UNIVERSITY OF MILANO-BICOCCA
DEPARTMENT OF PSYCHOLOGY

DOCTORAL PROGRAM IN EXPERIMENTAL PSYCHOLOGY, LINGUISTICS
AND COGNITIVE NEUROSCIENCE
XVIII CYCLE

Moving in tune: neuro-functional mechanisms mediating auditory-motor associations

Doctoral thesis by:
CARLOTTA LEGA
711094

Tutor: Dott.ssa ZAIRA CATTANEO
Coordinator: Prof.ssa MARIA TERESA GUASTI

Academic year 2015/2016

Table of Contents

Abstract

In everyday life, our senses are constantly surrounded by many different signals. Our cognitive system may combine information from different sensory modalities in order to choose the appropriate responses and ultimately promote a more adaptive behavior. A large body of research now shows that individuals exhibit consistent crossmodal correspondences between many stimulus features in different sensory modalities. Thus, the human brain shows a systematic tendency to preferentially associate certain features of stimuli across senses.

The first part of the present dissertation aims at investigating the human tendency to spontaneously map the pitch dimension in a spatial format, exploring the experiential bases that could modulate this correspondence. Specifically, in three different studies, we investigated whether and how musical expertise and prior visual experience impact on the pitch-space association, testing two particular populations: congenitally blind individuals and professional musicians. We found that sensorimotor experience due to musical training affects the mental representation of pitch in the horizontal space (Study 1). Conversely, neither musical expertise (Study 1) nor prior visual experience (Study 2) seem to be critical for the pitch-space association to develop in the vertical dimension. Furthermore, we showed that musical experience not only influences motor responses, but also impacts on the representation of peri-personal space, in both the visual and tactile modality (Study 3).

In the second part, we considered the association between auditory pitch and visual size, which refers to the finding that high-pitched sounds are perceptually associated with smaller visual stimuli, whereas low-pitched sounds are associated with larger ones. Pushing this mapping one step further, the goal of Study 4 was to verify whether this cross-modal correspondence, reported so far for perceptual processing, also modulates motor planning. We carried out a series of six different kinematic experiments to verify whether actions implying size processing are affected by auditory pitch. Our results provide evidence for a close link between musical cognition and motor control, by demonstrating an interaction between representation of pitch dimension and representations of action-coded information for grasping, partially modulated and strengthened by musical expertise.

In the last part of the thesis, we tested the possible causal role of different brain regions in mediating auditory-motor associations by means of Transcranical Magnetic Stimulation (TMS). Our findings demonstrated a crucial role of the cerebellum in pitch

processing, extending prior neuroimaging and neuropsychological evidence that suggests a cerebellum involvement in perceptual auditory tasks (Study 5). Furthermore, we showed that inhibitory TMS over premotor cortex impairs the ability to learn and apply auditory-motor associations, and that this effect is greater when a novel association has to be explicitly acquired (Study 6). Overall, the findings reported in this final section provide new evidence in favor of a strong link between the perceptual and the motor systems.

Riassunto

Nella vita di tutti i giorni, i nostri sensi sono costantemente circondati da una moltitudine di segnali sensoriali. Il nostro sistema cognitivo deve combinare informazioni da diverse modalità sensoriali con lo scopo di scegliere la risposta più appropriata al fine di promuovere il comportamento più adattivo. Numerose ricerche hanno mostrato che gli individui esibiscono costanti associazioni cross-modali tra diverse modalità sensoriali. Il cervello umano evidenzia pertanto una tendenza sistematica ad associare preferenzialmente certe caratteristiche degli stimoli tra i sensi.

La prima parte della tesi si focalizza sulla tendenza spontanea ad associare la dimensione dell'altezza sonora con la dimensione spaziale e sulle esperienze che possono modulare tale associazione. Nello specifico, abbiamo studiato se e come una normale esperienza visiva e l'expertise musicale abbiano un effetto sull'associazione altezza-spazio, testando due popolazioni particolari: ciechi congeniti e musicisti professionisti. I risultati hanno dimostrato che l'esperienza senso-motoria legata al training musicale influenza la rappresentazione dei suoni sul piano orizzontale (Studio 1). Al contrario, né l'expertise musicale (Studio 1), né l'esperienza visiva (Studio 2) sembrano essere critici per lo sviluppo dell'associazione suono-spazio nella dimensione verticale. Inoltre, abbiamo mostrato che l'esperienza musicale influenza non solo le risposte motorie, ma anche la rappresentazione dello spazio peri-personale, sia nella modalità visiva che in quella tattile (Studio 3).

In una seconda parte abbiamo invece considerato l'associazione tra frequenza del suono e grandezza, che si riferisce al fatto che i suoni acuti vengono solitamente associati a stimoli visivi più piccoli, mentre suoni gravi sono associati a stimoli più grandi. Lo scopo dello Studio 4 è stato quello di verificare se questa associazione cross-modale, fino ad oggi riportata esclusivamente nella dimensione percettiva, è in grado di modulare anche la programmazione motoria. Abbiamo pertanto effettuato una serie di esperimenti di cinematica, con lo scopo di verificare se le azioni che implicano l'elaborazione della dimensione di grandezza, siano influenzate dalla frequenza uditiva. I nostri risultati hanno evidenziato una robusta interazione tra l'elaborazione della frequenza del suono e la cinematica del movimento, dimostrando una forte associazione tra cognizione musicale e controllo motorio, associazione che viene parzialmente modulata e rinforzata dall'esperienza musicale.

Nell’ultima parte della tesi, abbiamo invece testato il possibile ruolo causale di due differenti regioni cerebrali nel mediare associazioni uditivo-motorie attraverso l’utilizzo della Stimolazione Magnetica Transcranica (TMS). I nostri risultati hanno dimostrato un ruolo cruciale del cervelletto nell’elaborazione dell’altezza tonale (*pitch),* estendendo evidenze neuropsicologiche e di neuroimmagine che suggerivano un coinvolgimento del cervelletto in compiti percettivi. Inoltre, abbiamo mostrano che la TMS inibitoria sulla corteccia pre-motoria ha un effetto negativo sulla capacità di apprendere e utilizzare associazioni uditivo-motorie, e questo effetto è maggiore quando un’associazione nuova deve essere esplicitamente appresa. Complessivamente, i risultati riportati in questa ultima sezione supportano con nuove evidenze un forte legame tra sistema percettivo e sistema motorio.

1.

General introduction

In everyday life, our senses are constantly surrounded by many different sensory signals. Our cognitive system may combine and disambiguate information from different sensory modalities in order to choose the appropriate responses. Moreover, integrating information from different modalities is fundamental to understand whether unisensory signals belong to different objects or events. In the present dissertation, I particularly focused on the neuro-functional mechanisms mediating auditory pitch processing and the interaction between pitch and other cognitive and motor dimensions. On the one hand, we analyzed the crossmodal pitch-space and pitch-size association. Although cross-modal correspondence is now a well-studied phenomenon, its origin is still controversial. With the present research project, we contributed to a better understanding of such associations, investigating whether and how life experiences modulate the association between pitch and space. Specifically, we focused on the role of musical expertise (**Study 1** and **Study 3**) and early visual experience (**Study 2**). Moreover, we moved beyond the perceptual domain, investigating whether processing of pitch information affects the preparation and execution of motor responses requiring size processing (**Study 4**). To address these issues, we used behavioral paradigms and motion capture analyzer. In a further development of the project, we employed Transcranical Magnetic Stimulation (TMS) to investigate the causal role of two key brain regions of the neural network mediating auditory-motor associations. Firstly, we aimed to shed light on the possible causal role of the cerebellum in pitch processing, in light of prior neuroimaging and patients' data suggesting a cerebellar involvement in processing of music and single sound features (**Study 5**). Then, we focused on the role of dorsal premotor cortex in learning new auditory-motor associations (**Study 6**). Overall, the findings reported in this final section provide new evidence in favor of a strong link between the perceptual and the motor systems.

1.1 CROSS-MODAL CORRESPONDENCES

The terms *crossmodal correspondences* (Gilbert et al., 1996; Mondloch & Maurer, 2004; Spence, 2011) or *crossmodal associations* (Martino & Marks, 2001) refer to the tendency to systematically associate certain dimensions of stimuli across senses (see Marks, 2004; Spence, 2011, for reviews). Cross-modal associations have been demonstrated by means of different paradigms and tasks. More commonly, speeded classification tasks are used to study the interaction between different dimensions. Typically, the paradigm requires participants to identify a stimulus in one modality while an irrelevant stimulus in another modality is presented. If the two modalities are associated, then it should be difficult to attend selectively to one of the pairs, leading to an effect on the response. Importantly, such correspondences are common and consistent in the general population. Furthermore, cross-modal associations have been documented between almost all possible combinations of sensory modalities. Accordingly, a large number of studies reported associations between vision and touch (Martino & Marks, 2000; Simner & Ludwig, 2009), sound and touch (Yau, Olenczak, Dammann & Bensmaia, 2009), audition and flavors (Crisinel & Spence, 2010, 2012) and many others. The origin of cross-modal correspondences is still controversial, but at least three different factors may be responsible for their development. One possibility is that cross-modal correspondences may have neural and structural bases and may arise from the natural organization of our perceptual system (Mondloch & Maurer, 2004; Marks, 1978). In this framework, recent studies suggest the idea of a generalized system in the parietal cortex for representing magnitude (Walsh, 2003), which could explain association between stimuli that are magnitude-based. Recently, this magnitude system has been proposed to include other dimensions (including the pitch dimension), that are also associated to space and quantity (see Bottini & Casasanto, 2013). Cross-modal associations may also reflect cross-modal correspondences on natural statistics of the environment, including simple co-occurrence of events (i.e. bigger/smaller objects producing lower/higher sounds, respectively) (Spence, 2011; Parise et al., 2014). Finally, the semantic hypothesis suggests that the cross-modal correspondences may arise because of different stimuli evoking similar semantic concepts (Martino & Marks, 1999; Walker & Walker, 2012). One of the most cited ‘semantic’ example resides in the use of spatial words (“high” and “low” for high and low-pitched sounds) to describe stimuli that vary in pitch, which would explain the pitch-elevation correspondences. Most frequently all these factors co-exist, so that it is difficult to isolate the possible contribution of each single factor. Furthermore, different cross-modal correspondences may operate in different manner and have different origins (some learned and some innate) (Spence, 2011; Spence & Deroy, 2012). This would also explain why some studies reported cross-modal correspondences at birth (Mondloch & Maurer, 2004; Walker et al., 2009) while other seem to be semantically mediated (Cohen Kadosh et al., 2009). As hinted earlier, cross-modal associations have been already demonstrated between different pairs of stimulus. Considering specifically the pitch dimension, auditory stimuli have been associated to visual elevation (Ben-Artzi & Marks, 1995, Evans & Treisman, 2009; Proctor & Cho, 2006, Rusconi et al., 2006; Lidji et al., 2007), direction of movement (Maeda et al., 2004; Sadaghiani et al., 2009; Deroy et al., 2016), brightness (Ludwig et al., 2011; Martino & Marks, 1999; Mondloch & Maurer, 2004) and size (Bien et al., 2012; Gallace & Spence, 2006; Mondloch & Maurer, 2004, Evans & Treisman, 2009).

1.1.1 Pitch - space correspondence

The cross-modal correspondences between auditory pitch and visual elevation represents one of the strongest association reported in cross-modal literature (Ben-Artzi & Marks, 1995; Maeda et al., 2004; Rusconi et al., 2006; Parise et al., 2014). Some authors explained this association as semantically mediated. Indeed, already Stumpf in 1883 claimed that several languages use spatial metaphors like ‘low’ and ‘high’ to describe stimuli that vary in pitch. According to this interpretation, Gallace & Spence (2006) demonstrated that the simple presentation of the words ‘low’ and ‘high’ had the same effect on participants’ performance as presenting high and low-pitched sounds. Nonetheless, subsequent investigations demonstrated that also pre-lingual infants (Dolscheid et al., 2014; Walker et al., 2009; see also Braaten, 1993) and remote populations who do not use spatial attributes for describing pitch (Parkinson et al., 2012) represented pitch along a vertical spatial dimension. Taking together, these results seem to go toward a structural or statistical interpretation of the pitch-elevation correspondence. Specifically, as a structural correspondence, pitch dimension may be interpreted as a part of a generalized system in the inferior parietal cortex for representing magnitude, which would include also number, space and time (Walsh, 2003; see Bottini & Casasanto, 2013 for a discussion about the pitch dimension). Such mechanism for coding magnitude may represent the neural underpinning of different cross-modal correspondences, including time-space, number-space and also pitch-space associations. Alternatively, as a statistical correspondence, pitch-elevation association may be interpreted as a consequence of learned statistical correlations in the environment (Ernst, 2006; Parise et al., 2014), which also represent an advantage for both our cognitive and neural systems (see for instance Kadosh et al., 2007). Accordingly, in nature the resonant frequency of an object is related to its size (Grassi, 2005; Spence & Zampini, 2006) and concurrently, larger objects produce louder sounds. To push it a step forward, the size of an object is directly correlated to its elevation, so that it is more common in nature that larger objects are also heavier than smaller objects and therefore is less frequent to see them flying or staying at high elevations. Taking together, these observations may account for both pitch-size and pitch-elevation correspondences.

Converging findings (Rusconi et al., 2006; Lidji et al., 2007; Nishimura & Yokosawa 2009; Cho et al., 2012; Stewart et al., 2013; Pitteri et al., 2015; Weis et al., 2016a) revealed that the mental pitch line affects manual motor responses: when an up or down response to a high or low pitched tone is required, individuals typically perform better for the mapping of the high tone to the up response (i.e., a key positioned in the upper part of the keyboard or response box) and the low tone to the down response (i.e., a key positioned in the lower part of the keyboard or response box) than for the opposite mapping. This effect has been labelled the *Spatial Musical Association of Response Codes* (SMARC) and it has been documented along both horizontal and vertical dimensions. In particular, in the vertical space, the SMARC effect refers to the tendency to associate bottom responses with low-pitched tones, and top responses with high-pitched tones. In the horizontal space, the SMARC effect refers to the tendency to associate left responses with low-pitched tones, and right responses with high-pitched tones (see Rusconi et al., 2006).

The role of musical expertise. Compared to the vertical representation of pitch, the horizontal mapping may be less automatic. Indeed, it has been suggested that the association between horizontal space and pitch is dependent on musical training, as it is typically stronger in musicians than in non-musicians. For instance, Rusconi and colleagues (Rusconi et al., 2006) reported that, when pitch was processed implicitly, only musicians showed a facilitation for the congruent mapping, in terms of both reaction times and accuracy (i.e., high pitch/right space, low pitch/left space). These results were further corroborated by Lidji et al.'s study (2007), which reported an automatic association of pitch with the horizontal space only in trained musicians. In turn, musically naïve participants tend to associate tones with the horizontal space only when the pitch dimension has to be explicitly processed (Rusconi et al., 2006; Lidji et al., 2007). Furthermore, the pitch-space correspondence in the horizontal plane seems to be modulated by the presence of a reference tone. In particular, Cho and colleagues (2012) demonstrated that even non-musicians show a horizontal SMARC effect in a timbre-judgment task, but only when a referent tone was present. Furthermore, Weis et al., (2016a) showed an influence of both musical timbre and pitch range only in the horizontal space, whereas no modulation of these features was reported along the vertical dimension. At the present juncture, therefore, scattered evidence suggests that differences observed between vertical and horizontal pitch-space correspondences may rely on musical training.

However, no study has so far directly investigated the impact of instrumental expertise on the SMARC effect. In fact, previous studies comparing musicians and non-musicians on the SMARC effect did not control for the instrument played by the musicians' group (Rusconi et al., 2006; Lidji et al., 2007). This could represent a possible confound, especially considering that for some musicians the pitch-space correspondence may be preferentially aligned along the horizontal axis (such as in the case of pianists), whereas for others it may be preferentially aligned along the vertical axis (such as in the case of clarinetists). Moreover, because in a recent study musical timbre was found to interfere with pitch-space correspondences (Weis et al., 2016a), it might be possible that the spatial representation of pitch would be modulated by the timbre of sounds, with piano and clarinet tones preferentially inducing horizontal and vertical mappings, respectively.

The role of vision. An interesting issue is whether the association between pitch and space is mediated by visual experience. Indeed, although the origin of the association between pitch and space is still under debate, some authors suggest that it may be a consequence of internalizing naturally occurring audio-visual correlations in the environment (Pratt, 1930; Proctor & Vu, 2002; Parise & Spence, 2012; Parise et al., 2014; Spence & Deroy, 2012; Deroy & Spence, 2013). In this framework, the cross-modal pitch-space (but also pitch-size, see below correspondence would reflect natural correlations occurring between the physical properties of the external world (Parise et al., 2014). For instance, larger objects are usually heavier than smaller objects and therefore is less frequent to see them placed at high elevations, hence possibly explaining the cross-modal correspondence between pitch and elevation. According to this hypothesis, visual experience would be important in mediating the formation of these cross-modal associations.

Previous studies demonstrated that the lack of prior visual experience affects spatial representations (for a review, see Cattaneo et al., 2008). For instance, blind individuals tend to rely more on body-centered spatial reference frames compared to sighted people, and to use different strategies to process and maintain spatial information in memory (e.g., Cattaneo et al., 2011; Monegato et al., 2007; Noordzij et al., 2006; Röder et al., 2007; Postma et al., 2007). Nonetheless, blind individuals likely use mental spatial representations similar to those generated by sighted individuals when representing abstract concepts, like numerical magnitude and temporal order. Specifically, blind individuals also mentally represent numbers as aligned along a left-to-right oriented mental number line, with low numbers spatially collocated on the left side of the space and high numbers on the right, as sighted people typically do (e.g., Castronovo & Seron, 2007; Cattaneo et al., 2010, 2011; Szűcs & Csépe, 2005; Rinaldi et al., 2015; but see Pasqualotto et al., 2014). Moreover, past and future events are likely to be associated to the left and right side of space respectively in blind individuals as well (Bottini et al., 2015), as it is the case of sighted individuals (Casasanto & Bottini, 2014; Fuhrman & Boroditsky, 2010; Santiago et al., 2007; Ulrich & Maienborn, 2010). Hence, early blindness does not seem to prevent the use of spatial representations to refer to concepts like magnitude and time, although there is also evidence that these representations may be accessed differently by the blind than by the sighted (Crollen et al., 2013; Pasqualotto et al., 2014). Also, blind individuals may rely less than sighted individuals on spatial representations when maintaining verbal information in short-term memory (Bottini et al., 2016).

Interestingly, a recent study by Deroy et al., (2016) explored the possible role of prior visual experience in mediating associations between pitch and direction of tactile movement, in a group of sighted, early blind and late blind individuals. Interestingly, while sighted individuals showed some systematic associations between direction of movement and changes in pitch (for instance, responding faster to tones increasing in pitch when they were associated with an outward movement than with an inward movement), early and late blind individuals did not show any association (Deroy et al., 2016), suggesting that vision may be critical in mediating pitch-movement direction correspondences. To date though, no studies have directly investigated whether early blind individuals also organize pitch along a music mental line.

1.1.2 Pitch - size correspondence and motor planning

Robust evidence supports the hypothesis that pitch-size correspondence is grounded in the statistics of the external world. Specifically, in nature a correlation exists between the size of an object and its relative resonant frequency: the lower the frequency of the sound, the larger the object that is generating the sound (Coward & Stevens, 2004; Grassi, 2005; Grassi, Pastore, & Lemaitre, 2013). A widely acknowledged account of crossmodal correspondence suggests that humans would refer to the natural mapping between auditory pitch and visual size when processing and integrating new audiovisual information (Ernst, 2006; see for a discussion Spence, 2011). Accordingly, when hearing a high/low-pitched sound, one would expect a small/large-size object to have produced it (Grassi, 2005; Grassi et al., 2013; see Parise & Spence, 2013). Critically, although vision and audition are known to play a critical role in the planning and control of action (Goodale & Humphrey, 1998; Sedda, Monaco, Bottini, & Goodale, 2011), pitch-size correspondence has been so far documented only for perceptual processing. It is therefore possible that the motor system might be as well affected by pitch-size correspondence. Indeed, humans can estimate the size of objects dropped on a surface by the frequency of the sound they produce (Grassi, 2005; Grassi et al., 2013). In turn, the frequency of this sound can be informative for the planning of a subsequent reach-to-grasp movement (see Sedda et al., 2011).

According to “A Theory Of Magnitude” (ATOM) model, magnitude-related information would be processed by a generalized system located in the inferior parietal cortex (Bueti & Walsh, 2009; Walsh, 2003). In this view, prothetic dimensions (i.e., concerned with quantitative variation; Stevens, 1957), such as quantity, space and time, all share a magnitude code. These dimensions would mutually operate on similar magnitude representations, because of the need to learn about the environment for acting on it (Walsh, 2003). Accordingly, increasing evidence has shown that symbolic number processing influences action planning (Andres, Davare, Pesenti, Olivier, & Seron, 2004; Lindemann, Abolafia, Girardi, & Bekkering, 2007). More recently, however, some authors have proposed an even broader scope of ATOM to include metathetic (i.e., concerned with qualitative variation, Stevens, 1957) dimensions as well, such as auditory pitch (see Bottini & Casasanto, 2013, for a discussion). Indeed, auditory pitch has been found to be consistently associated not only with size, but also with space (see Spence, 2011, for a review). For instance, in the vertical plane individuals associate high pitches with high positions in space (Chiou & Rich, 2012; Pratt, 1930; Rusconi, Kwan, Giordano, Umilta’ & Butterworth, 2006). Likewise, auditory pitch is also represented along a horizontal spatial dimension, with high tones preferentially mapped to rightward spatial positions, while low tones to leftward positions (e.g., Lega, Cattaneo, Merabet, Vecchi, & Cucchi, 2014; Lidji, Kolinsky, Lochy & Morais, 2007; Rusconi et al., 2006; but see also Trimarchi & Luzzatti, 2011). Nevertheless, no study has so far explored whether auditory pitch shapes size processing in action planning and whether musical expertise impacts on that processing.

Indeed, in musicians the auditory and motor system are closely interconnected. Making and perceiving music requires indeed the integration of sensory and motor information, as well as a precise monitoring of the performance. Expert musicians strongly tend to associate the sounds from their instrument with the movements that produce those sounds and vice versa. This strong sound-action association is a direct consequence of years of training, thank to which musicians learned the contingencies between sounds and movements (Chen et al., 2012; Elsner & Hommel, 2001; Lahav et al., 2007). Interestingly, Drost and colleagues (Drost et al., 2005a,b) demonstrated that the perception of sounds from their instrument facilitates and primes the corresponding action in musicians, even when sounds are irrelevant to the task. Auditory and motor influences are bidirectional, so that not only hearing musical sounds can automatically trigger the corresponding actions (see Drost et al., 2005a,b; Keller & Koch, 2008), but also performing movements can alter sound perception (Repp & Knoblich, 2007). Furthermore, previous studies already suggested that musicians and non-musicians differ in their ability to automatically activate a mental representation of pitch, when this dimension is irrelevant to the task (Rusconi et al., 2006; Lidji et al., 2007). These results are also supported by studies demonstrating that musicians process some dimensions of music in an automatic way (Koelsch et al., 1999; Pantev et al., 2001; Tervaniemi, Just, Koelsch, Widmann, & Schröger, 2005).

1.2 BRAIN NETWORK OF AUDITORY-MOTOR INTEGRATION

Auditory-motor integration is crucial for the learning and production of music and speech. Based on work in both animals and humans, the network of brain regions engaged in linking sound and action is thought to involve the auditory dorsal stream, including the posterior auditory, inferior parietal and premotor cortices (Brown et al., 2013; Brown, Zatorre, & Penhune, 2015; Chen, Penhune, & Zatorre, 2009; Chen, Rae, & Watkins, 2012; Hickok & Poeppel, 2004; Zatorre, Chen, & Penhune, 2007). Hickok & Poeppel (2004) proposed a dual–stream model specific for speech processing, in which the ventral auditory stream is responsible for mapping sounds onto meaning and the dorsal stream is critical for mapping sounds onto representations of articulatory movements. Lately, Warren et al. (2005) proposed a general model for auditory-motor transformations and pointed out the critical role of the dorsal auditory pathway, where the planum temporale analyzes sounds that are relevant for the motor system, such as speech and melodies, which are then transformed into motor representations in prefrontal, premotor and motor regions.

In this second part of the research project presented here we tested the critical role of cerebellum in auditory processing and the role of dorsal premotor cortex (dPMC) in auditory-motor associations learning by means of Transcranical Magnetic Stimulation (TMS). Compared to neuroimaging techniques, TMS allows to investigate the possible causal relevance of the targeted brain region in a specific cognitive function (Miniussi, Harris, & Ruzzoli, 2013; Pascual-Leone, Walsh, & Rothwell, 2000; Sack et al., 2009; Silvanto & Pascual-Leone, 2012; Walsh & Cowey, 2000).

