 5.653, filter order 1812), and sections containing non-stereotypical artefacts were manually rejected from the continuous file. Subsequently, we applied a binary compiled version of runica (binica) using the logistic infomax ICA algorithm (Onton & Makeig, 2006). The extracted ICA weights were then applied to the original unfiltered raw datasets and components related to eye movements and muscle activity were pruned. On average 8 components (range: 5 - 11) were rejected per participant. After ICA correction, the EEG was bandpass filtered from 0.5 to 30 Hz (windowed sinc FIR filter, Kaiser window, Kaiser beta 5.653, filter order 1812). Subsequently, epochs of 700 ms were obtained around each button press (-200 to +500 ms). No baseline correction was applied. Data from broken electrodes were interpolated from surrounding electrodes using spherical interpolation. Epochs still containing other large artefacts were rejected by applying a 75 µV maximal signal change per epoch threshold.

Epochs including button presses separated by less than 600 ms from either the previous or the subsequent button press, and those separated by more than 1200 ms from the previous button press were excluded from the ERP analysis. The first five presses of each button for each block and all the trials immediately following an unexpected event were also excluded from the ERP analysis.

ERP Analysis

We were mainly interested in investigating the presence of the expectation effects on the waveforms elicited by specific events of interest (Figure 1). Thus, we contrasted the waveforms recorded in the experimental condition to those recorded in the control condition. These comparisons allowed us to focus on four possible expectation effects: 1) the effects of matching a sound expectation (ES-CS); 2) the effects of mismatching a sound expectation (UO-CO); 3) the effects of matching an omission expectation (EO-CO); 4) the effects of mismatching an omission expectation (US-CS).

Analysis based on components of interest (hypothesis-driven)

ERP literature reports the existence of exogenous and endogenous ERP components which are modulated by expectation-match and -mismatch effects. Specifically, the effects of matching an expectation have been indicated mainly by the attenuation of the exogenous components P1, N1 and P2, and by the presence of the endogenous component termed RP. The effects of mismatching an expectation have been indicated mainly by the presence of N2/MMN-like and P3 responses. Typically these expectation-related effects have only been investigated separately, in different studies; however, we hypothesize that all these effects are intimately related and might reflect the same underlying process. Thus we investigated, within the same study, all the exogenous and endogenous ERP components which have been typically reported in the literature as affected by expectation-related effects, in order to test which of the expectation-match and expectation-mismatch effects included in our design contribute to the modulation of each of the ERP components of interest. To do so, first we identified the time-windows pertaining to each component of interest following the typical procedures for each component (see Table 1, Figure 2). Subsequently, the mean amplitudes measured in the identified time-windows, and on the electrodes where each of the components of interest are typically more prominent, were contrasted between the experimental and control conditions in 4 double sided t-tests, one per each of the comparisons described above: ES-CS, UO-CO, EO-CO and US-CS (see Table 1).

Table 1. *ERP components of interest with detailed information regarding the difference waveforms used for component identification, electrodes and time-windows for analyses.*

ERP components	**Difference wave for Identification**	**Electrodes for analysis (ROI)**	**Time windows for analyses (ms.)**
P1	CS - CO	Cz	52 ± 10
Na	CS - CO	T7, T8	106 ± 10
N1b	CS - CO	Cz	116 ± 10
N1bT	CS - CO	M1, M2	118 ± 10
oN1	UO - CO	FT7, FC5, T7, C5 FT8, FC6, T8, C6	120 ± 20
Tb	CS - CO	T7, T8	160 ± 10
P2	CS - CO	Cz	184 ± 20
oN2	UO - CO	Fz, FCz, Cz	188 ± 10
N2/MMN	US - CS	Fz, Cz	190 ± 10
RP	ES - CS	Cz, C3, C4	178 ± 30
P3	US - CS	Fz, Cz, Pz	294 ± 40
oP3	UO - CO	Fz, FCz, Cz	380 ± 40

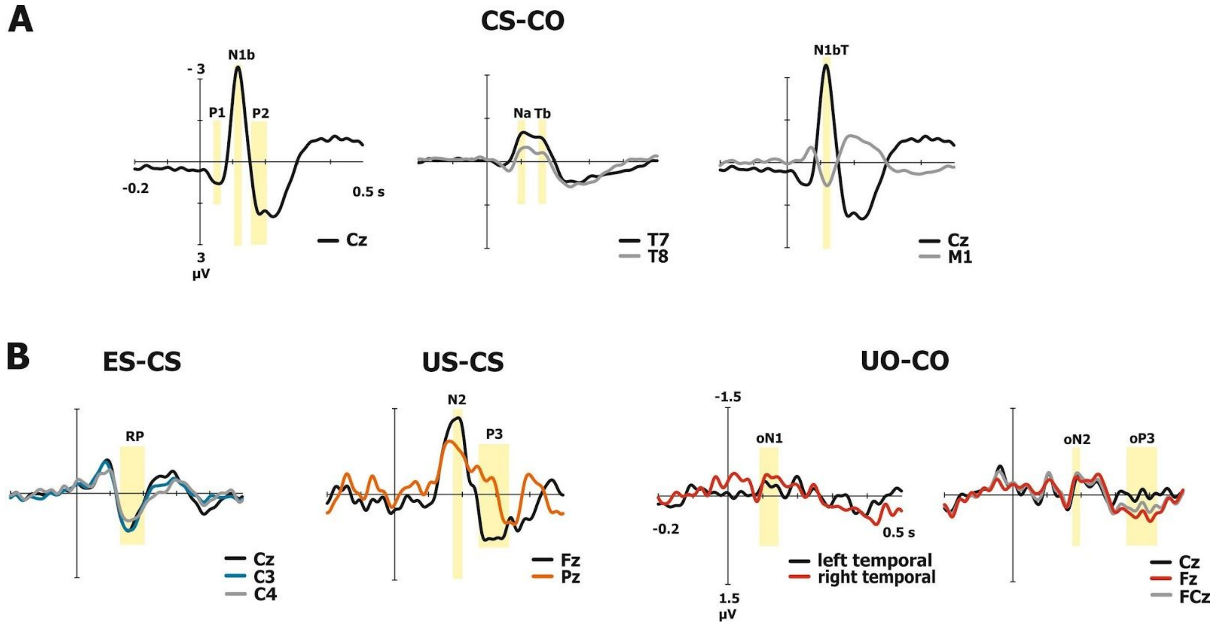

***Figure 2**. Difference waveforms used for component identification. Each panel illustrates the electrodes and time windows used for each component. **Panel A.** Exogenous auditory ERPs (P1, N1b, P2, Na, Tb, N1bT) identified on the Control Sound (CS) - Control Omission (CO) difference waveform, so that the motor response was eliminated. **Panel B.** Endogenous ERPs. Omission related components (oN1, oN2, oP3) were identified on the UO-CO difference waveform; deviance related components identified on the US-CS difference waveform; repetition positivity (RP) identified on the ES-CS difference waveform.*

For exogenous ERPs we focused on the auditory P1 and P2 components (Knolle et al., 2012; Horváth & Burgyán, 2013; SanMiguel et al., 2013 b; Saupe et al., 2013; van Elk et al., 2014), as well as in the auditory N1 subcomponents: Na, N1b, N1bT, and Tb (see SanMiguel et al., 2013b; Joos et al., 2014). Since our recorded auditory ERPs in the experimental condition were elicited in response to a motor command, in order to isolate and visualize the auditory response for clear and non-contaminated component identification, we eliminated the motor activity from the auditory response by subtracting the omission ERP in the control condition (CO) from the sound ERP in the control condition (CS) (Figure 2, panel A). On this waveform reflecting only the auditory response, the P1 peak was identified as the largest peak visible at Cz in the 0-100 ms range. N1 subcomponents Na and Tb were identified at T7 and T8 as the first and second largest negative peaks, respectively. The tangential component of the auditory N1 (N1bT) was identified as the positive peak occurring in the N1 latency range (70 to 150 ms) at mastoid electrodes displaying polarity reversal at Cz. Finally, the auditory P2 was identified at Cz as the largest positive peak occurring immediately after the N1.

For endogenous components elicited when the expectation driven by the motor-auditory association was either matched or mismatched (Figure 2, panel B), we identified the following ERP components. First, regarding match-related responses we isolated the repetition positivity (RP, Haenschel et al., 2005; Costa-Faidella et al., 2011). Typically, the RP is identified subtracting the activity elicited by an initial standard stimulus from that elicited by a final standard stimulus of the same sequences. As this component is assumed to reflect a progressive increase of the expectation or the expectation-match effect, here, to identify the RP, we subtracted the sound ERP in the control condition (CS) from the expected sound ERP in the experimental condition (ES). The RP peak was identified as the largest positive response on the midline central ROI (i.e., at Cz, C3, C4) in the 50-250 ms range. In the second place, two mismatch-related responses were isolated, namely the N2/MMN and the P3 (Escera & Corral, 2007; Folstein & Van Petten, 2008; Escera & Malmierca, 2014). Typically, these components are isolated in passive oddball paradigms by computing the difference waveform between the deviant (unexpected) sound and either the standard (expected) sound or control sound (a sound presented in a block without any auditory regularities). Here, these components were isolated by subtracting the sound ERP in the control condition (CS) from the unexpected sound ERP in the experimental condition (US). The MMN/N2 peak was identified as the largest negative response on the frontocentral midline ROI (i.e., Fz, Cz) in the 90-250 ms range. The P3 peak was identified as the largest positive response on the frontocentral midline ROI (i.e., Fz, FCz, Cz), in the 200-500 ms range. Mismatch-related responses were also retrieved from three different omission-related components, namely oN1, oN2, oP3 as defined in SanMiguel et al. (2013 a). Typically, these omission responses are isolated by comparing the omission of self-generated sounds to the equivalent silent button-press in a motor-alone control condition. In the present study, however, since we did not have a motor-alone condition, the omission responses were identified by subtracting the omission ERP in the control condition (CO), where no expectation was possible, from the unexpected omission ERP in the experimental condition (UO). The oN1 peak was identified as the first negative response between 0 and 100 ms over frontotemporal scalp locations (i.e., FT7, FC5, T7, C5, FC6, FT8, C6, T8); the oN2 peak as the second negative response, between 100 and 250 ms, maximal over the frontocentral midline (i.e., Fz, FCz, Cz); and the oP3 peak as a broadly distributed positive deflection occurring between 200 and 400 ms over the frontocentral midline (i.e., Fz, FCz, Cz).

Analysis based on sustained modulations (data-driven)

The analysis based on the components of interest suggested the existence of sustained expectation effects overlapping several of the well-known components, rather than phasic modulations of the individual components. Therefore, we defined larger time-windows which capture the sustained expectation-related effects of interest. That is, the sustained modulations visible on the four difference waveforms resulting from the four main comparisons (ES-CS, UO-CO, EO-CO, and US-CS).

Thus, we selected and analyzed four wider time-windows, in order to study the time-course of the expectation-related effects. The first time-window encompassed the P1, Na, N1b, N1bT and oN1 components (30-126 ms); the second time-window the Tb, P2, oN2, N2/MMN and RP components (134-208 ms); the third time-window covered the P3 component (206-330 ms); and the fourth time-window included the oP3 component (350-410 ms). To analyze these sustained modulations, a repeated measures ANOVA with the factors Expectation (Sound, Omission) times Match (Match, Mismatch) times Region of Interest (i.e., electrodes) was carried out on each of the four time-windows. The ROIs were focused on the areas that, by visual inspection, were showing the most prominent effects in each of the time-windows. Specifically, in the first and second time window the analysis included three fronto-central regions of interest (left: F7, F5, F3, FT7, FC5, FC3; center: F1, Fz, F2, FC1, FCz, FC2; right: F4, F6, F8, FC4, FC6, FT8); in the third time window one frontal and one central region were included (frontal: F3, Fz, F4; central: C3, Cz, C4) and in the fourth time window one frontal region was included (frontal: F3, Fz, F4).

Time-window preceding the button press (data-driven)

Additionally, we decided to explore the ERP time-region preceding the button press (from -0.2 to 0 s), in order to investigate if any expectation-related effect was present before the elicitation of the subsequent event (i.e., sound, omission). Thus, we firstly created an ERP for button A and button B, for each condition, averaging: in the Experimental condition 1)

Expected Sounds with Unexpected Omissions for studying Button A (preEA); 2) Expected Omissions with Unexpected Sounds for studying Button B (preEB); in the Control condition, 3) all the events elicited by button A (preCA); 4) all the events elicited by button B (preCB). Then, we computed an ANOVA with the factors Condition (Experimental, Control) and Button (A, B).

Results

Behavioral Results

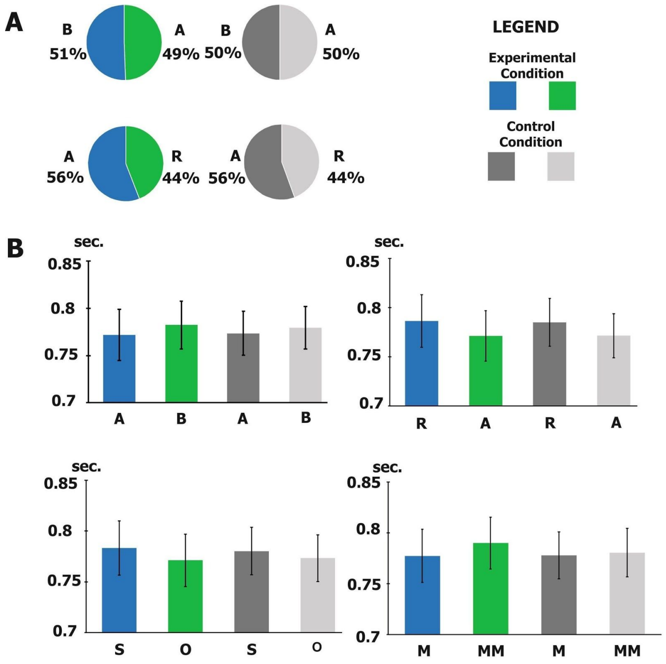

Figure 3*. Behavioral results.* ***Panel A****. Button press distribution. Top: distribution of presses of button A vs. B. Bottom: distribution of alternations (A) and repetitions (R).* ***Panel B****. Inter-press intervals (IPI). Top left: intervals preceding the press of button A or button B. Top right: intervals preceding a repetition (R) or an alternation (A). Bottom left: intervals*

after a sound (S) or an omission (O). Bottom right: intervals after a match (M) or a mismatch (MM).

Participants pressed button A (Sound expectation) on 49.33 (*M*) ± 1.82% (*SD*) of the trials in the Experimental condition (Figure 3, Panel A, Top left), and on 49.98 ± 1.40% of the trials in the Control condition (Figure 3, Panel A, Top right). The results from the ANOVA testing the influence of the expectation effects on the distribution of presses did not show any significant difference between buttons ($F_{(1,24)} = 1.362$, $p = .255$, $\eta p^2 = .054$). However, participants were more slightly biased towards button B in the E condition compared to the C condition ($F_{(1,24)} = 5.312$, $p = .03$, $\eta p^2 = .181$) (Figure 3, Panel A, Top). Moreover, participants displayed a significant bias towards producing more alternations (A) (E: 55.9 ± 13.2 %; C: 55.6 ± 12.4%) than repetitions (R) ($F_{(1,24)} = 5.111$, $p = .033$, $\eta p^2 = .176$). This bias was equally strong in both conditions (E: 11.9%; C: 11.2%; $F_{(1,24)} = 0.419$, $p = .524$, $\eta p^2 = .017$) (Figure 3, Panel A, Bottom).

The results regarding the button press timing showed that, overall, participants were able to keep the requested pace between presses. A mean of 1.5% (range: 0.5% - 2.9%) of the trials were rejected, because the participants were too slow (> 1200 ms) or too fast (< 600 ms) in pressing the button. In the remaining trials, the mean pace between button presses was of 0.78 ± 0.05 s, across conditions. A t-test revealed no significant differences in the overall IPI between conditions ($t_{(24)} = 0.131$, $p = .897$). Moreover, participants took more time to press button B ($F_{(1,24)} = 13.392$, $p = .001$, $\eta p^2 = .358$), in both conditions ($F_{(1,24)} = 0.015$, $p = .905$, $\eta p^2 = .001$) (Figure 3, Panel B, Top left). Participants were slower when the previous event was a sound than an omission ($F_{(1,24)} = 9.763$, $p = .005$, $\eta p^2 = .289$), in both conditions ($F_{(1,24)} = 0.012$, $p = .912$, $\eta p^2 = .001$), although this post-sound slowing effect was significantly larger in the Experimental condition than in the Control condition ($t_{(24)} = 2.627$, $p = .015$) (Figure 3, Panel B, Bottom left). Furthermore, participants were slower after an expectation-mismatch than an expectation-match ($F_{(1,24)} = 33.578$, $p < .001$, $\eta p^2 = .583$), and this effect was different between conditions ($F_{(1,24)} = 25.442$, $p < .001$, $\eta p^2 = .515$). Indeed, this post-mismatch slowing effect was larger in the experimental condition than in the control condition ($t_{(24)} = -5.044$, $p < .001$) (Figure 3, Panel B, Bottom right), confirming that in the Experimental condition each button was associated with a specific expectation, whilst in the control condition each button was not associated with any specific expectation. Finally, participants were slower in repeating the same button than in alternating the buttons ($F_{(1,24)} = 22.679$, $p < .001$, $\eta p^2 = .486$), and this effect was equally present in both conditions ($F_{(1,24)} = 0.019$, $p = .891$, $\eta p^2 = .001$) (Figure 3, Panel B, Top right).

Overall, the behavioral results show that the only difference between the two conditions was the post-mismatch slowing effect, and that all other effects were equally present in both conditions. Considering that we have rejected all the trials which were following a mismatch event for the ERP analysis, these findings excluded the presence of any confounding effect confirming that the behavioral responses were similar in the two conditions, and hence that the events elicited by these responses were comparable in the ERP analyses, between conditions.