1.2.1 The role of Cerebellum in auditory processing

The cerebellum is traditionally considered a “motor controller” and its role in the acquisition of motor skills is well established (Manto et al., 2012; Shiffrin & Schneider, 1984). Nevertheless, accumulating evidence suggests that the cerebellum may play a critical role in non-motor functions, contributing to cognitive and emotional processing (Ackermann, 2008; Schmahmann & Caplan, 2006; Stoodley, 2012, Strick et al., 2009; for reviews, D'Angelo & Casali, 2013). In particular, a cerebellar involvement has been found in high-order cognitive processes such as language (Ackermann, 2008), working memory (Ben-Yehudah, Guediche, & Fiez, 2007), and spatial processing (Lee et al., 2005; Molinari et al., 2008). Furthermore, the cerebellum seems to play a role in the generation of sensory predictions (Blakemore, Frith, & Wolpert, 2001; Miall, 1997), optimizing perception (Roth et al., 2013). Accordingly, patients with cerebellar lesions are often impaired in processing visual information, such as in deciding speed and direction of moving stimuli (Ivry & Diener, 1991; Jokisch et al., 2005; Thier, Haarmeier, Treue, & Barash, 1999). Neuroimaging evidence also shows that the cerebellum is significantly active in neurologically unimpaired individuals during sensory discrimination, such as visual (and auditory) motion discrimination (Baumann & Mattingley, 2010). Also, interfering with cerebellar activity via brain stimulation has been found to impact on visual processing (Cattaneo et al., 2014, Renzi et al., 2014). The cerebellum is also involved in auditory processing, where it plays a critical role in timing (e.g., Grube, Lee, Griffiths, Barker, & Woodruff, 2010; Kawashima et al., 2000; Teki, Grube, Kumar, & Griffiths, 2011). Cerebellar activations have also been observed in healthy subjects during passive listening of both music (Brown, Martinez, & Parsons, 2004; Petacchi, Kaernbach, Ratnam, &

Bower, 2011) and speech (Callan et al., 2006; Callan, Kawato, Parsons, & Turner, 2007). Moreover, pitch discrimination and melody discrimination (Gaab, Gaser, Zaehle, Jancke, & Schlaug, 2003; Griffiths, Johnsrude, Dean, & Green, 1999; Parsons, 2001; Petacchi et al., 2011), as well as sound intensity and duration discrimination (Belin et al., 2002), activate cerebellar regions. For instance, Petacchi and colleagues (2011) showed that cerebellar activity significantly increased during a pitch discrimination task compared to passive listening, with the cerebellum responding more when the difficulty of the discrimination to be performed increased. Importantly, cerebellar activity during auditory discrimination tasks has been consistently observed even in the absence of any motor or cognitive component (Petacchi et al., 2005). Patients' findings support evidence obtained in healthy individuals: in fact, cerebellar disorders are often associated to deficits in melody recognition (Tölgyesi & Evers, 2014), in discriminating small differences in sound duration (Frings et al., 2006; Ivry & Keele, 1989) and in pitch discrimination (Parsons, Petacchi, Schmahmann, & Bower, 2009). Although there is considerable evidence that the cerebellum contributes to auditory perception (see Petacchi et al., 2005), the precise role of the cerebellum in different aspects of sound processing is not completely clear. In particular, whilst neuroimaging (Petacchi et al., 2005, 2011) and patients' (Parsons et al., 2009) studies converge in indicating a role of the cerebellum in pitch processing, whether the cerebellum also contributes to other sound features such as timbre (i.e., the property of a sound which allows a person to distinguish musical instrument when pitch, loudness and duration remain identical), is less clear. Indeed, whereas some neuroimaging studies reported significant cerebellar responses to sound timbre (Alluri et al., 2012; Reiterer, Erb, Grodd, & Wildgruber, 2008), in other studies investigating timbre processing cerebellar activations were not considered (Halpern, Zatorre, Bouffard, & Johnson, 2004).

1.2.2 The role of Premotor cortex (PMC) in auditory-motor associations

The premotor cortex (PMC) has traditionally been divided into ventral (vPMC) and dorsal (dPMC) subregions, both of which represent a crucial node in the auditory-motor network. In the visual domain it has been proposed that these regions are involved in direct and indirect visuo-motor associations, respectively (Davare, Andres, Cosnard, Thonnard, & Olivier, 2006; Hoshi & Tanji, 2006, 2007). Direct associations are those that encode a fixed, one-to-one mapping between an object and an action, for example a cup and the hand shape needed to grasp it. In contrast, indirect sensorimotor associations are more abstract and flexible and, once established, a sensory stimulus may represent a conditional rule indicating which response to select among different possible responses (Amiez, Hadj-Bouziane, & Petrides, 2012; Hoshi & Tanji, 2007; Petrides, 2005; Zatorre et al., 2007). In parallel with the visual system, previous neuroimaging studies (Chen, Penhune, & Zatorre, 2009; Chen, Rae, & Watkins, 2012; Herholz, Coffey, Pantev, & Zatorre, 2015) suggest a similar functional dissociation in the auditory domain. In particular, the dPMC is hypothesized to be important for learning flexible and/or arbitrary associations between sounds and actions (Chen et al., 2009; Herholz et al., 2015; Zatorre et al., 2007) and for selecting movements in the appropriate context (Davare et al., 2006). Consistent with its role in abstract and higher-order aspects of sensorimotor integration (Cisek & Kalaska, 2004; Hoshi & Tanji, 2006; Petrides, 2005; Zatorre et al., 2007), activity in the dPMC has been shown to be sensitive to a rhythm's metric structure (Chen, Zatorre, & Penhune, 2006), and inactivation of the dPMC impairs conditional motor behaviors (Kurata & Hoffman, 1994).

In research on music, studies in both trained and untrained individuals show that dPMC is engaged when people listen to, learn, play or imagine musical sequences. Learning to play an instrument can be considered a special case of sensory-motor association learning, in which action and perception become intricately interconnected. In all of these cases, individuals are learning, or have an established association between sound and action. For example, in musicians the dPMC is engaged when either listening to music without playing or when playing without auditory feedback (Bangert et al., 2006; Baumann et al., 2007). Further, the dPMC is engaged when musicians listen to and learn to play new melodies (Brown et al., 2013). Relatively few studies have investigated the specific brain networks involved in learning new auditory-motor associations. In an early study, Bermudez and Zatorre (2005) showed increased activity in rostral dPMC after non-musicians had been trained to associate four chords to four button presses. Similarly, Lahav et al., (2007) demonstrated the activation of the PMC when subjects listened to melodies that they had been trained to play, but not when they listened to un-trained melodies composed of different notes. This result is in line with previous studies in musicians showing premotor activation when they listened to a rehearsed musical piece (Bangert et al., 2006; Lotze, Scheler, Tan, Braun, & Birbaumer, 2003). In a recent fMRI study, non-musicians were trained to play short melodies on the piano. After training, dPMC showed greater activation when people listened to or imagined playing the melodies they had learned (Herholz et al., 2015). Similarly, a recent study in non-musicians showed that dPMC was engaged during learning to play short piano melodies, and that the degree of engagement was related to improvements in performance on a pitch matching task (Chen et al., 2012).

1.3 AIMS OF THE RESEARCH PROJECT

Cross-modal correspondences and multisensory integration are a florid field of research where many questions are still unsolved. The contribution of this thesis project should be interpreted in three main directions.

Study 1, 2 and 3. Firstly, we wanted to better understand the mechanisms at the bases of the cross-modal correspondences between pitch and space. Analyzing the cognitive, perceptual and cultural factors which modulate such association might help to clarify its origin and function. To this aim, in **Study 1** we specifically focused on the role of musical expertise. Indeed, in musicians the association between space and auditory pitch becomes stronger and stronger thanks to years of training. On the one hand reading score and musical notation is itself a “spatial” language, which seems to be able to modify spatial mapping in musicians (see Stewart et al., 2004). On the other hand, expert musical performance requires precise spatial organization of instrument-dependent movements. For this reason, we decided to compare a group of pianists with the group of clarinet players, to assess the possible influence of instrumental expertise on the spatial representation of pitch, given that pianists have a stronger association of pitches along the horizontal plane, whereas clarinet players along the vertical plane. In **Study 2** we analyzed the role of early visual experience on pitch-space correspondence, in light of previous evidence suggesting that it may be a consequence of internalizing naturally occurring audio-visual correlations in the environment (Pratt, 1930; Proctor & Vu, 2002; Parise & Spence, 2012, Parise et al., 2014; Spence & Deroy, 2012; Deroy & Spence, 2013). To investigate the role of visual experience in shaping the association between pitch and elevation, we tested a group of early blind individuals. Finally, in **Study 3** we explored whether the pitch dimension was able to impact on the representation of peri-personal space, by means of a tactile and visual bisection task. So far, the effect of pitch on the representation of space has only been investigated in the visual modality (e.g., Rusconi et al., 2006; Lidji et al., 2007; Nishimura

& Yokosawa, 2009; Cho et al., 2012; Ishihara et al., 2013). Once more, we were interested in verifying whether musical expertise is able to modulate the representation of external space when this is explored visually or haptically.

Study 4. The second main goal of the thesis was to investigate the functional connection between pitch dimension and action planning. In the last few decades, many authors have emphasized that cognitive representation of perceptual (and semantic as well) information cannot be fully understood without considering their impact on action (see Gallese & Lakoff, 2005). In this framework, not much is known about the role of pitch dimension in the coupling between motor control and cognitive process. To this aim, in **Study 4** we conducted six different kinematic experiments. In experiments 1 to 4 we investigated the effects of auditory pitch on motor planning, by requiring participants to perform different manual actions primed by sounds. In experiment 5 we verified the automaticity of this effect by using an implicit task, were participants' attention was no more directed to the pitch dimension. Finally, given the strong sound-action association in expert musicians we also investigated the possible effect of instrumental expertise. Indeed, musicians strongly tend to associate the sounds from their instrument with the movements that produce those sounds, possibly influencing motor planning in response to musical tones.

Study 5 and 6. Finally, in the third part of this dissertation we mostly focused on the neural aspect mediating pitch processing and auditory-motor associations, by means of Transcranical Magnetic Stimulation (TMS). TMS is a technique that allows to directly investigate the causal involvement of brain regions in cognitive and motor processing. Furthermore, participants in TMS experiments act as their own controls overcoming some of the limitations intrinsic in patients' studies, such as potential differences in pre-morbid ability, and variability depending on high heterogeneity of lesions' sizes and gravity. The aim of **Study 5** was to verify the critical role of cerebellum in pitch processing, in consideration of a growing literature, claiming for a role of cerebellum beyond motor control. In this framework, a large amount of studies showed cerebellar involvement in high cognitive and perceptual functions, providing new evidence in favor of a strong link between the perceptual and the motor systems. Finally, **Study 6** aimed to shed further light on the role of dorsal premotor cortex in auditory-motor integration and learning. Indeed, the premotor cortex and specifically, the dorsal premotor cortex (dPMC), seems to play a key role in mapping the association between a sound and the movement used to produce it, as a part of the network mediating sensori-motor interactions.

2.

Study 1:

The role of instrumental expertise and musical timbre on the spatial representation of pitch

2.1 INTRODUCTION

In the present study, we compared a group of professional pianists with a group of professional clarinetists in a typical horizontal and vertical SMARC paradigm, to assess the possible influence of instrumental expertise on the spatial representation of pitch. Furthermore, by manipulating the timbre of the auditory tones (i.e., piano and clarinet timbre), we aimed at assessing the role of musical timbre on the SMARC effect in musicians with different instrumental expertise. We hypothesize that if the sensorimotor training related to instrumental expertise would affect the representation of pitch, then different pitch-space correspondences should emerge between pianists and clarinetists, in the vertical and in the horizontal space. In particular, a stronger association in the horizontal plane should be observed in pianists than in clarinetists, whereas a stronger association in the vertical plane should be observed in clarinet players. Yet, because some evidence speaks in favor of a universal mapping between auditory frequency and elevation it might be likely as well that in the vertical space no difference between groups will be observed. Finally, we hypothesize that if musical timbre would interfere with the spatial representation of pitch, then a modulation of timbre over the SMARC effect should be found, with piano tones activating preferentially a horizontal mapping and clarinet tones a vertical mapping.

2.2 METHOD AND MATERIALS

2.2.1 Participants

Twenty-six professional musicians took part in the study. All musicians had comparable levels of musical experience (average years of musical training: 19.38 years; SD: 8.06) and attended the Conservatory or Academic music program. Half of the participants (*N*=13; 7 males; mean age: 25.76 years; *SD*: 10.11) were pianists. Among them, three reported to have some experience with the guitar and other three were also singers. The other half of the participants were professional clarinetists (*N*=13; 10 males; mean age: 30.07 years; *SD*: 4.51). All of them reported to have experience with piano. Indeed, at least 3 years of piano training are mandatory in the classical Italian conservatory program. All participants were right-handed, as assessed by means of the Edinburgh Inventory (Oldfield 1971). Written informed consent was obtained from all participants before the experiment. The protocol was approved by the local ethical committee. All procedures performed were in accordance with the Declaration of Helsinki.

2.2.2 Stimuli

The auditory stimuli consisted of two low-pitched tones (C1 and E2, 32.7 Hz and 41.20 Hz, respectively) and two high-pitched tones (C6 and E5, 1046.5 Hz and 659.3 Hz, respectively). Each tone was synthesized with a piano timbre and a clarinet timbre, for a total of 8 auditory stimuli. They were all normalized in loudness at 0 db, by means of the software Audacity (http://audacity.sourceforge.net/). In particular, we adopted the peak normalization procedure, wherein the gain is changed to bring the highest signal peak to a given level (in our case 0 db, the loudest level allowed). Although all the sounds used were normalized at 0 db, auditory stimuli may differ in terms of perceived loudness. To rule out this possible confound, we conducted a preliminary experiment, which demonstrated that sounds were perceived as equal in terms of loudness (for a similar procedure, see Rinaldi et al. 2016).

2.2.3 Procedure

Participants seated comfortably in a dimly lit room and stimuli were binaurally delivered through professional headphones (Sennheiser HD 280 Pro headphone). Figure 1 shows the timeline of an experimental trial. Each trial consisted of a black fixation cross presented for 250 ms, and followed by a silent interval of 400 ms, preceding the presentation of the target auditory stimulus (500 ms tone). Participants were asked to judge whether the tone was played by a piano or by a clarinet (timbre task). The instructions emphasized both speed and accuracy. Each block was preceded by 8 practice trails, to help participants familiarize with the task. Participants took part in four experimental blocks: two with the keyboard aligned along the horizontal plane and two with the keyboard aligned along the vertical plane. In the first horizontal/vertical condition, the piano response was assigned to the left/bottom side and the clarinet response was assigned to the right/top side, whereas in the second condition a reversed assignment was given. Within each block, trials were presented in a pseudorandom order, to avoid the same pitch played by the same instrument occurring twice in a row. Each of the four tones were presented 12 times for a total of 96 trials in each block. The order of blocks was counterbalanced across participants. In the horizontal condition, participants had to press a left and a right key (button Q and P, respectively) of a computer keyboard. In the vertical condition, the computer keyboard was rotated into the radial plane, so that participants had to press the same keys of the horizontal condition, which were now in a top-bottom alignment. Right hand was always in the right/top position in the horizontal/vertical condition. The software E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA) was used for stimuli presentation and data collection. The whole experiment lasted about 25-30 min.

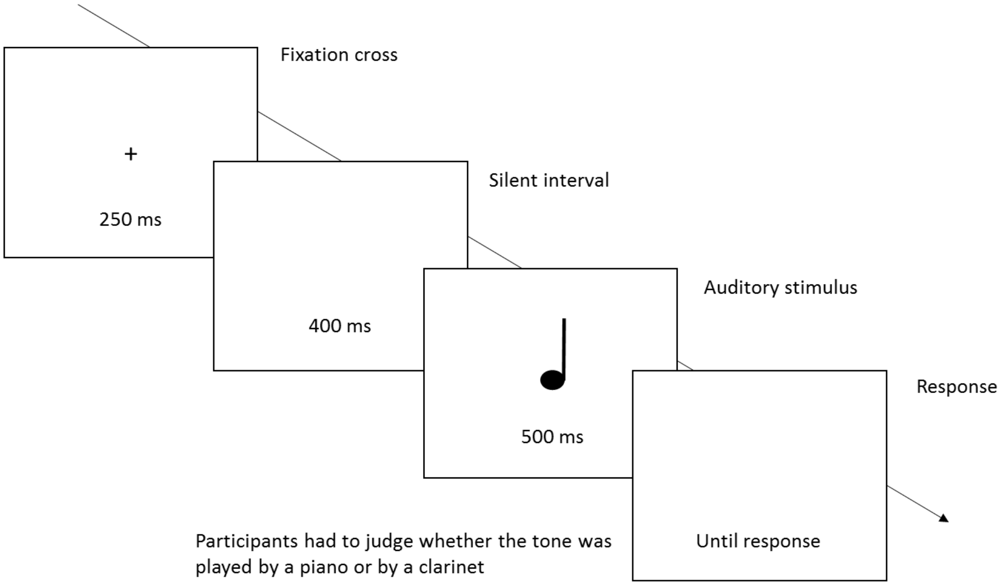

Figure 1. The experimental timeline. Each trial consisted of a black fixation cross presented for 250 ms and followed by a silent interval of 400 ms preceding the presentation of the target auditory stimulus (500 ms tone). Participants were asked to judge whether the target tone was played by a piano or by a clarinet (timbre task).

2.3 RESULTS

Because our study equally emphasized accuracy and processing speed, we combined both response speed and accuracy into a single score performance. Participants' performance was thus analyzed by computing inverse efficiency scores (IES), which are typically used to discount possible criterion shifts or speed-accuracy trade-offs in participants' performance (Akhtar & Enns 1989; Goffaux, Hault, Michel, Vuong, & Rossion, 2005). Specifically, IES were obtained by dividing the median response times (RT) by correct response rates (Townsend & Ashby, 1978), so that lower values on this measure (expresses in ms) indicate better performance.

In accordance with previous studies on the SMARC effect (Lidji et al. 2007; Weis et al. 2016a), the statistical analysis was performed by separately computing the difference between right and left responses in the horizontal plane, and between top and bottom responses in the vertical plane, for the IES values (dIES). Accordingly, positive values indicate faster responses for the left/bottom key, whereas negative values indicate faster responses for the right/top key. The SMARC effect would be, therefore, reflected by a main effect of pitch and, specifically, by more positive values for low pitches compared to high pitches, in both the horizontal and vertical response settings.

For the horizontal response setting, the dIES were submitted to a three-way ANOVA with timbre (clarinet tones and piano tones) and pitch (low and high) as within-subjects variables and group (clarinettists and pianists) as between-subjects variable. Results yielded a significant main effect of pitch (F(1, 24)=8.58, p<.01, η_p^2=.26), with more positive dIES for low pitches (M=19.50 ms) compared to high pitches (M= -26.16 ms), indicating the occurrence of the SMARC effect. Critically, the main effect of the pitch was modulated by both the timbre and the group, as revealed by the significant interaction between pitch and group (F(1, 24)=4.86, p=.03, η_p^2=.16) and the significant interaction between pitch and timbre (F(1, 24)=6.50, p=.01, η_p^2=.21). Specifically, planned t-test comparisons showed a significant difference between high (M=-37.74 ms) and low (M=-42.29 ms) pitches in the group of pianists (t(12)=3.15, p=.008). On the contrary, no difference was found between high (M=-14.58 ms) and low (M=-3.27 ms) pitches for the group of clarinet players (t(12)<1, p=.54) (see figure 2A). Moreover, t-test comparisons also yielded a significant difference between high (M=-43.18 ms) and low (M=35.29 ms) pitches, indicating the presence of a SMARC effect, for the piano timbre (t(25)=-4.14, p<.01). On the contrary, no difference was found between high (M=-9.14 ms) and low (M=3.72 ms) pitches for the clarinet timbre (t(25)<1, p=.58) (see figure 2B). None of the other main effects or interactions reached the significance.

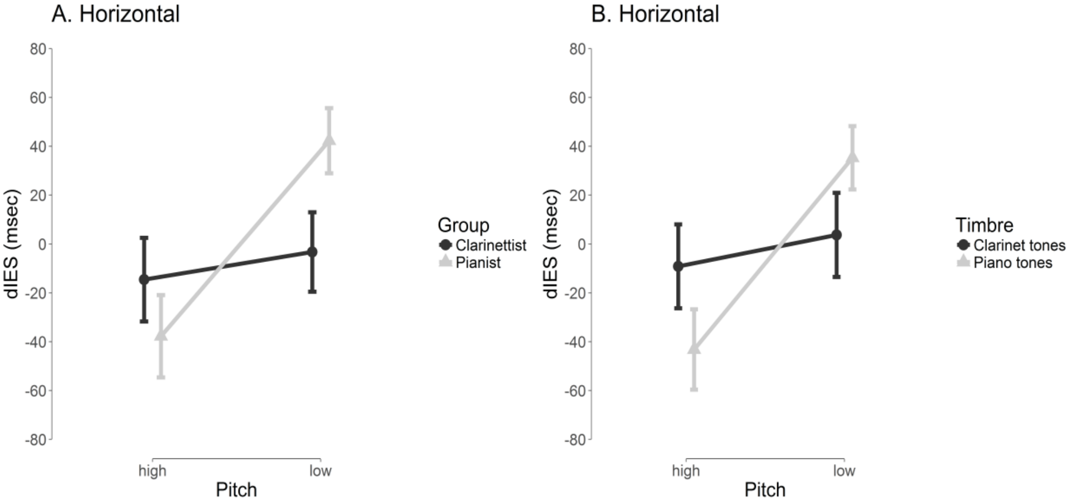

Figure 2. Differences in Inverse Efficiency Scores (dIES) between right and left responses in the Horizontal response setting. Positive dIES indicate faster responses for the left-side key compared to the right-side key, and vice versa for negative dIES**. A.** dIES as a function of Pitch (high, low) and Group (pianists, clarinetists): a significant SMARC effect was found only for the group of pianists, but not for the group of clarinetists. Error bars represent ± 1 SEM. **B.** dIES as a function of Pitch (high, low) and Timbre (clarinet tones, piano tones): a significant SMARC effect was found only for the piano tones, but not for the clarinet tones. Error bars represent ± 1 SEM.

In analogy to the horizontal space, for the vertical response setting, the dIEs were submitted to a three-way ANOVA with timbre (clarinet tones and piano tones) and pitch (low and high) as within-subjects variables and group (clarinettists and pianists) as between-subjects variable. Results showed a main effect of pitch, (F(1, 24)=27.31, p<.01, η_p^2=.53), with more positive dIES for low pitches (M=42.05 ms) compared to high pitches (M= -33.10 ms), indicating the occurrence of the SMARC effect also along the vertical plane (see figure 3A and B). None of the other main effects, nor the interactions were significant.

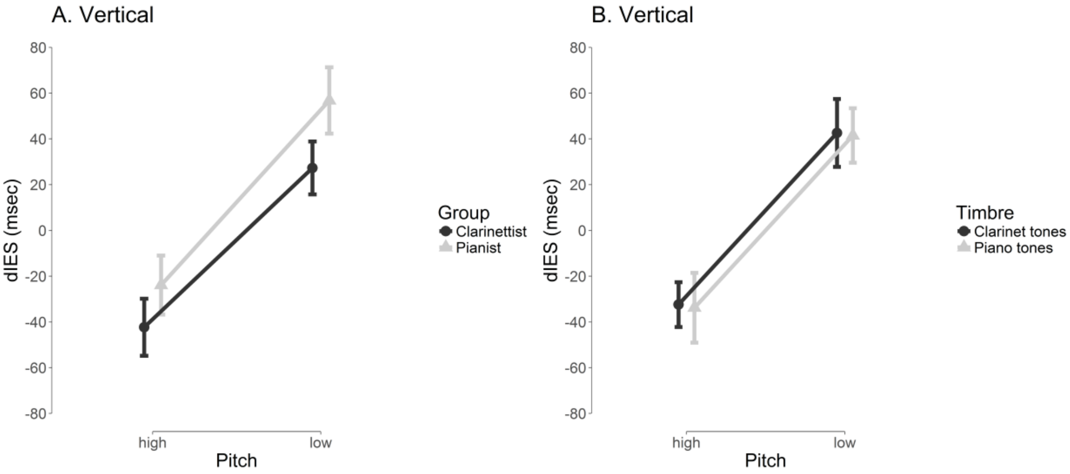

Figure 3. Differences in Inverse Efficiency Scores (dIES) between right and left responses in the Vertical response setting. Positive dIES indicate faster responses for the left-side key compared to the right-side key, and vice versa for negative dIES**. A.** dIES as a function of Pitch (high, low) and Group (pianists, clarinetists): a significant SMARC effect was found in both the group of pianists and the group of clarinetists. Error bars represent ± 1 SEM. **B.** dIES as a function of Pitch (high, low) and Timbre (clarinet tones, piano tones): a significant SMARC effect was found for both the piano tones and the clarinet tones. Error bars represent ± 1 SEM.

To substantiate the difference observed in the two distinct planes (i.e., horizontal and vertical), the dIES were also submitted to a 2 x 2 x 2 x 2 ANOVA with response setting (Horizontal, Vertical), timbre (Clarinet tones, Piano tones) and pitch (High, Low) as within-subjects variables and group (Clarinetists, Pianist) as between-subjects variable. Analysis showed first a main effect of pitch (F(1, 24)=19.87, $p<.01$, $\eta_p^2=.45$) and a significant interaction between pitch and response setting (F(1, 24)=5.03, $p=.031$, $\eta_p^2=.18$). Critically these effects were modulated by instrumental expertise, as revealed by the significant three-way interaction between pitch, response setting and group (F(1, 24)=5.03, $p=.034$, $\eta_p^2=.15$). The effect of timbre was also confirmed, as revealed by the significant interaction between timbre and pitch, (F(1, 24)=4.30, $p=.04$, $\eta_p^2=.15$). This was further modulated by space, as revealed by the three-way interaction between response setting, timbre and pitch, that showed a trend toward significance (F(1, 24)=3.60, $p=.07$, $\eta_p^2=.12$). None of the other main effects, nor the interactions were significant.

2.4 DISCUSSION

The present study investigated the influence of instrumental expertise and musical timbre on the spatial representation of pitch. For this purpose, a group of pianists and a group of clarinetists were required to perform an implicit timbre judgment task in the horizontal and in the vertical space. Results showed that both instrumental expertise and musical timbre interact in modulating the horizontal SMARC effect, pointing to a critical role of musical experience in pitch-space correspondence. In particular, results along the horizontal dimension yielded a significant SMARC effect, i.e. faster responses on the left side of the space for low tones and faster responses on the right side of the space for the high tones, only for pianists, but not for clarinet players. In addition, results further showed a horizontal SMARC effect for piano tones, but not clarinet tones. On the contrary, results in the vertical space indicated that all musicians, regardless of musical timbre, showed a typical SMARC-like effect, i.e., a facilitation in responding to low tones with the response key located in the bottom space and to high tones with the response key located in the top space. Together, these findings provide novel evidence for a key role of sensorimotor experience, linked to musical training, at the origin of the spatial representation of pitch along the horizontal space. Furthermore, by unveiling that the representation of pitch on the vertical space does not depend on musical practice, the present study brings further support to the view that the vertical mapping of pitch would be mainly learned from the statistics of natural auditory scenes.

With respect to the horizontal dimension, previous studies on the SMARC effect suggested that pitch-space association would rely on musical training (Rusconi et al., 2006; Lidji et al., 2007). Nonetheless, these previous studies did not consider the possible effect of instrumental training, as they mainly involved pianists (Lidji et al., 2007; Rusconi et al., 2006). On these grounds, here we reasoned that if instrumental expertise would affect the horizontal representation of pitch, then the horizontal SMARC effect should be stronger in pianists than other musicians who have less training on a piano keyboard, which is aligned horizontally. The idea that the spatial representation of musical pitch could be influenced by the structure of the instrument played, is in fact not completely new: for instance, Mikumo (1998), using a visual tracking task, demonstrated that pianists have a horizontal representation of tones, whereas other instrumentalists, like cellists, rather have a vertical representation. More recently, Timmers and Shen (2016) showed that instrumental expertise modulates the association between space and pitch in a spatial localization task, with stronger association for more experienced musicians, especially after performing their instrument. Accordingly, in the present study only pianists showed a horizontal representation of pitch.