ERP results

Analysis based on components of interest

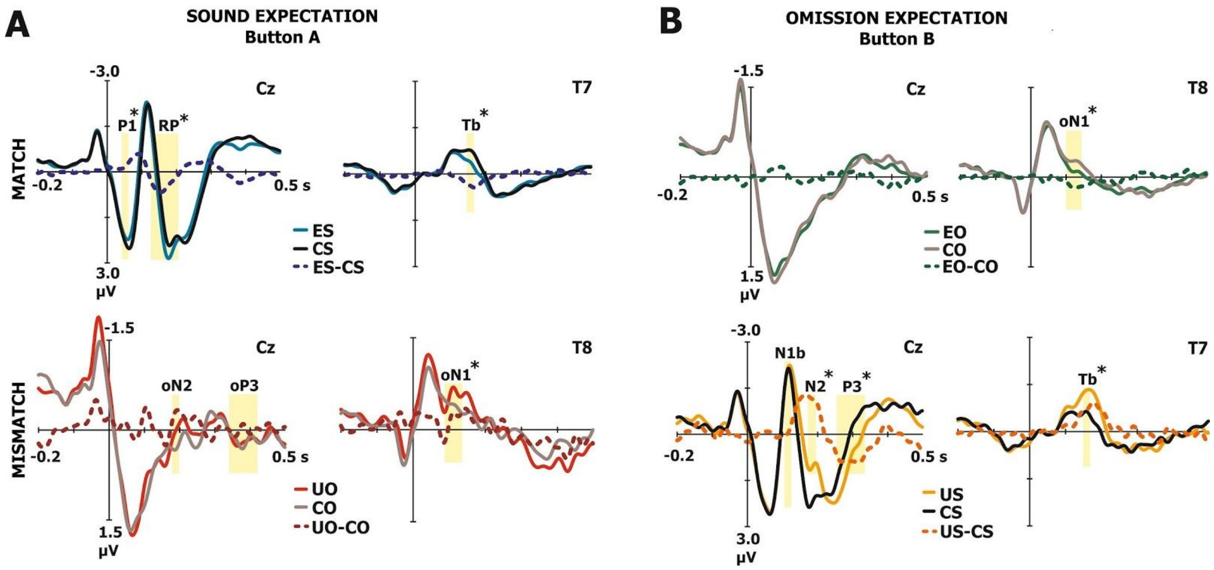

Figure 4. *ERP results on the components of interest. Sound expectation-related effects are depicted in Panel A, and Omission expectation-related effects are depicted in Panel B. Match-related effects are depicted in the first line, and Mismatch-related effects are depicted in the second line. The representation illustrates the main electrodes and time windows of interest. Significant values: * $p \leq 0.05$*

Table 2. *ERP results on the components of interest related to each expectation effect: 1) matching a sound expectation (ES-CS); 2) mismatch a sound expectation (UO-CO); 3) match an omission expectation (EO-CO); 4) mismatch an omission expectation (US-CS). Significant values: *$p \leq 0.05$.*

Auditory components (ROI)	**Expectation Effects**							
	Sound Expectation				**Omission Expectation**			
	1. Match		**2. Mismatch**		**3. Match**		**4. Mismatch**	
	$t_{(24)}$	*P*	$t_{(24)}$	*p*	$t_{(24)}$	*p*	$t_{(24)}$	*p*
P1	-2.148	**0.042***	-0.306	0.762	-0.949	0.352	-0.512	0.613
Na (left)	0.753	0.459	-1.383	0.179	0.256	0.8	-1.214	0.237
Na (right)	-1.95	0.063	-1.483	0.151	1.906	0.069	-1.533	0.138
Na (average)	-0.676	0.506	-1.768	0.09	1.463	0.156	-1.569	0.13
N1b	-0.379	0.708	-1.071	0.295	0.26	0.797	-0.606	0.55
N1bT	1.599	0.123	2.421	**0.023***	-0.017	0.987	0.104	0.918
oN1 (left)	0.947	0.353	-1.759	0.091	1.512	0.143	-2.509	**0.019***
oN1 (right)	-0.723	0.477	-2.645	**0.014***	3.559	**0.002***	-2.178	**0.039***
oN1 (average)	0.119	0.906	-2.668	**0.013***	3.242	**0.003***	-2.517	**0.019***
Tb (left)	6.2	**< 0.001***	0.056	0.956	1.165	0.255	-6.416	**< 0.001***
Tb (right)	1.832	0.079	-1.157	0.259	1.832	0.079	-3.451	**0.002***
Tb (average)	4.585	**< 0.001***	-0.794	0.435	1.808	0.083	-5.56	**< 0.001***
P2	3.002	**0.006***	-0.859	0.399	0.361	0.721	-4.336	**< 0.001***
oN2	2.399	**0.025***	-1.957	0.062	0.378	0.709	-4.259	**< 0.001***
N2/MMN	2.429	**0.023***	-1.754	0.092	0.419	0.679	-4.331	**< 0.001***
RP	4.207	**< 0.001***	-0.874	0.391	0.907	0.374	-5.237	**< 0.001***
P3	-2.172	**0.04***	-0.448	0.658	-0.986	0.334	2.504	**0.019***
oP3	2.344	**0.028***	1.344	0.191	0.825	0.418	1.381	0.18

Table 2 shows the results from all the comparisons of interest. On the following we highlighted the most relevant results.

1. Effects of Matching a Sound expectation (ES-CS)

The response to the ES was significantly attenuated compared to CS in the P1 time-window ($t_{(24)}$ = -2.148, p = .042). The results did not show any difference between conditions in the N1 time-window (N1b: $t_{(24)}$ = -0.379, p = .708). Nevertheless, the response to ES was significantly attenuated compared to CS in the Tb time-window, only on the left ROI (Tb left: $t_{(24)}$ = 6.2, $p < .001$), and significantly enhanced in the P2 time-window ($t_{(24)}$ = 3.002, p = .006). Moreover, the results showed that the ES elicited a significantly more positive response compared to CS in the Repetition Positivity time-window ($t_{(24)}$ = 4.207, $p < .001$).

2. Effect of Mismatching a Sound expectation (UO-CO)

The response to the UO was significantly more negative compared to the CO in the oN1 time window, (oN1 avg: $t_{(24)}$ = -2.668, p = .013; oN1 right: $t_{(24)}$ = -2.645, p = .014), but they did not differ in the oN2 ($t_{(24)}$ = -1.957, p = .062), nor in the oP3 ($t_{(24)}$ = -1.344, p = .191) time-windows.

3. Effects of Matching an Omission expectation (EO-CO)

The response to the EO was significantly more positive compared to the CO in the oN1 time window (oN1 avg: $t_{(24)}$ = -3.242, p = .003; oN1 right: $t_{(24)}$ = -3.559, p = .002). No other differences were found between the two responses.

4. Effect of Mismatching an Omission expectation (US-CS)

The response to the US was significantly enhanced compared to the CS in the Tb time window (Tb avg: $t_{(24)} = -5.56, p < .001$; Tb right: $t_{(24)} = -3.451, p = .002$; Tb left: $t_{(24)} = -6.416$, $p < .001$). Moreover, the US elicited a N2/MMN-like response ($t_{(24)} = -4.331$, $p < .001$), followed by a P3 response ($t_{(24)} = 2.504$, $p = .019$).

Analysis based on sustained modulations

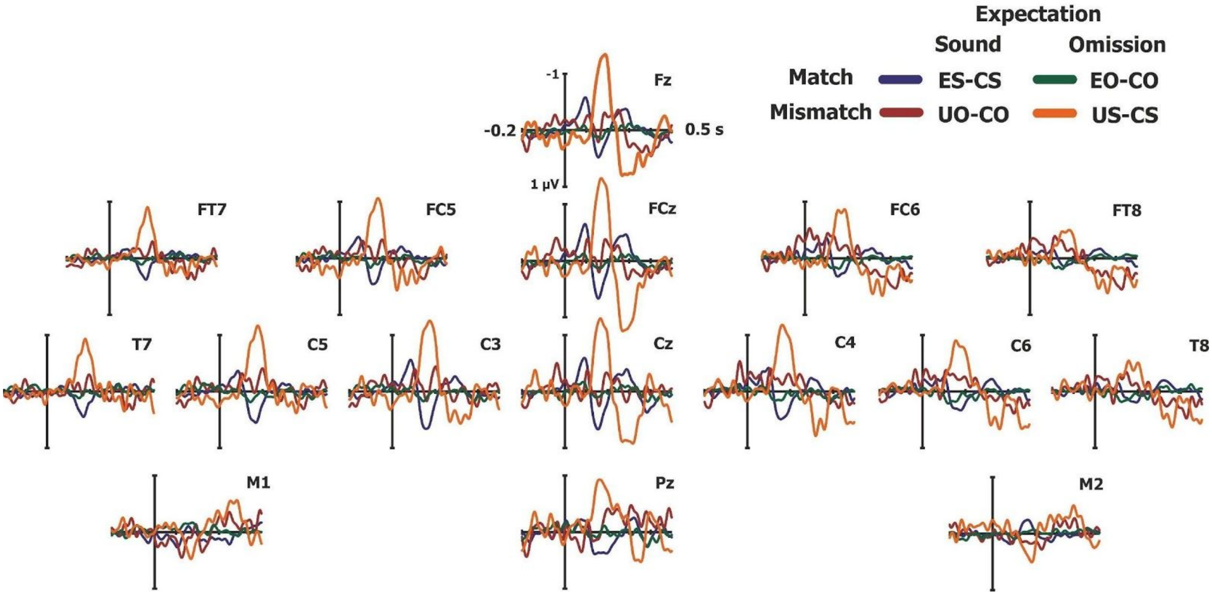

Figure 5*. Overview of the difference waveforms representing the sustained modulations: 1) matching a sound expectation (blue line); 2) mismatching a sound expectation (red line); 3) matching an omission expectation (green line); 4) mismatching an omission expectation (orange line).*

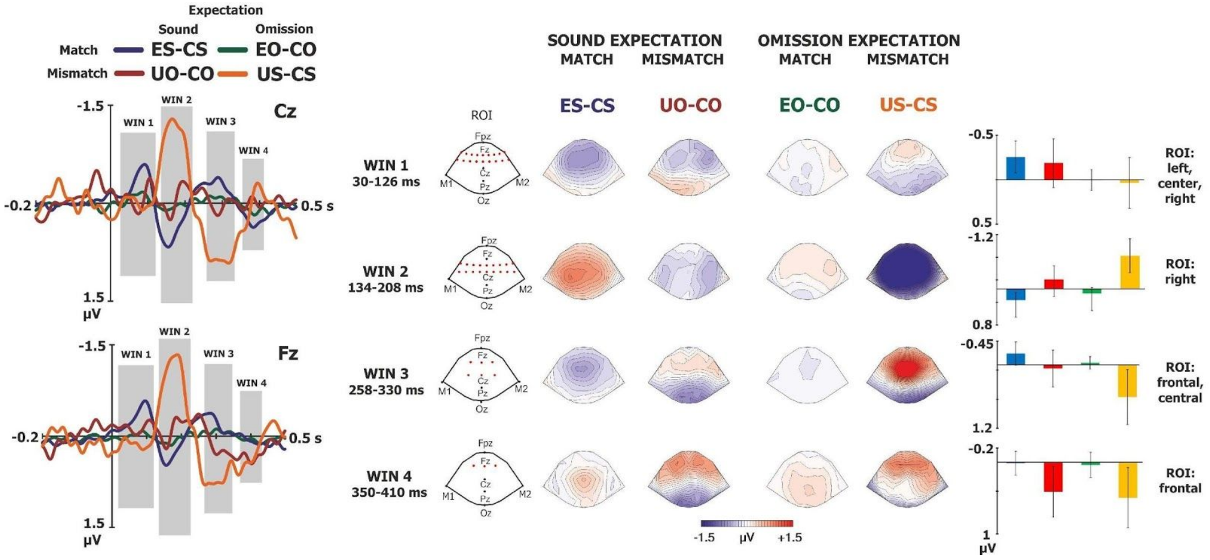

Figure 6*. ERP results on sustained modulations. The representation illustrates the ERPs (at Cz and Fz), scalp maps and bar plots for each of the time-windows of interest. The red channels marked in the white scalp maps are the specific electrodes used for the ERP analysis in each time-window. The bar plots represent only the regions (ROIs) specified on the right side of the figure. Those are, left (F7, F5, F3, FT7, FC5, FC3), center (F1, Fz, F2, FC1, FCz, FC2) and right (F4, F6, F8, FC4, FC6, FT8) for the first time-window; right (F4, F6, F8, FC4, FC6, FT8) for the second time-window; frontal (F3, Fz, F4) and central (C3, Cz, C4) for the third time-window; and frontal (F3, Fz, F4) for the fourth time-window.*

First time-window (30-126 ms)

The results indicate that the content of the expectation (sound vs. omission) was driving the neuronal response. Indeed, we observed a fronto-central negativity when participants expected to hear a sound, but not when they expected to hear an omission ($F_{(1,24)} = 8.456$, $p = .008$, $\eta p^2 = .261$) (Figure 6; ES-CS and UO-CO). The results did not show any main effect related to the degree of match (match vs. mismatch) ($F_{(1,24)}$ = .371, p = .548, ηp^2 = .015), nor an interaction between the content of the expectation and the degree of match ($F_{(1,24)}$ = .040, p = .844, ηp^2 = .002).

Second time-window (134-208 ms)

The results indicate that both the content of the expectation, the degree of match and the interaction of these two factors were modulating the neuronal response. Indeed, we observed a broadly distributed negativity in case of mismatch (UO-CO and US-CS) and positivity in case of match (ES-CS and EO-CO) ($F_{(1,24)}$ = 32.276, p < .001, ηp^2 = .574), which was larger when there was a sound (US-CS and ES-CS) than when there was an omission (UO-CO and EO-CO) ($F_{(1,24)}$ = 12.047, p = .002, ηp^2 = .334). However, we also found that the responses were more positive in case of sound expectation (ES-CS and UO-CO), and more negative in case of omission expectation (EO-CO and US-CS) ($F_{(1,24)}$ = 19.686, p < .001, ηp^2 = .451).

Third time-window (258-330 ms)

The results show that both the content of the expectation and the degree of match were driving the neuronal response, but there was no interaction between these factors ($F_{(1,24)}$ = 3.016, p = .095, ηp^2 = .112). Here, we observed a frontocentral positivity in case of mismatch (UO-CO and US-CS) and frontocentral negativity in case of match (ES-CS and EO-CO) ($F_{(1,24)}$ = 7.127, p = .013, ηp^2 = .229). Furthermore, we found that the responses were more positive in case of omission expectation (EO-CO and US-CS), than in case of sound expectation (ES-CS and UO-CO) ($F_{(1,24)}$ = 10.073, p = .004, ηp^2 = .296).

Fourth time-window (258-330 ms)

The results indicate that only the degree of match was driving the neuronal response. Indeed, we observed a very frontal positivity in case of mismatch (UO-CO and US-CS) ($F_{(1,24)}$ = 10.612, p = .003, ηp^2 = .307). The results did not show any main effect related to the content of the expectation ($F_{(1,24)}$ = .224, p = .640, ηp^2 = .009), nor an interaction between the content of the expectation and the degree of match ($F_{(1,24)}$ = .056, p = .816, ηp^2 = .002).

Time-window preceding the button press

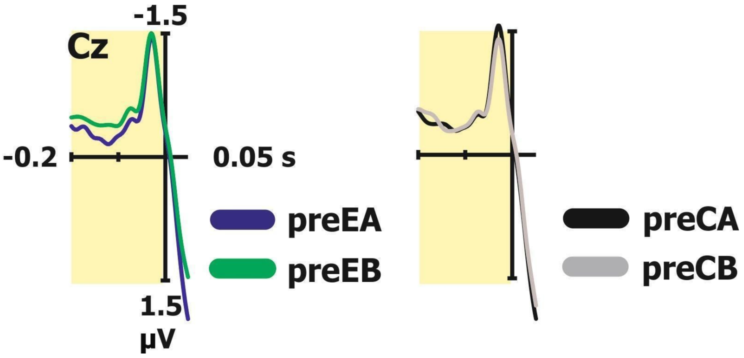

***Figure** 7. ERP results on the time-region preceding the button press (from -0.2 to 0 s). Left side: button A (preEA, blue line) and button B (preEB, green line) in the experimental condition. Right side: button A (preCA, black line) and button B (preCB, grey line) in the control condition.*

The results from the comparison between the two conditions, showed a significant difference between the two buttons only in the experimental condition, where the response is significantly more positive for button A compared to button B (E: $t_{(24)}$ = 2.167, p = .040; C: $t_{(24)}$ = 0.882, p = .386).

Discussion

Across different research lines and experimental paradigms, previous literature has reported various expectation-match and -mismatch effects on behavioral, and electrophysiological responses (for a review see Bendixen et al., 2012). So far, it has not been possible to give a parsimonious unified explanation to these various expectation-related effects, which have been described independently from each other. The predictive coding theory provides a useful framework that could offer a unified explanation, interpreting these effects as different manifestations of the same underlying process (Friston, 2009; Friston & Kiebel, 2009). Indeed, according to the predictive coding theory, prediction modulates the neuronal response through a top-down process, and as a result all the ERPs reflect mainly the prediction error, that is the difference resulting from the comparison between the prediction and the actual sensory input (Bastos et al., 2012; Jackson & Bolger, 2014). Hence, prediction-match and -mismatch effects are the manifestations of the same process, which can result in minimal prediction error (match effect) or large prediction error (mismatch effect).

We used predictive coding (Friston, 2005) as the framework for our study. We firstly examined the behavioral and electrophysiological responses which have been reported in the literature as modulated by expectation-related effects, and then we analyzed their relationships in order to uncover whether these phenomena are the manifestation of the same predictive processes or not. We proposed a novel paradigm through which to investigate the expectation-match and -mismatch effects, studying differences and similarities between expecting the presence of a sound and expecting the omission of a sound, and the violation and confirmation of each of these expectations. Specifically, we manipulated the event generated by the button press (sound vs. omission) and the predictability of the situation (experimental vs. control condition) altering the press-effect contingencies. Then, we compared the responses recorded in a predictable condition, in which each button was reliably associated to a high probability of causing a particular event (i.e., sounds vs. omissions in the experimental condition), to the responses recorded in a non-predictable condition in which no reliable press-effect associations existed (i.e., sounds vs. omissions in the control condition). We hypothesized to detect behavioral (inter-press intervals) and electrophysiological changes (exogenous and endogenous ERPs), reflecting top-down processing and the degree of match (match vs. mismatch) between the expectation and the event, both in the case of the presence of the sound, allowing us to observe the modulation of the stimulus-evoked response, and in the case of the omission of the sound, allowing us to observe purely endogenous responses related to the predictive processing.

In line with our hypotheses, the behavioral results showed differences in the inter-press intervals, with slowing effects after inputs mismatching the expectation (Barcelo et al., 2006; Fogelson et al., 2009; Notebaert et al., 2009; Iwanaga & Nittono, 2010). The ERP results, from the replication of previous studies, partially supported our hypothesis, since the expectation-related effects were not present for all the ERP components of interest. Indeed, in line with our hypotheses, we observed a P1 attenuation and a RP in case of matching a sound expectation (i.e., expected sound), a more positive response in case of matching an omission expectation (i.e., expected omission) and a more negative response in case of mismatching a sound expectation (i.e., unexpected omission) in the oN1 time window, and an N2/MMN-like and the P3 responses only in case of mismatching an omission expectation (i.e., unexpected sound), but not in case of mismatching a sound expectation (i.e., unexpected omission). However, the results from the analysis on the relationships between expectation effects showed that separate expectation-match and -mismatch effects coexist in the same time-windows, suggesting that similar predictive processing could underlie and explain the reported expectation-match and -mismatch effects.

In the following sections, we first discuss each expectation-related effect measured on the components of interest (from the hypothesis-driven analysis), and subsequently we discuss the sustained modulations and the time-window preceding the button press (from the data-driven analysis), interpreting them as manifestations of the same predictive processing.

Interpreting each expectation-related effect on the ERPs

In order to isolate the effects of matching a sound expectation, we compared the auditory response to the Expected sound to the response to the Control sound. The results show that the P1 response to the Expected Sound was significantly attenuated compared to the Control Sound. Our results are in line with previous studies reporting matching effects in the P1 time-window in response to self-generated stimuli (Baess et al., 2009; Lasaponara et al., 2011) and also in passive listening (Lü et al., 1992; Boutros et al., 1995; Lijffijt et al., 2009). However, we did not find any N1 attenuation for the Expected Sound compared to the Control Sound. This finding is not in line with previous studies which have reported an N1 attenuation in response to self-generated stimuli compared to externally-generated ones (Schäfer & Marcus, 1973; Martikainen et al., 2005; Baess et al., 2008, 2009, 2011; Knolle et al., 2013). One possible explanation for this result is related to the timing between button presses which was of around 800 ms. Indeed, a previous study showed the absence of N1 attenuation when the inter-press interval was shorter than 1200 ms (SanMiguel et al., 2013 b). A further explanation is related to our study design. Indeed, the N1 attenuation to self-generated sounds is typically found in contingent paradigms in which the auditory response extracted from self-generated sounds recorded in a motor-auditory condition (100% probability of the sound being elicited by the button press) is compared to the purely auditory response to externally-generated sound recorded in an auditory-alone condition (Horváth, 2015). Conversely, in our study we are comparing expected self-generated sounds (having 88% probability of being elicited by the button press) to control self-generated sounds (having 50% probability of being elicited by the button press). Hence, the results from this analysis suggest that the N1 attenuation may be more related to the self-generation variable (which differentiates the conditions in the typical self-generation paradigms), than to the degree of predictability (which differentiates the conditions in our paradigm). Moreover, contrary to our hypothesis, we found a significantly enhanced P2 response to the Expected Sound compared to the response to the Control Sound. Although a few studies investigating the motor-auditory interaction have reported a P2 attenuation following the N1 attenuation (Sowman et al., 2012; Knolle et al., 2013; Saupe et al., 2013; Horváth, Timm et al., 2014; van Elk et al., 2014), it has also been suggested that the modulations of the N1 and P2 components are probably related to different processes (Knolle et al., 2013), and hence that the suppression of both components not always occurs in tandem. A further possible explanation for the enhanced P2 could be the overlap between the P2 time-window (184 ± 20 ms) and the RP time-window, indeed there are studies showing a RP and an increased P2 (Baldeweg et al., 2004; Haenschel et al., 2005; Baldeweg et al., 2006). Particularly, we found a significantly more positive response to the Expected Sound compared to the Control Sound in the RP time-window (178 ± 30 ms), which confirmed our hypothesis about the presence of expectation match-effects in this time-window. Taken together, our results seem to confirm the idea that the RP reflects the modulation of several exogenous ERPs (e.g., P50, P2, and N1; Baldeweg et al., 2004; Haenschel et al., 2005; Baldeweg et al., 2006). Moreover, here we show that the RP time-window overlaps also with the time-windows of other mismatch-related endogenous ERPs (e.g., N2/MMN-like and oN2 time windows), pointing to a relationship between the match- and mismatch-related effects.