Critically, previous studies have interpreted the horizontal representation of pitch in terms of an orthogonal mapping, translated from the vertical to the horizontal space (Cho & Proctor, 2003, Proctor & Cho, 2006; Lidji et al., 2007). Specifically, it has been suggested that lower tones, which are perceived as lower in space, would be remapped into the left space, and, similarly, higher pitches would be remapped into the right space (Cho & Proctor, 2003, Proctor & Cho, 2006; Lidji et al., 2007). Yet, because in our study we found that this mapping was partially limited to pianists, it is likely that this effect relies on an external referent familiar to musicians, such as the piano keyboard. In other words, the present findings suggest that the horizontal representation of pitch would be influenced by the structure of the instrument played, rather than being the consequence of an orthogonal mapping (Proctor & Cho, 2006; Stewart et al., 2004). Another possible explanation is that the orthogonal remapping is partially influenced by the expertise in left/right action (i.e., like in the case of piano players) that should increase the degree to which horizontal re-mappings of polar responses can occur. Moreover, it is important to note that in the instrumental practice, pianists use left and right hands in left and right space, respectively, whereas clarinettists use both hands in overlapping space. This different saliency on different spatial frame of reference may have, therefore, limited the horizontal SMARC effect to the group pianists. Future studies should ideally test whether the SMARC effect relies onto an anatomical or an external frame of reference, by asking participants to perform the task also in a crossing hands posture.

More generally, the effect of instrumental expertise along the horizontal space may also be interpreted in relation to the view that stimuli experienced as action effects (in our case, musical stimuli) may thereafter activate the associated action (in our case, a key press in a specific spatial location; see Hommel, 2009 for a review). Compelling evidence, indeed, suggests that this is the case for musicians during music processing (Drost et al., 2005a, 2005b; Drost et al., 2007; Keller & Koch, 2006, 2008; Trimarchi & Luzzatti, 2011). For instance, Drost and colleagues (2005a) asked their participants (i.e., pianists and non-musicians) to play chords on a piano keyboard in response to visual stimuli, while listening to task irrelevant auditory stimuli, which could be congruent or incongruent with the chord to be played. Results yielded slower response times in pianists, but not in non-musicians, when the auditory stimulus was incongruent with the required response (Drost et al., 2005a). Moreover, Taylor and Witt (2015) showed that pianists, but not novices, automatically activate the spatial representation that correspond to the learned musical action, while listening to music. Similarly, in a magnetoencephalography study, pianists - but not musically naïve participants- showed involuntary motor activity, while they listened to piano pieces, pointing to a key role of piano practice at the origin of such sensorimotor association (Haueisen & Knösche, 2001). Taking together, these studies indicate that in musicians listening to music involves the participations of both cognitive and neural processes designed for action. Finally, various neuroimaging studies have shown an activation of the parietal cortex, a brain region notoriously involved in spatial processing and in the transformation of acoustic or visual information into motor representation (Stewart et al., 2003; Warren et al., 2005; Zatorre et al., 2007; Brown et al., 2013), when non-musicians were trained to play simple melodies on the piano (Stewart et al., 2003; Chen et al., 2012; Herholz et al., 2015). In this sense, sensorimotor practice linked to instrumental expertise, may promote and consolidate the use of spatial representations in piano players and may account for the observed effect along the horizontal space in the present study.

It is worth considering that, in the horizontal condition, results also yielded a significant influence of musical timbre. Indeed, a horizontal SMARC effect was found only for piano tones, but not for clarinet ones. It is therefore likely that only piano tones were able to activate the horizontal representation of pitches, since tones on the piano keyboard are horizontally aligned, whereas tones on the clarinet are not. In line with these findings, Weis and colleagues (2016a) stressed the role of the timbre and of pitch range on the occurrence of the SMARC effect. They found that in non-musicians, piano tones induce a SMARC effect in both the horizontal and the vertical plane, regardless of whether the pitch range is small or large, whereas for the vocal sounds the effect seems to depend on the pitch range (Weis et al. 2016a). Interestingly, Drost and colleagues (2007) asked a group of pianists and a group of guitarists to play chords on their instrument, while they were presented with task-irrelevant auditory distractors (i.e., congruent or incongruent with the chord to be played) in different timbres. Results showed that musicians had slower RTs for incongruent trials selectively for their own instrument (Drost et al., 2007). This suggests that, for musicians, action representations are activated when musical sounds are processed, and that this effect mainly emerges when musicians process sounds of their own instrument. In accordance with this interpretation, previous studies demonstrated an enhanced cortical representation in musicians, for the musical timbre of their own instrument (Pantev et al., 2001). This may apply, as well, to the response-side compatibility effect found in the horizontal space for piano tones in the present study.

On the contrary, results in the vertical space demonstrated that all musicians, independently of the instrument played, showed a facilitation (i.e., for both accuracy and reaction time) when the low tones were responded with the key located in the bottom part of the space and when the high tones were responded with the key located in the top part of the space. This finding is in line with previous studies reporting a vertical SMARC effect in musicians (Rusconi et al., 2006; Lidji et al., 2007). More critically, these results indicate that the spatial representation of pitch along the vertical plane is not influenced by instrumental expertise and musical timbre. These findings, therefore, corroborate the view according to which the mapping between pitch height and vertical space would reflect an intrinsic spatial characteristic of pitch, as confirmed by its presence in pitch-space synesthetes (Linkovski et al., 2012), as well as in the general population (Ben-Artzi & Marks, 1995; Rusconi et al., 2006; Lidji et al., 2007; Walker et al., 2009; Spence, 2011; Pitteri et al., 2015) and in pre-lingual infants (Roffler & Butler, 1968; Walker et al., 2009; Dolscheid et al., 2014; Nava et al., 2015). In fact, the vertical representation of pitch is supposed to originate from a natural mapping between auditory frequency and elevation, given that high-frequency sounds tend to originate from more elevated sources than low-frequency sounds (Parise et al., 2014). Nonetheless, it is worth noting that, in our paradigm the vertical mapping was assessed as transposed along the radial (i.e., depth) plane. In this specific context, therefore, the vertical (i.e., altitudinal) plane would have probably better simulated the gesture of playing the clarinet and would have maximized the possibility to observe differences due to instrumental expertise. Accordingly, future investigations should ideally address this issue by testing pitch-space correspondence along the altitudinal space, especially in light of the differences emerged in previous studies on the neural mechanisms subserving attentional control over horizontal, radial and vertical planes (Mennemeier et al., 1992; McCourt & Olafson, 1997; for a review see Jewell & McCourt, 2000).

3.

Study 2:

The Spatial Musical Association of Response Codes does not depend on a normal visual experience: a study with early blind individuals

3.1 INTRODUCTION

Previous studies demonstrated that the lack of prior visual experience can affect spatial representations. So far, no studies have directly investigated with a classical SMARC paradigm the spatial format of the pitch representation in the blind population. Indeed, one might wonder whether the pitch-space correspondence assessed in sighted subjects has to be related or not to the visual experience. Although the origin of the association between pitch and space is currently still under debate, some authors suggested that it might be a consequence of internalizing naturally occurring audio-visual correlations in the environment (Pratt, 1930; Proctor & Vu, 2002; Parise & Spence, 2012; Parise et al., 2014; Spence & Deroy, 2012; Deroy & Spence, 2013). In this framework, the cross-modal pitch-space and pitch-size correspondences would reflect the natural correlation between the physical properties of the external world (Parise et al., 2014). This experiment aimed to investigate the role of prior visual experience on the cross-modal correspondence between auditory pitch and vertical spatial location. In order to investigate this issue, we asked a group of early blind and sighted individuals to perform an implicit (timbre judgment) SMARC task. Specifically, the task aimed at assessing the mapping of pitch height along a bottom-to-top dimension in blind individuals, this association being quite robust in the sighted population (Linkovski et al., 2012; Spence, 2011; Rusconi et al., 2006). Hand assignment to top/bottom response-keys was also manipulated in light of prior findings showing that blind individuals tend to rely more on proprioceptive cues compared to sighted people (e.g., Crollen et al., 2013; Pasqualotto et al., 2014; Rinaldi et al., 2015; Noordzij et al., 2006). In light of prior evidence, we expected sighted participants to be faster in categorizing instruments' timbre when low pitches are responded to with bottom response-key pressing and high-pitches with up response-key pressing, even if pitch is irrelevant to the task (e.g., Rusconi et al., 2006; Lidji et al., 2007). If the mapping between pitch height and position in space is rooted in normal visual development, responses of early blind individuals should not be affected by the combination response key/pitch of the tone (i.e., SMARC effect).

3.2 METHOD AND MATERIALS

3.2.1 Participants

Twenty early blind participants (10 males; mean age = 39.75, *SD* = 12.79, range: 18-65, mean education: 14.80 years, *SD* =3.56) and twenty sighted participants (12 males; mean age = 39.45, *SD* = 13.04, range: 22-65, mean education: 14.95 years, *SD* = 2.81) took part in the experiment. All participants were right-handed (Oldfield, 1971). No participant had any neurological disorders apart from optic nerve damage in case of some blind participants (for further details, see Table 1). Moreover, none of them reported to have motor dysfunction. In the blind population, all participants were proficient Braille readers.

Gender	Age	Education	Cause of Blindness	Light perception present
M	48	18	Leber congenital amaurosis	-
F	65	8	Optic nerve damage	+
F	46	16	Optic nerve damage	+
F	40	18	Microphthalmos	-
F	27	16	Congenital glaucoma	-
M	44	18	Optic nerve damage	-
M	45	8	Congenital glaucoma	-
M	33	15	Retinitis	-
F	18	13	Congenital glaucoma	-
F	36	18	Retinal detachment	-
F	39	18	Optic nerve damage	+
M	60	16	Optic nerve damage	-
M	42	13	Optic nerve damage	+
M	28	13	Optic nerve damage	+
M	32	13	Optic nerve damage	-
F	20	13	Leber congenital amaurosis	+
M	54	18	Retrolental fibroplasias	+
F	33	18	Optic nerve damage	-
M	55	8	Optic nerve damage	+
F	28	18	Congenital glaucoma	+

Table 1. Characteristic of early blind participants in this study

3.2.2 Stimuli

The stimuli consisted of two low-pitched tones (C2 and E2, respectively, 65.41 Hz and 82.41 Hz) and two high-pitched tones (F5 and A5, respectively, 698.46 Hz and 880.00 Hz). Each tone was synthesized with piano and piano-keyboard timbre (belonging to the keyboard instrument family) and clarinet and saxophone timbre (belonging to the wind instrument family), for a total of 16 different stimuli. All stimuli lasted 500 ms. They were all normalized in loudness at 0 db, by means of the software Audacity (http://audacity.sourceforge.net/). Although all the sounds used were normalized at 0 db, auditory stimuli may differ in terms of perceived loudness. To rule out this possible confound, we conducted a preliminary experiment on 10 participants (not taking part in the main experiment), which demonstrated that sounds were perceived as equal in terms of loudness (for a similar procedure, see Rinaldi et al., 2016).

3.2.3 Procedure

Participants seated comfortably in a dimly lit room. Sighted participants were blindfolded throughout the entire experiment. Stimuli were binaurally delivered through professional headphone (Sennheiser HD 280 Pro headphone). Figure 1 shows the experimental procedure. In each trial, the target auditory stimulus was presented for 500 ms. Participants were asked to judge whether the tone was played by a wind or a keyboard instrument (timbre judgment task) by pressing two response keys, vertically aligned, one at the bottom (the space bar key) and one at the top (the key corresponding to the number “6”). There was no time limit for response but instructions emphasized both speed and accuracy. After participants’ response, a silent interval of 650 ms preceded the presentation of the next auditory stimulus.

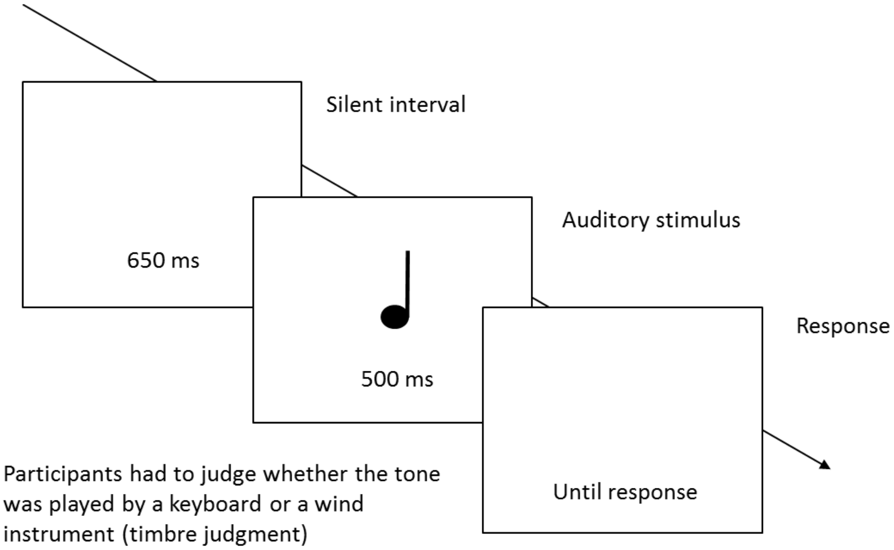

Figure 1. The experimental timeline. A silent interval of 650 ms preceded the presentation of the 500 msec-auditory stimulus. In each trial, participants were asked to judge whether the tone was played by a keyboard or by a wind instrument (timbre task).

Participants took part in two experimental blocks, in which the association instrument category (wind vs. keyboard) and response key was kept constant. However, in one block participants used their right hand to press the up-response key and the left hand to press the bottom-response key; in the other block the hand position was reversed (see figure 2). The order of the two experimental blocks were counterbalanced across participants. The response key-timbre assignment was also counterbalanced across participants: in half participants, keyboard instruments had to be responded to with the bottom response-key and wind instruments with the top response-key, in the other half of participants the reversed assignment was given. Each of the 16 stimuli (4 tones played by four different instruments, see above) was presented 6 times for a total of 96 trials in each block. Within each block, trials were presented in random order, with the only constraint that the same identical tone (same pitch played by the same instrument) never occurred consecutively. Before the experiment, participants listened to two tones of each instrument used during the task in order to familiarize with the auditory stimuli and performed 8 practice trials with tones of instruments not used in the real experiment (but still belonging to the keyboard or wind family). The software E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA) was used for stimuli presentation and data collection. The whole experiment lasted about 1 hour and 45 min, including instructions, eventual short breaks between the two blocks, and debriefing.

Figure 2. Hand assignment. Participants performed two experimental blocks: in one block participants used their right hand to press the up-response key and the left hand to press the bottom-response key; in the other block the hand position was reversed.

3.3 RESULTS

As in prior studies on the SMARC effect (Lidji et al., 2007; Weis et al., 2016a), the dependent variables were the difference in mean correct reaction times (dRTs) and percentage of error rates (dErrors) between top-key and the bottom-key responses. Accordingly, positive values indicate faster responses/ fewer errors for bottom-key responses and negative values indicate faster responses/ fewer errors for the top-key responses. The dependent variables were analyzed via a mixed repeated-measures ANOVA with pitch (low and high) and hand assignment (right hand top/left hand bottom vs. right hand bottom/left hand top) as within-subjects variables and group (blind and sighted) as between-subjects variable.

Trials in which participants' response latencies (RT) were 3SD over their own average response time were excluded from the analyses (.35% of trials were excluded following this criterion). The error rate averaged over the forty participants was 11.81% (blind participants: 10.01%; sighted participants: 13.61%, with the two groups performing similarly in terms of accuracy, $p>.31$). Importantly, there was no speed–accuracy trade-off, as indicated by the positive correlation, $r=.49$, $p<.005$, between RTs and error rate.

Analysis on the dErrors showed no significant main effects for either hand, *F*(1,38)=1.02, *p*=.31, η_p^2= .02, pitch, *F*(1,38) <1, *p*=.41, η_p^2= .01, or group, *F*(1,38)<1, *p*=.63, η_p^2= .00. None of the interactions reached significance: hand by group, *F*(1,38)=2.0, *p*=.16, η_p^2=.05, pitch by group, *F*(1,38)=1.88, *p*=.17, η_p^2= .04, hand by pitch, *F*(1,38) <1, *p*=.36, η_p^2= .02, and pitch by group by hand, *F*(1,38)<1, *p*=.70, η_p^2= .00.

The ANOVA on the dRTs yielded a significant main effect of pitch, *F*(1,38)=7.48, *p*=.009, η_p^2=.16, with positive dRTs in response to low tones (64 ms) and negative dRT in response to high tones (-52 ms), indicating the occurrence of a SMARC effect (see Figure 3). Neither the main effect of hand assignment, *F*(1,38)<1, *p*=.64, η_p^2= .00, nor the main effect of group, *F*(1,38)<1, *p*=.99, η_p^2= .00, reached significance. None of the interactions reached significance: hand assignment by group, *F*(1,38)<1, *p*=.60, η_p^2= .00; pitch by group, *F*(1,38)<1, *p*=.33, η_p^2= .02; hand assignment by pitch, *F*(1,38)<1, *p*=.34, η_p^2= .02; and hand assignment by pitch by group, *F*(1,38)<1, *p*=.58, η_p^2= .00.

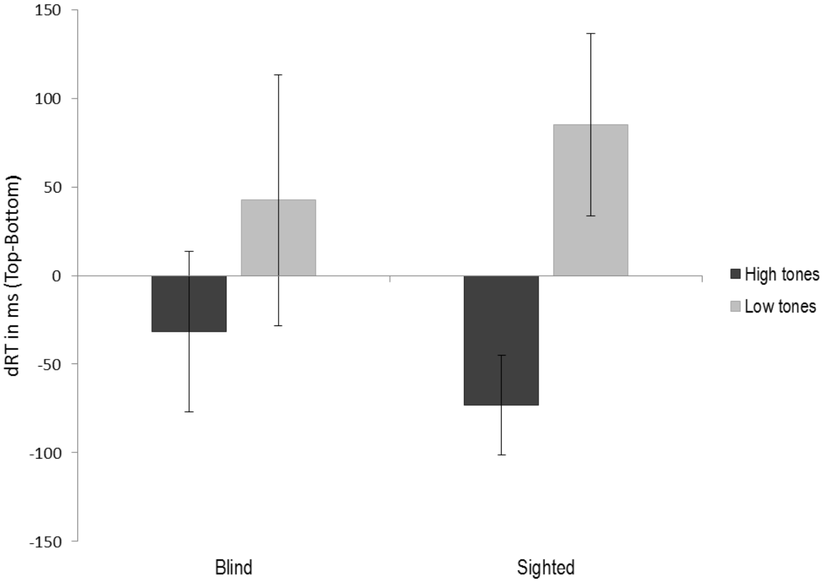

Figure 3. Differences in reaction times (dRTs) between bottom and top responses as a function of pitch (high, low) and group (blind, sighted). Positive dRTs indicate faster responses for the bottom-sided key compared to the top-sided key. A significant SMARC effect was found regardless of group. Error bars represent ±1 SEM.

3.4 DISCUSSION

We found that both early blind and sighted individuals showed a SMARC effect: Indeed, all participants responded faster to low tones when they were associated to the bottom response-key and to high tones when they were associated to the top response-key. Importantly, in neither group the SMARC effect was modulated by hand assignment (i.e., left or right hand used to press the bottom or top key). Overall, these results extend previous findings about the automaticity of the spatial representation of tones along a vertical spatial continuum, demonstrating that this pitch-space association does not require a normal visual development to occur.

A possible explanation of this pattern of results is that the spatial representation of auditory pitch in blind individuals could raise from their haptic perception and multimodal experience of the external world. This idea is in line with research showing that congenitally blind individuals are able to generate spatial images on the bases of tactile, motor, linguistic and also auditory information (Cattaneo et al., 2008; Kaski, 2002; Noordzij, Zuidhoek & Postma, 2006, 2007; Vecchi, 2001). Furthermore, both sighted and blind individuals use comparable visuospatial linguistic metaphors in describing musical pitch (e.g., “high” and “low” words for high and low-pitched tones; see for instance Antović et al., 2013; Eitan et al., 2012; Walker, 1985; Welch, 1991). Similarly, we cannot exclude that the robust mapping of pitch onto vertical space is grounded in body regions that resonate differently with pitch range, such as the chest (resonating with lower pitches) versus the head (resonating with higher pitches). In the same direction is the proprioceptive experience of our larynx, which rises and lowers to produce higher tones and lower tones respectively (see Shayan et al., 2011). This interpretation is in line with the embodied cognition, which claims in favor of a close link between cognition and bodily experience (Lakoff & Johnson, 1999; Fauconnier & Turner, 2002; Gibbs, 2003). Being independent from visual experience, this view may also give a plausible interpretation of why such correspondence is present even in the absence of early visual experience.

It is worth considering that this explanation is not in contrast with the hypothesis that the mapping between pitch and elevation is primarily learned through audio-visual regularities of the external world (Parise et al. 2014; Deroy & Spence, 2013; Parise & Spence, 2008, 2012). Although our results showed that vision may not be strictly necessary for pitch-space associations to be established, blind individuals may experience such statistical regularities in a different multi-modal way, especially considering that hearing is an extremely important source of information about the external world in case of loss of vision. Thus, it is possible that in the presence of a normal early visual experience, vision is the predominant modality to convey statistical regularities, whereas in case of blindness such experiences could be mediated by different senses and body/motor information.

Our results could be also explained within the framework of the ATOM theory (A Theory of Magnitude; Walsh, 2003), according to which space, time, and quantity would operate on the same magnitude system. Interestingly, it has been recently proposed that metathetic (i.e., concerned with qualitative variation; Stevens, 1957) dimensions, such as auditory pitch, could be part of this system as well (see Bottini & Casasanto, 2013, for a discussion). In line with this explanation, recent findings demonstrated that pitch height and numbers are not judged as separate magnitudes but as an integral one, suggesting a common mechanism for magnitude estimation (Weis et al., 2016b).

Our findings extend previous studies exploring the role of vision on the spatial conceptualization of abstract concepts, like numbers or temporal and serial order (Bottini et al., 2015; Bottini et al., 2016; Crollen et al., 2013; Pasqualotto et al., 2014; Castronovo & Seron, 2007; Cattaneo et al., 2010; Rinaldi et al., 2015). For instance, blind and sighted individuals show the same pattern in the way they mentally associate numbers (Castronovo & Seron, 2007; Cattaneo et al., 2010; Szűcs & Csépe, 2005) and time (Bottini et al., 2015) with spatial coordinates. The current study extends this literature, suggesting that also the origin of the spatial representation of musical tones could be independent from vision. A recent work on blind individuals failed in finding cross-modal correspondences between auditory pitch and tactile direction of movement in early blind participants, therefore suggesting that vision could play a role in mediating this audio-tactile association (Deroy et al., 2016). The differences between our results and those of Deroy and colleagues could be explained by the fact that Deroy and colleagues’ (2016) stimuli consisted of pure tones of linear increasing and decreasing pitch, whereas in our experiment single tones of different pitches were presented. Furthermore, we used a speeded classification task while Deroy and colleagues focused on tactile movement. Different cross-modal correspondences may have different origins and, even within the same dimension (like the pitch dimension), correspondences do not necessarily operate in the same manner (Deroy & Spence, 2013). Thus, we cannot exclude that different mechanisms mediate different types of correspondences, between pitch and elevation in our study and between rising-descending sounds and elevation in Deroy’s study. In accordance with this idea, Eitan and colleagues (2012) demonstrated that sighted, early blind, and congenitally blind adults reported few differences in pitch conceptualization, except for the mapping of pitch directions with vertical motion in the group of blind individuals.

In our study, no differences emerged related to hand assignment. Converging evidence demonstrated that sighted and blind individuals may rely on different spatial frames of reference when conceptualizing abstract concepts (Crollen et al., 2013; Pasqualotto et al., 2014; Rinaldi et al., 2015). Specifically, early blind individuals seem to rely preferentially on an anatomical (hand-based) frame of reference rather than an external one (Collignon et al., 2009; Röder et al., 2007; Noordzij et al., 2006; Cattaneo et al., 2008). Our results suggest that both sighted and blind individuals may rely on the same external coordinate system when mapping sounds onto vertical space. This is line with previous studies showing that the frame of reference used by blind individuals can differ according to the dimension considered (e.g., number or time; see Bottini et al., 2015). For instance, crossing hands is observed to reduce (Crollen et al., 2013; Wood et al., 2006) or even reverse (Müller & Schwarz, 2007) the mental number line in blind participants; conversely, both blind and sighted individuals represent time concepts from left to right, independently from hands position (Bottini et al., 2016). Overall, these findings show that early visual experience could impact differently on the way we use spatial coordinates system to map abstract concepts onto space, suggesting different experiential origins (Casasanto & Bottini, 2014; Bottini et al., 2015, 2016). Concerning our study, this point deserves further consideration, especially in light of previous studies assessing that blind and sighted individuals can rely on different reference frame when both horizontal and vertical plane were considered (see Struiksma, Noordzij & Postma, 2011). Furthermore, horizontal plane might also be more sensitive compared to vertical plane to hand position manipulation. Indeed, when hands are crossed along the horizontal plane, reaction times are longer than when hands are uncrossed (see Berlucchi et al., 1994). Therefore, while in the horizontal plane two distinct conditions occurred (crossed versus uncrossed hands), this is not true for the vertical alignment.

Ideally, future investigation should test the impact of vision on the mental musical line along the horizontal plane, in consideration of previous studies demonstrating that horizontal spatial association of pitch height is not automatic although it might be recruited upon specific task demands and cultural factors (Rusconi et al., 2006; Lidji et al., 2007; Weis et al., 2016a; Pitteri et al., 2015). Furthermore, although a direct correlation between the horizontal SMARC effect and writing/reading habits has not been demonstrated yet, some authors suggested that its origin may represent a directional bias from left to right in people from a culture in which writing direction follows this pattern (Lidji et al. 2007). If this was the case, then we should expect a similar pitch-space congruity effect across posture for both sighted and blind participants, given that Braille text is conventionally written from left to right.

4.

Study 3:

The effect of musical expertise on the representation of peri-personal space[*]

---

[*] This study is based on: Lega, C., Cattaneo, Z., Merabet, L. B., Vecchi, T., & Cucchi, S. (2014). The effect of musical expertise on the representation of space. *Frontiers in human neuroscience*, *8*, 250.

4.1 INTRODUCTION

Up to now, the possible effects of pitch height on the perception and representation of external space have not been investigated in musicians compared to non-musicians. Moreover, the effect of pitch on the representation of space has only been investigated in the visual modality (i.e. employing tasks requiring motor responses to visual stimuli (e.g., Rusconi et al., 2006; Lidji et al., 2007; Nishimura & Yokosawa, 2009; Cho et al., 2012), or in a visual line bisection task (e.g., Ishihara et al., 2013). Whether pitch height also affects the representation of space in other sensory modalities (e.g., tactile/haptics) has not been previously investigated. In this study, we investigated the effects of pitch height on the representation of external space by using a cross-modal sensory paradigm. We investigated whether pitch height differently affects the performance of musicians and non-musicians in a line bisection task performed either visually or haptically (i.e., implying tactile exploration of the rods while wearing a blindfold). During the task, participants were presented with tones of high vs. low pitch, that could be either relevant or irrelevant for a judgment participants had to make following bisection (i.e., pitch vs. timbre of a note). Only piano players were included in the experiment (as in Stewart et al., 2013): although the occurrence of a SMARC effect does not depend on the specific instrument played in musicians (Rusconi et al., 2006; Lidji et al., 2007; Cho et al., 2012), we preferred to have a more homogeneous group, also in light of previous evidence suggesting that the SMARC effect may be stronger in pianists due to learned action–effect associations (Lidji et al., 2007; see also Stewart et al., 2004; Stewart et al., 2013). We expected pitch height to modulate musicians’ bisection performance in both conditions when pitch height was relevant and when it was irrelevant for the task, in light of previous evidence suggesting that this cue is automatically represented by musicians in a horizontal spatial dimension (Rusconi et al., 2006; Lidji et al., 2007; Cho et al., 2012). We expected pitch height to also consistently modulate non musicians’ performance when pitch had to be intentionally processed (Rusconi et al., 2006; Lidji et al., 2007). In turn, we expected pitch to have no (Rusconi et al., 2006; Lidji et al., 2007) or at most, a weak effect (Ishihara et al., 2013) on non- musicians’ bisection performance when pitch was task irrelevant.

4.2 METHOD AND MATERIAL

4.2.1 Participants

Twelve skilled piano players (4 males, mean age=24.4 years, SD=4.33, range 20-36; mean years of piano experience=14.5, SD=2.94, range 10-20) and twelve non musicians (4 males, mean age=23.25 years, SD=1.42, range 21-26) took part in the experiment. All participants were right-handed (Oldfield, 1971) and had normal hand function and no auditory problems. All musicians had passed at least the fifth exam of instrument proficiency in an Italian Conservatory of Music. The inclusion criterion for non-musicians was to have no musical experience beyond a basic level acquired during middle school (typically implicating basic practice with a recorder instrument).

4.2.2 Stimuli

Stimuli consisted of wooden rods of five different lengths (30, 35, 40, 45, 50 cm) all with a diameter of 14 mm. The rods were positioned horizontally, with respect to the midline and presented on a table at a fixed distance of 38 cm. Each rod was fixed with Velcro strips horizontally onto a wooden panel. The rods could thus be haptically explored without being moved and a constant alignment between the participant's mid-sternum and the midpoint of the rod could be maintained. (see Baek et al., 2002; Cattaneo et al., 2011). auditory stimuli were presented and varied according to the experimental condition (see Procedure). In the "height judgment" condition, auditory stimuli consisted of two pure low tones (C3, 131 Hz and G3, 196 Hz) and two pure high tones (E5, 659 Hz and B5, 988Hz) (for further details see Lidji et al. (2007)). In the timbre judgment condition, auditory stimuli consisted of two pure tones (C3 and B5), and two distorted tones (C3 and B5 distorted; i.e., in which the physical shape of the original tone was changed leaving pitch height unaffected). In the control condition, participants were presented with white noise in order to provide a baseline measure for neutral auditory stimulation (see also Ishihara et al., 2013). All the auditory stimuli were created using the software Audacity and reproduced with Quick Time Player using a MacBook computer.

4.2.3 Procedure

The task required participants to carry out a haptic bisecting task while listening to different auditory stimuli. The same task was presented: 1) in two modalities: haptic and visual and 2) under two task conditions: a pitch judgment task (requiring to focus attention to the pitch of the tones) and a timbre judgment task (for which pitch height was irrelevant). The bisection task was performed with the right hand only and in the haptic condition, participants were blindfolded throughout the entire experiment.

In the visual condition, participants were instructed to look at the rods and indicate the midpoint of the bar using their right index finger. A maximum response time of 6 sec was allowed. During error measurement, the measuring tape was positioned in such a way that numbers were visible only to the experimenter. In the haptic condition participants were instructed to explore the length of the rod (either left-to-right or right-to-left) and as many times as they wished, during a 6 sec limit.

During the tactile and visual exploration, participants were simultaneously presented (via headphones) with an auditory stimulus lasting 6 seconds. In the visual trials, the auditory stimulus started 3 sec before the start of the bisection task. A wooden panel was placed in front of the to-be-explored rod to prevent participants from seeing it, and it was then removed 3 sec after the sound was presented. This ensured that participants had processed the sounds before performing the bisection task (see Cattaneo et al., 2012a, for a similar procedure). In the haptic trials, the stimulus started with the beginning of the exploration. Six sec were given for haptic exploration and the sound duration covered the entire exploration period. Previous findings have shown that auditory cues presented concurrently with the tactile exploration were effective in modulating the bisection bias (Cattaneo et al., 2012 a,b). Participants were instructed to pay attention to the sounds. Following bisection, in the height judgment task, participants had to verbally indicate whether the auditory stimulus was a low tone, a high tone, or consisted of white noise. In the timbre judgment task, participants had to verbally indicate whether the auditory stimulus was a normal tone, a distorted tone, or consisted of white noise. In each condition, the auditory stimuli, as well as different lengths of rods, were presented in random order. In each task (height and timbre judgment) and for each modality (visual and haptic), each of the five rods was presented 5 times (once for each different sound). Hence, there were 50 trials in the visual modality (25 for the height and 25 for the timbre task), and 50 trials in the haptic modality (25 for the height and 25 for the timbre task). Trials were presented in blocks for task condition (height vs. timbre) and modality (visual vs. haptic). The order of task and modality was counterbalanced across participants. The entire experiment lasted approximately two hours (included breaks between conditions).

Before performing the real experiment, participants were presented with a practice session (results not included in the analyses), in which they were instructed to bisect each of the five lengths rods within the 6 sec limit both in the visual and in the haptic modality. They were also taught which tones were identified as low and high, and the distinction between a pure and distorted tone. No feedback on performance was given to the participants during the testing.

4.3 DATA ANALYSIS

For the data analysis, deviations from the veridical center were converted into signed percentage scores (positive if bisections were to the right and negative if bisections were to the left) by subtracting the true half-length of the rod from the measured distance of each setting from the left extremity of the rod and then dividing this value by the true half-length and multiplying the quotient by 100 (see Laeng et al., 1996; Cattaneo et al., 2011). Signed percentage deviations for the five different lengths were collapsed together in the following analyses.

One sample *t*-tests were used to compare the average bisection bias in the control (white noise) conditions against zero (i.e., the true midline, absence of bisection directional bias), collapsing across the two tasks (height and timbre). A mixed repeated measures ANOVA, with condition (low tone, high tone, white noise) as within-subjects variables and group (musicians vs. non musicians) as a between-subjects variable was performed on the signed percentage scores and on the variable error (i.e., standard deviations) reported in each task (pitch height vs. timbre) separately. Bonferroni correction was applied to all post-hoc comparisons.

Accuracy in the height and timbre judgments was at ceiling in both groups of participants in both the visual and the haptic conditions (mean accuracy >98% in all conditions), and was not further considered in the analysis. The high performance in the auditory tasks confirmed that participants did indeed pay attention to the sounds.

4.4 RESULTS

Visual bisection. Figure 1 shows musicians and non-musicians' mean bisection bias in the different experimental conditions of the visual bisection task. In the baseline white-noise condition, a significant rightward bias was reported in both musicians, $t(11)=2.86$, $p=.016$, and non-musicians, $t(11)=2.95$, $p=.013$. The overall mean variable error (i.e., standard deviations of the bias scores) was comparable in musicians and non-musicians, $t(22)<1$, $p=.416$, suggesting comparable precision in the two groups.

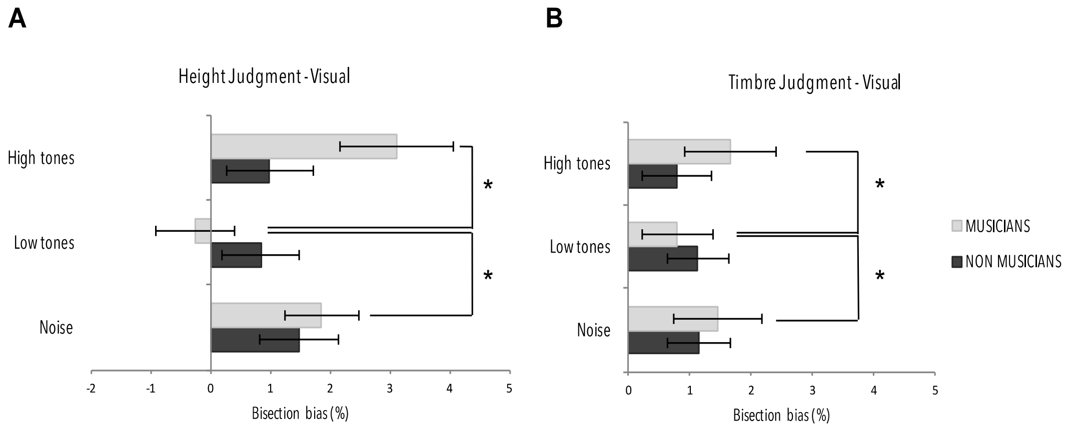

Figure 1. Mean percentage visual bisection bias in the height A) and in the timbre B) judgment tasks in visually bisecting rods in the control (white noise), low tones and high tones conditions. Overall, listening to low tones shifted musicians' bisection significantly to the left compared to listening to white noise or high tones; pitch of the tones did not affect non musicians' performance. Error bars represent ± 1 SEM. Asterisks indicate significant differences between task conditions.

In the *height judgment task* (Figure 1A), the analysis revealed a significant main effect of condition, $F(2,44)=10.00$, $p<.001$, $\eta^2_p=.31$, and a significant interaction condition by group, $F(2,44)=7.65$, $p=.001$, $\eta^2_p=.26$. The main effect of group was not significant, $F(1,22)=.28$, $p=.61$, $\eta^2_p=.012$. The main effect of condition was further analyzed in light of the significant interaction condition by group. An analysis of the simple effect of condition within each group showed that condition was not significant for non-musicians, $F(2,22)=2.49$, $p=.11$, $\eta^2_p=.19$, whereas it was significant in the musician group, $F(2,22)=9.79$, $p=.001$, $\eta^2_p=.47$. Pairwise comparisons revealed that in musicians the low tones shifted the perceived midline significantly to the left compared to the white noise condition, $t(11)=4.28$, $p=.004$, and to the high tones condition, $t(11)=3.32$, $p=.021$.

Conversely, the bisection bias shown in the high tone condition was not significantly different from that shown in the white noise condition, $t(11)=1.77$, $p=.31$.

In the *timbre judgment task* (Figure 1B), the main effect of condition was not significant, $F(2,44)=1.14$, $p=.33$, $\eta^2_p=.049$; however, the interaction condition by group reached significance, $F(2,44)=3.35$, $p=.044$, $\eta^2_p=.13$. The main effect of group was not significant, $F(1,22)=.12$, $p=.74$, $\eta^2_p=.005$. An analysis of the simple effect of condition within each group showed that condition was not significant for non-musicians, $F(2,22)=.50$, $p=.61$, $\eta^2_p=.044$, whereas it was significant in the musician group, $F(2,22)=7.36$, $p=.004$, $\eta^2_p=.40$. Pairwise comparisons revealed that in musicians, the low tones shifted the perceived midline significantly to the left compared to both the white noise condition, $t(11)=3.14$, $p=.028$, and the high tones condition, $t(11)=3.19$, $p=.026$, whereas no differences in the bisection bias were reported between the white noise condition and the high tone condition, $t(11)<1$, $p=.37$.

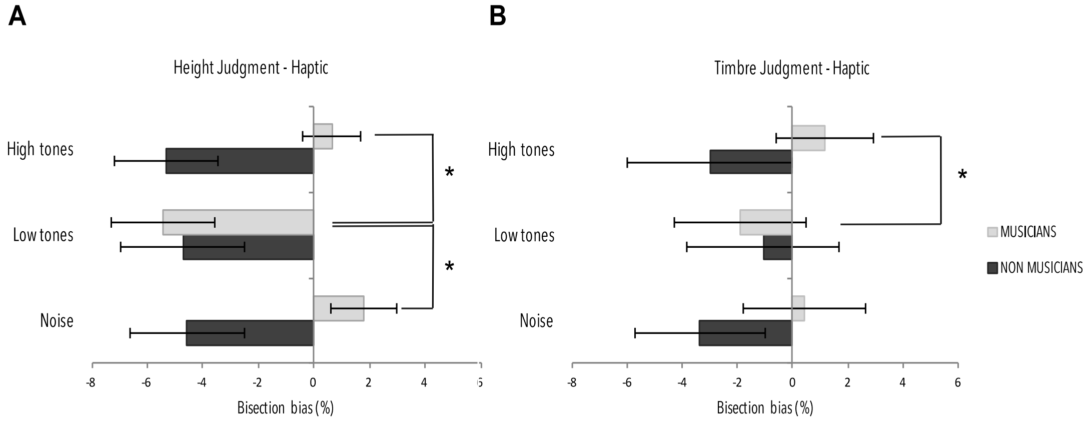

Figure2. Mean percentage haptic bisection bias in the height A) and in the timbre B) judgment tasks in bisecting rods in the control (white noise), low tones and high tones conditions. Overall, listening to low tones shifted musicians' bisection significantly to the left compared to listening to white noise or high tones. pitch of the tones did not affect non musicians' performance. Error bars represent ± 1 SEM. Asterisks indicate significant differences between task conditions.

Haptic bisection. Figure 2 shows musicians and non-musicians' mean bisection bias in the different experimental conditions of the haptic bisection task. Overall, in the baseline white-noise condition, musicians showed a tendency to bisect to the right of the veridical midpoint, but this deviation was not significant, $t(11)<1$, $p=.45$. Non musicians significantly bisected to the left of the veridical midpoint, $t(11)=2.30$, $p=.042$. Musicians and non-musicians did not differ in their overall variable error, $t(22)<1$, $p=.441$, suggesting comparable precision in the two groups (as in the visual task). Moreover, musicians and non-musicians did not significantly differ in their haptic exploration strategy (i.e., number of scanning movements), $t(22)<1$, $p=.860$ (mean number of explorations for musicians=4.17; for non-musicians=4.25), thus ruling out a possible role of this factor in contributing to differences observed in the bisection bias.

In the *height judgment task* (Figure 2A), the analysis revealed a significant main effect of condition, $F(2,44)=9.00$, $p=.001$, $\eta^2_p=.29$, and a significant interaction condition by group, $F(2,44)=9.64$, $p<.001$, $\eta^2_p=.31$. The main effect of group was not significant, $F(1,22)=2.89$, $p=.10$, $\eta^2_p=.12$. The main effect of condition was further analyzed in light of the significant interaction condition by group. An analysis of the simple effect of condition within each group showed that condition was not significant for non-musicians, $F(2,22)<1$, $p=.70$, $\eta^2_p=.03$, whereas it was significant in the musician group, $F(2,22)=12.48$, $p<.001$, $\eta^2_p=.53$. Pairwise comparisons revealed that in musicians the low tones shifted the perceived midline significantly to the left compared to the white noise condition, $t(11)=6.74$, $p<.001$, and to the high tones condition, $t(11)=3.21$, $p=.025$. Conversely, the bisection bias shown in the high tone condition was not significantly different from that shown in the white noise condition, $t(11)<1$, $p=.48$.

A similar ANOVA performed for the *timbre judgment task* (Figure 2B) revealed no significant main effect of condition, $F(2,44)=.31$, $p=.74$, $\eta^2_p=.01$; however, the interaction condition by group reached significance, $F(2,44)=5.96$, $p=.005$, $\eta^2_p=.21$. The main effect of group was not significant, $F(1,22)<1$, $p=.48$, $\eta^2_p=.02$. An analysis of the simple effect of condition within each group showed that condition was not significant for non-musicians, $F(2,22)=2.18$, $p=.14$, $\eta^2_p=.17$, whereas it was significant in the musician group, $F(2,22)=4.25$, $p=.028$, $\eta^2_p=.28$. Pairwise comparisons revealed that in musicians the low tones shifted the perceived midline significantly to the left compared to the high tones condition, $t(11)=3.67$, $p=.012$, whereas no differences in the bisection bias were reported between the white noise condition and either the low tone, $t(11)=2.44$, $p=.29$, or the high tone condition, $t(11)<1$, $p=.53$.

Possible differences in the initial scanning direction induced by the auditory stimuli were also analyzed. Percentage of trials in which exploration started to the left vs. to the right was computed. One-sample *t*-tests against 50% (i.e., no preferential initial scanning direction) were carried out to verify whether in the white-noise baseline condition (collapsed for timbre and height task) musicians and non-musicians showed a preferential initial scanning direction. Non musicians showed a tendency to start exploring the rod to the left (this was the case in 70% of the trials), but it did not reach full significance, $t(11)=1.72$, $p=.11$. In turn, musicians showed an opposite tendency starting exploration preferentially to the right (this was the case in 66.7% of the trials), but again this tendency failed to reach full significance, $t(11)=1.82$, $p=.096$. In particular, seven non musicians and four musicians always started exploration towards the same side of the rod. A pairwise comparison performed on the other participants who varied their initial scanning direction across trials, revealed no difference in the directional bias depending on the initial scanning direction, $t(12)<1$, $p=.983$. Hence, a repeated measures ANOVA with condition (low pitch, high pitch, white-noise) as a within-subjects variable and group as a between-subjects variable was performed on the starting scanning directions values for each task (height and timbre). In the pitch judgment task, the analysis revealed a significant main effect of group, $F(1,22)=9.30$, $p=.006$, reflecting musicians tendency to starting exploration to the right, and non-musicians tendency to starting exploration to the left. Neither the main effect of condition ($p=.77$) nor the group by condition interaction ($p=.47$) were significant, indicating that the different auditory stimuli did not significantly affect the initial scanning direction. In the timbre judgment task the analysis revealed an almost significant main effect of group $F(1,22)=3.99$, $p=.058$, reflecting the opposite tendency in the starting direction found in the baseline noise condition. The main effect of condition ($p=.40$) and the interaction group by condition ($p=.93$) were not significant.

4.5 DISCUSSION

The current experiment demonstrated that pitch height influences the representation of visual and haptic horizontal space (measured via a bisection paradigm) in musicians, but not in non-musicians. This was the case both when pitch height was relevant for the task, and when it was irrelevant (although in latter case, the effect of pitch on the bisection bias was weaker in the haptic modality). Notably, the effect of pitch height on space representation was limited to low-tones that induced a significant leftward shift in the bisection bias of musicians, whereas listening to high-pitch tones did not affect the bisection bias differently than listening to a neutral auditory condition (white-noise). Hence, although previous studies have shown that sound frequency is represented in a spatial format and that the “music mental line” can affect bimanual motor responses

(SMARC and SPARC effect) (Rusconi et al., 2006; Lidji et al., 2007; Nishimura & Yokosawa, 2009; Cho et al., 2012), our study provides the first evidence that pitch height influences the allocation of spatial attention crossmodally in tactile and visual peripersonal space in musicians.

Previous studies comparing behavioral performance of musicians and non-musicians have revealed an influence of pitch on motor responses, as an indexing the association between low tones and left responses and high tone and right responses (Rusconi et al., 2006; Lidji et al., 2007; Nishimura & Yokosawa, 2009; Cho et al., 2012; see also Vu et al., 2013). However, in non-musicians, this effect was reported only when pitch height had to be attended to intentionally (but see Cho et al., 2012), whereas in musicians, the association occurred even when pitch height was irrelevant for the task. Our findings in musicians appear in line with this previous evidence (Rusconi et al., 2006; Lidji et al., 2007; Cho et al., 2012) suggesting that in musicians, tones are likely to automatically activate a "music spatial line" that is oriented left to right. Indeed, pitch height is likely to be treated by the brain as an ordinal sequence, as with the order of days of the week, months of the year, or the alphabet, with ordinal sequences being mentally represented in a left-to-right direction (Gevers et al., 2003). In the case of musicians who are piano players, this may be even more strongly evident due to low notes being produced by left keys and high notes being produced by right keys on the piano keyboard (see Stewart et al., 2004, 2013). Our data show that the activation of such spatial representation of tone height in musicians (or at least in piano players such as in our participants) is able to interfere with the representation of external space, visually or haptically perceived, resembling previous evidence reported in the numbers' domain (see Cattaneo et al., 2012a). The modulation of pitch height on the bisection bias was observed in musicians both when pitch height had to be attended to, and when it was irrelevant (timbre judgment condition), although in the latter condition effects were less pronounced in the haptic modality. These findings suggest that pitch height automatically activates a "music mental line" in musicians (although the effect may be stronger when pitch is intentionally processed), supporting previous evidence (Rusconi et al., 2006; Lidji et al., 2007), and again showing strong resemblance with the effects exerted by numbers on spatial representation (e.g., Dehaene et al., 1993; Fias et al., 1996; Bonato et al. 2008; Cattaneo et al., 2012a).

Critically in musicians, the effect of pitch tones was reported only for low pitches that shifted the bisection bias significantly leftward, whereas listening to high tones did not significantly shift the pre-existing rightward bias further to the right. White noise has been previously used as a control auditory stimulus (e.g., Ishihara et al., 2013; Mendonça et al., 2013). Although all frequencies are equally represented in white noise, the sound is perceived as higher-pitched to human observers, partly because the perception of pitch is not linear, and partly because human ears are more sensitive to higher frequencies (Plack, 2005). This may partly explain why in our experiment the high-tones did not produce a significant modulation of the response bias compared to the baseline (white-noise) condition. Moreover, it is possible that only low pitch tones influenced bisection errors in musicians because they have a tonic rightward bisection error which is "counteracted" (i.e. moved leftward) by the low pitch tones. A similar argument has been made with respect to the influence of left and right visual cues on visual line bisection (McCourt et al., 2005), where the effect of cues delivered to the end of lines was more effective if they counteracted an existing bias (see also Tamietto et al., 2005; Cattaneo et al., 2012a, 2013). Indeed, when a cue is added to the pre-existing bias in bisection (as in our experiment high pitch tones that are likely to occupy the right portion of the putative music line in musicians), a threshold point may be reached at which errors are no further tolerated by the perceptual system and corrections are taken (McCourt et al., 2005; Laeng et al., 1996).

There was no evidence for the influence of sound pitch over spatial representation in the non-musician group we tested. This finding may appear in contrast to previous findings (Rusconi et al., 2006; Lidji et al., 2007; Nishimura & Yokosawa, 2009; Cho et al., 2012) who reported that there is an association between pitch height and the horizontal space when pitch had to be attended intentionally. However, in these previous studies this pattern was evident only when response latencies were considered, with no effect on participants' accuracy (with the exception of Cho et al., 2012, that also reported an effect on accuracy in one of their experiments). Conversely, an association between the vertical space and pitch height was more consistently observed across different measures (i.e., accuracy and reaction times, the latter being affected even when the pitch height was irrelevant for the task) (see Rusconi et al., 2006) suggesting that a left-right remapping of the low-high height dimension of tones is weaker than a more direct remapping of tone height into a down-up direction in non-musicians. Moreover, in this study, we used a line bisection paradigm, which represents a direct estimate of the external space while previous studies (Rusconi et al., 2006; Lidji et al., 2007; Nishimura & Yokosawa, 2009; Cho et al., 2012) measured response compatibility effects that may be more vulnerable to the influence of a simultaneously activated spatial representation. Indeed, in a line bisection task, the effect of a concurrently activated spatial mental representation needs to overcome a physically perceived space (that may be more “robust” to interference), whereas this is not the case when a fast motor response in space is required. Finally, in a recent study using visual line bisection in non-musicians, Ishihara et al. (2013) reported that the concurrent presentation of tones of different pitches (high or low) modulated performance with low tones shifting the bisection bias leftward, and high tones shifting the bisection bias rightward. However, the effect of tones pitch on the vertical bisection bias was weaker than in case of horizontal bisection (Ishihara et al., 2013) in contrast with prior evidence suggesting that the SPARC effect is stronger in the vertical plane (Rusconi et al., 2006; Lidji et al., 2007; see also Vu et al., 2013). Moreover, the modulatory effect of pitch on bisection was found only when low and high tones were precisely intermixed in a repeated order (i.e. low, high, low, high,…) and not when they were presented in blocks. The fact that the tone presentation was not randomized (as it was in our study) in the alternate presentation, together with the fact that only one line length was used and that participants’ music experience was not controlled (as stated by the authors Ishihara et al., 2013, see footnote 2) may have somehow affected task sensitivity and overall performance. Finally, a baseline auditory condition was not included in that paradigm (Ishihara et al., 2013, Experiment 1) so it is not clear whether the effects were driven by the low or by the high pitch. Overall, these differences in the paradigm may account for the discrepancy of our results with those by Ishihara et al. (2013).

In the baseline (white-noise) condition, non-musicians showed a rightward deviation in the visual modality and a leftward deviation in the haptic modality. The rightward deviation may appear surprising since individuals who are non-musicians typically bisect to the left of the veridical midpoint. However, listening to white-noise *per se* is known to induce a shift in the bisection bias compared to a silent condition due to alertness effects affecting hemispheric imbalance (Cattaneo et al., 2012b). The white-noise may have simply reduced the leftward bias in the haptic modality (as in Cattaneo et al., 2012b) and even reversed it to an opposite bias in the visual modality. This cannot be directly assessed in our data, since we did not include a silent baseline condition given that our goal was to have a baseline “auditory” condition to control for unspecific effect of auditory stimulation.

5.

Study 4:

Grasping the sound: auditory pitch influences size processing in motor planning[*]

---

[*] Part of this study is based on: Rinaldi, L., Lega, C., Cattaneo, Z., Girelli, L., & Bernardi, N. F. (2016). Grasping the sound: Auditory pitch influences size processing in motor planning. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(1), 11.

5.1 INTRODUCTION

In six different experiments, we investigate whether pitch-size correspondence modulates motor planning. In the first experiment, participants were required to reach, grasp and lift either the smaller or the larger section of a target object, depending on the auditory pitch of musical tones. In the second experiment, participants had to merely point to a specific section of the target object, an action requiring no size processing. In the third experiment, participants were required to reach, grasp and lift an object of uniform size (i.e., either small or large), so that now auditory pitch was irrelevant to the programming of the grip aperture. Furthermore, in the forth experiment we evaluated whether auditory pitch influences manual gestures conveying abstract concepts; to this end, participants were required to perform gestures expressing “small” and “large”. In Experiment 5 participants performed the same reach-to-grasp movements in response to a timbre judgment (implicit task) instead of a direct pitch judgment. Finally, in the last experiment the same implicit task was performed by a group of professional musicians.