In order to investigate the effects of mismatching a sound expectation, we compared the endogenous response to the Unexpected omission to the response to the Control omission. The results showed a significantly more negative response to the Unexpected omission compared to the response to the Control omission in the oN1 time-window. This finding was in line with our hypothesis and with previous studies (Raij et al., 1997; Hughes at al., 2001; Bendixen et al., 2009; Wacongne et al., 2011; SanMiguel et al., 2013 a, c) observing a difference between experimental and control omission responses, despite the absence of inputs, although this analysis does not offer the possibility to clarify whether this difference represents the top-down modulatory activity per se or the prediction error. However, conversely from our hypothesis, we did not find an oN2/MMN-like response followed by an oP3 response to the Unexpected omission. Probably, this could be due to the weakness of the neuronal response related to the absence of any physical input, or to the interval between button presses, which as been found to play an important role both in passive listening (Tervaniemi et al., 1994; Yabe et al., 1998, 2001) and in motor-auditory interactions (SanMiguel et al., 2013 b). A further possible explanation could be associated with our study design, which differs from the previous studies that have investigated omissions of self-generated sounds. Indeed, one difference is that here we are comparing the omission responses recorded in the experimental condition (where the omission probability is 12%) to the omission responses recorded in the control condition (where the omission probability is 50%), and hence the two conditions had a different degree of predictability about the content elicited by the button press (sound or omission). Conversely, the previous omission studies have used a purely motor response recorded in a motor-alone condition (having 100% probability of producing no sound) as control. However, the absence of oN2 and oP3 responses should not be due to our study design, since a previous study has reported the absence of expectation-related effects in a condition, similar to our control, where the probability of generating a sound was 50%, compared to a motor-alone condition (SanMiguel et al., 2013 c), and since we found a MMN and P3 in case of mismatching an omission expectation (Unexpected Sound minus Control Sound). Moreover, since we used only one sound, conversely to previous studies (SanMiguel et al., 2013 a), we cannot explain the absence of omission N2 and P3 responses as due to the impossibility of having a specific sound representation. Thus the only factor which further differentiates our study from previous ones is that in our study participants are pressing one of two buttons and perhaps this has a negative effect on the strength with which specific predictions are formulated per each button, and hence on the presence of omission N2 and P3 responses.

In order to study the effect of matching an omission expectation, we compared the endogenous response to the Expected omission to the response to the Control omission. In line with our hypothesis, we found that the endogenous response to Expected omission was significantly more positive compared to the response to the control omission in the oN1 time-window. The presence of a matching effect on the omission response suggests that predictive processes are modulating the endogenous response (Bendixen et al., 2009; Wacongne et al., 2011) although, as in the case of the modulation observed on the Unexpected omission, from this analysis it is not possible to understand whether it is due to the top-down processing or to the prediction errors. However, we can notice that in the oN1 time window the effect of matching an omission expectation shows a reversed polarity than the effect of mismatching a sound expectation, pointing a possible relation between these opposite expectation-related effects.

Finally, in order to examine the effects of mismatching an omission expectation, we compared the response to the Unexpected Sound to the response to the Control Sound. In line with literature, the Unexpected Sound elicited an N2/MMN-like response, followed by a P3 response, which mimic the deviance responses typically registered in passive listening oddball paradigms (Näätänen et al., 2007; Escera & Malmierca, 2014). This finding confirms that an association between the button press B and the omission of the sound was created along the experiment and that participants were able to detect a violation of this press-omission association. Furthermore, our findings suggest that probably the negative (Falkenstein et al., 1990; Gehring et al., 1993; Falkenstein et al., 2000) and positive (Falkenstein et al., 2000; Nittono & Ullsperger, 2000; Nittono, 2004; Overbeek et al., 2005; Nittono, 2006; Waszak & Herwing, 2007; Katahira et al., 2008; Band et al., 2009; Iwanaga & Nittono, 2010) components which have been observed in previous studies, in response to error-feedback elicited by an action, could be similar to the MMN and P3 responses that have been observed in previous passive listening studies, since all these responses are characterized by being unexpected (Folstein & Van Petten, 2008). Moreover, the increased negativity observed for the Unexpected Sound compared to the Control Sound was also found in the P2, oN2 and RP time windows, revealing the existence of a wider time-window (134-208 ms) in which both the MMN and the RP were coinciding. Thus, probably, the modulation usually reported on the P2 could be correlated to the overlapping effects observed in this wide time-window, and specifically to the mismatch-related negativity.

In summary, the findings from this analysis replicate the expectation-match and -mismatch effects reported in the ERP literature on the endogenous responses (e.g., on the RP, oN1, N2/MMN-like, and P3 responses), but they are less consistent with the data reported in the literature on the exogenous responses (e.g., P1, N1, P2).

Interpreting expectation effects as manifestations of a unitary phenomenon

The ERP literature has generally examined and described the modulation of the expectation-related effects on the exogenous (P1, N1, P2) and endogenous (RP, MMN and P3) components as unrelated phenomena, and thus studying them in separated time-windows. However, examining each of the single ERP components typically reported in literature as affected by expectation effects (i.e., hypothesis-driven analysis) we noticed that opposite expectation effects (i.e., match- and mismatch-related effects) were coexisting and overlapping in time-windows which were very close to one other (e.g., P2, RP, N2/MMN and oN2). This finding motivated us to select four wide time-windows and to carry on further data-driven analysis, in order to study more carefully the relation and the time-course of these sustained modulations related to the predictive processing.

In the first time window (30-126 ms), the results showed that the content of the expectation was modulating the neuronal response, indeed scalp maps and bar plots seem to reveal a frontocentral negativity in case of sound expectation (i.e., ES-CS and UO-CO) and a frontocentral positivity in case of omission expectation (i.e., EO-CO and US-CS), independently from the actual presence of the sound. In this time-window, it has been typically reported the presence of the exogenous component P1/P50 and N1, related respectively to preattentive (Jerger et al., 1992; White & Yee, 1997, 2006) and early attentive processing (Näätänen, 1992). Probably, taken together with the lack of effects related to the degree of match (match vs. mismatch effects), we could speculate that in this first time-window we are observing a top-down modulatory activity of the neuronal response, allocating attentive resources.

In the second time window (134-208 ms), also the degree of match between the event and the content of the expectation was subsequently driving the neuronal response. Scalp maps and bar plots revealed that although the magnitude of the modulation of the signal elicited by the omissions (UO-CO and EO-CO) was smaller, their match and the mismatch-related effects followed the same modulatory patterns of the signal elicited by the sounds (ES-CS and US-CS), with a positive displacement in the case of match between expectation and input, and a negative displacement in the case of mismatch. In this time-window, we found both the RP (ES-CS) and the MMN-like response (US-CS), indicating respectively the case of matching a sound expectation and the case of mismatching an omission expectation. Previous studies have described RP (Haenschel et al., 2005; Baldeweg, 2007; Costa-Faidella et al., 2011; Recasens et al., 2015) and MMN (Näätänen et al., 2007; Escera & Malmierca, 2014) as separated phenomenon, related respectively to the confirmation and to the violation of an expectation however, as pointed-out by Baldeweg (2007), it would be interesting to uncover the processing underlying RP and MMN. Our results suggest that these two endogenous responses are probably the manifestation of the same process, which here could be the comparison between expectation and current event, and that they differ in the amount of prediction error: larger and negative in the case of mismatching (MMN), and minimum and positive in the case of matching (RP).

In the third time-window (258-330 ms), the effects related to the content of the expectation and to the degree of match remained constant. However, scalp maps and bar plots showed that in this time-window there was a positive displacement when the sensory input mismatched with the expectation and a negative displacement when the sensory input matched with the expectation. In line with previous studies, typically reporting the presence of the P3 response in this time-window (Escera et al., 1998, Escera et al., 2000; Schröger et al., 2000; Friedman et al., 2001; Polich, 2007; Joos et al., 2014), we can observe a frontocentral positivity typical of the P3 response both in case of mismatching an omission expectation (US-CS), and in case of mismatching a sound expectation (UO-CO), confirming the idea that this component depends on the degree of expectation (Holroyd & Coles, 2002) and that it can indicate the updating of internal models triggered by stimuli mismatching an expectation (Donchin & Coles, 1988), independently from the content of the expectation (sound or omission).

Finally, in the fourth time window (350-410 ms), the degree of the match between the input and the expectation maintains an important role. Indeed, scalp maps and bar plots revealed a strong frontal activity not only in case of mismatching an omission expectation (US-CS), but also in case of mismatching a sound expectation (UO-CO). Probably, the similar frontal distribution observed in the third and fourth time window in case of mismatching suggests that the P3 response is still present up to around 400 ms, in order to better predict the content associated with the next button press (Friedman et al., 2001) and decrease uncertainty levels (Sutton et al., 1965; Friston, 2010).

Furthermore, we were interested in studying the time-regions preceding the button press (from -0.2 to 0 s), in order to verify the presence of anticipatory expectation effects in the two conditions. The results showed a difference between the two buttons only in the experimental condition, which was the only predictable situation, confirming that only in this condition it was possible to generate a robust association and hence expectation for each button. This finding is quite novel because to our knowledge, literature has reported the expectation effects elicited only after a specific action, but it has never been proved the existence of these effects before the action. Most likely, the difference which has emerged in the experimental condition among the two buttons could be the result of the whole process previously described, and aimed to reduce the uncertainty levels elicited by the sensory consequences of incoming events (Sutton et al., 1965; Friston, 2010).

Based on the findings from this analysis, we could speculate that the modulation of the predictive processing on the neuronal response is characterized by four main stages: a first one, in which the content of the expectation modulates the neuronal response through top-down processing (represented by P1 and N1); a second stage characterized by the amount of prediction error resulting from the comparison between expectation and current event (represented by RP and MMN-like response); a third stage necessary to update the internal model about the consequences of the button press (indicated by the presence of the P3-like response) and a last stage of preparation for the subsequent button press (still indicated by the presence of the P3-like response).

The hypothesis about the existence of different stages of predictive modulation on the neuronal response is in agreement with the idea of a hierarchical predictive system (Friston, 2005), which has been empirically shown also in ERP studies (Winkler et al., 2005; Bekinschtein et al., 2009; Wacongne et al., 2011), and fMRI studies (Diekhof et al., 2009).

In summary, the study on wider time-windows has offered the possibility of examining the time-course and the relation between these sustained modulations on the ERPs, suggesting that they are interrelated with each other, and differentiated by the stage (i.e., time-course) in which the predictive processing is modulating the response: 1) initial top-down modulation, 2) comparison between expectation and input, resulting in a minimal prediction error in case of matching and a larger prediction error in case of mismatching, 3) updating of the internal model, and 4) preparation to the incoming event.

Conclusions

The present study investigated the relationships between expectation-match and expectation-mismatch effects related to the presence and to the omission of a particular auditory stimulus, at behavioral and electrophysiological level.

The findings of this study suggest that at behavioral level the expectation-mismatch causes slowing effects on the action-timing, and at electrophysiological level the expectation-match and -mismatch effects appear as manifestations of a unique phenomenon which vary based on the time-course.

The strength of this study resides in the original experimental design and analyses which allowed to examine different expectation-related effects and their relationships within the same study. However, further studies will be necessary in order to have the confirmation that these effects are the reflection of the same predictive processing.

Overview

Several researchers propose that auditory event-related potentials (ERPs) are modulated by expectation-related effects arising from internal models about auditory regularities. Generally, the ERP literature reports an enhancement of the sensory response mainly resulting in mismatch negativity (MMN), followed by a P3 response (mismatch-related effect), and an attenuation of the auditory P1, N1 and P2 components and the presence of repetition positivity (RP) (match-related effect). However, expectation-mismatch and -match effects have been mostly studied across different research lines and paradigms and their relation is still unclear. We aimed to clarify the relationship between these expectation-related effects and the related exogenous and endogenous responses. We recorded ERPs while participants performed button presses that were associated with either sounds or silences (i.e., “omissions” of the sounds), and with occasional violation of these associations. We examined behavioral and electrophysiological responses reflecting the degree of match (match vs. mismatch) between expectations and events. The four events of interest recorded in the predictable experimental condition (expected sounds, unexpected sounds, expected omissions, and unexpected omission) were compared to the corresponding events recorded in a non-predictable control condition. Behaviorally, we observed slowing effects after events mismatching-expectation. Electrophysiologically, we found overlapping manifestation of opposing match- and mismatch-related effects in the same ERP time-windows (e.g., RP and MMN) suggesting that these modulations reflect the same predictive processing.

STUDY II

The second study of this thesis is entitled “An auditory ERP study tracking the buildup of action-effect associations”, by Vittoria Spinosa, Bence Neszmélyi, János Horváth, and Iria SanMiguel.

Introduction

In everyday life, each motor act is associated with feedback which allows to adjust the actions and to direct our goal with less efforts. For instance, by experience, humans learn that a dedicated button on a bus allows to reach a specific goal (i.e., to request the next stop), and its press elicits a specific sensory consequence. Indeed, often, this action elicits a sound, which indicates if the action was successful or not. If the press does not elicit the expected consequence (i.e., no sound) it could indicate that the button was not working properly or that the force (i.e., the pressure) used to press it was not enough. Hence, based on the sensory consequence, it is possible to decide whether to repeat and to adjust the action or not. But, how do humans learn to generate and modify the internal associations between actions and effects? Which are the behavioral and electrophysiological indices of this learning process?

The interest in the formation of a relation between two events nourished ancient philosophical debates. Indeed, already the philosopher David Hume (1739/1969) suggested that two of the main cues to determine this relation are the contingency (that is, causation) and the temporal or spatial contiguity. Later on, other philosophers, interested in the action-perception relation, proposed the ideomotor theory (Herbarth, 1816, 1825; Lotze, 1852; Harless, 1861), which James (1890/1950) finally brought to the attention of psychologists (Stock & Stock, 2004; Shin et al., 2010).

The ideomotor (IM) theory suggests that action and effect share a common internal representation, thus when a given action-effect association has been learned, this relationship results in bidirectional effects (Greenwald, 1970; Prinz, 1987; Elsner & Hommel, 2001;

Hommel et al., 2001; Kunde et al., 2004). That is, actions modulate sensory processing, but sensory consequences also have an impact on the physical characteristics of the actions themselves.

More recently, the predictive coding theory has proposed that the brain continuously generates and keeps updating an internal representation of the sensory consequences of each motor act in order to maintain an optimal level of energy (Friston, 2005). That is, the brain is seen as a bayesian inference machine which adjusts its internal models based on the degree of match between the prediction and the current sensory consequence of each action.

In experimental psychology, several studies have investigated the action-effect association, and found behavioral and electrophysiological indices reflecting its establishment. For example, studies manipulating the sensory feedback of the actions in order to compare expected versus unexpected effects have described a slowing effect on the action timing in response to unexpected events (Fogelson et al., 2009; Notebaert et al., 2009; Iwanaga & Nittono, 2010). Moreover, by measuring event-related potentials (ERPs) researchers found a negative component similar to the auditory mismatch negativity (i.e., MMN-like wave), and a positive component (i.e., P3-like wave). The negative component was reported, at around 80 ms in response to negative feedback presentation (Folstein & Van Petten, 2008), such as the error negativity, Ne (Falkenstein et al., 1990; 2000) and the error-related negativity, ERN (Gehring et al., 1993), and was interpreted as error detection. The positive component was reported between 200-500 ms after the feedback, and was interpreted as related to post-error processing. Moreover, this positive component tends to be larger for unexpected self-generated stimuli than for stimuli externally-generated (Falkenstein et al., 2000;

Ullsperger, 2000; Nittono, 2004; Nittono & Overbeek et al., 2005; Nittono, 2006; Waszak & Herwing, 2007; Katahira et al., 2008; Band et al., 2009; Iwanaga & Nittono, 2010). Although there is an exiguous number of works investigating the relation between ERN and MMN (Folstein & Van Petten, 2008), it has been suggested that the factor triggering these components is not the feedback valence per se, but the expectancy degree (Holroyd & Coles, 2002).

Other studies aimed to compare the neuronal response related to self-generated and externally-generated sounds, using the so-called contingent paradigm, in which participants are instructed to listen to the sounds that are generated by each of their button presses in one condition, and then to the same sounds replayed passively to them in another condition. In order to compare the auditory response between these two conditions, an additional motor-control condition is also included, in which participants perform the same button presses, but these presses never produce a sound. Thus, this motor activity is subtracted from the sensory response recorded in the motor-auditory condition. These studies show an attenuation of neural sensory responses elicited by self-generated stimuli. Indeed, the ERPs elicited by self-generated sounds, in particular the auditory N1 component and also the P2 component, are strongly attenuated when compared to the ERPs elicited by externally-generated sounds (Marcus, 1973; Martikainen et al., 2005; Baess et al., 2008; Aliu et al., 2009; Baess et al., 2009; Hesse et al., 2010; Baess et al., 2011; Horváth, 2013 a, b; SanMiguel et al., 2013 b; Saupe et al., 2013; Schäfer & Timm et al., 2013; van Elk et al., 2014; Mifsud & Whitford, 2017; Neszmélyi & Horváth, 2017). Some studies also found the attenuation of the auditory N1 subcomponent, Tb (SanMiguel et al., 2013 b; Saupe et al., 2013).

To date, there are numerous studies focusing on the auditory-alone regularities to investigate more comprehensively the formation of internal models and the electrophysiological indices associated with this process. Most of these studies used the typical oddball paradigm which consists in replacing a repeated standard stimulus by a novel, or deviant, sound (Escera et al., 2014). A paradigm similar to the oddball one is the roving standard paradigm which consists in delivering different trains of tones, in order to study the repetition-related effects (Cowan et al., 1993; Haenschel et al., 2005; Baldeweg et al., 2006; Costa-Faidella et al., 2011). In the roving paradigm, the first tone of a train acts as deviant stimulus, because of its lower-probability compared to the previous tone-train, and the last tone of a train acts as a standard tone because of its higher-probability within the same tone train (Costa-Faidella et al., 2011). Specifically, these paradigms allow to isolate the mismatch negativity (MMN) which occurs at around 150 ms as the results from the difference waveform obtained in response to the standard sound and in response to the deviant sound (Näätänen et al., 2007; Escera & Malmierca, 2014). Moreover, the deviance detection is also indicated by the presence of a positive response, the so-called P3 component, which occurs at around 300 ms after the deviant presentation, and it is characterized by a frontocentral positivity (Escera et al., 1998; Escera et al., 2000; Schröger et al., 2000; Friedman et al., 2001; Polich, 2007; Joos et al., 2014). The P3 response has been related to attentional orienting (Escera et al., 2000; Nittono, 2006; SanMiguel et al., 2010), task relevance (Nittono, 2006) but also expectancy, indeed, the amplitude of P3 is larger when the stimuli are more unexpected (Squires et al., 1976; Johnson & Donchin, 1980; Polich & Bondurant, 1997). According to the "context-updating" model, the presence of the P3 response can indicate the necessity of revising internal models of the environment (Donchin & Coles, 1988), and of resolving the uncertainty of unpredictable events (Sutton et al., 1965), considering that deviant sounds break down an internal system of predictions (Parmentier et al., 2011).

Moreover, some of the studies focusing on the auditory domain describe the suppressed sensory responses and the repetition positivities as effects of stimuli repetition (Haenschel et al., 2005; Baldeweg et al., 2007), proposing the learning processes as one of the possible explanations (Grill-Spector et al., 2006).

Predictive coding explains the attenuation of the sensory response as a match between the prediction and the actual sensory stimulus (e.g., N1 suppression), and the enhancement of the sensory response as an index of the mismatch (e.g., N2/MMN response), that is, of the difference between the prediction and the actual sensory consequence (Baldeweg, 2006).

However the computational accounts, such as the predictive coding theory (Friston, 2005), seem to explain only the cause-effect contingency Humean principle, but the causal learning process can also derived by an alternative associative learning, such as contiguity (Einhorn & Hogarth, 1986; Young, 1995; Gallistel & Gibbon, 2001; Elsner & Hommel, 2004; Buehner & McGregor, 2009; Greville & Buehner, 2010).