In particular, we expect that if participants exploit the audiovisual natural mapping to integrate current multisensory information for action, they should be facilitated in selecting the appropriate motor response when the information is congruent. This should be reflected, for instance, by faster reach-to-grasp movements toward a small object when the action is primed by a (congruent) high-pitched sound, compared to when it is primed by a (incongruent) low-pitched sound. This natural pitch-size correspondence would, therefore, facilitate action when crossmodal information is congruent, even if the auditory stimulus and the visual object are not part of the same event. Thus, this may happen regardless of whether the auditory information is perceived as originating from the object (see Sedda et al., 2011) or not, as in the present study, since learnt statistical properties of the environment would facilitate (i.e., faster reaction times) integration of congruent multisensory information. In fact, humans judge the size of a visual stimulus more rapidly when the frequency of a simultaneous irrelevant sound is congruent (i.e., high pitch tone with a small visual stimulus; Gallace & Spence, 2006). Hence, prior experience with the acoustic resonance properties of stimuli varying in size can influence current integration of multisensory information, even when visual and auditory stimuli are not apparently related (Gallace & Spence, 2006; Grassi, 2005). Similarly, we also reasoned that auditory pitch should be capable of influencing movement scaling, with higher pitches associated to relatively smaller grip apertures, and lower pitches to larger grip apertures. More specifically, whenever hearing a high/low-pitched sound, participants should expect a small/large object to be associated with it. This possibility should, in turn, affect grip scaling, by modulating the contribution of visual information (i.e., the real object size) in motor planning. Finally, in light of previous experiments demonstrating a more automatic processing of pitch by musicians compared to non-musicians, we expected musical expertise to differently impact on kinematic parameters compared to musically naïve participants.

5.2 EXPERIMENT 1

Experiment 1 investigated whether auditory pitch influences initiation times and kinematic parameters of grasping movement. Participants had to judge the pitch, i.e., high *vs.* low, of auditory stimuli, by means of two different reach-to-grasp movements toward the smaller *vs.* the larger section of a target object. Half of the participants performed the task with a wooden object consisting of two cylindrical sections, a larger section at the bottom and a thinner section on top of it (standard orientation). The other half of participants performed the same task with the object tilted upside-down, i.e., with the thinner section at the bottom and the larger one at the top of it (tilted orientation). This manipulation was introduced in light of prior evidence showing that auditory pitch is mentally represented in a spatial format, with high pitches consistently associated with higher spatial positions than low pitches (Chiou & Rich, 2012; Pratt, 1930; Rusconi et al., 2006).

Grasping an object relies on estimates of various object proprieties, such as size and shape (Jeannerod, Arbib, Rizzolatti & Sakata, 1995). Accordingly, we hypothesized that if pitch is associated to object size, small grips should be initiated faster in response to high tones, whereas large grips should be initiated faster in response to low tones, for both object orientations. This pattern should therefore extend previous compatibility effects between auditory pitch and size to more complex motor planning. Conversely, if actions are influenced by pitch-space associations along the vertical dimension, we should expect movements toward the higher part of the target object to be initiated faster in response to high pitches and *vice versa* for low pitches. Importantly, we also explored whether auditory pitch might affect kinematic parameters. More specifically, if auditory pitch plays a critical role in size processing, we hypothesized that it might influence the relative scaling of the hand pre-shaping, with high tones associated to smaller grip aperture.

5.2.1 Method

Participants. Participants included 28 students, randomly assigned to 2 groups. A first group of participants (*N* = 14, *M* age = 29 years, *SD* = 2.1; 5 females) performed the experiment with the object in the standard orientation, while a second group (*N* = 14, *M* age = 26.6 years, *SD* = 3.7; 7 females) completed the experiment with the object tilted upside-down. Handedness was assessed by means of the Edinburgh Inventory (Oldfield, 1971). All participants were classified as right-handers. All participants expressed written informed consent to participate in the study. The study protocol was approved by the ethics board of the University of Milano-Bicocca.

Apparatus and stimuli. Participants sat in front of a table on which they rested their right hand with the thumb and index finger in pinch position, in front of the body, centered relative to body midline (starting position) (see Figure 1). A first group of participants performed the task with a wooden object consisting of two cylindrical sections, a larger section (diameter: 5 cm; height: 5 cm) at the bottom and a thinner section (diameter: 1 cm; height: 5 cm) on top of it, was placed in front of the participant, 21 cm distant from the starting position. The starting position was represented by a blue sticker placed on the table. Participants were explicitly instructed to move back to the sticker at the end of the required movement. A second group performed the task with the object tilted upside-down. Since the thinner section was placed on the table surface, making the whole object instable, a squared base (side: 2.5 cm; height: 0.5 cm) was fastened below the object in the tilted orientation. This was done to prevent participants to have to handle an instable object, which could have altered their motor performance. The auditory stimuli consisted of a piano low pitched tone (C1, Hz = 32.7), a piano high pitched tone (C6, Hz = 1046.5) and a white noise, that were all normalized by loudness at 0 db[1].

---

[1] Although all the sounds used were normalized at 0 db, auditory stimuli may differ in terms of perceived loudness. To control for this possible confound, 8 participants of Experiment 1 were required to match in loudness the auditory stimuli. In particular, participants were first presented with the white noise to be used as standard, and subsequently, they had to adjust a target stimulus (i.e., the high or low pitched tone) till it matched the loudness of the standard (see for a similar method, Parise et al., 2014). Varying systematically loudness levels of the standard and of targets, a total of 36 trials was presented. Results showed no difference between the high and the low tones, hence indicating that the sounds were perceived as equal in terms of loudness.

Procedure. Participants performed a go/no-go grasping task. Each trial started with the presentation of an auditory stimulus, lasting for 1500 ms, delivered through headphones (Sennheiser HD 280 Pro headphone). As fast as possible after the onset of the sound, participants had to perform a precision grip movement or no movement at all, depending on the sound perceived (see Figure 1). In particular, for both the object orientations, in “go” trials participants had to reach, grasp and lift the object from its larger section when hearing a low tone and from its thinner section when hearing a high tone, or *vice versa* depending on the experimental block. In turn, the white noise always signaled that no movement had to be performed (“no go” trials). In “go” trials, after completion of the movement, participants had to place their right hand back to the starting position. After 5 seconds from the presentation of the sound, the next trial began. To avoid artifacts induced by linguistic correspondence (see Dolscheid, Shayan, Majid & Casasanto, 2013), words referring to size (e.g., “small” or “large”) or to space (e.g., “high” or “low”) were not used for the instruction in any of the experiments.

Participants took part in two experimental blocks of 30 trials each (10 presentations for each auditory stimulus - low tone, high tone, and white noise - in a pseudo-randomized order), one with the high tone assigned to grasping the object from its smaller section and the low tone assigned to grasping the object from its larger section (Congruent condition), and one with the reversed assignment (Incongruent condition). The order of blocks was counterbalanced across participants.

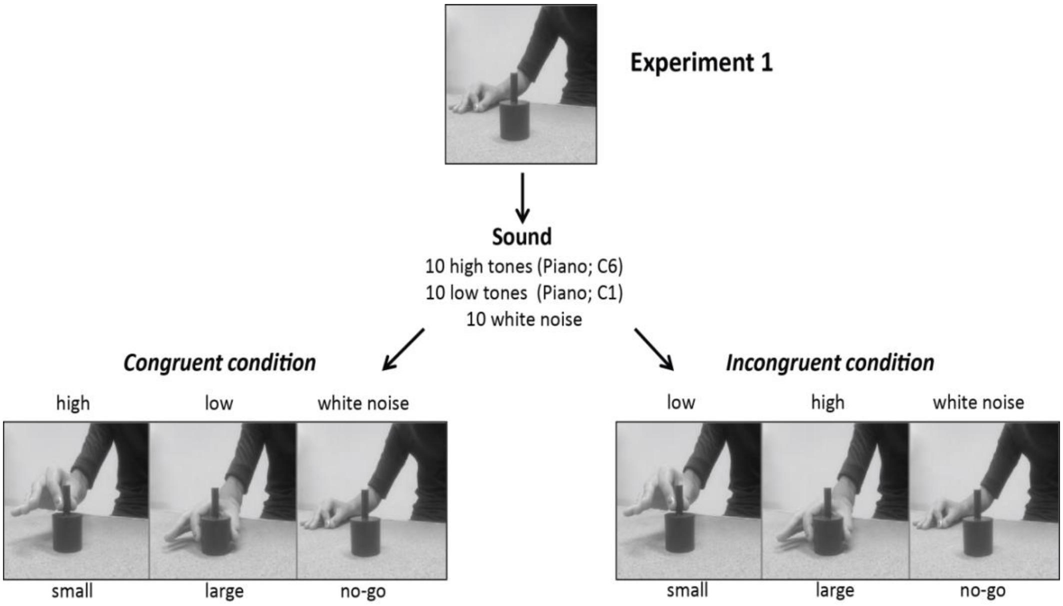

Figure 1. Apparatus and procedure of Experiment 1 with the standard object orientation. Each trial started with a sound (high tone, low tone or white noise), presented for 1500 ms via headphones. Upon sound presentation, participants had to perform a precision grip movement toward the object (*go*-trials) or refrain from moving (*no-go* condition). In *go*-trials, participants were required to grasp the object by its smaller *vs*. larger section depending on the auditory pitch (Congruent condition: high pitch = small section and low pitch = large section *vs*. Incongruent Condition: low pitch = small section and high pitch = large section). A second group of participants performed the same go/no-go task with the object tilted upside-down.

Data acquisition and analysis. A 3D-optoelectronic motion analyzer (SMART system, sampling rate of 120 Hz, spatial resolution ≈ 0.3 mm) recorded the 3D spatial position of three passive reflective markers fixed respectively on the tip of the right thumb (marker 1), on the tip of the right index finger (marker 2) and on the styloid process of the ulna (marker 3) of the participant.

Marker 3 was used to compute the reaction time (RT), defined as the time elapsed between the onset of the sound and the onset of the reaching movement. The beginning of the reach was measured as the first frame during which the displacement of the wrist marker along any Cartesian body axis increased more than 0.3 mm, with respect to the previous frame. The detection of movement onset was performed automatically via software and for each movement was visually checked and manually corrected when necessary. Markers 1 and 2 were used to compute the maximum grip aperture (MGA), defined as the maximum distance between marker 1 and 2 between reach onset and offset.

Data were analyzed offline for each trial and then averaged across trials for each experimental condition and participant, with a custom software written in MATLAB version 7.7 (R2008b). All variables showed normal distribution, as confirmed by the Kolmogorov–Smirnov test (all *ps* > .05). A repeated measures ANOVA with auditory pitch (high, low) and object section (small, large) as within-subjects variables and with object orientation (standard, tilted) as between-subjects factor was performed on each variable. For the RTs analysis, the presence of a congruency effect between auditory pitch and object section was tested by the interaction in the ANOVA.

5.2.2 Results and discussion

Incorrect motor responses were excluded from the analysis, resulting in the removal of 1.7% of the trials. Figure 2a reports the reaction time data. A repeated measures ANOVA on mean RTs with auditory pitch (high, low) and object section (small, large) as within-subjects variables and with object orientation (standard, tilted) as between-subjects factor, revealed no significant main effects for either auditory pitch, $F(1, 26) < 1$, $p = .51$, or object orientation, $F(1, 26) < 1$, $p = .49$. A significant main effect of object section was found, $F(1, 26) = 6.08$, $p <.05$, $\eta^2_p = .19$, power = .66, indicating faster initiation of movements directed to the small section compared to the large section of the target object. Importantly, the interaction auditory pitch *by* object section was significant, $F(1, 26) = 56.89$, $p <.001$, $\eta^2_p = .69$, power = 1 (Figure 2a). Post-hoc analysis showed that movements toward the small section of the target object were initiated faster in response to high pitches than to low pitches ($p < .001$), whereas movements toward the large section were initiated faster in response to low pitches than to high ones ($p < .001$), thus indicating the presence of a congruency effect. Finally, the interaction auditory pitch by object section by object orientation was significant, $F(1, 26) = 4.88$, $p <.05$, $\eta^2_p = .16$, power = 1. Post-hoc analysis indicated that for both object orientations the congruency effect was significant (all *ps* <.001), although maximized in the standard orientation. Neither the interaction object section by object orientation, $F(1, 26) < 1$, $p = .35$, nor the interaction auditory pitch by object orientation reached significance, $F(1, 26) < 1$, $p = .81$.

The grand-averaged profiles of grip aperture in different conditions are reported in Figure 3. The ANOVA on mean MGA revealed a main effect of object section, $F(1, 26) =$

1974, *p* <.001, η^2_p = .99, power = 1, with a larger MGA for grasping the large section of the target object than the small one. Importantly, the main effect of auditory pitch was also significant, *F*(1, 26) = 23.42, *p* <.001, η^2_p = .47, power = .99 (Figure 2b), indicating that pitch influenced the relative scaling of the hand pre-shaping, with high pitches being associated with smaller grip aperture compared to lower ones. Conversely, the main effect of object orientation was not significant, *F*(1, 26) < 1, *p* = .45. Neither the interaction object section by object orientation, *F*(1, 26) = 2.88, *p* = .1, nor the interaction auditory pitch by object section reached significance, *F*(1, 26) < 1, *p* = .46. The interaction auditory pitch by object orientation showed a trend toward significance, *F*(1, 26) = 3.76, *p* =.063, η^2_p = .13, power = .46. Post-hoc analysis, showed that high pitches were always associated with smaller grip aperture compared to lower ones, although this effect was maximized in the tilted orientation (*p* < .001 vs. standard orientation, *p* < .05). Finally, the interaction auditory pitch by object section by object orientation was not significant, *F*(1, 26) < 1, *p* = .91.

Overall, the results of Experiment 1 support the hypothesis of an influence of auditory pitch on grasping movements in the context of size processing. Indeed, the observed effect of auditory pitch on initiation times is in agreement with the stimulus-response compatibility effect so far reported in pitch-size perceptual processing (Gallace & Spence, 2006). The present results show for the first time that this compatibility effect extends to more complex motor processing. More critically, grasping kinematic was also influenced by auditory pitch. Indeed, auditory pitch modulated the grip aperture independently from the object size and prior to any interaction with it. This means that auditory pitch was *per se* informative about size in motor planning. Notably, this compatibility effect was mainly driven by the pitch-size association, since object orientation did not impact on the reported tendency (although at a descriptive level the congruency effect was more accentuated with the object in the standard orientation than in the reversed orientation).

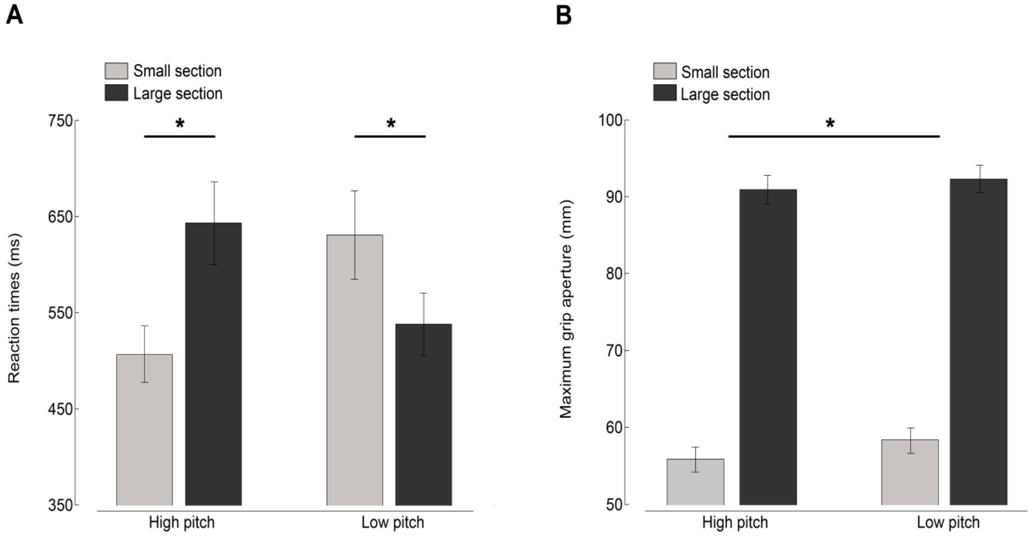

Figure 2. Results of Experiment 1. Panel A: movements toward the small section of the target object were initiated faster in response to high pitches, whereas movements toward the large section were initiated faster in response to low pitches. Panel B: pitch influenced the relative scaling of the hand pre-shaping, with high pitches associated with a smaller grip aperture, compared to low pitches, irrespective of the object-section to be reached. Error bars indicate ±1 standard error of the mean. *: $p < 0.05$.

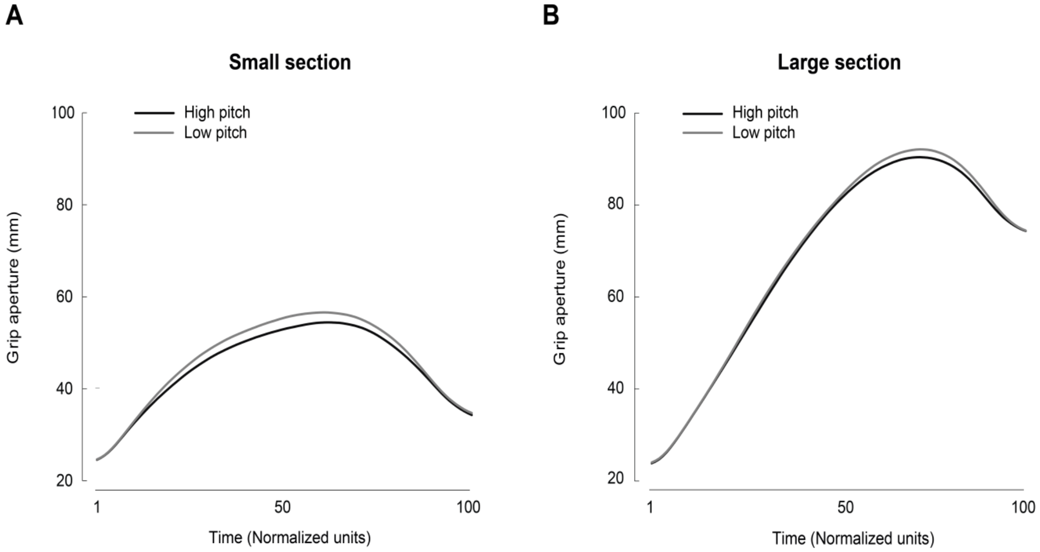

Figure 3. Grand-average of grip apertures profiles of Experiment 1. The time dimension was normalized to 100 units. Black and grey lines represent the grip aperture averages for the high and low auditory pitch, respectively. Panel A: grip apertures toward the small section of the target object. Panel B: grip apertures toward the large section of the target object.

5.3 EXPERIMENT 2

Results from Experiment 1 provide evidence for an influence of auditory pitch in actions requiring size processing. However, it is possible that the observed effects in terms of initiation movements were driven by a crossmodal association at a visual level, rather than by the size processing required when grasping a tool. Accordingly, Experiment 2 explored whether the effects of auditory pitch were exclusively determined by the visual object size or whether grip preparation was a necessary context for these effects to manifest. To address this possibility, in Experiment 2 we required participants to point to the object sections, thus reaching the object without grasping it. Indeed, grasping requires the translation of the object size into an appropriate grip aperture and this size processing might be critical for observing the reported effects (see Lindemann et al., 2007). Consequently, if pitch influences actions only when size processing is required, we should expect no effect of auditory pitch when merely pointing to the target object. Contrarily, if the effects found in the previous experiments reflect a perceptual crossmodal correspondence between auditory pitch and visual size, movement initiation times should be affected by the frequency of the sound.

5.3.1 Method

Participants. A new group of 13 right-handed students (*M* = 28.2 years, *SD* = 6.1; 12 females) took part in Experiment 2.

Stimuli and procedure. We used the same apparatus as in Experiment 1. Participants performed the same go/no-go task of Experiment 1. However, this time participants were required to perform a pointing movement toward the target without reaching it (i.e., stopping at a distance of about 1 cm). Specifically, participants had to point to either the small or to the large section of the target object, depending on the auditory pitch.

Data acquisition and analysis. Two passive reflective markers were fixed on the styloid process of the ulna (marker 1) and on the tip of the right index finger (marker 2). Marker 2 was used to compute the reaction time (RT), defined as the time elapsed between the onset of the sound and the onset of the pointing movement. All variables showed normal distribution, as confirmed by the Kolmogorov–Smirnov test (all *ps* > .05). A two-way repeated measures ANOVA on auditory pitch (high, low) and object section (small, large) as within-subjects variables, was performed on RT. For the RTs analysis, the presence of a congruency effect between auditory pitch and object section was tested by the interaction in the ANOVA.

5.3.2 Results and discussion

Incorrect motor responses were excluded from the analysis, resulting in the removal of 1.7 % of the trials. A two-way repeated measures ANOVA on mean RTs with auditory pitch (high, low) and object section (small, large) as factors, revealed no significant effects for either auditory pitch, *F*(1, 12) = 1.8, *p* = .20, object section, *F*(1, 12) < 1, *p* = .44, or for their interaction, *F*(1, 12) <1, *p* = .54 (Figure 4). In short, no compatibility effect was found when merely pointing to the object, indicating that auditory pitch only influences motor control of actions requiring size processing.

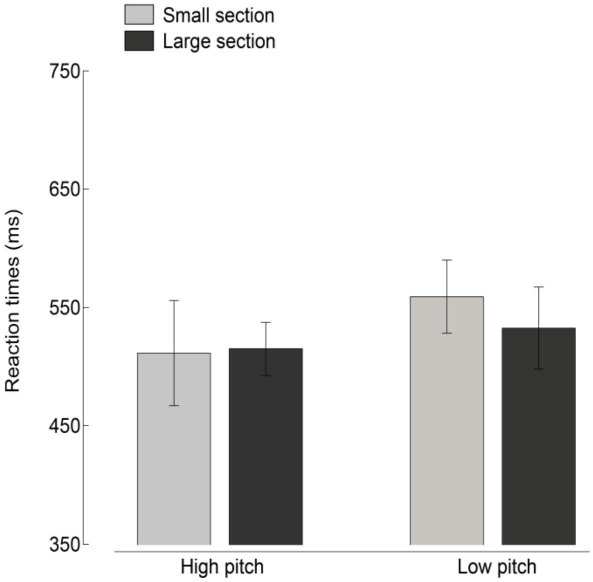

Figure 4. Results of Experiment 2. Reaction times for pointing movements toward the target object were not affected by either auditory pitch or by the section of the object they had to be directed (small vs. large). Error bars indicate ±1 standard error of the mean.

5.4 EXPERIMENT 3

Experiment 1 revealed that reach-to-grasp initiation movements were affected by the compatibility between the size of the object sections to be grasped and the auditory pitch. However, in Experiment 1 pitch modulated grip aperture irrespective of the size of the object (i.e., there was no interaction between pitch and size). Therefore, from

Experiment 1 it is not clear whether size processing is necessary to modulate grip aperture. Indeed, it may be possible that pitch alone is sufficient to systematically alter grip aperture. In order to verify whether size processing is a precondition for pitch to modulate grip aperture, we carried out an additional experiment based on a paradigm similar to Experiment 1, but in which the object was kept constant in size. In Experiment 3 we tested whether the effects of pitch on maximum grip aperture might arise even when auditory pitch is irrelevant to the programming of the grip aperture.

5.4.1 Method

Participants. Fourteen right-handed students (*M* = 26.3 years, *SD* = 3.3; 11 females) took part in Experiment 3. None of them had participated in Experiments 1 and 2.

Stimuli and procedure. We used the same apparatus and task of Experiment 1. However, two new objects were used, both consisting in a unique piece (and not made of different sections): a small object and a large object. The small object and the large objects corresponded in shape and size to the small and large sections respectively of the composite object used in Experiment 1. In one block participants were only presented with the large object, in another block only with the small object. The order of blocks was counterbalanced across subjects.

Data acquisition and analysis. Data acquisition and analysis were identical to Experiment 1. All variables showed normal distribution, as confirmed by the Kolmogorov–Smirnov test (all *ps* > .05). A two-way repeated measures ANOVA with auditory pitch (high, low) and object size (small, large) as within-subjects variables, was performed on each variable. For the RTs analysis, the presence of a congruency effect between auditory pitch and object size was tested by the interaction in the ANOVA.

5.4.2 Results and discussion

Incorrect motor responses were excluded from the analysis, resulting in the removal of 0.9 % of the trials. The analysis on mean RTs revealed no significant effects of either auditory pitch, $F(1, 13) < 1$, $p = .62$, object size, $F(1, 13) < 1$, $p = .77$, or their interaction, $F(1, 13) = 2.2$, $p = .16$ (Figure 5a). The same ANOVA on mean MGA revealed that the main effect of object size was significant, $F(1, 13) = 668.5$, $p < .001$, $\eta^2_p = .98$, power = 1, with a larger MGA for grasping the large object than the small one. Neither auditory pitch, *F*(1, 13) = 1.43, *p* = .25, nor the interaction pitch by object size, *F*(1, 13) <1, *p* = .54, were significant (Figure 5b). Results of Experiment 3 indicate that auditory pitch does not affect grasping movements when it does not convey information about the type of grasping to be performed.

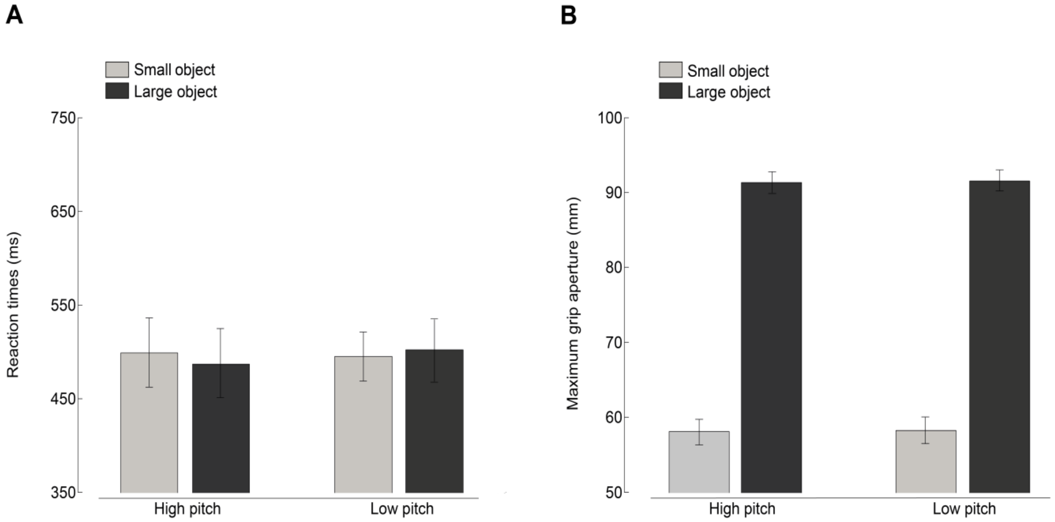

Figure 5. Results of Experiment 3, in which two different sized objects were presented in two separated blocks. Panel A: Reaction times for reaching movements toward the target object were not affected by either auditory pitch or by object size (small vs. large). Panel B: Auditory pitch did not influence the relative scaling of the hand pre-shaping, either when grasping the small or the large object. Error bars indicate ±1 standard error of the mean.

5.5 EXPERIMENT 4

In Experiment 4 we extended the investigation of the pitch-size associations to symbolic manual gestures. In particular, we reasoned that auditory pitch might not only interact with object-directed grasping, but also with communicative actions conveying size. Indeed, in many every-day life situations individuals refer to size by spontaneously gesturing about quantity (Winter, Perlman & Matlock, 2013). More specifically, speakers' metaphorical conceptualizations of size are often translated in gesture: when emphasizing that a certain quantity is a "large quantity" speakers might move their arms away from their body, thus increasing the space between hands; contrarily, speakers might move their arms close to each other, to emphasize a "small" quantity (Winter et al., 2013). Similar gestures are also exploited in the American Sign Language to express size-related concepts. Thus, in

Experiment 4 we explored whether auditory pitch might influence manual gestures conveying abstract concepts about size. Consequently, we required participants to perform manual gestures, adapted from American Sign Language, expressing “small” and “large” concepts.

5.5.1 Method

Participants. A new group of 13 right-handed students (*M* = 28.3 years, *SD* = 2.6; 8 females) took part in Experiment 4.

Stimuli and procedure. Participants were comfortably seated and were blindfolded throughout the experiment in order to avoid any motor adjustment based on visual feedback of the hands. The same auditory stimuli of Experiment 1 were used.

Participants performed a go/no-go task similar to Experiment 1. In experiment 4, they were required to keep their arms attached to the trunk with their elbows at 90° angle and their hands open and parallel out in front of the body (i.e., starting position) (see Figure 6). Participants were then required to move their hands away (“large” gesture) or close (“small” gesture) to each other, depending on the auditory pitch, avoiding rotation of the hands and without touching the palms (Figure 6). At the beginning of the experiment, participants practiced the required actions. Participants took part in two experimental blocks, one with the high tone assigned to “small” gesture and the low tone assigned to “large” gesture (Congruent condition), and one with the reversed assignment (Incongruent condition). The order of blocks was counterbalanced across participants.

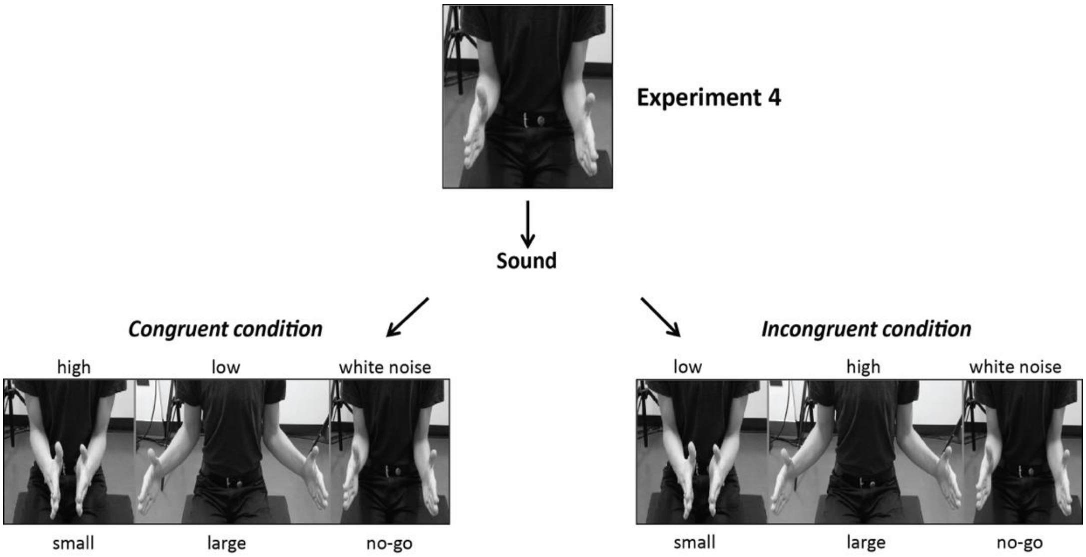

Figure 6. Procedure of Experiment 4. Similarly to Experiment 1, participants had to act in response to the auditory pitch of auditory stimuli. When the sound was presented participants had to move their hands close to each other (“small” gesture) or away (“large” gesture). The movement depended on the auditory pitch of the sound (Congruent condition: high pitch = small and low pitch = large vs. Incongruent Condition: low pitch = small and high pitch = large).

Data acquisition and analysis. Four passive reflective markers were fixed respectively on the right (marker 1) and on the left (marker 2) styloid process of the ulna and on the right (marker 3) and on the left (marker 4) tip of the index finger. Markers 1 and 2 were used to compute the reaction time (RT), defined as the time elapsed between the onset of the sound and the onset of the hands movement. As in the previous experiments, the beginning of the movement was measured as the first frame during which the distance along any Cartesian body axis between marker 1 and 2 increased (“large” gesture) or decreased (“small” gesture) more than 0.3 mm, with respect to the previous frame. The detection of movement onset was performed automatically via software and for each movement was visually checked and manually corrected when necessary.

Markers 1 and 2 were used to compute hands aperture (HA). For the “large” gesture, HA was defined as the maximum distance between hands. Specifically, we subtracted, in each trial, the final hands aperture from the initial hands aperture and we considered the maximum value during the hands movement. Conversely, for the “small” gesture, HA was defined as the minimum distance between hands. In this case, we subtracted the initial hands aperture from the final hands aperture, and we considered the minimum value during hands movement.

All variables showed normal distribution, as confirmed by the Kolmogorov–Smirnov test (all *ps* > .05). A two-way repeated measures ANOVA on auditory pitch (high, low) and type of gesture (“small”, “large”) as within-subjects variables, was performed on each variable. For the RTs analysis, the presence of a congruency effect between auditory pitch and type of gesture was tested by the interaction in the ANOVA.

5.5.2 Results and discussion

Incorrect motor responses were excluded from the analysis, resulting in the removal of 2% of the trials. A two-way repeated measures ANOVA was carried out on mean RTs with auditory pitch (high, low) and type of gesture (“small”, “large”) as within-subjects factors. The analysis revealed no significant main effects for auditory pitch, $F(1, 12) = 1.1$, $p = .33$, or gesture, $F(1, 12) = 2.62$, $p =.13$. The interaction auditory pitch *by* type of gesture was significant, $F(1, 12) = 5.71$, $p <.05$, $\eta^2_p = .32$, power = .59 (Figure 7a). In particular, “small” gestures were initiated faster in response to high pitches than to low pitches (p <.05), whereas “large” gestures were initiated faster in response to low pitches than to high pitches (p <.05), thus indicating the presence of a congruency effect.

The same ANOVA on mean HA revealed a trend for auditory pitch, $F(1, 12) = 4.16$, $p = .06$, $\eta^2_p = .26$, power = .47 (Figure 7b), indicating that high pitches tended to be associated with smaller HA than low pitches. Furthermore, the main effect of type of gesture was significant, $F(1, 12) = 37.7$, $p <.001$, $\eta^2_p = .76$, power = 1, indicating larger HA for larger gestures than smaller gestures. The interaction pitch by gesture was not significant, $F(1, 12) <1$, $p = .48$. These results, therefore, partially extend the previous finding to manual gestures conveying symbolic size information.

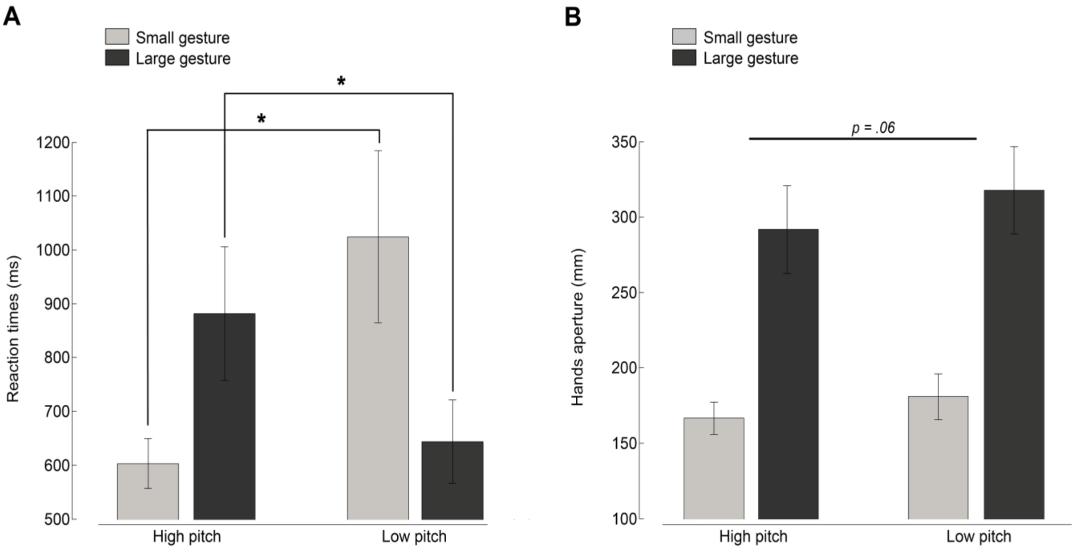

Figure 7. Results of Experiment 4. Panel A: Reaction times were faster for "small" gestures that had to be performed in response to high pitches, and for "large" gestures that had to be performed in response to low pitches. Panel B: Analysis of the hands aperture showed a statistically non-significant trend, with high pitches associated with a smaller hands aperture, compared to low pitches, suggesting that manual gestures expressing symbolic size information were affected by auditory pitch. Error bars indicate ±1 standard error of the mean. *: $p < 0.05$.

5.6 EXPERIMENT 5

In Experiment 5 we extended the investigation of the pitch-size association, exploring whether the effects reported in the previous experiments were detectable even when participants did not directly pay attention to the pitch of sounds. In the numerical domain, previous studies (Andres et al., 2004; Lindemann et al., 2007) indicated the existence of a strong connection between numerical magnitude and action planning, while participants had to judge the parity of Arabic number of both small and large magnitude, i.e. during an implicit judgment. The presence of an influence of pitch high on the execution of grasping movement during an implicit judgment would indicate an automatic and strong activation of the coupling between motor control and pitch processing. Furthermore, it would exclude the possibility that the effects previously reported were driven by a more linguistic semantic association. Indeed, we did not use words referring to size (e.g., "small" or "large") or to space (e.g., "high" or "low") in the instruction in order to avoid artifacts induce by linguistic correspondence (see Dolscheid, Shayan, Majid &

Casasanto, 2013). Nonetheless, we cannot exclude the hypothesis that a more explicit association between pitch and size was responsible of the outcomes previously described. Consequently, in experiment 5 we required participants to perform the reach-to-grasp movements toward the smaller *vs.* the larger section of the object, based on a timbre judgment (piano vs clarinet tones).

5.6.1 Method

Participants. A new group of 13 right-handed students (*M* = 27.5 years, *SD* = 3.2; 9 females) took part in Experiment 5.

Stimuli and procedure. The auditory stimuli consisted of a piano low-pitched tone (C1, Hz = 32.7), a piano high pitched tone (C6, Hz = 1046.5) and a white noise (already used in Experiment 1-4). Moreover, we now add a clarinet low-pitched tone (Db2, Hz = 69.30) and a clarinet high-pitched tone (Db5, Hz = 554.37). We used the same apparatus and task of Experiment 1. Participants took part in two experimental blocks of 50 trials each (10 presentations for each auditory stimulus - low tone, high tone, and white noise - in a pseudo-randomized order), one with the piano tones assigned to grasping the object from its smaller section and the clarinet tones assigned to grasping the object from its larger section and one with the reversed assignment. The order of blocks was counterbalanced across participants.

Data acquisition and analysis. Data acquisition and analysis were identical to Experiment 1. A two-way repeated measures ANOVA with auditory pitch (high, low) and object size (small, large) as within-subjects variables, was performed on each variable. In Experiment 5 we also analyzed the variable pick lift. Trials in which participants' reach response latencies (RT) were 2SD over their own average response time were excluded from the analyses. For the RTs analysis, the presence of a congruency effect between auditory pitch and object size was tested by the interaction in the ANOVA.

5.6.2 Results and discussion

A two-way repeated measures ANOVA was carried out on mean RTs with auditory pitch (high, low) and object section (“small”, “large”) as within-subjects factors. The analysis revealed no significant main effects for auditory pitch, *F*(1, 12) > 1, *p* = .55, η^2_p = .03 or object section, *F*(1, 12) > 1, *p* =.57, η^2_p = .02. The interaction auditory pitch *by* object section also did not reach the significance, *F*(1, 12) = 1.38, *p* = .26, η^2_p = .10, indicating the absence of the congruency effect observed in the explicit task (Experiment 1).

The same ANOVA on the MGA revealed a significant main effect of object section, *F*(1, 12) = 1518.35, *p* > .01, η^2_p = .99, with a larger MGA for grasping the large section of the target object than the small one. Importantly, the main effect of auditory pitch was also significant, *F*(1, 12) = 7.66, *p* = .01, η^2_p = .39 (Figure 8), indicating that pitch influenced the relative scaling of the hand pre-shaping, with high pitches being associated with smaller grip aperture compared to lower ones. The interaction between auditory pitch and object section was not significant, *F*(1, 12) < 1, *p* = .46, η^2_p = .04.

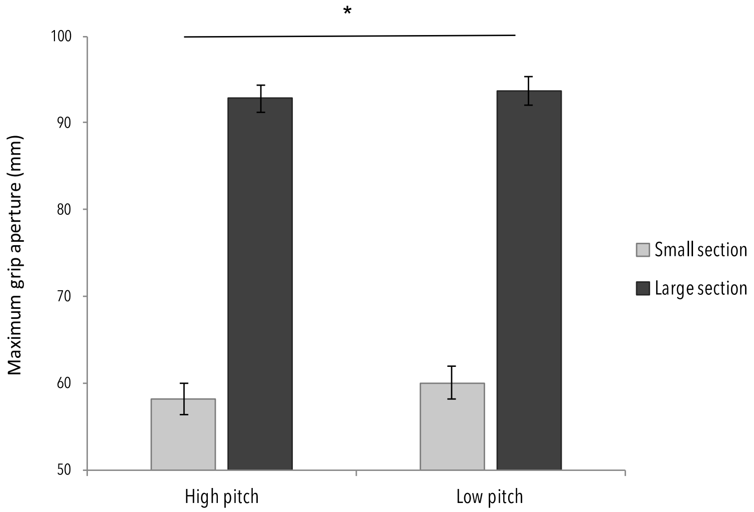

Figure 8. Results of Experiment 5. pitch influenced the relative scaling of the hand pre-shaping, with high pitches associated with a smaller grip aperture, compared to low pitches, irrespective of the object-section to be reached. Error bars indicate ±1 standard error of the mean. *: $p < 0.05$.

Finally, the pick lift was analyzed with the same ANOVA as used for the RT and MGA data. Analysis showed a significant main effects of object section, *F*(1, 12) = 77.26, *p* < .01, η^2_p = .86, indicating higher elevation when participants grasped the smaller section of the object compared to the bigger one. The main effect of pitch was not significant, *F*(1, 12) < 1, *p* = .64, η^2_p = .01, as well as the pitch *by* object section interaction, *F*(1, 12)< 1, *p* = .42, η^2_p = .05. In order to compare the pick lift results under implicit vs explicit judgment we re-analyzed data from Experiment 1, performing the same analysis on the pick lift dependent variable. Analysis revealed a significant main effect of object section, *F*(1, 12) =

134.93, $p < .01$, $\eta^2_p = .91$. The main effect of pitch was not significant, $F(1, 12) < 1$, $p < .58$, $\eta^2_p = .02$. Interestingly the interaction pitch *by* object section was also significant, $F(1, 12) = 6.07$, $p < .03$, $\eta^2_p = .33$, indicating higher elevation in case of a congruent mapping (high tones/small object section and low tones/big object section), compared to the incongruent mapping.

Results of experiment 5 demonstrated a pitch priming effect on grasping kinematics, i.e. the maximum grip aperture was enlarged when the object was grasped in the presence of a low pitched tones, even when the pitch was irrelevant to the task at hand. Interestingly, the reaction times facilitation effect, observed in the explicit task (Experiment 1), was not found in this experiment. Finally, we did not observe any effect of auditory pitch on the pick lift variable. On the contrary, a re-analysis of experiment 1 yielded an effect of pitch on the pick lift. In particular, we observed that participants tended to lift the object higher when the high tone was associated to the small part of the object and vice versa. This pattern, however, was likely induced by the congruency effect reported for the reaction times (for details see Experiment 1). Participants were indeed faster in the congruent condition and this may have led them to lift the object higher, compared to the incongruent condition. This interpretation may also account for the null effect of pitch in this experiment on the pick lift variable, where indeed no congruency effect on the reaction time was found. Together, these results seem to suggest that the processing of pitch information biased the processing of size-related motor features when preparing the grasping movement, despite being implicit and irrelevant to the task at hand. At the same time, no influence was observed on initiation times movement, suggesting that the explicit processing of pitch is crucial in order to impact on this dimension, whereas the processing of size possibly arises at a more implicit level. Accordingly, the MGA measure is more strongly linked to size processing compared to RT measure, conceivably explaining such dissociation. It is worth noting that previous studies reported an effect on both MGA and RT between the implicit representation of numerical magnitude and representations of action-coded magnitude (Lindemann et al., 2007). Nonetheless, numerical quantities are more directly associated to the physical size of object, compared to pitch dimension. This holds as well in language. For instance, in many languages, we tend to use words associated to the size for describing numerical magnitude (numbers can be “large” and “small”), whereas pitch is more often described by spatial metaphors (pitch can be “high” and “low”). Thus, we cannot exclude that the effect reported in experiment 1 on the RT was at least partially semantically mediated, reflecting a higher-level association between pitch and size dimensions.

5.7 EXPERIMENT 6

In 5 different experiments we demonstrated that pitch influenced the execution of manual actions when they implied size processing. In experiment 6 we investigated whether the effects previously reported were modulated by instrumental practice. Indeed, in musicians the auditory and motor systems are closely interconnected. Expert musicians strongly tend to associate the sounds from their instrument with the movement that produce those sounds and vice versa. Through years of training, musicians learned the contingencies between sounds and movements (Chen et al., 2012; Elsner & Hommel, 2001; Lahav et al., 2007), leading to a strong sound-action association. Interestingly, Drost and colleague (Drost et al., 2005a,b) demonstrated that the perception of their instrument facilitates and primes the corresponding action in musicians, even when sounds are irrelevant to the task. Auditory and motor influences are nonetheless bidirectional: hearing musical sounds can automatically trigger the corresponding actions (see Drost et al., 2005a,b; Keller & Koch, 2008), and performing movements can alter sound perception (see Repp & Knoblich, 2007). Recently, converging findings claiming in favor of an embodied approach to music cognition, emphasizing the close coupling of action and perception and demonstrating how the musical mind can be shaped by the motor system (Phillips-Silver & Trainor, 2007; Repp & Knoblich, 2009; Maes & Leman, 2013; Iordanescu et al., 2013, see also Maes et al., 2014).

Furthermore, previous studies already suggested that musicians and non-musicians differ in their ability to automatically activate a mental representation of pitch, when this dimension is irrelevant to the task (Rusconi et al., 2006; Lidji et al., 2007). Overall, these findings suggested that musical practice promotes a robust association between a specific sound and the specific movement used to produce it. Moreover, musicians tend to automatically process musical stimuli, especially of their own timbre (Pantev et al., 2001). In this experiment we asked a group of professional piano and clarinet players to perform the reach-to-grasp movements toward the smaller *vs.* the larger section of the object, based on a timbre judgment (piano vs clarinet tones) with both the right and left hand. If indeed professional players automatically process musical features, then we should expect a stronger influence of tones compared to non-musicians (Experiment 5) when pitch dimension is not explicitly processed. Furthermore, we hypothesized specific timbre-related effect, with musicians showing faster initiating movements primed by sounds of their own instrument. Finally, we also expected specific hand-related effect, with high and low tones differently impacting on right and left hand depending on the instrument training. Indeed, we reasoned that in pianists, association between high tones/right hand and low tones/left hand is stronger compared to clarinet players, for whom frequently the sound-hands association is the reverse.

5.7.1 Method

Participants. Thirteen skilled piano players (*M* = 25.76 years, *SD* = 4.51; 6 females) and thirteen skilled clarinet players (*M* = 28.69 years, *SD* = 10.16; 4 females) took part in Experiment 6. All participants were professional musicians (group of pianists: *M* = 17.07 years of formal musical training, *SD* = 7.19; group of clarinetist: *M* = 17.07 years of formal musical training, *SD* = 6.93) and there was no significant difference between groups in term of years of musical practice (*p*=1). All clarinet players also had some experience with the piano, whereas in the pianists group two participants also played the guitar. Handedness was assessed by means of the Edinburgh Inventory (Oldfield, 1971). All participants were classified as right-handers, with exception of one participants belonging to the group of clarinet players. All participants expressed written informed consent to participate in the study. The study protocol was approved by the ethics board of the University of Milano-Bicocca.

Stimuli and procedure. The auditory stimuli and the apparatus were the same used in Experiment 5. Participants took part in two experimental blocks for each hand consisted of 50 trials each (10 presentations for each auditory stimulus - low tone, high tone, and white noise - in a pseudo-randomized order), for a total of four experimental blocks. In one block the piano tones were assigned to grasping the object from its smaller section and the clarinet tones assigned to grasping the object from its larger section and one with the reversed assignment. The order of the two blocks and the order of responding hand were counterbalanced across participants.

Data acquisition and analysis. Data acquisition was identical to Experiment 1 and Experiment 5. In order to analyze effect specifically linked to musical and instrumental expertise we firstly performed a 2 x 2 x 2 x 2 repeated measures ANOVA on participants' reach response latencies (RT) with responding hand (left, right), timbre (piano, clarinet) and auditory pitch (high, low) as within-subjects variables and group (pianist, clarinetist) as between-subject variable. Furthermore, in light of previous results on non-musicians, a 2 x 2 x 2 repeated measures ANOVA with auditory pitch (high, low) and object size (small, large) as within-subjects variables and group (pianist, clarinetist) as between-subject variable was performed on each variable (reach RT, MGA and pick lift) and separately for each hand. Trials in which participants' reach response latencies (RT) were 2SD over their own average response time were excluded from the analyses. For the RTs analysis, the presence of a congruency effect between auditory pitch and object size was tested by the interaction in the ANOVA.

5.7.2 Results and discussion

A 2 x 2 x 2 x 2 repeated measures ANOVA was performed on participants’ reach response latencies (RT) with responding hand (left, right), timbre (piano, clarinet) and auditory pitch (high, low) as within-subjects variables and group (pianist, clarinetist) as between-subject variable. Analysis revealed no significant main effect of responding hand, $F(1, 24) = 2.56$, $p = .12$, $\eta^2_p = .09$. Analysis also showed a significant main effect of pitch, $F(1, 24) = 8.33$, $p < .01$, $\eta^2_p = .21$, indicating faster initiation response time for high tones (615 msec) compared to low tones (634 msec). Interestingly, the effect of pitch was modulated by the responding hand, as revealed by the significant interaction pitch *by* responding hand, $F(1, 24) = 5.02$, $p = .03$, $\eta^2_p = .17$, indicating faster reach reaction times for high tones selectively for the right hand (see Figure 9). Accordingly, planned t-test comparisons yielded a significant faster reach reaction times for high tones compared to low tones when participant responded with their right hand, $t(25)=3.18$, $p<.01$, but not when they responded with their left hand, $t(25)<1$, $p=.39$. Analysis also revealed a significant main effect of timbre, $F(1, 24) = 20.69$, $p < .01$, $\eta^2_p = .46$, expressing faster reaction times for piano tones (606 msec) compared to clarinet tones (643 msec). Interestingly, the main effect of timbre was modulated by the group, as indicated by the significant interaction timbre *by* group, $F(1, 24) = 6.40$, $p = .01$, $\eta^2_p = .21$. Planned t-test comparisons indicated that pianists were significantly faster in responding to piano tones compared to clarinet tones, $t(12)=7.75$, $p<.01$. Conversely, this difference was not significant for the group of clarinetist, $t(12)=1.13$, $p=.27$. There was also a significant interaction timbre *by* pitch, $F(1, 24) = 4.57$, $p = .04$, $\eta^2_p = .16$, meaning that participants were faster for high pitch tones compared to low tones exclusively for piano timbre, *t*(25)=3.79, *p*<.01, but not for clarinet timbre tones, *t*(25)<1, *p*=.95. Finally, this effect was influenced by the group, as revealed by the significant three-way interaction pitch *by* timbre *by* group, *F*(1, 24) = 4.82, *p* = .03, η^2_p = .16. To further analyzed the three-way interaction, we computed a new 2 x 2 x 2 ANOVA separately for the two groups. Analysis showed for the group of clarinetist a significant interaction timbre *by* pitch, *F*(1,12) = 10.47, *p* < .01, η^2_p = .46: despite not fully significant clarinet players showed a trend to be faster in responding to low tones compared to high tones with the clarinet timbre, *t*(12)=1.85, *p*=.08 and conversely they were significantly faster in responding to high tones compared to low tones with piano timbre, *t*(12)=2.96, *p*=.01. Instead, pianists were overall always faster in responding to high tones, even though this effect was greater for piano timbre, *t*(12)=2.33, *p*=.03, compared to clarinet timbre, *t*(12)=1.73, *p*=.10.

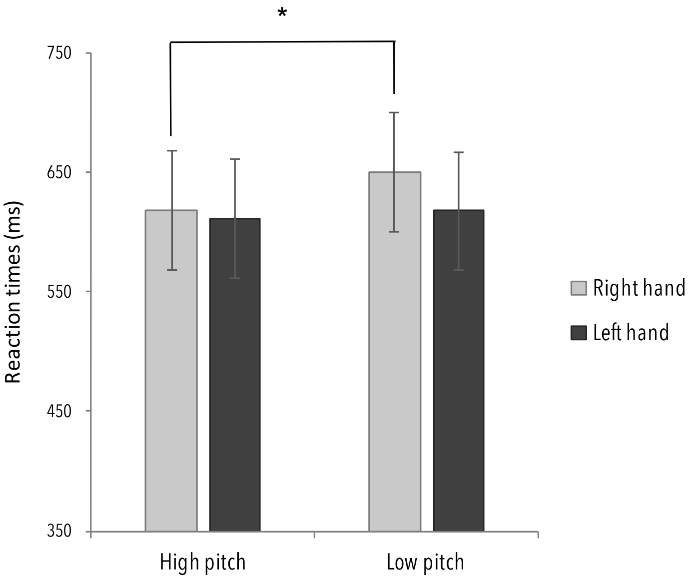

Figure 9. Results of Experiment 6. Movement with right hand were initiated faster in response to high pitches compared to low pitches, whereas there was no difference between high and low tones with left hand. Error bars indicate ±1 standard error of the mean. *: $p < 0.05$.