Indeed, a parallel line of research has yielded results that are inconsistent with the predictive model explanation, suggesting that responses to sounds may be unspecifically gated during movement (Hazemann et al., 1975; Tapia et al., 1987; Makeig et al., 1996; Horvath et al., 2012; Horvath, 2013a,b; Horvath & Burgyan, 2013; Horvath, 2014), and not as the result of a specific prediction stemming from an internal model of the learned action-effect association.

In particular, Horváth, Maess, Baess and Tóth (2012) designed an N1-suppression paradigm in which, as opposed to the contingent paradigm, the series of button presses and the series of sounds were uncorrelated. Despite the lack of contingent associations between the presses and the sounds, random temporal coincidences between button presses and sounds resulted in an attenuation of the auditory N1 and P2 elicited by the sounds coinciding with the button press. These results pose a challenge to the predictive framework, and strongly suggest that the temporal proximity between motor behavior and sound could be responsible for learning processes.

Conversely, in favor of the ideomotor principles, several studies endorse the idea that the presence or absence of a reliable sensory feedback affects how movements are planned and executed (Prinz, 1990; Elsner & Hommel, 2001; Hommel et al., 2001; Kunde et al., 2004; Todorov, 2004; Hommel, 2009; van der Wel et al., 2013; Neszmélyi & Horváth, 2017, 2018). For instance, recent studies demonstrate that actions eliciting sounds have different physical parameters than actions without sensory consequences (Neszmélyi & Horváth, 2017; Horváth et al., 2018). These findings demonstrate the existence of bidirectional motor-sensory influences and as suggested by Neszmélyi and Horváth (2017), it is fundamental to consider the interdependency between an action and its sensory consequence in order to study predictive processes, particularly in the contingency paradigms in which it is supposed to compare the same physical events among the different conditions.

Most of the studies presented so far adopted conditions with extreme arrangements (Horváth, 2015): with either total action-effect correlation in the contingent paradigm (i.e., sounds are generated always and exclusively by the actions), or without any action-effect correlation in the coincidence paradigm (i.e., sounds are always externally-generated and only randomly coincide on occasion with the actions), or without considering the motor contribution in the oddball and roving standard paradigms.

However, to date, the studies investigating the association between an action and its auditory consequence (i.e., motor-auditory regularities) share similar results with studies investigating the effects elicited by the repetition of the same auditory stimulus over time (i.e., auditory regularities). Hence, we hypothesized that these studies are investigating mechanisms underlying the same learning processes. Overall, there is a lack of studies tracking the buildup of the action-effect associations (i.e., internal models), from the initial detection to the consolidation, especially in dynamic contexts in which contingency and coincidence effects coexist and no stable action-sound association exists.

The present work aimed to disentangle the learning processes which allow the buildup of internal representations about motor-auditory regularities. We will employ externally-generated and self-generated sounds in order to examine, at a behavioral and electrophysiological level, the contributions of the motor action in this learning process.

Materials and Methods

Participants

All participants gave written informed consent for their participation after the nature of the study was explained to them. Twenty healthy young adults (16 women, 4 men, 4 left-handed) ranging in age 18–25 years (mean = 25 years) participated in the experiment. The subjects were recruited through a student job agency and their collaboration was compensated by a monetary payment. All participants had normal or corrected-to-normal vision, and reported no hearing impairment or history of neurological disease. The experiment was conducted in accordance with the Declaration of Helsinki and approved by the Ethical Research Committee in Psychology (Hungary).

Stimuli and procedure

The experimental task was presented in Octave (Eaton, Bateman, Hauberg, & Wehbring, 2014). Participants sat comfortably inside an electrically shielded chamber and fixated on a fixation cross displayed on a screen placed at a distance of approximately 100 cm from their eyes. In order to collect force measurements and to avoid any mechanical or transient sound production, participants held a Force Sensitive Resistor (FSR; FSR Model 400, Interlink Electronics, Westlake Village, CA, USA), mounted on a thin plastic sheet, between the thumb and the index or middle finger of their dominant hand. The main task consisted in pinching this FSR every 1.5 seconds.

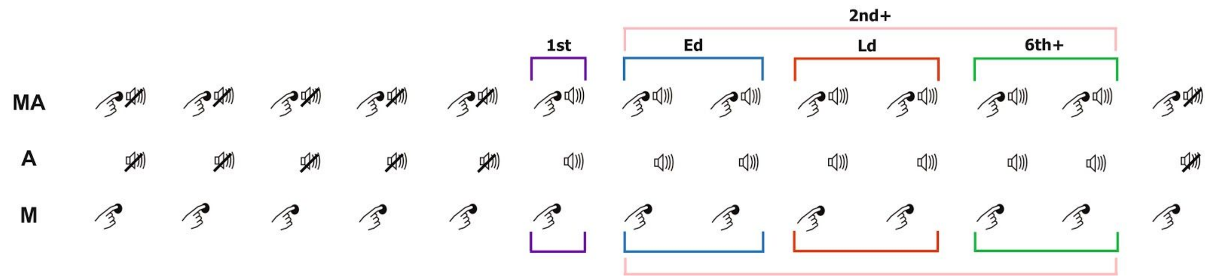

Figure 1. *Schematic representation of the experimental design. There were three conditions: Motor-Auditory (MA), Auditory (A) and Motor (M). For each condition, there were five main events of interest: the first event (1st), the average of the second and third event (Early duplet, Ed), the average of the fourth and fifth event (Late duplet, Ld), the events after the fifth event (6th+), and the events after the first event (2nd+).*

The experimental design consisted of three conditions: motor-auditory condition (MA), auditory condition (A) and motor condition (M). In the motor-auditory condition every pinch could either elicit the presentation of a tone or it could have no auditory consequences (from now on referred to as "omission" of the tone). Specifically, most of the times the pinch resulted in an omission, but occasionally consecutive pinches elicited a tone micro-sequence (Figure 1). In every block the first five events were omissions, and these omissions were associated with visual feedback about the timing between pinches. After these omissions, there was a 10% chance of generating a tone sequence with the next pinch. When a tone sequence was started, five pinches in a row always elicited tones. After the five-tone sequence, for each subsequent pinch there was a 50% chance that the tone sequence ended and that a sequence of omissions started. If a new omission was generated, the 10% rule was implemented again. A block lasted until there were seven tone sequences completed that came after omission sequences of at least five actions.

The tones were sinusoidal waves with a duration of 50 ms, including 10-ms rise and fall ramps. There were two types of tones: standard tones with a pitch of 1002.273 Hz (low pitch), and target tones with a pitch of 1774 Hz (high pitch). To ensure that participants paid a similar level of attention in motor-auditory and auditory conditions, at the end of each block, the participant had to report the total number of target tones. The target tones were only included in the tone-sequences following less than five omissions. The events generated during the motor-auditory condition were recorded, and replayed in the auditory condition. Tones were presented through headphones (Sennheiser HD 600) at a fixed intensity of 50 dB above the individual threshold that was measured for each participant, through an audiometric test, prior to the start of the experiment. In the motor condition, participants had to pinch the FSR every 1.5 seconds as in the motor-auditory condition, but the pinches never elicited any tones. The average number of events per block was 170 (min: 107; max: 269). The average length of the sequences was 6 for the tones (min: 5; max: 14) and 10 for the omissions (min: 1; max: 82). On average one target tone was present both in MA and in A (min: 0; max: 6). The general probability of generating a tone, across the whole experiment, was around 37.6%. The probability of generating target tones was of 0.075% in tone sequences following omission sequences with less than five events.

The start of the experiment was preceded by training blocks, with the aim of habituating to the use of the FSR and attaining the correct 1.5s timing between pinches (inter-press interval, IPI). During the trainings, visual feedback regarding the force applied to the pinch and the IPI was presented on every trial. The experiment started when the participant felt confident with the FSR and the requested timing between pinches. The first block was always a motor-auditory block, followed by an auditory block. Overall, there were 7 blocks for the motor-auditory condition, 7 blocks for the auditory condition and 2 blocks for the motor condition. There was a motor block after the second auditory block and the fifth motor-auditory block. Each motor-auditory and auditory block had a duration of around 5 minutes. The duration of each motor block was of around 2 minutes. Participants could repeat the training if needed and take short breaks between blocks. Total duration of the experiment was around 1 hour, excluding breaks and preparation.

Force data acquisition

The FSR-signal was recorded by using the high level input of a SynAmps2 EEG amplifier (Compumedics NeuroScan, NC, USA), with a sampling rate of 1000 Hz, and online low-pass filtered at 200 Hz. The FSR-pinch was successfully triggering an event when it was exceeding 1.22 V, after a period under the threshold of at least 10 ms. Before analysis, a log-linear function was used in order to establish the applied force - FSR signal relationship (Interlink Electronics, FSR 400 Series Data Sheet). Thus, the force values were calculated from the signal by applying an exponential transformation.

Force processing and analysis

For each pinch the applied force was determined by the maximal peak in the corresponding FSR signal. An interval from 0 to 800 ms was set to search for the peak.

In the force analysis, we registered force-related response separately for pinches eliciting tones and pinches eliciting omissions in MA. In particular, we were interested in the force-related response to the first pinch (from now on, named “1st”), and in the force related to the second pinch and the third pinch (from now on, named Early Duplet, “Ed”), to the fourth pinch and fifth pinch (from now on, named Late Duplet, “Ld”), to the pinches following the fifth pinch (from now on, named “6th+”), and in the force belonging to the last event of each sequence (i.e., “last tone”, “last omission”; Figure 1).

Specifically, with the aim of analyzing the same events both in the force-related analysis and in the event-related potentials (ERPs) analysis, we included in the analysis the force values belonging only to tone sequences following five or more consecutive omission events and the force values belonging only to omission sequences following at least five standard tone repetitions in a row. Consequently, we disregarded target tones, and omission sequences following tone sequences containing target tones.

In order to investigate whether the pinch force changes gradually, as the event (i.e., tone, omission) is repeated within the micro-sequences, we performed a 1-way ANOVA, with a Position factor with 4 levels (1st, Ed, Ld, 6th+), separately both on the tones and on the omissions. The data used for the ANOVAs were normally distributed. Wilcoxon signed-rank tests were used to test for any difference between events of interest (i.e., 1st, Ed, Ld, 6th+, last event of the sequence), both within the MA condition and between MA and M conditions (e.g., for the contrast between omissions).

The first five pinches in each block were excluded. Moreover, the actions were considered potentially erroneously registered, as separated actions during the pinch-release, when they were registered within one second of the following action. These specific actions were manually rejected and not included in the analysis.

Electroencephalogram (EEG) acquisition

The EEG was continuously acquired at a sampling rate of 500 Hz from 61 Ag/AgCl scalp electrodes mounted on an elastic nylon cap (Easy-Cap, Herrsching, Germany) according to the 10% system (Nuwer et al., 1998). The signal was amplified by SynAmps 2 (Compumedics NeuroScan) amplifiers and recorded with Scan 4.4 (Compumedics NeuroScan), applying an online 0.05-100 Hz bandpass filter. An electrode placed on the tip of the nose was used as reference, and a ground electrode was placed in the middle of the forehead. In addition, an electrode was placed on each mastoid. Bipolar montages were used to record eye movements. The horizontal electrooculogram was recorded with two electrodes placed at the outer canthi of each eye, and the vertical electrooculogram was recorded with one electrode below the left eye and the Fp1 electrode.

EEG processing

EEG analysis was performed with EEGlab (Delorme & Makeig, 2004), and ERPs were visualized with Eeprobe (ANT). In order to run an independent component analysis (ICA) (Delorme & Makeig, 2004) the raw EEG data was first filtered with a 0.5 Hz high pass (windowed sinc FIR filter, Kaiser window, Kaiser beta 4.53351, filter order 734), and sections containing non-stereotypical artefacts were manually rejected from the continuous file. Subsequently, we applied a binary compiled version of runica (binica), that uses the logistic infomax ICA algorithm (Onton & Makeig, 2006). The extracted ICA weights were then applied to the original unfiltered raw datasets and components related to eye movements and muscle activity were pruned. On average, nine components (range: 6 - 13) were rejected per participant. After the ICA correction, the EEG was bandpass filtered from 0.5 to 30 Hz (windowed sinc FIR filter, Kaiser window, Kaiser beta 4.53351, filter order 734). Subsequently, epochs of 600 ms, time-locked to the FSR pinch-threshold, were defined (-200 to 400 ms). A baseline correction of 200 ms was applied (-200 to 0 ms). Epochs still containing large artefacts were rejected by applying a 60 μV maximal signal change per epoch threshold. Rejection rates ranged 2.14 - 4%. On average 79 trials were retained per event type, namely 1st, Ed, Ld, 6th+ (range: 51.7 - 106.4). Only the tone sequences following a sequence of at least 5 omissions were included in the analysis, therefore the target tone-sequences were not included in the ERP analysis.

ERP Analysis

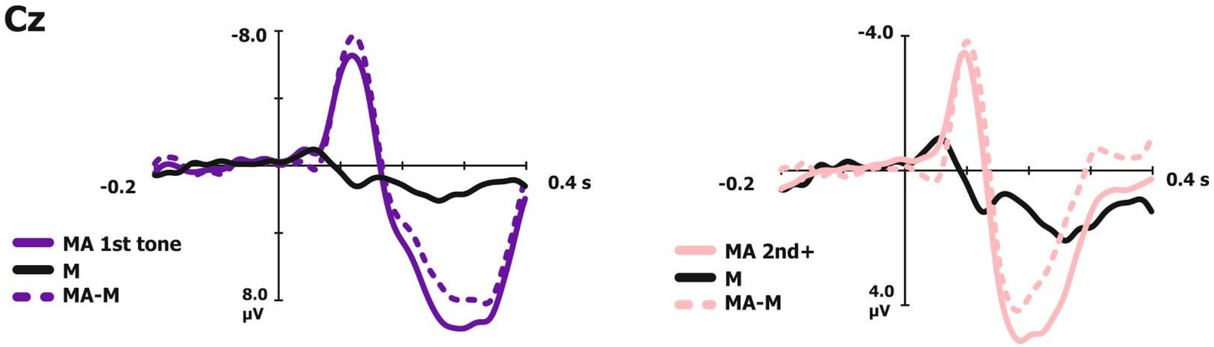

Figure 2. *ERPs indicating the motor-correction, related to the 1st and 2nd+ tones, which results from the subtraction between the MA and M conditions.*

We extracted ERPs, time-locked to each action (i.e., a pinch), separately for actions eliciting tones and omissions in MA, for tones in A, and for pinches in M.

In order to track the development of the action-tone association, we examined events at different tone-positions within the tone sequences separately. The ERPs in the MA condition were motor-corrected (from now on, the abbreviation MA will refer to the motor-corrected ERPs), subtracting the motor-ERPs recorded in the M condition from the motor-auditory response recorded in the MA condition (Figure 2). Subsequently, we compared the auditory ERPs for self-generated tones (MA) to the ERPs to externally-generated tones (A).

We focused on three auditory components, N1, P2 and P3, which might be influenced by the discovery and acquisition of new action-sound regularity representations. For each ERP we analyzed the mean amplitude values for the time-windows and electrodes indicated in Table 1. The auditory sensory components of interest, N1 and P2, were identified on the ERP elicited by the tone in the A Condition. The N1 component was identified as the largest negative peak occurring in the N1 latency range at the electrode Cz. The N1 subcomponent, Tb, was identified as the second largest negative peak identifiable after the tone onset on electrodes T7 and T8. The P2 component was identified at Cz as the largest positive peak occurring after the N1 component. The P3 component was identified on the difference waveform derived from the first tone of the sequence minus the 2nd+ tones of the sequence.

Typical self-generation effects

The first analysis aimed to check whether it was possible to observe the typical self-generation effects grouping all tones together except for the first one. Thus, we averaged all the tones following the 1st tone (2nd+) (Figure 1). We ran a t-test contrasting the 2nd+ in the two conditions (i.e., MA, A) on the amplitude values for each of the time-windows of interest (Table 1).

Table 1. *ERP components of interest with detailed information regarding the waveform used for their identification and electrodes and time-windows for the analysis.*

Auditory components	**Waveform for Identification**	**Electrodes for analysis (ROI)**	**Time-window for analysis (ms.)**
N1 (for the 1st)	A1st	Cz	116 ± 10
N1 (for the 2nd+)	A2nd+	Cz	102 ± 10
P2	A1st	Cz	200 ± 20
P3	A1st - A2nd+	Cz	306 ± 40
Tb	A2nd+	FT7, T7, TP7, T8, FT8, TP8	136 ± 10

Tracking the action-effect association

The aim of the second analysis was to track the different auditory ERP components along the tone-sequence focusing on the first tone (1st), the subsequent two duplets of tones, that are the Early duplet (Ed) averaging the 2nd and 3rd tone, and the Late duplet (Ld) averaging the

4th and 5th tone, and the tones following the 5th tone (6th+) (Figure 1). Then, we ran an ANOVA with the factors Condition (MA, A) and Position (1st, Ed, Ld, 6th+) on the amplitude values for each of the time-windows of interest (Table 1).

Time-windows selection

In order to test if there was any progressive latency-shift among tone positions, we ran an ANOVA with the factors condition (MA, A) x position (1st, Ed, Ld, 6th+), on the N1 peak latencies. We did this analysis applying the jackknifing method (Miller et al., 1998). Specifically, we extracted the latency values of the most negative peak in the time-window between 80-160 ms from the jackknife-derived ERPs. The results from the analysis on the jackknife values showed a significant N1-latency shift to repeated tone presentation (position: $F_{(3,60)} = 16.9$, $p < .001$) which was equally present in both conditions (cond: $F_{(1,20)} = 1.20$, $p = .28$; int: $F_{(3,60)} = 0.36$, $p = .78$).

Post-hoc analysis between the First tone and the rest of positions (Ed, Ld, 6th+), showed that the latency of the N1 peak in response to the first tone was significantly longer compared to the rest of tone positions, in both conditions (all $p \leq .0015$) (in A: 1st - Ed: $t = 3.8918$, $p < .001$; 1st - Ld: $t = 4.2105$, $p < .001$; 1st - 6th+: $t = 4.1885$, $p < .001$; in MA: 1st - Ed: $t = 3.9429$, $p < .001$; 1st - Ld: $t = 6.5927$, $p < .001$; 1st - 6th+: $t = 3.7149$, $p = .0015$). For this reason we defined a specific N1 time-window for the 1st tone response (106-126 ms) centered around the peak at Cz on the 1st response.

The analysis between the Early Duplet and the rest of tone positions (Ld, 6th+) did not reveal any significant difference in the N1 peak latency (all $p > .05$). (in A: Ed - Ld: $t = 0.7347$, $p = .4715$; Ed - 6th+: $t = 0$, $p = 1$; Ld - 6th+: $t = -1.2850$, $p = 0.2142$; in MA: Ed - Ld: $t = 1.8446$, $p = 0.08$; Ed - 6th+: $t = 0.4457$, $p = 0.66$; Ld - 6th+: $t = -0.4764$, $p = 0.6392$). For this reason we selected a unique N1 time-window for all the other tone-position responses (92-112 ms), centered around the peak at Cz on the 2nd+ tone-position responses.

N1 peak latencies were also compared between the two conditions (MA and A) for each tone position of interest. The results did not indicate a significant difference between conditions for any of the tone positions (1st: $t_{(19)} = 1.3680$, $p = .1873$; Ed: $t_{(19)} = 0.8241$, $p = .4201$; Ld: $t_{(19)} = 0$, $p = 1$; 6th+: $t_{(19)} = 0$, $p = 1$).

Regarding the P2 component, it was not possible to differentiate the most positive peak in the P2 time interval from the most positive peak in the P3 time interval in response to the 1st. For this reason, we used the same P2 time-window for all the tone positions of interest, which was identified as the largest positive peak occurring in the typical P2 time range in response to the A2nd+ (180-220 ms) (Figure 4, panel B).