In order to analyzed the congruency, effect we computed a 2 x 2 x 2 repeated measures ANOVA with auditory pitch (high, low) and object size (small, large) as within-subjects variables and group (pianist, clarinetist) as between-subject variable on each variable (reach RT, MGA and pick lift) and separately for each hand. In these second analysis the left-handed participant was not included in the analysis.

Right hand. Figure 10A reports the reaction time data. A two-way repeated measures ANOVA was carried out on mean RTs with auditory pitch (high, low) and object section (“small”, “large”) as within-subjects factors and group as between-subjects variable. Analysis revealed a main effect of pitch, $F(1,23) = 8.92$, $p < .01$, $\eta^2_p = .28$. The main effect of pitch was modulated by the object size, as revealed by the significant interaction pitch *by* object size, $F(1,23) = 11.16$, $p < .01$, $\eta^2_p = .32$. Accordingly, post-hoc analysis showed that movements toward the small section of the target object were initiated faster in response to high pitches than to low pitches ($p = .01$), whereas movements toward the large section were initiated faster in response to low pitches than to high ones ($p < .001$), thus indicating the presence of a congruency effect. Conversely, both the main effect of object size, $F(1, 23) = 1.07$, $p = .31$, $\eta^2_p = .04$ and the main effect of group, $F(1, 23) <1$, $p = .37$, $\eta^2_p = .03$, were not significant. Neither the interaction object section by group, $F(1, 23) < 1$, $p = .72$, $\eta^2_p = .00$ nor the interaction auditory pitch by group reached significance, $F(1, 23) =1.80$, $p = .19$, $\eta^2_p = 07$. Finally, also the three-way interaction was not significant, $F(1, 23) <1$, $p = .77$, $\eta^2_p = 00$.

The ANOVA on mean MGA revealed a main effect of object section, $F(1, 23) = 1686.71$, $p <.001$, $\eta^2_p = .99$, with a larger MGA for grasping the large section of the target object than the small one. Importantly, the main effect of auditory pitch was also significant, $F(1, 23) = 4.72$, $p = .04$, $\eta^2_p = .17$, (Figure 10B), indicating that pitch influenced the relative scaling of the hand pre-shaping, with high pitches being associated with smaller grip aperture compared to lower ones. Conversely, the main effect of group was not significant, $F(1, 23) <1$, $p = .66$, $\eta^2_p = .00$. Neither the interaction object size *by* group, pitch *by* group and object size *by* pitch, respectively, $F(1, 23) <1$, $p = .34$, $\eta^2_p = .04$, $F(1, 23) =1.08$, $p = .30$, $\eta^2_p = .04$ and $F(1, 23) <1$, $p = .48$, $\eta^2_p = .02$. Finally the three-way interaction pitch *by* object size *by* group did not reach the significance, $F(1, 23) <1$, $p = .43$, $\eta^2_p = .02$.

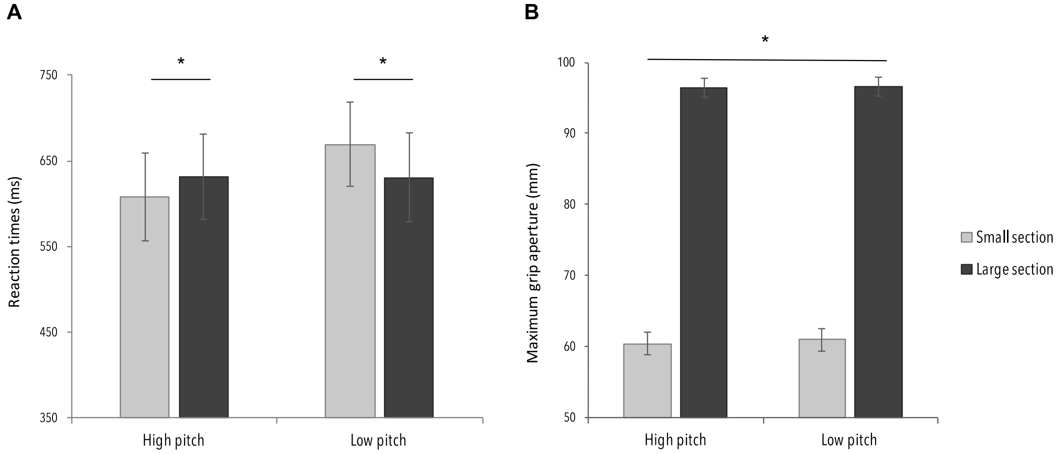

Figure 10. Results of Experiment 6, **right hand**. Panel A: movements toward the small section of the target object were initiated faster in response to high pitches, whereas movements toward the large section were initiated faster in response to low pitches. Panel B: pitch influenced the relative scaling of the hand pre-shaping, with high pitches associated with a smaller grip aperture, compared to low pitches, irrespective of the object-section to be reached. Error bars indicate ±1 standard error of the mean. *: $p < 0.05$.

The ANOVA on the pick lift variable yielded a significant main effect of object size, $F(1, 23) = 193.83$, $p < .01$, $\eta^2_p = .89$, indicating higher pick elevation for the small object size (214.75 mm) compared to the large one (183.38 mm). Interestingly, analysis also revealed a significant main effect of pitch, $F(1, 23) = 110.39$, $p = .02$, $\eta^2_p = .19$. The main effect of group was not significant, $F(1, 23) <1$, $p = .96$, $\eta^2_p = .00$. Again, neither the interaction object size *by* group, $F(1, 23) <1$, $p = .56$, $\eta^2_p = .01$, nor the interaction auditory pitch *by* group, $F(1, 23) = 1.30$, $p = .26$, $\eta^2_p = .05$, nor the interaction object size *by* auditory pitch, $F(1, 23) <1$, $p = .65$, $\eta^2_p = .00$ were significant. Finally, the interaction object size *by* pitch *by* group was also not significant, $F(1, 23) <1$, $p = .93$, $\eta^2_p = .00$.

Left hand. The same analysis was performed on the left hand. Figure 11A reports the reaction time data. ANOVA on the mean reach RT showed a significant interaction auditory pitch *by* object size, $F(1, 23) = 12.52$, $p < .01$, $\eta^2_p = .35$, indicating the presence of a congruency effect. Post-hoc analysis showed that movements toward the small section of the target object were initiated faster in response to high pitches than to low pitches ($p < .01$), whereas movements toward the large section were initiated faster in response to low pitches than to high ones, even though this difference did not reach the significance ($p = .14$). The main effects of auditory pitch, $F(1, 23) <1$, $p = .40$, $\eta^2_p = .03$, object size, $F(1, 23)$

<1, p = .36, η^2_p = .03 and group, F(1, 23) = 1.45, p = .24, η^2_p = .05 were not significant. Neither the interaction object size *by* group, F(1, 23) <1, p = .37, η^2_p = .03 nor the interaction auditory pitch *by* group, F(1, 23) = 1.33, p = .25, η^2_p = .05 reached the significance. The three-way interaction pitch *by* group *by* object size was not significant, F(1, 23) <1, p = .84, η^2_p = .00.

The ANOVA on the mean MGA showed a significant main effect of object size, F(1, 23) =571.82, p < .01, η^2_p = .96. None of the other main effects, auditory pitch, F(1, 23) <1, p = .50, η^2_p = .02 and group, F(1, 23) =1.12, p = .29, η^2_p = .04 reached the significance. All the interactions were also non-significant: object size *by* group, F(1, 23) <1, p = .50, η^2_p = .02, auditory pitch *by* group, F(1, 23) <1, p = .46, η^2_p = .03, pitch *by* object size, F(1, 23) <1, p = .70, η^2_p = .00 and pitch *by* object size *by* group, F(1, 23) =1.57, p = .22, η^2_p = .06 (see figure 11B).

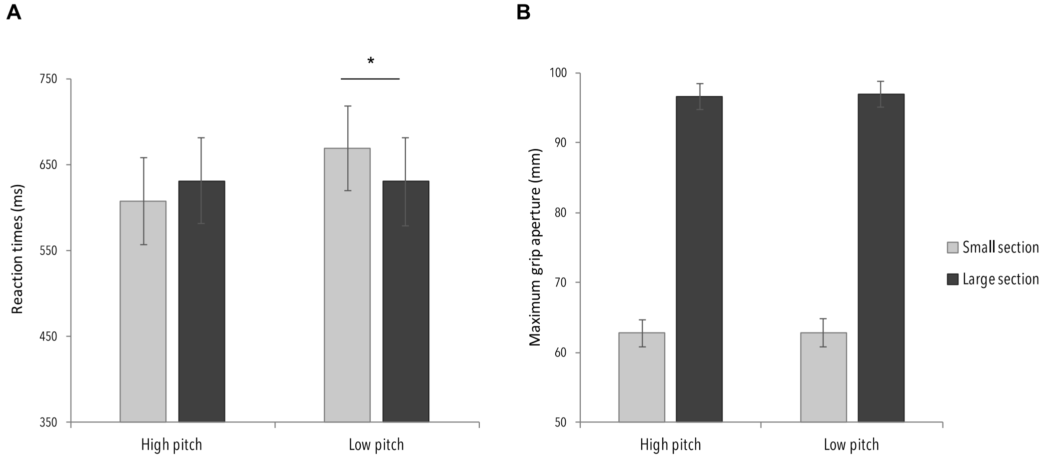

Figure 11. Results of Experiment 6, **left hand**. Panel A: movements toward the small section of the target object were initiated faster in response to high pitches, whereas movements toward the large section were initiated faster in response to low pitches. Panel B: auditory pitch did not influence the relative scaling of the hand pre-shaping, either when grasping the small or the large object. Error bars indicate ±1 standard error of the mean.

Finally, the ANOVA on the mean pick lift yielded a significant main effect of object size, F(1, 23) =138.61, p < .01, η^2_p = .85, indicating higher pick elevation for the small object size (216.20 mm) compared to the large one (186.90 mm). Differently from the right hand, the main effect of pitch was not significant, F(1, 23) <1, p = .42, η^2_p = .02. Also, the main effect of group was also not significant, F(1, 23) <1, p = .76, η^2_p = .00. All the interactions did not reach the significance: object size *by* group, $F(1, 23) < 1$, $p = .66$, $\eta^2_p = .00$, auditory pitch *by* group, $F(1, 23) < 1$, $p = .85$, $\eta^2_p = .00$, pitch *by* object size, $F(1, 23) < 1$, $p = .56$, $\eta^2_p = .01$ and pitch *by* object size *by* group, $F(1, 23) = 1.70$, $p = .20$, $\eta^2_p = .06$.

The aim of experiment 6 was to verify whether the effect of auditory pitch on motor planning, assessed in the previous kinematic experiments on the general population, was modulated by musical and instrumental expertise. Firstly, we showed a strong pitch-size cross-modal correspondence, as reflected by faster movement initiation for both right and left hand, which was not observed for non-musicians at the implicit level (Experiment 5). Secondly, an effect of pitch was also found for the kinematic variables considered. Indeed, pitch modulated both the maximum grip aperture and the pick lift, although only for the right (dominant) hand. Thirdly, despite all musicians were faster with piano tones, this effect was stronger for piano players compared to clarinet players and maximized for high tones. Finally, a generalized facilitation for initiating movements primed by high tones with the right hand, independent from the specific instrumental expertise, was found.

Overall, results of experiment 6 corroborated the view that musicians tend to process pitch dimension in an automatic way. The difference between non-musicians and musicians in an implicit task (Experiment 5 and 6) may be explained by the influence of musical expertise on processing musical stimuli (see Peretz & Zatorre, 2005 for a review). Accordingly, it has been demonstrated that musical expertise can enhance sensitivity in pitch variations in both music (Kishon-Rabin et al., 2001) and spoken sentences (Schön et al., 2004). Differences between musicians and non-musicians in the ability to extract relevant information from the incoming auditory stimuli have been studied using the mismatch negativity (MMN), which reflects the neural ability to detect change in pattern. Accordingly, larger MMN reflects a greater perceived distance between two sounds (Winkler et al., 2009). Interestingly, musicians (but not musically naïve participants) show a MMN to slightly impure chords even when instructed to ignore these auditory stimuli (Koelsch et al., 1999). Moreover, musicians exhibit stronger MMN to linguistic pitch contours (Chandrasekaran & Kraus, 2010) and to abstract sound features compared to non-musicians (van Zuijen et al., 2005).

Interestingly, the effect of pitch-size congruency effect on kinematic variable seems to be dependent on the hand used to perform the grasping movement. Specifically, pitch dimension consistently impacted on reaction times movements for both right and left hand, whereas it affected the maximum grip aperture and the pick lift variables selectively for the right hand. As all of our participants were right handers, we hypothesize handedness (and manual fluency) as a difference possible candidate beyond such dissociation. Accordingly, the association between pitch and size may be exaggerated for the hand that we more often used to interact with the surrounding environment. In addition to this, previous studies demonstrated manual asymmetries in peak velocity, movement time and variability of maximum grip aperture, with reach-to-grasp actions faster and more accurate when performed with the right dominant hand, compared to the left one (Flindall et al., 2014; Mieschke et al., 2001; Elliott & Chua, 1996; Elliott et al., 1993). This may have therefore further contribute to the absence of any crossmodal correspondence with the left hand. At the same time, finding an effect on reaction times in both hands, strengthened the interpretation that musicians processed pitch dimension more automatically compared to non-musicians, for whom an effect did not emerge with their right dominant hand (Experiment 5).

Experiment 6 also yielded specific effects linked to musical and instrumental expertise. Overall, all musicians were faster in initiating movements with piano tones compared to clarinet tones. This effect may be partially explained by the common familiarity of piano timbre for both groups of musicians. Moreover, piano tones were higher in pitch than clarinet tones, and we reported an overall advantage for high tones compared to low tones. Thus, we cannot exclude that these two effects were partially dependent on each other. Critically, this timbre-related effect was modulated by the instrumental expertise, as revealed by the fact that pianists were overall faster in responding to piano tones compared to clarinettists. This result is in line with previous studies demonstrating a selective enhancement in the brain in response to auditory signals, as a function of participants' musical expertise. For instance, musicians showed enhanced cortical representation for the musical timbre associated with their usual instrument compared with another instrument (Pantev et al., 2001). Similarly, conductors show superior peripheral spatial auditory processing relative to pianists (Nager et al., 2003). Critically, Drost and colleagues (Drost et al., 2005a,b) demonstrated that the perception of their own instrument facilitates and primes the corresponding action in musicians, even when sounds are irrelevant to the task.

Finally, we also showed a general effect of pitch on hands initiating movement, with right hand movement facilitated by high tones compared to low tones. We initially hypothesized a difference between hands as a function of instrumental groups. Specifically, we expected a stronger association between high tones and right hand for pianists compared to clarinet players, due to the long-term training with piano keyboard structure. Indeed, expert musicians strongly tend to associate the sounds from their instrument with the movement that produce those sounds. Accordingly, converging evidence demonstrated that hearing musical sounds can automatically trigger the corresponding actions, especially in musicians (see Drost et al., 2005a,b; Keller & Koch, 2008). Our results may be explained, at least partly, by the general expertise of piano keyboard which is common by both piano and clarinet players. In fact, conservatory musicians who are not first-study pianists nevertheless had good piano keyboard skills. This would explain the general facilitation of right hand movement in response to high tones for both groups of musicians. This interpretation is in line with the finding supporting that merely listening to sounds or music automatically activate motor responses, as a function of their previously established association (Rusconi et al., 2006; Lidji e al., 2007; Trimarchi & Luzzatti, 2011; Stewart et al., 2013; Baumann et al., 2004; D'Ausilio et al., 2004; Lahav et al., 2007). In this sense, simply listening to music become a kinaesthetic experience (see Maes et al., 2014). Critically, we did not test non-musicians with both hands, thus excluding any possible comparison with Experiment 6. Hence, it would be interesting for future investigations testing this issue, thus strengthening the hypothesis that the effect we reported was selectively dependent on musical expertise. It is nonetheless possible that the paradigm used was not sensitive enough for our purposes. Firstly, the fact that it is impossible to find professional musicians who do not also have piano experience may represent a possible confound. Secondly, musicians were tested with right and left hand in two separate blocks and not simultaneously, thus possibly reducing effects linked to the response-side mapping. Finally, using the non-dominant hand for reach-to-grasp movements may have obscured or weakened possible effects related to hand in response to different pitched-tones.

In sum, results of experiment 6 corroborate and extend results obtained from experiments 1 to 5, showing that the pitch-size cross-modal compatibility effect was modulated (and specifically strengthened) by musical expertise. Furthermore, musical and instrumental expertise impact on motor planning even when pitch dimension is irrelevant to the task, promoting faster movement for their own instrument compared to other timbre. Finally, high tones speeded movements with right hand compared to left hand, independently form the instrument played, possibly due to the general familiarity with the piano keyboard structure.

5.8 GENERAL DISCUSSION

In six kinematic experiments, we investigated the effects of auditory pitch on motor planning, by requiring participants to perform different manual actions primed by sounds. We found that pitch influenced the execution of manual actions when they implied size processing.

First, movement initiation times revealed that pitch-size compatibility effect, so far reported for perceptual processing (see Marks, 2004, for a review), holds for motor processing too. In particular, actions directed to small objects were facilitated by the presentation of high-pitched tones, while actions directed to larger objects were facilitated by low-pitched tones (Experiments 1). This pattern of results extends previous findings in the literature (e.g., Gallace & Spence, 2006) to more complex motor behavior. Pointing movements were not affected by auditory pitch (Experiment 2), likely because in this case the size of the target was irrelevant for the purpose of the action. Moreover, no effect on initiation times was found when auditory pitch was irrelevant to the programming of the grip aperture. A pitch-size compatibility effect was also found for manual gestures expressing "small" and "large" concepts, where actions were not aimed at grasping an object (Experiment 4). When the pitch dimension was irrelevant to the task, no effect was reported on the initiation times for musically naïve participants, but still influenced the initiating movements in skilled professional musicians (Experiment 5 and Experiment 6).

Second, and more importantly, we found that auditory pitch influenced the size of the hand pre-shaping. In fact, high pitches prompted smaller grip aperture, while low pitches prompted larger grip aperture both when the pitch dimension was explicitly and implicitly processed (Experiment 1 and Experiment 5). Yet, this effect emerged only when auditory pitch was relevant to the task, i.e., when it conveyed information about the type of grasping to be performed (Experiment 3). Overall these results suggest that auditory pitch *per se* is informative about size in motor planning, and this influence can be seen both in the context of interacting with a real object and when referring to abstract concepts.

The present study contributes to the current debate concerning how humans process magnitude-related information (Bueti & Walsh, 2009). According to the ATOM theory (Walsh, 2003), space, time and quantity (i.e., prothetic dimensions) would mutually operate on similar magnitude representations, because of the need to learn about the environment for acting on it. In line with this, compelling evidence suggests that action planning and symbolic number processing share a cognitive representation of magnitude (Andres et al., 2004; Lindemann et al., 2007). In particular, Lindemann et al. (2007) investigated the effect of numerical processing on the planning and control of reach-to-grasp movements, reporting that small grip movements were initiated faster in response to small numbers, while large grip movements were initiated faster in response to large numbers. Moreover, grip aperture was influenced by number magnitude, with a larger maximum grip aperture in response to larger numbers (Lindemann et al., 2007). However, according to a more comprehensive view of ATOM, metathetic dimensions concerned with qualitative variations, such as auditory pitch, should be included as well, since they are also associated with space and quantity (see Bottini & Casasanto, 2013, for a discussion). In the present study, both initiation times and grip aperture revealed that the influence of auditory pitch on size processing is not limited to perception, rather it extends to actions. Consequently, the present study shows that auditory pitch processing and action planning share a cognitive representation of magnitude.

Taken together, our findings extend prior literature on pitch-size correspondence (see Spence, 2011, for a review). In nature, object resonant frequency is related to object size and experiencing such correspondence is known to bias, through repetition, our perception of the surrounding world (see Parise, Knorre & Ernst, 2014, for a statistical account of pitch-space association). For instance, if we are listening to a high-frequency sound, we expect the sound to be generated by a small object (Grassi et al., 2013). Given the tight link between perception and action, pitch-size natural correspondence might also bias our actions in the environment. The present study offers support to this view. In particular, our results suggest that auditory information biases our action, even when the sound does not originate from the object to be manipulated (see Gallace & Spence, 2006). Hence, participants might have exploited the natural mapping between auditory pitch and visual size to integrate current multisensory information for actions, as reflected by the congruency effect in movement initiation. Furthermore, the influence of auditory pitch on grip aperture suggests that participants' prior experience with the acoustic resonance properties of objects varying in size modulated action planning. Critically, this occurred with non-natural sounds and in presence of visual feedback, indicating that prior pitch-size experience competes with online perceptual processing (i.e., the real object size) for driving action.

In an evolutionary perspective, implementing pitch-size correspondence at the motor level might have relevant advantages. Specifically, auditory pitch might be linked with size to speed up the programming of object-directed actions (see Morton, 1977).

Accordingly, the motor facilitation on initiation times reported here might favor interaction with the environment, with pitch that timely signals the motor system about the action to be executed.

More generally, our results provide novel evidence supporting the role of auditory information in driving actions. Many object-related actions can be inferred by their sounds (Kohler, Keysers, Umiltà, Fogassi, Gallese & Rizzolatti, 2002) and, in turn, natural sounds have been shown to drive grasping movements (Castiello, Giordano, Begliomini, Ansuini & Grassi, 2010; Sedda et al., 2011). Furthermore, individuals can rapidly learn to plan reach-to-grasp movements directed to different sized objects from the frequency of an auditory cue (Säfström & Edin, 2006). In particular, humans can establish a novel audiomotor map that allows them to properly grasp different sized objects, on the basis of sound frequency (Säfström & Edin, 2006). In line with these findings, we found that an association between object size and auditory pitch already exists, likely because of the repeated interaction with the surrounding environment. Yet, our results show that auditory pitch, a specific feature of sound, manipulated under controlled conditions, may as well affect motor planning. Moreover, these findings extend prior literature supporting multisensory integration of tactile (Patchay, Castiello & Haggard, 2003), olfactory (Castiello, Zucco, Parma, Ansuini & Tirindelli, 2006) and auditory modalities (Sedda et al., 2011) for acting in the environment.

Further support about a tight connection between the motor and the auditory system comes from evidence showing that the kinematics of grasping movements can influence the kinematics of speech (see Gentilucci & Corballis, 2006, for a review). In particular, Gentilucci and colleagues (Gentilucci, Benuzzi, Gangitano & Grimaldi, 2001; Gentilucci, Santunione, Roy & Stefanini, 2004) found that grasping objects of different size and bringing them to the mouth induced significant modulations in voice spectra of syllables simultaneously pronounced. Interestingly, an influence was also found during the mere observation of the movements (Gentilucci, 2003; Gentilucci et al., 2004). A similar correspondence has been found between spatial perception and speech production (Shintel, Nusbaum, & Okrent, 2006). In particular, individuals tend to spontaneously change the fundamental frequency of their voice to better describe the direction of motion along the vertical space, creating an analogical mapping between vocal frequency and the conveyed direction of motion. For instance, speakers raised and lowered their voice pitch to describe objects moving upward and downward, respectively (Shintel et al., 2006). In line with these findings, the present study provides additional evidence for a strong interaction between manual actions and auditory pitch processing, extending this link also to non-speech sounds.

In Experiment 4, we showed that auditory pitch modulated manual actions expressing abstract concepts of “small” and “large”. To explain how auditory pitch could influence manual gestures conveying abstract concepts, we first notice that in communication humans frequently exploit manual gestures to further emphasize speech (Bernardis & Gentilucci, 2006). Interestingly, manual gestures are also used to reinforce concepts related to size (Winter et al., 2013). For instance, people can emphasize size-related words by modifying the space between hands (see Winter et al., 2013). Hence, it might be possible that auditory pitch could modulate manual gestures by altering in the first place the mental representation of the size to be expressed, and this altered representation would then translate into a correspondingly larger or smaller gesture. Accordingly, the reported effects might arise from a natural correlation between vocal frequency and manual gestures, while people communicate concepts of size. In fact, a tendency to use high pitched vocal segments for words referring to the meaning “small” and low pitched vocal segments for those referring to the meaning “large” has been documented in different languages (see Ohala, 1983). Yet, this natural correlation would also account for the motor facilitation reported here. Future research using linguistic stimuli will be useful to gain further insights on these hypotheses.

In terms of which brain areas could be responsible for the effect of auditory pitch on motor planning, existing evidence suggest the involvement of a fronto-parietal network in the audio-visual integration guiding action (Rizzolatti & Sinigaglia, 2010). Indeed, some neurophysiological studies have shown that audio-visual mirror neurons in the premotor area F5, respond to both actions accompanied by sounds and by the presentation of sounds alone (Kohler et al., 2002). Moreover, recent neuroimaging studies have demonstrated an activation in the intraparietal sulcus, an area involved in grasping movements, during both pitch processing (Foster & Zatorre, 2010) and size processing (Pinel, Piazza, Le Bihan & Dehaene, 2004). Overall, this shared neuroanatomical resource adds to behavioral evidence pointing to the relevance of audio-visual integration for acting efficiently in the environment.

6.

Study 5:

A TMS investigation on the role of cerebellum in pitch and timbre discrimination[*]

---

[*] This study is based on: Lega, C., Vecchi, T., D'Angelo, E., & Cattaneo, Z. (2016). A TMS investigation on the role of the cerebellum in pitch and timbre discrimination. *Cerebellum & ataxias*, *3*(1), 1.

6.1 INTRODUCTION

The aim of this study was to analyze the role of the cerebellum in pitch and timbre processing using transcranial magnetic stimulation (TMS). While neuroimaging techniques provide correlational evidence regarding the activation in a specific brain region during an ongoing cognitive process, TMS allows establishing the causal role of specific cortical areas in a given task (Parkin et al., 2015; Halpern et al., 2004). Moreover, participants in TMS experiments act as their own controls overcoming some of the limitations intrinsic in patients' studies, such as potential differences in pre-morbid ability, and variability depending on high heterogeneity of lesions' sizes and gravity. Specifically, in this study we applied off line low frequency repetitive TMS to induce transient suppression of cerebellar activity (Chen et al., 1997; Oliveri et al., 2005) before participants' performance in a pitch and timbre discrimination tasks. If the cerebellum is causally involved in pitch and timbre processing, participants should perform worse following real than sham (faked) stimulation.