Results

Behavioral results

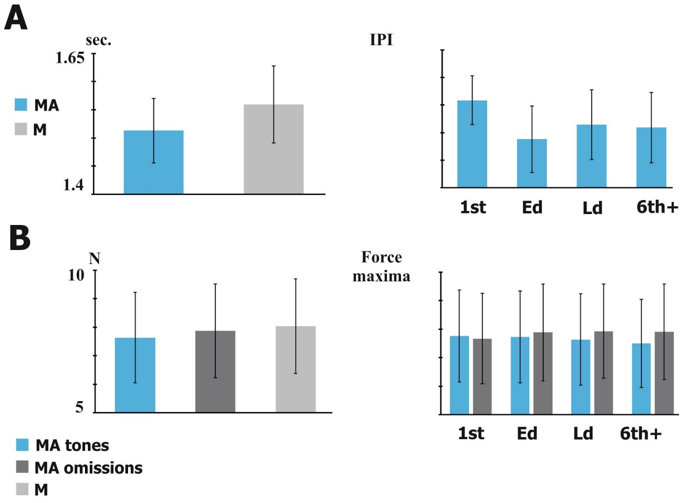

Figure 3. *Behavioral results.* ***Panel A.*** *Inter-pinch interval (IPI) results. On the left: bar plots related to the Motor-Auditory (MA) and Motor (M) conditions. On the right: bar plots representing the four tone positions of main interest in MA.* ***Panel B.*** *Results from the force-related analysis. Left: bar plots representing the tones and omissions in MA and the omissions in M condition. Right: bar plots representing the four positions of main interest in the MA condition.*

Inter-pinch interval

Participants were able to maintain a stable inter-pinch interval (IPI) during the experiment, as required by the experimenter in the instructions. Indeed, the mean IPI was of 1.53 ± 0.12 sec, across conditions. A one-way ANOVA with a 3 levels Condition factor (MA, A, M) revealed no significant difference between conditions (Cond: $F_{(2,38)} = 2.846$, $p = .107$) (Figure 3; panel A, on the left).

We intended to check whether the tone position within the tone-sequence had an impact on the IPI, thus, we first compared the 1st with the 2nd+, within the tone-sequence. A t-test revealed a significant difference between the IPI related to 1st and 2nd+ ($t_{(19)} = 3.012$, $p = .007$). Then, we checked more into details for the IPI related to the different tone-positions along the tone-sequence, through further t-tests revealing that the IPI related to the 1st tone was significantly larger compared to the following tone-positions (1st - Ed: $t_{(19)} = 3.820$, $p = .001$; 1st - Ld: $t_{(19)} = 2.198$, $p = .041$; 1st - 6th+: $t_{(19)} = 2.604$, $p = .017$; Ed - Ld: $t_{(19)} = -2.724$, $p = .013$; Ed - 6th+: $t_{(19)} = -1.437$, $p = .167$; Ld - 6th+: $t_{(19)} = 0.425$, $p = .675$). These results confirm that a robust pinch-omission association was made before the beginning of the tone-sequence. Assuming that when the 1st tone was following a sequence of at least 5 omissions, most likely the participants were predicting to keep eliciting an omission. Indeed, the IPI related to the 1st was significantly larger compared to the IPI related to each of the following tone-positions, but we did not find any difference between the IPI related to the 1st pinch-tone in MA and the IPI related to pinches generating omissions in the M condition (M - 1st tone: $t_{(19)} = 0.033$, $p = .974$) (Figure 3; panel A, on the right).

Force

The results of the force-related analysis are depicted in Figure 3 (panel B) and are reported in Table 2. Moreover, Table 3 shows the means and standard deviations for each of the events of interest.

Table 2. *Results from the force-related analysis. First column: main comparisons of interest; Second column: results from force maxima values.*

Comparison	***Force results***
MA Standard Tones - MA Om.	*Z* = -2.427, ***p* = .015***
MA Om. - M Om.	*Z* = -0.821, *p* = .411
Last Om. - 1st tone	*Z* = -2.053, ***p* = .040***
Last Om. - Ed tones	*Z* = -0.859, *p* = .391
Last Om. - Ld tones	*Z* = -1.195, *p* = .232
Last Om. - 6th+ tones	*Z* = -1.867, *p* = .062
Last Tone - 1st omission	*Z*= -0.784, *p* = .433
Last Tone - Ed omissions	*Z* = -1.792, *p* = .073
Last Tone - Ld omissions	*Z* = -1.979, ***p* = .048***
Last Tone - 6th+ omissions	*Z* = -2.053, ***p* = .040***

Table 3. *Results from the analysis on force maxima values, with means (M) and standard deviations (SD).*

Events of interest	***Force results***
MA Standard Tones	$M = 7.64$, $SD = 3.18$
MA Om.	$M = 7.87$, $SD = 3.29$
M Om.	$M = 8.04$, $SD = 3.33$
1st tone	$M = 7.76$, $SD = 3.22$
Ed tones	$M = 7.73$, $SD = 3.23$
Ld tones	$M = 7.64$, $SD = 3.22$
6th+ tones	$M = 7.50$, $SD = 3.10$
last tone	$M = 7.60$, $SD = 3.20$
1st omission	$M = 7.67$, $SD = 3.17$
Ed omissions	$M = 7.89$, $SD = 3.40$
Ld omissions	$M = 7.93$, $SD = 3.30$
6th+ omissions	$M = 7.91$, $SD = 3.35$
last omission	$M = 7.93$, $SD = 3.35$

In the MA condition, overall pinch force within the tone micro-sequences was significantly reduced, compared with the force registered overall in the omission sequences ($Z = -2.427$, $p = .015$). As expected, no significant difference between conditions was found comparing all the omissions in the MA condition with all the omissions in the M condition ($Z = -0.821$, $p = .411$).

The 1-way ANOVA on the omissions, aimed to check for any gradual force change, did not show a position effect ($F_{(3,57)} = 1.039$, $p = .347$) nor a linear trend ($F_{(1,19)} = 1.316$, $p = .266$).

The 1-way ANOVA on the tones did not show a position effect ($F_{(3,57)} = 2.936$, $p = .081$), nor a significant linear trend ($F_{(1,19)} = 3.728$, $p = .069$).

In order to check for any force change due to the beginning of a new event sequence, we compared the last omission of the sequence and the following tone positions of interest. We found a significant difference only between the last omission and the first tone position (1st tone: $Z = -2.053$, $p = .040$; Ed: $Z = -0.859$, $p = .391$; Ld: $Z = -1.195$, $p = .232$; 6th+: $Z = -1.867$, $p = .062$). We also compared the last tone of the sequence and the following omission positions of interest. We found a significant difference only starting from the late duplet of omissions (1st om.: $Z = -0.784$, $p = .433$; Ed om.: $Z = -1.792$, $p = .073$; Ld om.: $Z = -1.979$, $p = .048$; 6th+ om.: $Z = -2.053$, $p = .040$).

ERP results

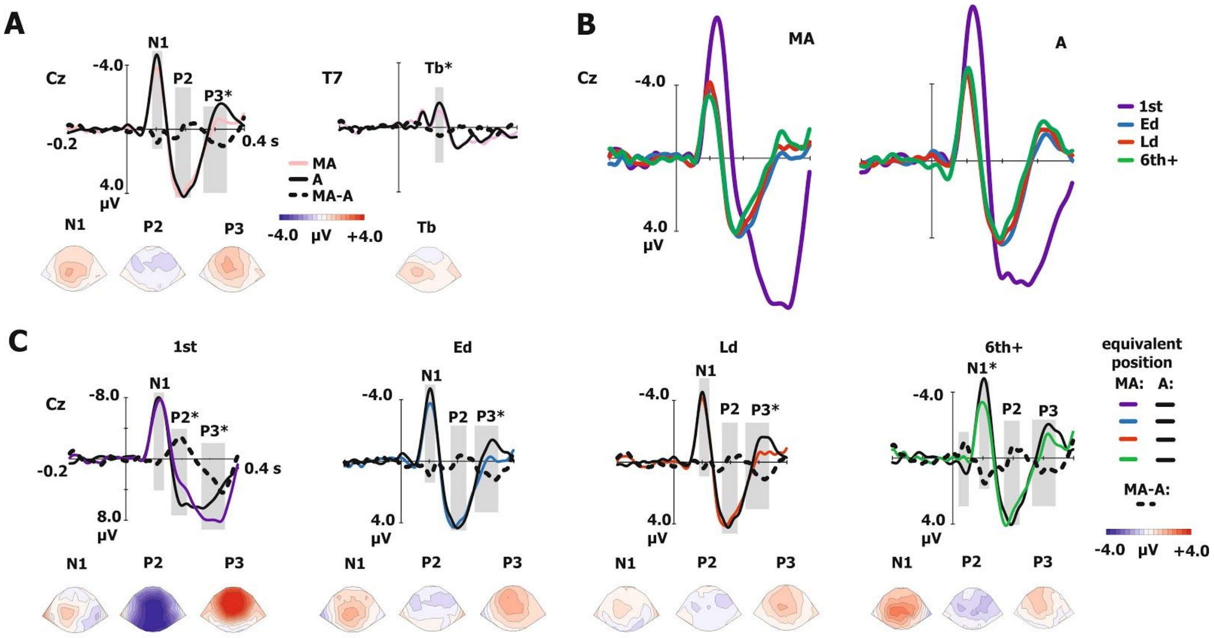

Figure 4. *Results from the EEG analysis, showing the ERPs and the scalp maps on the difference between MA and A conditions (N1, P2, P3 time-windows). Significant values: * p ≤ 0.05.* ***Panel A.*** *Focusing on self-generation effects. ERPs in response to the tones following the first one (2nd+) at Cz and at T7.* ***Panel B.*** *Focusing on latency differences between the MA and A condition. ERPs at Cz.* ***Panel C.*** *Focusing on each tone position of interest (i.e., respectively to 1st, Ed, Ld, 6th+) in MA, on the equivalent tone position in A, and on the difference wave between conditions (MA-A).*

Typical self-generation effects

Contrasting the 2nd+ response from the MA and from the A condition, the results did not show any suppression effects for self-generated tones in the N1 time-window (N1: $t_{(19)} = -2.032$, $p = .056$), nor any significant difference between conditions in the P2 time-window (P2: $t_{(19)} = 0.624$, $p = .540$). The results showed a statistically significative enhanced P3 and reduced Tb response for the self-generated tones compared to the externally-generated tones (P3: $t_{(19)} = 2.565$, $p = .019$; Tb: $t_{(19)} = -2.290$, $p = .034$) (Figure 4, panel A).

Tracking the action-effect association

The results from the second analysis are depicted in Figure 4 (panel C).

N1

The results from the ANOVA did not reveal an N1 attenuation for self-generated tones (cond: $F_{(1,19)} = 3.652$, $p = .071$). The N1 amplitude decreased with tone repetition (pos: $F_{(3.57)} = 33.797$, $p < .001$), equally in both conditions (int: $F_{(3.57)} = 1.857$, $p = .147$). Single t-tests on each position of interest revealed that the N1 attenuation was significantly different between conditions only in response to the 6th+ (tones after the 5th position) (Ed: $t_{(19)} = -1.904$, $p = .072$; Ld: $t_{(19)} = - .808$, $p = .429$; 6th+: $t_{(19)} = -2.475$, $p = .023$).

Tb

The results from the ANOVA in the Tb time-window showed a significant difference between positions (Condition: $F_{(1,19)} = 2.645$, $p = .120$; Position: $F_{(3,57)} = 31.675$, $p < .001$; Interaction: $F_{(3,57)} = 0.525$, $p = .589$). Single t-tests between conditions did not show any significant result, although Tb seems to be descriptively smaller in MA than in A, in all cases except the first tone.

P2

The results from the ANOVA showed a condition x position interaction (cond x pos: $F_{(3,57)} = 5.335$, $p = .007$). The linear trend for the cond x pos interaction was also significant ($F_{(1,19)} = 6.972$, $p = .016$). Post-hoc t-tests showed significant condition effects only for the 1st (1st: $t_{(19)} = 3.621$, $p = .002$; Ed: $t_{(19)} = .375$, $p = .711$; Ld: $t_{(19)} = .497$, $p = .625$; 6th+: $t_{(19)} = .850$, $p = .406$).

P3

The condition x position ANOVA in the P3 time-window revealed a significantly larger P3 response in the MA condition than in the A condition ($F_{(1,19)} = 16.938$, $p = .001$). P3 response was significantly different across positions ($F_{(3,57)} = 67.276$, $p < .001$). We found an interaction between condition and position effects ($F_{(3,57)} = 5.294$, $p = .017$). We ran post-hoc analysis in order to check whether the condition effect was significant at each position. This analysis revealed that the self-generated tones elicited a more positive response in the MA condition for all positions except the 6th+ (1st: $t_{(19)} = -3.729$, $p = .001$; Ed: $t_{(19)} = -3.089$, $p = .006$; Ld: $t_{(19)} = -2.391$, $p = .027$; 6th+: $t_{(19)} = -1.590$, $p = .128$). Further post-hoc analysis revealed that the response to the 1st elicited a larger difference between conditions, than in the Ed, Ld & 6th+ (1st - Ed: $t_{(19)} = 2.530, p = .020$; 1st - Ld: $t_{(19)} = 2.435, p = .025$; 1st - 6th+: $t_{(19)} = 2.551, p = .020$; Ed - Ld: $t_{(19)} = .456, p = .654$; Ed - 6th+: $t_{(19)} = .735, p = .472$; Ld - 6th+: $t_{(19)} = .410, p = .686$).

Discussion

This study aimed to investigate and track the buildup of motor-auditory regularities, in a dynamic context in which there was no stable regularity. To this aim we ideated a novel experimental paradigm, simulating a "broken device" with which it was possible to generate sounds intermittently. We were particularly interested in disentangling the role of the motor action and repetition in this learning process.

We hypothesized to find bidirectional effects due to the buildup of action-effect associations. Consequently, we expected to observe behavioral (i.e., inter-pinch interval and force pressure) and electrophysiological changes (i.e., in N1, P2 & P3 time-windows) as indices of this evolution. Specifically, at a behavioral level, we expected to find a motor adaptation to the sensory consequence of each action, resulting in slowing effects due to the first tone generation and gradual force changes (i.e., reduction by tone repetition & enhancement by omission repetition). At an electrophysiological level, we hypothesized to find effects indicating the detection of a new action-effect association after sequences of at least five consecutive pinch-omissions in a row (i.e., sensory enhancements to the first tone), then indicators of the updating of the internal model (i.e., a gradual sensory attenuation with tone-repetitions along the tone sequence), and finally the proof of the action-effect consolidation (i.e., a significant sensory attenuation).

In line with our hypotheses, we observed bidirectional effects of the pinch-effect association. Indeed, behaviorally the first tone after the omission sequence was affecting the timing of the following action (i.e., faster action after the first tone) and force online adjustments (i.e., force reduction for tones). Concomitantly, as electrophysiological response to the first tone, we registered an MMN-like and a P3 response which were particularly enlarged for self-generated tones compared to externally generated tones. Subsequently, we observed an enlarged P3 response also in response to the tones following the first tone, which stopped to be significantly different across conditions only after the five tone micro-sequence, when we observed an N1 suppression for self-generated tones. Overall, the results suggest that the amount of omission-repetitions (around 10 per sequence) was enough to generate a strong pinch-omission association, and expectation, indicating that the participants were expecting to generate a further omission when actually the pinch was generating the beginning of a tone-sequence.

Detection

The first tone position of the tone micro-sequence played a peculiar role within each block. Indeed, this tone position interrupts the previous and consolidated pinch-omission association.

Behaviorally, we found a significant shorter action timing after the first pinch of the tone sequence compared to the pinch timing related to the subsequent tone positions. This finding does not go in line with the literature which showed slowing effects due to unexpected events (Fogelson et al., 2009; Notebaert et al., 2009; Iwanaga & Nittono, 2010). The force results revealed a significant reduction in force pressure in response to the first tone compared to the last omission of the sequence. This difference in force could reflect an online force adjustment due to the detection of a new pinch-effect association. Indeed, as described by Neszmélyi and Horváth (2018) only impulse values are not affected by online adjustments. To be noted that, contrasting the force related to the last tone with the force related to the first omission we did not find any significant difference. These findings could indicate that the novel pinch-tone association needs further repetitions in order to start generating significant force pressure changes.

The electrophysiological measurements showed results in line with the behavioral ones. Indeed, similarly to oddball and roving standard studies, the first tone could be interpreted as the classical "deviant tone" within sequences of standard tones. Assuming that a strong association between pinches and no tone was generated before the beginning of the tone-sequence (similarly to the standard events), the difference between the MA and A condition in the P2 time-window could be interpreted as due to the "surprise effects" deriving by the first tone. Indeed, we found a delay and an MMN-like response followed by an enlarged P3 response to the first tone of the sequence, which suggest the detection of the violation of an established action-effect association (i.e., previous pinch-omission association). Additionally, the scalp topographies show a central negativity in the P2 time-window and a typical fronto-central positivity in the P3 time-window in response to the first tone. This interpretation of the MMN-like response would be in line with the interpretational framework proposed by Winkler in 2007, affirming that an effect of the deviant tone is also to mismatch-predictions and to trigger the update of the predictive internal models. Furthermore, the presence of the P3 response confirms the context-updating model proposed by Donchin and Coles (1988), indicating the necessity of updating the incorrect internal model after the deviant detection.

Alternatively, we could interpret the self-generated first tone as more surprising compared to the externally-generated first tone, because of the existence of a strong association between action and no tone, formed during the long sequence of preceding silent presses. Thus, the first tone would be more surprising in the self-generated condition than in the externally-generated condition, as it violates this specific "no tone" expectation generated with the previous actions. This finding could indicate that the action in itself has an important role in fostering an action-effect association. This hypothesis would be in line with Nittono's interpretation of the action-effect anticipation (2006). Indeed, Nittono and colleagues found that, comparing self-generated and externally-generated stimuli in oddball task studies, the frontocentral P3 in response to deviant stimuli was more marked in motor-auditory than in auditory conditions (Nittono & Ullsperger, 2000; Nittono, 2004, 2006; Iwanaga & Nittono, 2010). A similar study from Waszak and Herwing (2007) confirmed this idea.

Alternatively, the externally-generated first tone could be interpreted as more surprising compared to the self-generated first tone because, in the Auditory condition, lacking the action, it is more difficult to predict the exact moment in which the sound will be elicited. Consequently, the enlarged externally-generated response in the P2 time-window could be interpreted as an early P3 component (i.e., P3a). However, this interpretation is not supported by the scalp maps.

Updating

Moreover, we hypothesized that the event repetition would have caused a gradual decrease of the force in the case of pinches eliciting tones and a gradual increase of the force in the case of pinches eliciting omissions. Contrasting with our expectation, the results from the ANOVAs did not show a significant gradual change in the force applied to the pinches generating omission repetitions nor to the pinches generating tone repetitions. These findings suggest that probably several repetitions of the same action-effect association are necessary in order to observe significant repetition-related effects, as the ones reported in a previous study (Neszmélyi & Horváth, 2017).

Comparing the force related to the last omission with the force related to the tone positions after the 1st tone (i.e., Ed, Ld, 6th+) we did not find any significant force decrease for the pinches generating tones than for the pinches generating omissions. However, contrasting the force related to the last tone with the force related to the omission positions after the first omission we found a significant force enhancement only starting from the late duplet position on. These findings suggest that a minimum number of five repetitions is needed in order to establish a new contingency that may result in significant force adjustments when broken.

From the ERP results of this study, we can also notice that the P3 response is still present in response to the tone repetitions following the first tone, at least until the late duplet, or rather until the end of the five tone micro-sequence. Indeed, the effect in the P3 time-window starts to not be significant only from the sixth position on. Examining the progression of the P3 response along the tone sequence, the ERP response in the P3 time-window becomes progressively attenuated with tone repetitions, but a positivity is still present along the tone sequence. Specifically, compared to externally generated tones, the presence of an enlarged P3 response to self-generated tones following the first position could confirm that the action facilitates the buildup of a more robust internal model compared to the externally generated ones. Assuming that the P3 response involves an updating of the internal model (Donchin & Coles, 1988), this finding could indicate that a certain amount of tone repetitions is needed in order to reinforce the pinch-tone association (local rule) and make this association as robust as the pinch-omission association (general rule).

Consolidation

An effect due to the repetitive absence of sensory consequences of the action was indicated by the force adaptation. That is, in line with our expectation and with literature (Neszmélyi & Horváth, 2017), the overall pinch force applied during the tone sequences was significantly reduced compared to the force used during the omission sequences.

Although the inter-pinch interval was constant across conditions, we observed a longer IPI in the Motor condition suggesting that, in general, the absence of auditory feedback tends to slow down the regular pace between pinches.

In the first analysis we averaged all the tones following the first tone, within the micro-sequence, with the main aim of checking for self-generation effects. In this regard, we only found a trend towards the N1 suppression in response to the 2nd+ in the MA, which is in line with the results reported in literature that overall the response to self-generated tones is attenuated compared to the response to externally-generated tones (Schäfer & Marcus, 1973; Martikainen et al., 2005; Baess et al., 2008; Aliu et al., 2009; Baess et al., 2009; Hesse et al., 2010; Baess et al., 2011; Horváth, 2013 a, b; SanMiguel et al., 2013b; Saupe et al., 2013; Timm et el., 2013; van Elk et al., 2014; Mifsud & Whitford, 2017; Neszmélyi & Horváth, 2017). However, we did not find a robust N1 attenuation, probably because in our study there is not a stable action-effect regularity and because we examined only short tone micro-sequences, contrariwise to the typical contingent paradigms.

The second analysis adds more information regarding the effects of the tone repetitions after the first tone. Indeed, we found an N1 suppression only in response to the tones following the fifth tone. This finding suggests that a minimum of five repetitions is needed in order to elicit significant self-generation effects (N1 attenuation at Cz), independently from the probability of eliciting further tones after the fifth. This result seems to confirm Shanks' proposal (1993) about the role played by the "relative predictiveness" between an action and the outcome in determining the learning about their relationship.

Taking into consideration both the findings on the N1 and on the P3 response it seems plausible that the updating of the internal model, represented by the presence of an enlarged P3 response, stops when the action-effect association is robust enough to elicit significant N1-suppression effects.

Assuming that each tone belonging to the five tone-sequence had an equal probability of being generated (i.e., 100%), we expected to find a gradual N1 suppression along the five tone sequence, with a maxima for self-generated tone in response to the late duplets (which includes the fifth tone). Differently from our hypothesis, the electrophysiological responses did not show this pattern. One possible explanation of this finding could be that the actual experience of generating a tone has a bigger impact than the probability in itself of eliciting tones. Or else, it could indicate a cognitive difficulty of tracking the actual probability of eliciting each tone with pinch repetition.

It is notable that in this task context, we did not observe any coincidence N1-suppression effects (Horváth et al., 2012). The lack of coincidence effects could be due to the fact that there is always a pinch-effect contingency in the motor-auditory condition, although most of the times the pinch does not elicit any “apparent” effect for the participant. That is, the main difference between this design and the coincidence designs is that here there are no tones presented in isolation during the motor-auditory condition. Thus, although the probability of a pinch eliciting a tone is low, all tones presented are always elicited by pinching. Moreover, the lack of coincidence effects can represent also a demonstration that the temporal contiguity is only one of the key elements determining the causal learning, but the N1-suppression is in reality the manifestation of a learning process which allow the creation of internal models related to an action-effect associations.

The experimental paradigm used in this study aimed also to overcome some of the limitations which usually bias the contingency arrangement (Horváth, 2015). Indeed, in the Motor-Auditory and Auditory condition, the inter-pinch interval was equal, because each auditory block was an exact reply of the previous motor-auditory one; in both conditions, the task consisted in counting target tones, implying the same cognitive commitment; in line with the instructions, the inter-pinch interval was constant in the Motor-Auditory and Motor condition (i.e., 1.5 sec.); lastly, through the selection of the force sensitive resistor it was possible to avoid any transient mechanical sound (Horváth, 2014).

Overall, this study helps to examine more attentively the evolution of the auditory regularities, taking into consideration also the contribution of the motor action.

Moreover, this study goes in line with the two-stage model proposed by Elsner and Hommel (2001), even adding further empirical evidence to explain more in details each step of the causal learning. Indeed, based on the findings collected by this research, we could propose a three-stage model (i.e., detection, updating, consolidation).

However, we cannot exclude the presence of confounding effects due to the presentation order, considering that, most of the times, the auditory blocks were following the motor-auditory ones. One of the limitations of this study is also represented by the number of tone repetitions (i.e., five tone repetitions). Indeed, we can only argue that a minimum of six repetitions is necessary to elicit a significant N1 suppression, but we cannot establish at which exact position, after the fifth position, the tone repetition starts to foster the action-effect association.

This study could represent a contribution to bridge the results from different paradigms used so far to investigate motor-auditory and auditory regularities. In general, we believe that the actual bridge (i.e., the common finding between all these studies) is represented by the formation of internal associative models about regularities which allow to actively perceive and interact with the surrounding world, keeping an optimal level of energy (i.e., "why" humans learn to generate and to modify action-effect associations). Specifically, the results confirm the key role of the repetition (Cowan et al., 1993) and of the motor action in fostering the learning process necessary to detect auditory regularities, and the interconnection between an action and its sensory consequence (i.e., "how" humans learn to generate and to modify action-effect associations). Along the bridge, the MMN can be seen as a common index to study the internal representations in different domains (Cowan et al., 1993). Further research works have to be carried out in order to keep disentangling the specific mechanisms at the base of the contingency and the contiguity, keeping in mind that probably both factors are fundamental in this learning process. Only afterwards, it will be possible to better understand the psychophysiological indices related to each stage of this process.

Overview

Studies investigating motor-auditory and auditory regularities converge in considering the mismatch negativity-like and P3-like responses as indicators of deviant detection, and the attenuation of neuronal responses as indicator of predictable auditory regularities. The aim of this study was to investigate and track the buildup of auditory regularities, and the motor act contribution, in dynamic contexts in which there was no stable regularity.

Participants pressed regularly a button that mostly produced no effect. Occasionally, the button presses elicited a micro-sequence of at least five tones time-locked to consecutive presses establishing a new regularity. We measured behavioral responses (i.e., inter-press interval and force values) and the auditory ERPs typically related to the deviant detection (MMN, P3) and the auditory regularity consolidation (N1 attenuation). Specifically, we looked at various positions within the tone sequences and compared the auditory ERPs recorded in the motor-auditory condition to those recorded in a passive replay of the same tone sequences.

The results show that the first tone caused shorter action timing and elicited an MMN-like, and a P3-like response. The P3-like response emerged also in response to the following tones and stopped to be significant only in response to the tones following the five self-generated tone repetitions, when we registered a maxima N1 attenuation. The results suggest that a robust action-no tone association was created during the long silent presses. Indeed, the first tone of the sequence acted as a deviant tone, which triggered the updating of the internal model along the whole five tone sequence, until the model consolidation. The results indicate the interconnection between an action and its sensory consequence.

5. GENERAL DISCUSSION

For a long time, it has been thought that the brain was passively registering external sensory inputs. Nevertheless philosophers, physiologists and psychologists have always nursed a great interest about human perception and about the contribution of previous interactions with the environment in modulating how humans perceive sensory inputs. More recently, the literature has shown the active role of the brain in modulating sensory processing. Action-effect associations and predictions about the causes of our sensory experiences seem to be particularly relevant for human survival, since they allow the formation of internal models which guarantee a stable perception of the physical world around us (Friston, 2005; Bar, 2007; Winkler et al., 2009; Albright, 2012). However, the empirical studies investigating perception have been carried out through different lines of research and paradigms.

A large body of studies has investigated auditory perception. The results, from studies focusing on the auditory-alone domain and on the motor-auditory interaction, converged in reporting several modulations of the internal models on human perception. However, the mechanisms behind the modulations found on behavioral and electrophysiological responses are still unclear. This thesis represents an attempt to shed some light on the phenomena underlying these modulations, providing a unified explanations of the results from previous studies.

Thus, the general aim of the present thesis was to uncover the mechanisms underlying associative and predictive processing, focusing specifically on action-effect pairings. We modified the typical contingent self-generation paradigm, in order to be able to bridge the results from different research lines, testing the interpretations provided by the predictive coding theory.

In Study I we focused on consolidated action-effect associations (predictive processing), and explored the electrophysiological indices related to the expectation-effects. Specifically, we first analyzed all the components which have been described in the literature as related to the extraction of auditory regularities, then we investigated consolidated expectation-match and -mismatch effects, in order to uncover if they were manifestations of the same underlying process. Furthermore, we were interested in studying the endogenous responses (driven by the omissions), recorded in a predictable and in an unpredictable condition, which should indicate respectively the presence or absence of expectation-related effects modulating the stimulus-evoked responses (driven by the sound). In Study II we focused on the buildup of action-effect associations (associative processing). Specifically, we explored how contingency- and coincidence-related effects modulate the electrophysiological indices. Moreover, we were particularly interested in examining both the contribution of the motor action and of action-effect repetitions in fostering an expectation, and the existence of bidirectional effects between an action and its sensory consequence. The expectation-related effects were investigated at behavioral level by means of action timing (Study I and Study II) and force maxima peaks (Study II), and at an electrophysiological level by means of ERPs (Study I and Study II). Overall, the results collected in this thesis show that associative and predictive processing modulate the behavioral and electrophysiological responses. Additionally, these findings suggest that the motor action and its sensory consequences are interrelated, and characterized by bidirectional effects, confirming the proposal launched by the ideomotor theory (Greenwald, 1970; Prinz, 1987; Hommel et al., 2001; Kunde et al., 2004). That is, not only an action can modulate the processing of a sensory consequence, but this consequence can affect the behavioral response itself. Therefore, we decided to discuss in this section separately the modulations found at a behavioral level (i.e., on action parameters), and the modulations found at an electrophysiological level (i.e., on sensory ERPs). Finally, we attempted to relate all the results collected in Study I & II.

MODULATIONS FOUND AT BEHAVIORAL LEVEL

Overall, the modulations found on the action parameters suggest that the sensory consequence of the action can affect the action itself. However, the results showed a different impact of the sensory consequence on the action timing (i.e., inter-press interval), depending on the specific task settings of each study. Indeed, in Study I we observed that unexpected sensory consequences of the action caused slowing effects on the subsequent action, confirming the results reported in literature (Fogelson et al., 2009; Notebaert et al., 2009; Iwanaga & Nittono, 2010), and indicating the attentional orientation to the "deviant stimulus" (Barcelo et al., 2006; Notebaert et al., 2009). However, in Study II we observed faster action timing after unexpected sensory consequences (i.e., first tones). This divergent finding could be related to the different aims and types of analysis computed in each study. Indeed, in Study I, we averaged all the unexpected events, including sounds and omissions, independently from their position; conversely, in Study II we focused only on the first tone after more than five consecutive omissions. Therefore, the contrasting results could be explained as due to the inclusion of the omissions in the analysis of Study I, but this hypothesis seems to be disconfirmed by the results suggesting that the action timing after sounds is longer, than after omissions (Study I).

A further possible explanation is that the tones were used by the participants as auditory feedback guiding towards a regular pace between actions, as required in the instructions (Study II: required interval between presses of ca. 1.500 ms). Thus, the results on the action timing observed in Study II could be interpreted as an attempt to maintain the right pace, which was more difficult in the repetitive absence of sensory inputs (i.e., in the case of omissions) associated with the action, characterizing Study II. Further behavioral results on the action force seem to support the idea that the sensory consequence elicited by the action can modify the action itself. Indeed, we found that in the case of the presence of a sound the overall force associated to the action was significantly decreased (Study II) compared to the case in which the auditory stimulation was omitted. Moreover, although participants were expecting a further omission as consequence of the action after long sequences of pinch-omission (Study II), we observed that the force related to the last pinch-omission was significantly higher than the force related to the first pinch-tone, suggesting the presence of online adjustments (Neszmélyi & Horváth, 2017; Horváth et al., 2018).

Moreover, in Study II, we aimed at studying the role of repetition in the buildup of an action-effect association, examining each position of interest. Although the results seem to point towards a gradual force attenuation in the case of pinch-tone associations and towards a gradual force enhancement in the case of pinch-omission associations, conversely to our expectations, we have not found a significant gradual force change within the micro-sequences. Probably a higher number of repetitions is needed in order to observe significant effects along the positions, indeed the repetition-related effects reported in a previous study were referring to positions within a block and not within a micro-sequence (Neszmélyi & Horváth, 2017), thus this could explain why we detected only an effect related to the actual presence or omission of the auditory input, and not to the repetition per se. The action timing did not show any modulation due to the repetition in the case of pinch-tone, confirming that participants were committed in maintaining a regular pace between pinches, independently from the repetition of this association.

MODULATIONS FOUND AT ELECTROPHYSIOLOGICAL LEVEL

The results from this thesis reflect the intrinsic characteristics of the experimental designs and contrasts we selected for each study. In Study I we focused on established action-effect associations (predictive processing), comparing a predictable versus an unpredictable condition. Indeed, the only difference between conditions should reflect the degree of predictability between conditions, since all the events were self-generated by a button-press. Whereas, in Study II we focused on the buildup of action-effect associations in a dynamic context, comparing self-generated versus externally generating sounds. Therefore in Study II the degree of predictability was not manipulated, the difference between conditions should reflect the motor contribution in fostering an expectation, hence also a higher degree of control about the onset of the sound elicited by the action.

In both studies, the results from the electrophysiological measurements suggest that the modulation of exogenous and endogenous ERPs could be related to expectation-related effects, although they seem to be more consistent with the modulation of endogenous responses, than of the exogenous responses, reported in literature. Moreover, the results seem to suggest the existence of different stages characterizing the modulations of the neuronal responses (Winkler et al., 2005; Baldeweg, 2006; Bekinschtein et al., 2009; Diekhof et al., 2009; Wacongne et al., 2011).

In both studies we found a mismatch negativity-like (MMN-like) response to stimuli violating the expectations (expectation-mismatch effects; Study I: unexpected sounds; Study II: first tones), confirming the idea that this endogenous component indicates the detection of stimuli breaking down a regularity (Escera & Malmierca, 2014), and that it is possible to observe this component also in cross-modalities (Winkler et al., 2009), such as in the case of motor-auditory associations, suggesting that the MMN is a key element to study the internal models in the auditory domain as much as in motor-auditory interactions.

Moreover, the findings reveal that the overlap between the P2 time-window and the MMN (Study I and II) and the RP time-windows (around 200 ms) (Study I) can impede to clearly isolate expectation-related effects separately for exogenous and endogenous components, suggesting that probably the modulation observed in previous studies in the P2 time-window was a reflection of the modulation of other endogenous components, such as the RP (Baldeweg et al., 2004; Haenschel et al., 2005; Baldeweg et al., 2006).

Indeed, in Study I, we isolated another endogenous response such as the repetition positivity (RP) (Haenschel et al., 2005), which was modulated by consolidated expectation-match effects (Study I: expected sound), but concomitant to the MMN-like. Particularly, selecting a wider time-window, we noticed that the RP and the MMN-like response were coinciding and mirroring each other, suggesting that they could be manifestations of the same process of comparison between event and expectation, and that the amount of prediction error is what differentiates them (Friston, 2005; Garrido, Kilner, Kiebel, & Friston, 2007; Garrido et al., 2009). Furthermore, in Study I we observed that the RP falls on the downslope of the N1, thus we could speculate that this endogenous response could be related to the N1-attenuation. In Study II, we did not observe a clear buildup of the RP in response to the first four repetitions of the same action-tone association. However, we observed an N1-attenuation in response to the tones following the fifth position, which seems to suggest that the co-occurence of a press-tone association is needed in order to consolidate a prediction and be observable on the sensory responses (e.g., in the RP time window). This result seems to confirm the hypothesis that the consecutive repetitions of the same press-tone association can elicit a RP (Cowan et al., 1993).

The expectation-mismatch effects were also visible on the P3-like response, which was subsequent to MMN-like response (Study I and Study II). In particular, in Study II, we observed the presence of the P3-like response up to the fifth tone repetition (i.e., late tone duplet), which was immediately followed by the maxima N1 attenuation in response to the self-generated tones following five tone repetitions (i.e., 6th+). This finding shows how repetition-related effects modulate the P3 response, suggesting that the need of updating the internal regularity model (Donchin & Coles, 1988) stops once the action-effect association is consolidated, and hence once the levels of uncertainty are low (Sutton et al., 1965; Friston, 2010).

Particularly, Study II allowed us to compare self-generated and externally-generated responses to the tones, in order to examine the contribution of the motor action and of repetitions in fostering an association. Here, indeed, we observed an attenuation of the exogenous N1 component in response to self-generated tones compared to externally-generated tones, and although this finding could be related to multiple causes (Horváth, 2015), it seems to suggest that the motor action has a key role in the buildup of an auditory regularity, facilitating the consolidation of internal models about auditory regularities. Moreover the findings from Study II showed that the N1 attenuation was maximal only in response to more than five consecutive repetitions of the same pinch-tone association, suggesting that repetitions are necessary in order to foster an internal model of the auditory regularity (Cowan et al., 1993), and hence an action-effect prediction, causing a modulation of the brain activity (Jiang et al., 2000; Sayres & Grill-Spector, 2006), and of the N1 component (Baum et al., 2017; Garrido et al., 2009). Conversely to our expectations, we have not found any gradual N1 attenuation as correlated of the action-effect repetition, although evidence similar to ours were reported also in a previous study showing an N1 enhancement to the third and fourth repetition of the same stimulus (Garrido et al., 2009).

However, in Study I we have not found any significant attenuation in the N1 time-window (i.e., no N1 attenuation for self-generated tones) comparing expected sounds with control ones. These divergent findings suggest that the N1 attenuation could be more associated to the origin of the sound (Study II: self-generated vs. externally-generated), than to the degree of predictability between conditions, and hence to expectation-match effects (Study I: 88% vs. 50% probability of eliciting sound). Furthermore, since the switch between the two buttons in Study I, we could speculate that probably the absence of a constant and repetitive action-effect contingency could have caused the absence of N1 attenuation in response to expected sounds. The lack of N1 attenuation in Study I could also be explained by the timing between presses, indeed the average inter-press interval was ca. 800 ms in Study I, and ca.

1.500 ms in Study II, confirming the importance of this factor, as emerged also in a previous study (SanMiguel et al., 2013 b).

However, in Study I, we observed modulations of the omission (i.e., endogenous) responses in the oN1 time-window. That is, when we compared the neuronal response elicited by the omissions in a predictable situation (experimental condition) and in an unpredictable situation (control condition), we found that the neuronal response was more positive to expected omission and more negative to unexpected omission responses (i.e., extracted from the experimental condition) compared to the response in an unpredictable situation. The modulation of the neuronal response in the absence of any sensory stimulation, suggests that expectation-match and -mismatch effects were taking place and modulating the electrophysiological responses (Bendixen et a., 2009; Wacongne et al., 2011), pointing to the idea that expectation-match and -mismatch effects are manifestations of the same phenomenon. Moreover, the results from the data-driven analysis of Study I suggest that only the expectation about the presence or omission of the input was driving the neuronal response in the initial wider time-window which included N1 (30-126 ms), thus we could speculate that the modulation on the N1 component was mainly related to the predictive top-down processing and not to the comparison between the current event and the expectation (i.e., to the prediction error).

Finally, in Study I, in the same wider time-window in which we observed the N1 modulation, we also observed that the expectation-related effects modulated also the P1 component, generally elicited concomitantly to the sound onset (Winkler et al., 2009). Indeed, we found a P1 attenuation in response to expected sounds compared to the control sounds, suggesting that the expectation-related effects can affect the middle-latency responses elicited at around 50 ms (Baess et al., 2009) in response to a sound, and that probably this is due to pre-attentive processing (Jerger et al., 1992; White & Yee, 1997, 2006) due to the sound expectation.

Moreover, in Study I we also observed that anticipatory expectation-related effects were visible before the subsequent button-press (e.g., from -0.2 to 0 s), suggesting that the buildup of internal models and the consolidation of a prediction are already guiding the action itself. This hypothesis would be in line with the ideomotor theory proposing that the motor-action and its sensory consequence share the same representation, which is activated simultaneously once the motor-action is planned.

Overall, the results from the electrophysiological measurements show that the expectation-related effects modulate the neuronal response, affecting mainly the endogenous signals. Indeed, in the case of sounds we found expectation-match effects resulting in the RP (Study I) and expectation-mismatch effects resulting in the MMN-like and P3-like responses (Study I and II), and in the case of omissions we found a modulation in the oN1 time window (Study I), with opposite polarities for expected and unexpected omissions. On the exogenous response, we only found a modulation of N1 in response to self-generated tones (Study II) and of P50 in response to expected sounds (Study I).

Furthermore, the results from both studies suggest that the electrophysiological responses are characterized by different stages, starting with the violation of a previous consolidated action-effect association (eliciting an MMN-like response), which triggers the update of the internal models (P3), up to the establishment of a new action-effect association, and hence the consolidation of an expectation (indicated by a N1 attenuation), due to the contingent repetition of an action-effect association.

Taking together behavioral and electrophysiological results, we could speculate about the existence of a close-loop characterizing the motor-auditory interaction. That is, the motor action causes a specific sensory consequence (e.g., a sound or omission) which in turn affects associative and predictive processing causing not only behavioral modulations (e.g., action timing, force peaks), but also electrophysiological modulations (e.g., MMN/RP, P3, N1/oN1) reflecting the preparation for the incoming sensory input elicited by the subsequent action.

STRENGTHS AND LIMITATIONS

Most of the studies reported in literature on motor-auditory interactions are characterized by a classical experimental design composed by a motor-auditory condition, an auditory-alone and a motor-alone condition (Horváth, 2015). So far, this classical experimental design did not allow to directly compare the expectation-match and -mismatch effects within the same study (differently from Study I). Moreover, most of the previous studies excluded from the analysis the first five events of interest, in which the building-up of an action-effect association is supposingly taking shape (differently from Study II).

In the two studies presented in this thesis, we adopted a contingent paradigm in which each sensory consequence is caused and contingent to a motor action (e.g., button press). However, our studies differ from the previous ones, because they are characterized by original experimental designs which allowed to study predictable and unpredictable self-generated sound and omissions (Study I), or which was characterized by the absence of a stable action-effect association (Study II); and by an original analysis, for example the analysis on wider time-windows and on the time-window preceding the button press (Study I); and the analysis per position of interest within a tone micro-sequence (Study II), which allowed to reach aims and findings so far uncovered. Indeed, Study I suggests that expectation-match and -mismatch effects are modulations of the same underlying process, and that this process varies based on the time-course. Study II demonstrated the importance of the first five action-effect associations in order to extract a regularity, and hence to generate an expectation.

The novelty of our experimental designs incorporates strengths, but also limitations. Indeed, in Study I, the possibility given by studying different expectation-related effects, such as match- vs. mismatch-related effects driven by the actual presence and by the omission of the sound within the same study, has excluded the possibility of comparing these responses to a purely auditory and/or motor response, since the inclusion of these conditions would have negatively increased the duration of the whole experiment, causing tiredness in the participants and prejudicing the quality of the measurements. Furthermore, in Study I, since the participants were using two different buttons, we cannot totally guarantee the exclusion of all possible interferences caused by the switch between the buttons (e.g., in fostering the prediction associated with each button-press). Parallelly, in Study II since we isolated the events of interest by position and by their sensory consequence (i.e., analyzing only tones), we had the necessity of reducing the number of repetitions per tone at five per micro-sequence, in order to guarantee an enough number of trials per each event of interest, without negatively increasing the duration of the whole experiment.

FUTURE DIRECTIONS

The two studies presented in this thesis leave open questions which could inspire future works. An interesting direction of Study I could be to investigate whether it is possible to generate an expectation-related not only to the presence or to the omission of a sound, but also to the intrinsic characteristics of the sound itself. That is, to use within the same study, two sounds which differ for a physical feature, and thus to investigate whether the expectation-related effects change based on the intrinsic characteristics of an auditory stimulus (e.g., manipulating the frequency of a tone). An interesting direction of Study II would be to manipulate different factors which can affect the action-effect association. For instance, to manipulate the degree of action-effect contingency (low contingency vs. high contingency, and vice versa) in order to explore the impact of action-effects contingency and coincidence in determining a new association (Horváth, 2015).

Furthermore, it would be interesting to carry on the studies presented in this thesis also on two particular samples of human subjects: the musicians and the schizophrenic patients. These samples could allow to deeply investigate the existence of relations between specific disorders and non-functional mechanisms underlying the associative and predictive processing. Indeed, in the case of musicians, it has been shown that the musical abilities depends on the associations between a specific motor action (e.g., the movement of the fingers on a keyboard) and a sensory consequence (e.g., sounds of different intensity or pitch) that the musicians develop along their musical career through repetitive action-effect associations (Chen et al., 2011). However, musicians can also be affected by neurological disorders, such as focal dystonia (Altenmüller, 2003), which can alter their psychophysiological responses to errors- and deviants-detection (Strübing et al., 2012), which are normally present in healthy musicians (Maidhof et al., 2009; Maidhof, 2013). On the other side, schizophrenic patients can present psychotic symptoms, such as delusions of control (DSM-V), during which the subject is not able to differentiate between self-generated vs. externally-generated motor acts (Frith, Blakemore, & Wolpert, 2000 a, b; Ford & Mathalon, 2005), and there are EEG studies which have shown differences between schizophrenic patients and healthy subjects, for instance in the sensory attenuation to self-generated sounds (Ford et al., 2001, 2007; Ford & Mathalon, 2005).

Therefore, further studies will be necessary in order to investigate mechanisms underlying associative and predictive processing, both in healthy and clinical populations.

6. CONCLUSIONS

1. The repeated association between an action and a sensory event facilitates the formation of internal models related to the action-effect regularity, which is observable through modulations both at electrophysiological level (sensory processing of the event), and at behavioral level (parameters related to the motor action).

2. Both associative and predictive processing contribute in modulating how the sensory consequence of the action affects the action itself. That is, actions eliciting sounds provide support in maintaining a regular pace between actions, cause slowing effects on the subsequent action, and are executed with less force, the force being adjusted instantly as the sensory feedback is presented. Additionally, unexpected consequences of the action can modify the timing of the subsequent action, and the direction of this modulation depends on the specific context.

3. Electrophysiologically, the buildup of the action-tone association is correlated by the modulation of different ERPs. The violation of a previous consolidated association is indicated by an MMN-like response. Then, by means of the repetition of the same action-effect contingency, the phase of updating of internal associations is indicated by the presence of the P3-like response. Finally, after around six repetitions, the internal action-tone association is consolidated and indicated by an attenuated N1 component.

4. An established action-effect association can generate consolidated expectations, and hence specific predictions about the presence or omission of a stimulus. Electrophysiologically, the expectation-related effects on the ERPs can be described as a unitary phenomenon which varies along its time-course: up to ca. 120 ms the brain signals are mostly driven by the expectations (P1 and N1), being negatively displaced when a sound is expected, compared to when its omission is expected. Subsequently, match and mismatch effects show opposite polarities based on the comparison between expectation and input, presenting a positive displacement in case of expectation-match (RP), and a negative displacement in case of expectation-mismatch (MMN-like response). At ca. 300 ms the brain signal appears more positive (P3) in case of expectation-mismatch than in case of expectation-match, indicating the necessity of updating the internal model, and finally preparing for the following action and incoming input.

5. Similar match- and mismatch-related effects are observable both on stimulus-evoked responses, driven by sounds, and on endogenous responses, driven by omissions of the sounds, suggesting that the endogenous signals reflect expectation-related effects, and underlie the modulations on the stimulus-evoked responses.

6. The modulations in the N1 time window are heterogeneous and context dependent, however they seem to be affected by associative and predictive processing. Indeed, the temporal action-effect coincidence, in dynamic contexts characterized by unstable regularities, does not seem to elicit the coincidence-like effects. However,

contingency-like effects on the N1 attenuation are observable after more than five consecutive repetition of the same action-effect association.

7. REFERENCES

Albright, T. D. (2012). On the perception of probable things: neural substrates of associative memory, imagery, and perception. *Neuron*, *74*(2), 227-245.

Aliu, S. O., Houde, J. F., & Nagarajan, S. S. (2009). Motor-induced suppression of the auditory cortex. *Journal of Cognitive Neuroscience*, 21, 791–802. doi: 10.1162/jocn.2009.21055

Altenmüller, E. (2003). Focal dystonia: advances in brain imaging and understanding of fine motor control in musicians. *Hand clinics*, *19*(3), 523-538.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

Baess, P., Horváth, J., Jacobsen, T., & Schröger, E. (2011). Selective suppression of self-initiated sounds in an auditory stream: An ERP study. *Psychophysiology*, 48, 1276–1283. doi: 10.1111/j.1469-8986.2011.01196.x

Baess, P., Jacobsen, T., & Schröger, E. (2008). Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: Evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology*, 70, 137–143. doi: 10.1016/j.ijpsycho.2008.06.005

Baess, P., Widmann, A., Roye, A., Schröger, E., & Jacobsen, T. (2009). Attenuated human auditory middle latency response and evoked 40-Hz response to self-initiated sounds. *European Journal of Neuroscience*, 29(7), 1514–1521. doi: 10.1111/j.1460-9568.2009.06683.x

Baldeweg, T. (2006). Repetition effects to sounds: evidence for predictive coding in the auditory system. *Trends in cognitive sciences*. doi: 10.1016/j.tics.2006.01.010

Baldeweg, T. (2007). ERP repetition effects and mismatch negativity generation: a predictive coding perspective. *Journal of Psychophysiology*, 21(3-4), 204-213. doi: 10.1027/0269-8803.21.34.204

Baldeweg, T., Klugman, A., Gruzelier, J., & Hirsch, S. R. (2004). Mismatch negativity potentials and cognitive impairment in schizophrenia. *Schizophrenia research*, *69*(2-3), 203-217.

Band, G. P., van Steenbergen, H., Ridderinkhof, K. R., Falkenstein, M., & Hommel, B. (2009). Action-effect negativity: Irrelevant action effects are monitored like relevant feedback. *Biological Psychology*, 82(3), 211-218. doi: 10.1016/j.biopsycho.2009.06.011

Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends in cognitive sciences*, *11*(7), 280-289.

Bar, M., Aminoff, E., Mason, M., & Fenske, M. (2007). The units of thought. *Hippocampus*, *17*(6), 420-428.

Barcelo, F., Escera, C., Corral, M. J., & Periáñez, J. A. (2006). Task switching and novelty processing activate a common neural network for cognitive control. *Journal of cognitive neuroscience*, *18*(10), 1734-1748.

Baum, F., Wolfensteller, U., & Ruge, H. (2017). Learning-Related Brain-Electrical Activity Dynamics Associated with the Subsequent Impact of Learnt Action-Outcome Associations. *Frontiers in human neuroscience*, *11*, 252.

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences*, *106*(5), 1672-1677.

Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: a review. *International Journal of Psychophysiology*, *83*(2), 120-131.

Bendixen, A., Schröger, 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.

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature neuroscience*, *1*(7), 635.

Blakemore, S. J., Wolpert, D., & Frith, C. (2000). Why can't you tickle yourself?. *Neuroreport*, *11*(11), R11-R16.

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.

Buehner, M. J., & McGregor, S. J. (2009). Contingency and contiguity trade-offs in causal induction. *International Journal of Comparative Psychology*, 22(1).

Bulf, H., Valenza, E., & Simion, F. (2009). The visual search of an illusory figure: A comparison between 6-month-old infants and adults. *Perception*, 38(9), 1313-1327.

Cacciaglia, R., Costa-Faidella, J., Żarnowiec, K., Grimm, S., & Escera, C. (2019). Auditory predictions shape the neural responses to stimulus repetition and sensory change. *NeuroImage*, *186*, 200-210.

Carpenter, W. B. (1852). On the influence of suggestion in modi- fying and directing muscular movement, independently of volition. *Proceedings of the Royal Institution*, 147–154.

Chen, J. L., Rae, C., & Watkins, K. E. (2012). Learning to play a melody: an fMRI study examining the formation of auditory-motor associations. *Neuroimage*, *59*(2), 1200-1208.

Condry, K. F., Smith, W. C., & Spelke, E. S. (2001). Development of perceptual organization. *Emerging cognitive abilities in early infancy*, 1-28.

Corlett, P. R., Honey, G. D., Krystal, J. H., & Fletcher, P. C. (2011). Glutamatergic model psychoses: prediction error, learning, and inference. *Neuropsychopharmacology*, *36*(1), 294.

Costa-Faidella, J., Baldeweg, T., Grimm, S., & Escera, C. (2011). Interactions between “what” and “when” in the auditory system: temporal predictability enhances repetition suppression. *Journal of Neuroscience*, 31(50), 18590-18597. doi: 10.1523/JNEUROSCI.2599-11.2011

Cowan, N., Winkler, I., Teder, W., & Näätänen, R. (1993). Memory prerequisites of mismatch negativity in the auditory event-related potential (ERP). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(4), 909. doi: 10.1037/0278-7393.19.4.909

Csibra, G. (2001). Illusory contour figures are perceived as occluding surfaces by 8-month-old infants. *Developmental Science*, *4*(4), F7-F11.

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

Deouell, L. Y., & Bentin, S. (1998). Variable cerebral responses to equally distinct deviance in four auditory dimensions: A mismatch negativity study. *Psychophysiology*, *35*(6), 745-754.

Diekhof, E. K., Biedermann, F., Ruebsamen, R., & Gruber, O. (2009). Top-down and bottom-up modulation of brain structures involved in auditory discrimination. *Brain research*, *1297*, 118-123.

Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating?. *Behavioral and brain sciences*, *11*(3), 357-374. doi: 10.1017/S0140525X00058027

Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin*, 99(1), 3. doi: 10.1037/0033-2909.99.1.3

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of experimental psychology: human perception and performance*, *27*(1), 229. doi: 10.1037/0096-1523.27.1.229

Elsner, B., & Hommel, B. (2004). Contiguity and contingency in action-effect learning. *Psychological research*, *68*(2-3), 138-154. doi: 10.1007/s00426-003-0151-8

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.

Escera, C., & Malmierca, M. S. (2014). The auditory novelty system: an attempt to integrate human and animal research. *Psychophysiology*, *51*(2), 111-123.

Escera, C., Alho, K., Schröger, E., & Winkler, I. W. (2000). Involuntary attention and distractibility as evaluated with event-related brain potentials. *Audiology and Neurotology*, 5(3-4), 151-166. doi: 10.1159/000013877

Escera, C., Alho, K., Winkler, I., & Näätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *Journal of cognitive neuroscience*, *10*(5), 590-604. doi: 10.1162/089892998562997

Escera, C., Leung, S., & Grimm, S. (2014). Deviance detection based on regularity encoding along the auditory hierarchy: electrophysiological evidence in humans. *Brain topography*, 27(4), 527-538. doi: 10.1007/s10548-013-0328-4

Falkenstein, M., Hohnsbein, J., Hoormann, J., Blanke, L., (1990). Effects of errors in choice reaction tasks on the ERP under focused and divided attention. In: Brunia, C.H.M., Gaillard, A.W.K., Kok, A. (Eds.), *Psychophysiological Brain Research*. Tilburg University Press, Tilburg, pp. 192–195.

Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: a tutorial. *Biological psychology*, *51*(2-3), 87-107. doi: 10.1016/S0301-0511(99)00031-9

Fogelson, N., Wang, X., Lewis, J. B., Kishiyama, M. M., Ding, M., & Knight, R. T. (2009). Multimodal effects of local context on target detection: evidence from P3b. *Journal of Cognitive Neuroscience*, *21*(9), 1680-1692. 10.1162/jocn.2009.21071

Folstein J.R., Van Petten C. (2008) Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology*, 45(1), 152–170. doi: 10.1111/j.1469-8986.2007.00602.x

Ford, J. M., & Mathalon, D. H. (2005). Corollary discharge dysfunction in schizophrenia: can it explain auditory hallucinations?. *International Journal of Psychophysiology*, *58*(2-3), 179-189.

Ford, J. M., Gray, M., Faustman, W. O., Roach, B. J., & Mathalon, D. H. (2007). Dissecting corollary discharge dysfunction in schizophrenia. *Psychophysiology*, *44*(4), 522-529.

Ford, J. M., Mathalon, D. H., Heinks, T., Kalba, S., Faustman, W. O., & Roth, W. T. (2001). Neurophysiological evidence of corollary discharge dysfunction in schizophrenia. *American Journal of Psychiatry*, *158*(12), 2069-2071.

Friedman, D., Cycowicz, Y.M. & Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neurosci. Biobehav. R.*, 25, 355–373. doi: 10.1016/S0149-7634(01)00019-7

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 360(1456), 815-836. doi: 10.1098/rstb.2005.1622

Friston, K. (2009). The free-energy principle: a rough guide to the brain?. *Trends in cognitive sciences*, 13(7), 293-301.

Friston, K. (2010). The free-energy principle: a unified brain theory?. *Nature Reviews Neuroscience*, 11(2), 127.

Friston, K., & Kiebel, S. (2009). Cortical circuits for perceptual inference. *Neural Networks*, 22(8), 1093-1104.

Frith, C. D., Blakemore, S. J., & Wolpert, D. M. (2000 a). Explaining the symptoms of schizophrenia: abnormalities in the awareness of action. *Brain Research Reviews*, *31*(2-3), 357-363.

Frith, C. D., Blakemore, S. J., & Wolpert, D. M. (2000 b). Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *355*(1404), 1771-1788.

Frith, C. (2013). *Making up the mind: How the brain creates our mental world*. John Wiley & Sons.

Gallistel, C. R., & Gibbon, J. (2001). Computational versus associative models of simple conditioning. *Current Directions in Psychological Science*, 10(4), 146-150. doi: 10.1111/1467-8721.00136

Garrido, M. I., Kilner, J. M., Kiebel, S. J., & Friston, K. J. (2007). Evoked brain responses are generated by feedback loops. *Proceedings of the National Academy of Sciences*, *104*(52), 20961-20966.

Garrido, M. I., Kilner, J. M., Kiebel, S. J., Stephan, K. E., Baldeweg, T., & Friston, K. J. (2009). Repetition suppression and plasticity in the human brain. *Neuroimage*, *48*(1), 269-279.

Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological science*, *4*(6), 385-390. doi: 10.1111/j.1467-9280.1993.tb00586.x

Gentsch, A., Ullsperger, P., & Ullsperger, M. (2009). Dissociable medial frontal negativities from a common monitoring system for self-and externally caused failure of goal achievement. *Neuroimage*, 47(4), 2023-2030.

Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: with special reference to the ideo-motor mechanism. *Psychological review*, 77(2), 73. doi: 10.1037/h0028689

Greville, W. J., & Buehner, M. J. (2010). Temporal predictability facilitates causal learning. *Journal of Experimental Psychology*: General, 139(4), 756. doi: 10.1037/a0020976

Griffin, J. D., & Fletcher, P. C. (2017). Predictive processing, source monitoring, and psychosis. *Annual review of clinical psychology*, *13*, 265-289.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in cognitive sciences*, 10(1), 14-23. doi: 10.1016/j.tics.2005.11.006

Grimm, S., Widmann, A., & Schröger, E. (2004). Differential processing of duration changes within short and long sounds in humans. *Neuroscience Letters*, 356(2), 83-86. doi: 10.1016/j.neulet.2003.11.035

Haenschel, C., Vernon, D. J., Dwivedi, P., Gruzelier, J. H., & Baldeweg, T. (2005). Event-related brain potential correlates of human auditory sensory memory-trace formation. *The Journal of Neuroscience*, 25(45), 10494-10501. doi: 10.1523/JNEUROSCI.1227-05.2005

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature neuroscience*, *5*(4), 382.

Harless, E. (1861). Der Apparat des Willens [The apparatus of the will]. In I. H. Fichte, H. Ulrici & I. U. Wirth (Eds.), *Zeitschrift für Philosophie und philosophische Kritik* [Magazine for philosophy and philosophical criticism], 38, 50–73. Halle, Germany: Pfeffer.

Herbart, J. F. (1816). Lehrbuch zur Psychologie [Textbook for psychology]. Königsberg, Germany: Unzer.

Herbart, J. F. (1825). Psychologie als Wissenschaft neu gegründet auf Erfahrung, Metaphysik und Mathematik. Zweiter, analytischer Teil [Psychology as science based on experience, metaphysics, and mathematics: Part 2]. Königsberg, Germany: Unzer.

Hesse, M. D., Nishitani, N., Fink, G. R., Jousmaki, V., and Hari, R. (2010). Attenuation of somatosensory responses to self-produced tactile stimulation. *Cereb. Cortex* 20, 425–432. doi: 10.1093/cercor/bhp110

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological review*, *109*(4), 679. doi: 10.1037//0033-295X.109.4.679

Hommel, B. (1996). The cognitive representation of action: Automatic integration of perceived action effects. *Psychological research*, 59(3), 176-186. doi: 10.1007/bf00425832

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and brain sciences*, 24(5), 849-878. doi: 10.1017/s0140525x01000103

Horváth J., (2013a). Action–sound coincidence-related attenuation of auditory ERPs is not modulated by affordance compatibility. *Biol. Psychol.*, 93 (1), 81–87. doi: 10.1016/j.biopsycho.2012.12.008

Horváth, J. (2015). Action-related auditory ERP attenuation: Paradigms and hypotheses. *Brain Research*, 1626, 54-65. doi: 10.1016/j.brainres.2015.03.038

Horváth, J., (2013b). Attenuation of auditory ERPs to action–sound coincidences is not explained by voluntary allocation of attention. *Psychophysiology*, 50 (3), 266–273. doi: 10.1111/psyp.12009

Horváth, J. (2014). The role of mechanical impact in action-related auditory attenuation. *Cognitive, Affective, & Behavioral Neuroscience*, 14(4), 1392-1406. doi: 10.3758/s13415-014-0283-x

Horváth, J., & Burgyán, A. (2013c). No evidence for peripheral mechanism attenuating auditory ERPs to self-induced tones. *Psychophysiology*, 50(6), 563-569.

Horváth, J., Bíró, B., & Neszmélyi, B. (2018). Action-effect related motor adaptation in interactions with everyday devices. *Scientific reports*, *8*. doi: :10.1038/s41598-018-25161-w

Horváth, J., Maess, B., Baess, P., & Tóth, A. (2012). Action-sound coincidences suppress evoked responses of the human auditory cortex in EEG and MEG. *Journal of Cognitive Neuroscience*, 24, 1919–1931. doi: 10.1162/jocn_a_00215

Hughes, G., Desantis, A., & Waszak, F. (2013). Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction. *Psychological bulletin*, *139*(1), 133.

Hughes, H. C., Darcey, T. M., Barkan, H. I., Williamson, P. D., Roberts, D. W., & Aslin, C. H. (2001). Responses of human auditory association cortex to the omission of an expected acoustic event. *Neuroimage*, 13(6), 1073-1089.

Hume, D. (1969). A treatise of human nature, ed. Ernest C. Mossner. *New York: Penguin Books*, 677, 17. (Original work published 1739)

Interlink Electronics. FSR 400 Series Data Sheet. Retrieved from http://www.interlinkelectronics.com (2016, October 10).

Iwanaga, M., & Nittono, H. (2010). Unexpected action effects elicit deviance-related brain potentials and cause behavioral delay. *Psychophysiology*, 47(2), 281-288. doi: 10.1111/j.1469-8986.2009.00937.x

James, W. (1890). The principles of psychology, Vols. I, II. *Cambridge*, MA: Harvard University Press.

James, W. (1950). The principles of psychology: Vol. 2. New York, NY: Dover. (Original work published 1890)

Jerger, K., Biggins, C., & Fein, G. (1992). P50 suppression is not affected by attentional manipulations. *Biological psychiatry*, *31*(4), 365-377.

Johnson Jr, R., & Donchin, E. (1980). P300 and stimulus categorization: Two plus one is not so different from one plus one. *Psychophysiology*, 17(2), 167-178. doi: 10.1111/j.1469-8986.1980.tb00131.x

Joos, K., Gilles, A., Van de Heyning, P., De Ridder, D., & Vanneste, S. (2014). From sensation to percept: the neural signature of auditory event-related potentials. *Neuroscience & Biobehavioral Reviews*, 42, 148-156. doi: 10.1016/j.neubiorev.2014.02.009

Kaiser, J., & Schütz-Bosbach, S. (2018). Sensory attenuation of self-produced signals does not rely on self-specific motor predictions. *European Journal of Neuroscience*, *47*(11), 1303-1310. doi: 10.1111/ejn.13931

Kanizsa, G., 1955. 'Margini quasi-percettivi in campi con stimolazione omogenea', *Rivista di Psicologia*, 49 (1) pp.7–30. English translation, 'Quasi-perceptual margins in homogeneously stimulated fields', in S. Petry and G. E. Meyer (Eds) 1987, The Perception of Illusory Contours pp. 40-49, Springer: NY.

Katahira, K., Abla, D., Masuda, S., & Okanoya, K. (2008). Feedback-based error monitoring processes during musical performance: an ERP study. *Neuroscience research*, 61(1), 120-128. doi: 10.1016/j.neures.2008.02.001

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, 36(14), 1.

Knolle, F., Schröger, E., & Kotz, S. A. (2013). Prediction errors in self-and externally-generated deviants. *Biological Psychology*, 92(2), 410-416.

Köhler, W. (1938). Physical Gestalten.

Kunde, W., Koch, I., & Hoffmann, J. (2004). Anticipated action effects affect the selection, initiation, and execution of actions. *The Quarterly Journal of Experimental Psychology Section A*, 57(1), 87-106. doi: 10.1080/02724980343000143

Lasaponara, S., Chica, A. B., Lecce, F., Lupianez, J., & Doricchi, F. (2011). ERP evidence for selective drop in attentional costs in uncertain environments: challenging a purely premotor account of covert orienting of attention. *Neuropsychologia*, *49*(9), 2648-2657.

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.

Loehr, J. D. (2013). Sensory attenuation for jointly produced action effects. *Frontiers in psychology*, *4*, 172. doi: 10.3389/fpsyg.2013.00172

Lotze, R. H. (1852). Medizinische Psychologie oder Physiologie der Seele [Medical psychology or physiology of the soul]. Leipzig, Germany: Weidmannsche Buchhandlung.

Lü, Z. L., Williamson, S. J., & Kaufman, L. (1992). Human auditory primary and association cortex have differing lifetimes for activation traces. *Brain research*, *572*(1-2), 236-241.

Maia, T. V., & Frank, M. J. (2017). An integrative perspective on the role of dopamine in schizophrenia. *Biological psychiatry*, *81*(1), 52-66.

Maidhof, C. (2013). Error monitoring in musicians. *Frontiers in human neuroscience*, *7*, 401.

Maidhof, C., Rieger, M., Prinz, W., & Koelsch, S. (2009). Nobody is perfect: ERP effects prior to performance errors in musicians indicate fast monitoring processes. *PLoS One*, *4*(4), e5032.

Makeig, S., Müller, M. M., & Rockstroh, B. (1996). Effects of voluntary movements on early auditory brain responses. *Experimental brain research*, *110*(3), 487-492.

Mangun, G. R., & Hillyard, S. A. (1991). Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming. *Journal of Experimental Psychology: Human perception and performance*, 17(4), 1057.

Martikainen, M. H., Kaneko, K., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex*, 15, 299–302. doi: 10.1093/cercor/bhh131

McGurk H., and MacDonald J. (1976). "Hearing lips and seeing voices.". *Nature*. 264 (5588): 746–8.

Messinger, A., Squire, L. R., Zola, S. M., & Albright, T. D. (2001). Neuronal representations of stimulus associations develop in the temporal lobe during learning. *Proceedings of the National Academy of Sciences*, *98*(21), 12239-12244.

Mifsud, N. G., & Whitford, T. J. (2017). Sensory attenuation of self-initiated sounds maps onto habitual associations between motor action and sound. *Neuropsychologia*, 103, 38-43. doi: 10.1016/j.neuropsychologia.2017.07.019

Miller, J., Patterson, T., & Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology*, 35(1), 99-115. doi: 10.1111/1469-8986.3510099

Mitchell, C. J., Livesey, E., & Lovibond, P. F. (2007). A dissociation between causal judgement and the ease with which a cause is categorized with its effect. *The Quarterly Journal of Experimental Psychology*, 60(3), 400-417. doi: 10.1080/17470210601002512

Miyashita, Y. (1993). Inferior temporal cortex: where visual perception meets memory. *Annual review of neuroscience*, *16*(1), 245-263.

Murray, E. A., Gaffan, D., & Mishkin, M. T. (1993). Neural substrates of visual stimulus-stimulus association in rhesus monkeys. *Journal of Neuroscience*, *13*(10), 4549-4561.

Näätänen, R. (1992). *Attention and brain function*. Routledge.

Näätänen, R., & Alho, K. (1995). Mismatch negativity - a unique measure of sensory processing in audition. *International Journal of Neuroscience*, 80(1-4), 317-337. doi: 10.3109/00207459508986107

Näätänen, R., Gaillard, A.W.K., & Mäntysalo, S. (1978). Early selective attention effect on evoked potential reinterpreted. *Acta Psychologica*, 42, 313–329. doi: 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. doi: 10.1016/j.clinph.2007.04.026

Neszmélyi, B., & Horváth, J. (2017). Consequences matter: Self‐induced tones are used as feedback to optimize tone‐eliciting actions. *Psychophysiology*, 54(6), 904-915. doi: 10.1111/psyp.12845

Neszmélyi, B., & Horváth, J. (2018). Temporal constraints in the use of auditory action effects for motor optimization. *Journal of Experimental Psychology: Human Perception and Performance*, 44(11), 1815. doi: 10.1037/xhp0000571

Nittono, H. (2004). The action-perception paradigm: A new perspective in cognitive neuroscience. In *International Congress Series* (Vol. 1270, pp. 26-31). Elsevier.

Nittono, H. (2006). Voluntary stimulus production enhances deviance processing in the brain. *International Journal of Psychophysiology*, 59(1), 15-21. doi: 10.1016/j.ijpsycho.2005.06.008

Nittono, H., & Ullsperger, P. (2000). Event-related potentials in a self-paced novelty oddball task. *NeuroReport*, 11(9), 1861-1864.

Nittono, H., Hamada, A., & Hori, T. (2003). Brain potentials after clicking a mouse: A new psychophysiological approach to human-computer interaction. *Human Factors*, 45(4), 591-600.

Notebaert, W., Houtman, F., Van Opstal, F., Gevers, W., Fias, W., & Verguts, T. (2009). Post-error slowing: an orienting account. *Cognition*, 111(2), 275-279. doi: 10.1016/j.cognition.2009.02.002

Nuwer, M. R., Comi, G., Emerson, R., Fuglsang-Frederiksen, A., Guérit, J. M., Hinrichs, H., ... & Rappelsburger, P. (1998). IFCN standards for digital recording of clinical EEG. *Clinical Neurophysiology*, 106(3), 259-261. doi: 10.1016/S0013-4694(97)00106-5

Onton, J., & Makeig, S. (2006). Information-based modeling of event-related brain dynamics. *Progress in brain research*, 159, 99-120. doi: 10.1016/S0079-6123(06)59007-7

Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical neurophysiology*, 112(4), 713-719.

Overbeek, T. J., Nieuwenhuis, S., & Ridderinkhof, K. R. (2005). Dissociable components of error processing: On the functional significance of the Pe vis-à-vis the ERN/Ne. *Journal of Psychophysiology*, 19(4), 319-329. doi: 10.1027/0269-8803.19.4.319

Parmentier, F. B., Elsley, J. V., Andrés, P., & Barceló, F. (2011). Why are auditory novels distracting? Contrasting the roles of novelty, violation of expectation and stimulus change. *Cognition*, *119*(3), 374-380. doi: 10.1016/j.cognition.2011.02.001

Partanen, E., Vainio, M., Kujala, T., & Huotilainen, M. (2011). Linguistic multifeature MMN paradigm for extensive recording of auditory discrimination profiles, *Psychophysiology*, 48, 1372–1380.

Pavlov, I. P. (1927). Conditional reflexes: an investigation of the physiological activity of the cerebral cortex. Translated and edited by Anrep, G.V., *Oxford University Press*, London

Plack, C. J. (2018). The sense of hearing. *Routledge*.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology*, 118(10), 2128-2148. doi: 10.1016/j.clinph.2007.04.019

Polich, J., & Bondurant, T. (1997). P300 sequence effects, probability, and interstimulus interval. *Physiology & Behavior*, 61(6), 843-849. doi: 10.1016/S0031-9384(96)00564-1

Pressnitzer, D., Suied, C., & Shamma, S. (2011). Auditory scene analysis: the sweet music of ambiguity. *Frontiers in human neuroscience*, 5, 158.

Prinz, W. (1987). Ideo-motor action. In H. Heuer & A. F. Sanders (Eds.), Perspectives on perception and action (pp. 47–76). Hillsdale, NJ: Lawrence Erlbaum. doi: 10.2307/1423061

Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann & W.

Prinz (Eds.), Relationships between perception and action (pp. 167–201). Berlin Heidelberg New York: Springer. doi: 10.1007/978-3-642-75348-0_7

Pritchard, W. S., Shappell, S. A., & Brandt, M. E. (1991). Psychophysiology of N200/N400: A review and classification scheme. In J. R. Jennings & P. K. Ackles (Eds.), Advances in psychophysiology: A research annual (Vol. 4, pp. 43–106). London: Jessica Kingsley.

Purves, D., & Lotto, R. B. (2011). Why we see what we do redux: A wholly empirical theory of vision. *Sinauer Associates*.

Raij, T., McEvoy, L., Mäkelä, J. P., & Hari, R. (1997). Human auditory cortex is activated by omissions of auditory stimuli. *Brain research*, 745(1-2), 134-143.

Ranganath, C., Cohen, M. X., Dam, C., & D'Esposito, M. (2004). Inferior temporal, prefrontal, and hippocampal contributions to visual working memory maintenance and associative memory retrieval. *Journal of Neuroscience*, *24*(16), 3917-3925.

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

Rock, I., & Palmer, S. (1990). The legacy of Gestalt psychology. *Scientific American*, *263*(6), 84-91.

SanMiguel, I., Morgan, H. M., Klein, C., Linden, D., & Escera, C. (2010). On the functional significance of Novelty-P3: facilitation by unexpected novel sounds. *Biological Psychology*, 83(2), 143-152. doi: 10.1016/j.biopsycho.2009.11.012

SanMiguel, I., Saupe, K., & Schröger, E. (2013a). I know what is missing here: electrophysiological prediction error signals elicited by omissions of predicted ”what” but not ”when”. *Frontiers in human neuroscience*, 7, 407.

SanMiguel, I., Todd J., & Schröger, E. (2013b). Sensory suppression effects to self-initiated sounds reflect the attenuation of the unspecific N1 component of the auditory ERP. *Psychophysiology*, 50, 334-343. doi: 10.1111/psyp.12024

SanMiguel, I., Widmann, A., Bendixen, A., Trujillo-Barreto, N., & Schröger, E. (2013c). Hearing silences: human auditory processing relies on preactivation of sound-specific brain activity patterns. *Journal of Neuroscience*, 33(20), 8633-8639. doi: 10.1523/JNEUROSCI.5821-12.2013

Saupe, K., Widmann, A., Trujillo-Barreto, N. J., & Schröger, E. (2013). Sensorial suppression of self-generated sounds and its dependence on attention. *International journal of psychophysiology*, 90(3), 300-310. doi: 10.1016/j.ijpsycho.2013.09.006

Schäfer, E. W., & Marcus, M. M. (1973). Self-stimulation alters human sensory brain responses. *Science*, 181(4095), 175–177. doi: 10.1126/ science.181.4095.175

Schröger, E., Giard, M.-H., & Wolff, Ch. (2000). Auditory distraction: Event-related potential and behavioral indices. *Clinical Neurophysiology*, 111, 1450–1460. doi: 10.1016/S1388-2457(00)00337-0

Shams L., Kamitani Y., Shimojo S. (2000). What you see is what you hear. *Nature* 408, 788–788. doi: 10.1038/35048669

Shanks, D. R. (1993). Human instrumental learning: A critical review of data and theory. *British Journal of Psychology*, 84(3), 319-354. doi: 10.1111/j.2044-8295.1993.tb02486.x

Shanks, D. R. (1993). Human instrumental learning: A critical review of data and theory. *British Journal of Psychology*, *84*(3), 319-354.

Shanks, D. R. (2007). Associationism and cognition: Human contingency learning at 25. *The Quarterly Journal of Experimental Psychology*, 60(3), 291-309. doi: 10.1080/17470210601000581

Shin, Y. K., Proctor, R. W., & Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychological bulletin*, *136*(6), 943. doi: 10.1037/a0020541

Sowman, P. F., Kuusik, A., & Johnson, B. W. (2012). Self-initiation and temporal cueing of monaural tones reduce the auditory N1 and P2. *Experimental brain research*, 222(1-2), 149-157.

Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and clinical neurophysiology*, 38(4), 387-401. doi: 10.1016/0013-4694(75)90263-1

Stein, B. E., Stanford, T. R., & Rowland, B. A. (2014). Development of multisensory integration from the perspective of the individual neuron. *Nature Reviews Neuroscience*, 15(8), 520.

Steinschneider, M., Fishman, Y. I., & Arezzo, J. C. (2003). Representation of the voice onset time (VOT) speech parameter in population responses within primary auditory cortex of the awake monkey. *The Journal of the Acoustical Society of America*, 114(1), 307-321.

Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., ... & Corlett, P. R. (2018). The predictive coding account of psychosis. *Biological psychiatry*, *84*(9), 634-643.

Stock, A., & Stock, C. (2004). A short history of ideo-motor action. *Psychological research*, *68*(2-3), 176-188. doi: 10.1007/s00426-003-0154-5

Strübing, F., Ruiz, M. H., Jabusch, H. C., & Altenmüller, E. (2012). Error monitoring is altered in musician's dystonia: evidence from ERP-based studies. *Annals of the New York Academy of Sciences*, *1252*(1), 192-199.

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*(3700), 1187-1188. doi: 10.1126/science.150.3700.1187

Tervaniemi, M., Saarinen, J., Paavilainen, P., Danilova, N., & Näätänen, R. (1994). Temporal integration of auditory information in sensory memory as reflected by the mismatch negativity. *Biological psychology*, 38(2-3), 157-167. doi: 10.1016/0301-0511(94)90036-1

Thorndike, E. L. (1931). Human learning. New York: Century.

Timm, J., SanMiguel, I., Keil, J., Schröger, E., & Schönwiesner, M. (2014). Motor intention determines sensory attenuation of brain responses to self-initiated sounds. *Journal of cognitive neuroscience*, *26*(7), 1481-1489.

Timm, J., SanMiguel, I., Saupe, K., & Schröger, E. (2013). The N1-suppression effect for self-initiated sounds is independent of attention. *BMC neuroscience*, *14*(1), 2. doi: 10.1186/1471-2202-14-2

Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature neuroscience*, 7(9), 907. http://dx.doi.org/ 10.1038/nn1309

Van der Wel, R. P. R. D., Sebanz, N., & Knoblich, G. (2013). Action perception from a common coding perspective. *People watching: Social, perceptual, and neurophysiological studies of body perception*, 101-120. doi: 10.1093/acprof:oso/9780195393705.003.0007

van Elk, M., Salomon, R., Kannape, O., & Blanke, O. (2014). Suppression of the N1 auditory evoked potential for sounds generated by the upper and lower limbs. *Biological psychology*, 102, 108-117. doi: 10.1016/j.biopsycho.2014.06.007

Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proceedings of the National Academy of Sciences*, *108*(51), 20754-20759.

Warren, R. M., & Obusek, C. J. (1971). Speech perception and phonemic restorations. *Perception & Psychophysics*, 9(3), 358-362.

Waszak, F., & Herwig, A. (2007). Effect anticipation modulates deviance processing in the brain. *Brain Research*, *1183*, 74-82. doi: 10.1016/j.brainres.2007.08.082

Webster, M. J., Ungerleider, L. G., & Bachevalier, J. (1991). Connections of inferior temporal areas TE and TEO with medial temporal-lobe structures in infant and adult monkeys. *Journal of Neuroscience*, *11*(4), 1095-1116.

White, P. M., & Yee, C. M. (1997). Effects of attentional and stressor manipulations on the P50 gating response. *Psychophysiology*, *34*(6), 703-711.

White, P. M., & Yee, C. M. (2006). P50 sensitivity to physical and psychological state influences. *Psychophysiology*, *43*(3), 320-328.

Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology*, 21(3-4), 147-163. doi: 10.1027/0269-8803.21.34.147

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. doi: 10.1016/j.tics.2009.09.003

Winkler, I., Takegata, R., & Sussman, E. (2005). Event-related brain potentials reveal multiple stages in the perceptual organization of sound. *Cognitive Brain Research*, *25*(1), 291-299.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, *269*(5232), 1880-1882.

Yabe, H., Tervaniemi, M., Sinkkonen, J., Huotilainen, M., Ilmoniemi, R. J., & Näätänen, R. (1998). Temporal window of integration of auditory information in the human brain. *Psychophysiology*, *35*(5), 615-619.

Yabe, H., Winkler, I., Czigler, I., Koyama, S., Kakigi, R., Sutoh, T., ... & Kaneko, S. (2001). Organizing sound sequences in the human brain: the interplay of auditory streaming and temporal integration. *Brain research*, *897*(1-2), 222-227.

Young, M. E. (1995). On the origin of personal causal theories. *Psychonomic Bulletin & Review*, 2(1), 83-104. doi: 10.3758/BF03214413

UNIVERSITAT DE
BARCELONA

Department of Clinical Psychology and Psychobiology
Faculty of Psychology
University of Barcelona