6.2 METHODS

6.2.1 Participants

Fourteen participants (9 F; mean age = 21.93 ys; SD = 1.86) took part in the experiment. All participants were right-handed (Oldfield, 1971) and had less than 3 years of formal musical training, as revealed by a self-reported history of musical experience. Prior to the experiment, each participant filled in a questionnaire (translated from Rossi et al. 2009) to evaluate compatibility with TMS. None of the volunteers reported neurological problems, familiarity for seizures nor was taking any medication that could interfere with neuronal excitability. Written informed consent was obtained from all participants before the experiment. The protocol was approved by the local ethical committee. Participants' treatment was conducted in accordance with the Declaration of Helsinki.

6.2.2 Stimuli

Stimuli used in the pitch discrimination task consisted of 21 pure tones (i.e., tones with a sinusoidal waveform, where the wave consists of a single frequency) of 200 ms generated through the software Audacity (http://audacity.sourceforge.net/). All tones had a frequency comprised between 1000 and 1200 Hz, and were presented at a level of 75 dB SPL. Stimuli used in the Timbre discrimination task consisted of 21 complex tones of 200 msec duration. Sound files in the timbre task were created from digitized samples of real musical instruments, with all instruments belonging to the wind or string family. Sound files used in the timbre task were taken from the University of Iowa Musical Instrument Samples (Lawrence Fritts, http://theremin.music.uiowa.edu/MIS.html).

6.2.3 Procedure

Figure 1 shows the experimental paradigm (Figure 1a) and the timeline of an experimental trial (Figure 1b). Participants seated comfortably in a dimly lit room and stimuli were binaurally delivered through professional headphone (Sennheiser HD 280 Pro headphone). Each subject took part in two different sessions (Real and Sham) that were separated by an average of 6 days (range 5-7). In each session, participants performed both the pitch discrimination task and the timbre discrimination task twice: once before, and once after receiving 15 minutes of off-line 1Hz rTMS over the right cerebellum. During TMS, no task was performed and participants were instructed to minimise movements and be silent. The post-stimulation task started immediately after the end of the stimulation. Both the pitch and timbre discrimination tasks required participants to indicate by left/right key pressing using their dominant hand whether two consecutively presented sounds (separated by 1 sec of silence interval) were identical or different. Intertrial interval was 2 sec. In each task, 42 sounds were presented: in half of the trials the two sounds to be compared were identical, in the other half they were different. In the different trials of the pitch task, the second pure tone presented could be 20, 30 or 40 Hz higher (ascending trials) or lower (descending trials) compared to the first one. The number of descending and ascending trials was counterbalanced. In the timbre task, the two sounds to be compared in each trial were identical in terms of frequency and intensity, but they had different timbre. In particular, two different string sounds may be presented, or two different wind sounds (wind and string sounds were never presented in the same trial to avoid ceiling effects in recognition). Task order (pitch and timbre discrimination), TMS condition order (Real vs. Sham), and the response key assignment for same/different response were counterbalanced across participants. The software E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA) was used for stimuli presentation, data collection and TMS triggering. Pre and post-stimulation task sessions lasted approximately 10 minutes (5 minutes for each task).

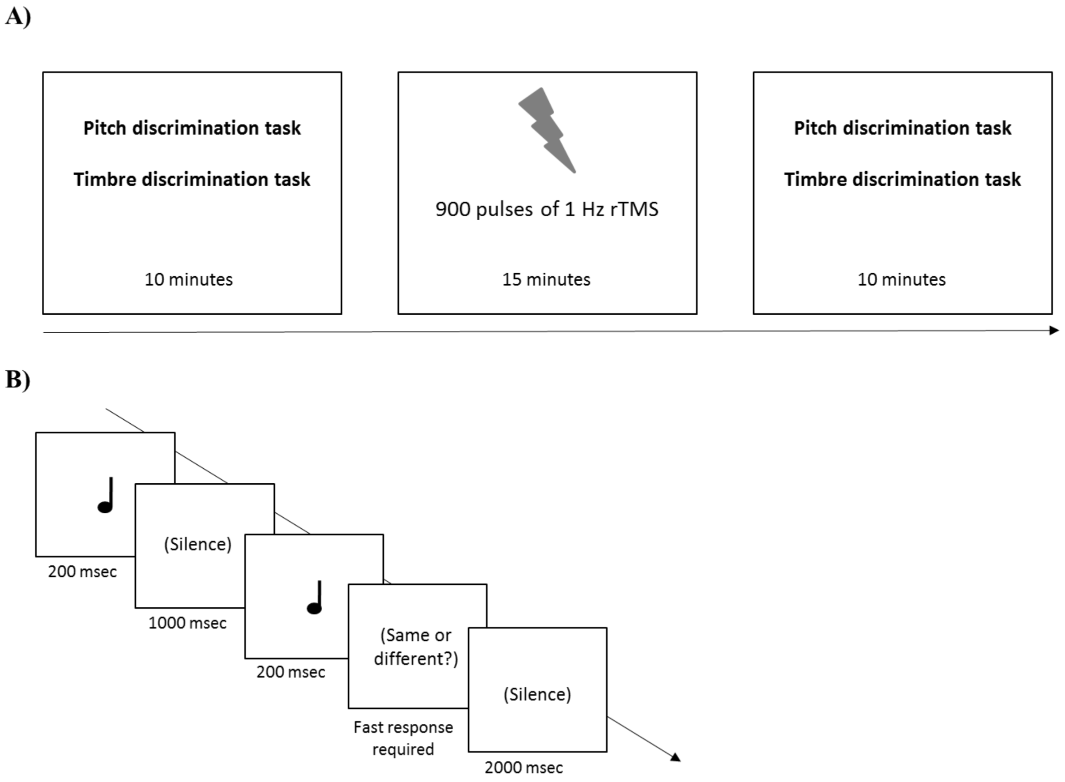

Figure 1. **A)** The experimental paradigm: participants underwent two experimental sessions, one with real TMS, and the other with sham TMS (order of sessions counterbalanced). In each session, participants performed the task twice, once before and once after receiving 1 Hz repetitive 15 minutes TMS over the right cerebellum. **B)** The timeline of an experimental trial. In the pitch discrimination task the two sounds were pure tones only differing for pitch. In the timbre discrimination task the two sounds were tones of the same instrumental family (wind vs. string) presented at the same high frequency and differing only in their timbre characteristic.

6.2.4 Transcranical magnetic stimulation

TMS was administered over the right cerebellum by means of a Magstim Rapid^2 machine (Magstim Co Ltd, Whitland, UK) with a 70 mm butterfly coil. An air-cooled coil was used, in order to avoid coil overheating. A fixed intensity of 45% of the maximum stimulator output was used, in line with prior studies (Schutter & van Honk, 2009). The right cerebellar hemisphere was targeted in light of previous evidence pointing to a right lateralized cerebellar activation during timbre processing (Reiterer et al., 2008), whereas pitch discrimination seems to induce bilateral cerebellar activations (Petacchi et al., 2011). The right cerebellum was localised in each participant as the region located 1 cm under and 3 cm lateral to the inion as in prior studies (Théoret et al., 2001; Torriero et al., 2004). Prior studies using neuronavigated TMS have demonstrated that this point lies over the cerebellar hemisphere (Renzi et al., 2014; Del Olmo et al. 2007). The coil was placed over the right cerebellum with the handle pointing upward, parallel to the inion-nasion line (Théoret et al., 2001; Oliver et al., 2011). Previous studies have shown that rTMS at 1 Hz temporarily reduces the excitability of the stimulated cortex for a time window that outlasts the period of stimulation (Chen et al., 1997; Oliveri et al., 2005). Sham stimulation was conducted with the coil held at a 90° position in order to ensure that the magnetic field did not stimulate the target area. The stimulation paradigm in the sham condition was the same as that of real rTMS stimulation.

6.2.5 Statistical analyses

Analyses were performed on mean accuracy scores and on reaction times (RT) for correct responses. Prior to analyses, reaction times 3 s.d.'s above or below the participants' mean were removed (this corresponded to 1.33% and 1.99% of the trials in the pitch and timbre discrimination task, respectively). A repeated-measures analysis of variance (ANOVA) with TMS condition (real vs. sham) and Session (pre-stimulation vs. post-stimulation) as within-subjects factors was performed separately for the pitch and the timbre discrimination task on accuracy scores and correct RT. Bonferroni-Holmes correction was applied to post-hoc comparisons.

6.3 RESULTS

Pitch discrimination task. Mean accuracy was above 73% (SD=9%) in all the experimental conditions. Analysis on accuracy scores revealed no significant main effects of TMS, $F(1,13)=3.75$, $p=.08$, $\eta_p^2=.22$, and of Session, $F(1,13)=1.47$, $p=.25$, $\eta_p^2=.10$. The interaction TMS by Session was not significant, $F(1,13)<1$, $p=.92$, $\eta_p^2=.00$. Mean correct RT are shown in Figure 2. The ANOVA on correct RT showed no significant main effect of either TMS, $F(1,13)=1.30$, $p=.27$, $\eta_p^2=.09$, or Session, $F(1,13)=2.86$, $p=.11$, $\eta_p^2=.18$. The interaction TMS by Session was significant, $F(1,13)=8.08$, $p=.01$, $\eta_p^2=.38$. Post-hoc *t*-

tests revealed that participants were significantly faster in responding in the post-sham stimulation session compared to the pre-sham stimulation session, *t*(13)= 3.52, *p*=.016, reflecting learning effects. In turn, RT were comparable between pre-real and post-real stimulation sessions, *t*(13)=.63, *p*=.54, suggesting that real TMS interfered with learning. Moreover, whilst RT were comparable in the sham and real pre-sessions, *t*(13)= 1.24, *p*=.24, indicating a similar level of baseline performance, participants tended to be slower following real, *t*(13)= 2.31, *p*=.09 (*p*=.03 uncorrected), than sham TMS.

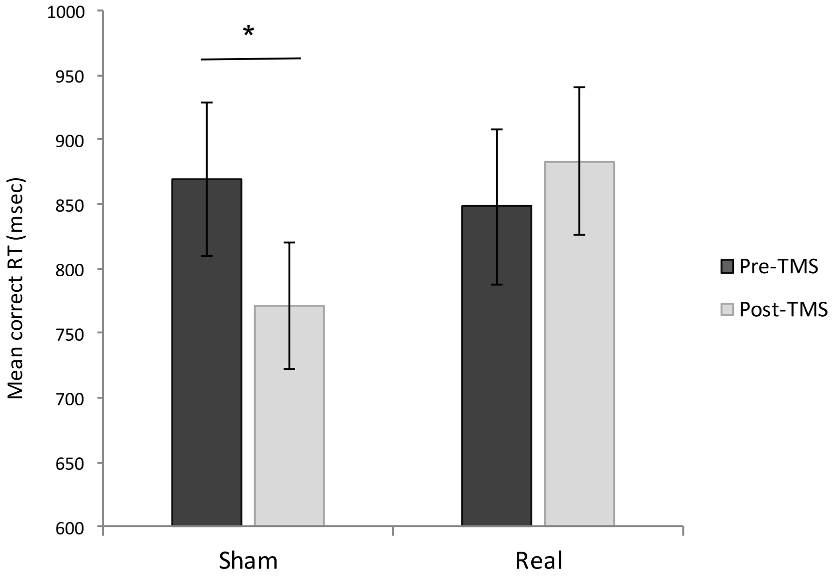

Figure 2. Mean response latencies for correct responses as a function of TMS (Real vs. Sham) and Session (Pre vs. Post stimulation) in the Pitch discrimination task. Participants were significantly faster (as indicated by the asterisk) in the post compared to the pre session when sham TMS was delivered, reflecting learning effects. RT were comparable in the pre and post real TMS sessions, suggesting that real TMS affected learning effects. Error bars represent ±1 SEM.

Timbre discrimination task. Mean accuracy was above 82% (SD=9%) in all the experimental conditions. The ANOVA on accuracy scores revealed no significant main effect of either TMS, *F*(1,13)=.50, *p*=.49, η_p^2 =.04, or Session, *F*(1,13)=.16, *p*=.22, η_p^2 =.12. The interaction TMS by Session was not significant, *F*(1,13)<1, *p*=.99, η_p^2 =.00. Figure 3 shows mean participants' correct RT. The ANOVA revealed a significant main effect of Session, *F*(1,13)=4.95, *p*=.04, η_p^2 =.28: participants were overall faster in the post-stimulation session (irrespective of stimulation being real or sham), reflecting learning effects. Neither the main effect of TMS, $F(1,13)=.03$, $p=.87$, $\eta_p^2=.00$, nor the interaction TMS by Session, $F(1,13)=.57$, $p=.46$, $\eta_p^2=.04$, reached significance.

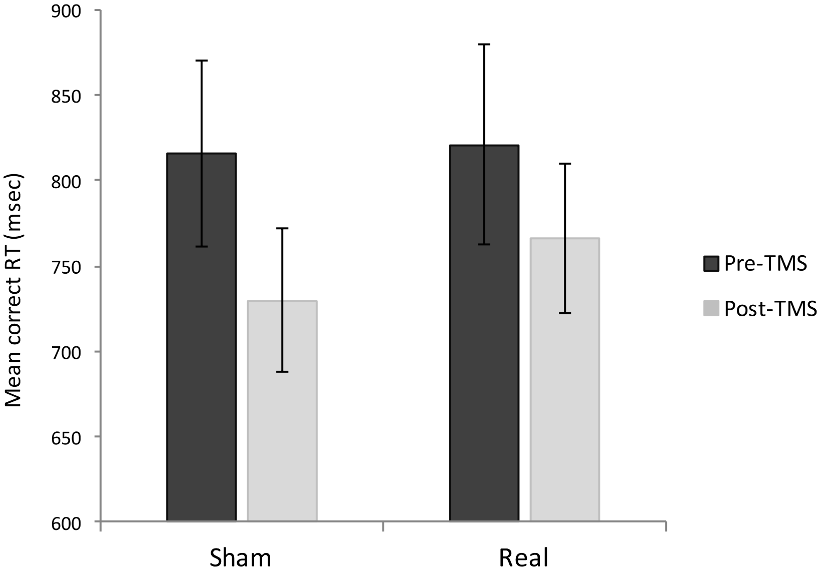

Figure 3. Mean response latencies for correct responses as a function of TMS (Real vs. Sham) and Session (Pre and Post stimulation) in the Timbre discrimination task. Participants were overall faster in the post compared to the pre session, reflecting learning effects. The type of stimulation (Real vs. Sham) did not affect performance. Error bars represent ±1 SEM.

6.4 DISCUSSION

In this study we aimed to shed light on the possible causal role of the cerebellum in auditory processing, and in particular in pitch and timbre discrimination, in light of prior neuroimaging and patients' data suggesting a cerebellar involvement in processing of music and single sound features (Tölgyesi & Evers, 2014; Ivry & Keele, 1989; Parsons et al., 2009; Frings et al., 2006). We found that interfering with cerebellar excitability via offline low frequency TMS significantly affected pitch discrimination, whereas it had no effect on timbre discrimination. In particular, in the pitch discrimination task real TMS counteracted learning effects that emerged in the other experimental conditions as faster responses in the post compared to the pre stimulation sessions.

Our findings are consistent with prior neuroimaging and neuropsychological evidence suggesting that the cerebellum is involved in perceptual tasks (Ben-Yehudah et al., 2007; Ackermann et al., 2007; Ivry & Schlerf, 2008), possibly monitoring the incoming sensory events to optimize perception (Bower, 1997; Roth et al., 2013). More specifically, our data add to previous studies that showed that discrimination of pitch and melody elicit activation of cerebellar regions (Gaab et al., 2003; Griffiths et al., 1999; Parsons et al., 2009), pointing to a causal role for the (right) cerebellum in processing pitch. Accordingly, resting state fMRI has shown functional connectivity between bilateral anterior cerebellum and the auditory cortex in the temporal lobes (He et al., 2004). In line with this, the cerebellum and the lateral anterior temporal lobe appeared to be bi-directionally interconnected during a rhyming judgement task (Booth et al., 2007). More in general, consistent evidence suggests that the cerebellum projects not only to motor, but also to somatosensory areas (D'Angelo and Casali, 2013). Still, it is important to consider that the type of stimulation we used may have only affected posterior cerebellar regions, given the deep position of the cerebellum (Hardwick et al., 2014). Indeed, coil geometry seems to be an important factor in determining effective stimulation of deep cerebellar regions, with the figure of eight coil likely being suboptimal in reaching motor areas (Hardwick et al., 2014). Nonetheless, cerebellar stimulation parameters similar to ours significantly affected perceptual (Cattaneo et al., 2014) and cognitive (for instance, linguistic) functions in prior studies [e.g., (Argyropoulos & Muggleton, 2013; Arasanz et al., 2012)], suggesting that the stimulation we used was able to interfere with neural activity in the cerebellar-cortical network subtending discrimination of sound features.

Although prior neuroimaging evidence also suggested a possible role for the cerebellum in timbre processing (Alluri et al., 2012; Reiterer et al., 2005), we did not find evidence for this in our study. On the one hand, the lack of TMS effect in the timbre discrimination task ensures that the effects we reported in the pitch discrimination task were not due to unspecific effects of TMS slowing down responses regardless of the specific task at play. On the other hand, it is possible that real TMS affects auditory discrimination task only when the task has a certain level of complexity. Performance accuracy was indeed overall higher in the timbre than in the pitch task. This is in line with prior literature showing that non-musicians are more sensitive in sound categorization to changes in timbre than to changes in pitch (Pitt, 1994). Interestingly, previous studies demonstrated a positive correlation between cerebellar activation and task difficulty (Baumann & Mattingley, 2010; Petacchi et al., 2011; Gaab et al., 2003; Holcomb et al., 1998). Moreover, the right cerebellum may be more important than the left in difficult auditory discrimination (Reiterer et al., 2005).

In interpreting our data, it is also worth mentioning that pitch and timbre processing may have a different degree of lateralization in the brain. There is evidence for a right hemisphere dominance in the temporal lobes for musical timbre discrimination (Crummer et al., 1994; Samson, 2003) related to a right hemispheric specialization in processing spectral sound features, that are critical for discriminating timbre differences (Menon et al., 2002; Zatorre et al., 2002; Warren et al., 2005). In line with these findings, the left cerebellum may be more important than the right in timbre processing (cerebral cortex fibers mainly projecting to the contralateral cerebellar cortex (Jansen et al., 2005). However, other studies in infants (Dehaene-Lambertz, 2000) and in adults (Reiterer et al., 2008; Deike et al., 2004) reported left hemispheric cerebral cortex engagement underlying perception of timbre change. Moreover, Reiterer and colleagues (Reiterer et al., 2008) showed a right cerebellar activation during timbre processing, speculating that pre-linguistic sound features (including timbre) may be represented by a complex network that connects Broca’s area and the right cerebellum. In turn, previous neuroimaging studies mainly indicate bilateral cerebellar activation during pitch processing (Petacchi et al. 2011; Griffiths et al., 1999; Parsons, 2001; Gaab et al., 2003; Petacchi et al., 2005; Alluri et al., 2012), although some degree of lateralization may occur depending on task complexity (Belin et al., 2002; Reiterer et al., 2005; Belin et al., 1998). Evidences are thus not entirely consistent regarding lateralization of timbre and pitch processing. Future studies may address this issue by comparing the effect of left and right cerebellar stimulation on auditory discrimination.

7.

Study 6:

Testing the role of dorsal premotor cortex in auditory-motor association learning using TMS[*]

---

[*] This study is based on: Lega, C., Stephan, M. A., Zatorre, R. J., & Penhune, V. (2016). Testing the Role of Dorsal Premotor Cortex in Auditory-Motor Association Learning Using Transcranical Magnetic Stimulation (TMS). *PloS one*, *11*(9), e0163380.

7.1 INTRODUCTION

Despite strong correlational evidence assessing the role of dorsal premotor cortex (dPMC) in auditory-motor association formation, there is no direct causal evidence that dPMC is required for learning new auditory-motor associations. Previous studies using inhibitory rTMS to disrupt dPMC function have shown that it perturbs auditory-motor synchronization (Giovannelli et al., 2014; Meehan et al., 2013), but its effect on auditory-motor learning has not been examined. Therefore, in the current study we used rTMS to disrupt dPMC function as non-musicians learned to associate a musical note with a key press. To test the effect of rTMS on auditory-motor learning, we used a variation of the paradigm developed by Chen and colleagues (Chen et al., 2012) in which non-musicians learn a set of auditory-motor associations through melody training. Further, we tested the role of dPMC in learning a new auditory-motor association in two contexts: first when the sound to key-press mapping was conventional and ordered (low to high mapped on to left to right key order), and then when it was scrambled. In the first context, when the mapping is conventional, even non-musicians may have pre-existing associations between pitch and spatial location. Indeed, previous studies demonstrate that non musicians tend to associate sounds in an ascending musical line with a spatial mapping from left to right (Rusconi et al., 2006; Lidji et al., 2007), and that this conventional, ordered mapping facilitates action planning and sequence learning (Hoffmann et al., 2001; Stöcker et al., 2003; Keller & Koch, 2008). In the second context, the mapping between pitch and key location was scrambled, guaranteeing that participants would have to learn an entirely new set of arbitrary auditory-motor associations. We hypothesized that if the dPMC is crucial in learning new auditory-motor association, then rTMS over dPMC should interfere with that learning, and that interference would be greater in the scrambled compared to the conventional pitch-to-key mapping.

7.2 METHODS

7.2.1 Participants

Fifty young, healthy participants took part in these experiments. Twenty-four were tested in Experiment 1 (7 M; mean age = 22.37; SD = 3.80) and twenty-six in Experiment 2 (10 M; mean age = 22.88; SD = 3.68). In each experiment half of the participants (12 in Exp 1 and 13 in Exp 2) were randomly assigned to the dPMC stimulation and the other half to the V1 stimulation. Participants were selected to have little musical training (Experiment 1: mean years of musical training = 0.58; SD = 0.77; Experiment 2: mean years of musical training = 0.92; SD = 1.01). They were all right-handed, according to the Edinburgh Handedness Inventory (Oldfield, 1971). Prior to the experiments, each participant filled out a questionnaire to assess whether it was safe for them to undergo TMS. None of the volunteers reported neurological or psychiatric problems, seizures, or was taking any medication that could interfere with neuronal excitability. All participants provided written informed consent. The local ethics committee (Comité d'éthique de la recherche en santé (CERES)) approved the protocol, and participants were treated in accordance with the Declaration of Helsinki.

7.2.2 Procedure, tasks and stimuli

The paradigm used for Experiments 1 and 2 was based on that developed by Chen and colleagues (Chen et al., 2012). The timeline for both experiments is shown in Figure 1. At the beginning of each experiment, participants were tested on a pitch-to-keypress matching task (referred to as “Pitch Matching”) to assess their ability to associate each of four pitches (C, D, E, G) with one of four keys on the computer keyboard. Participants heard single pitches and had to match them to the keys using the four fingers of the right hand (not including the thumb). The letters on the keys were covered so participants had to associate each pitch with a location or keypress, not with the letter on the key. Each of the four pitches was presented 10 times for a total of 40 trials. Trials were pseudo-randomly ordered (to avoid the same pitch occurring twice in a row), and no auditory feedback was given, in order to rule out possible learning effects (Chen et al., 2012; Lahav et al., 2007). Key-press responses and reaction times were recorded by the computer.

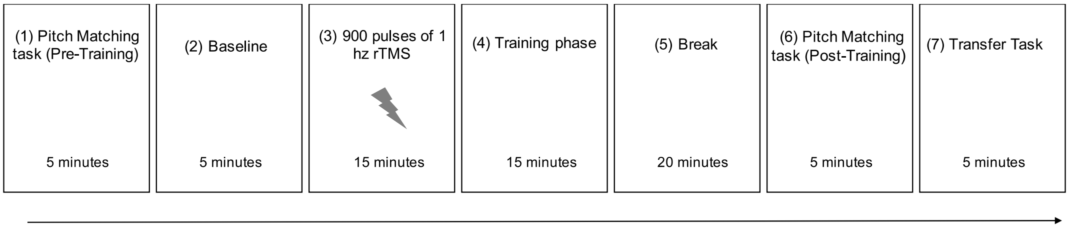

Figure 1. The timeline of the Experiment 1 and 2. Participants were asked to perform the Pitch-to-key press matching task (1) and the Baseline (2) task before the TMS stimulation. After that, 1 Hz rTMS was applied over the dorsal premotor cortex (dPMC) and over the primary visual cortex (V1) (3) before the 3 Blocks of training (4). At least 40 minutes after the end of the stimulation (5) participants performed again the pitch-to-key press matching task (6) and the Transfer task (7).

In Experiment 1 pitches were mapped to keys in a conventional, ordered low-to-high/left-to-right mapping: C=key 1, D=key 2, E=key 3, G= key 4. In Experiment 2 the mapping of pitches to keys was scrambled: E=key 1, C=key 2, G=key 3, D= key 4 (See Figure 2).

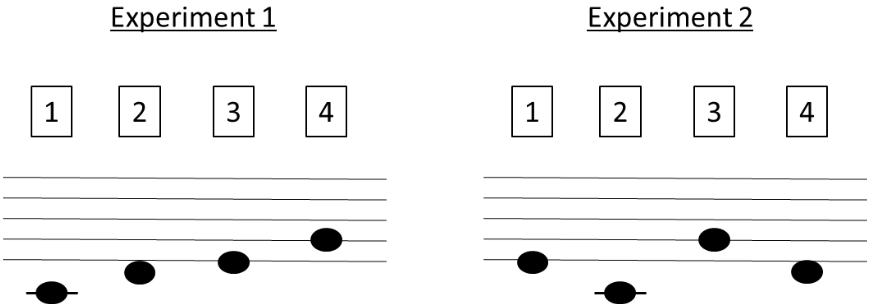

Figure 2. The auditory-motor mapping of Experiment 1 and Experiment 2. In Experiment 1 tones were ordered from low to high and corresponded to the spatial position from left to right on the computer keyboard (C=key 1, D=key 2, E=key 3, G= key 4). In Experiment 2 the order of notes was scrambled (E=key 1, C=key 2, G=key 3, D= key 4).

Participants were then trained on the pitch-to-key association through a melody playback task in which the same four pitches were used. The first block of training occurred immediately before application of the rTMS and served as a baseline (Pre-training). Participants then received 900 pulses of inhibitory 1 Hz rTMS over either the dPMC or over the control site (primary visual cortex, V1). For details of the rTMS stimulation see below. Participants then performed three additional blocks of melody playback (Training).

On each trial of training participants first listened to a melody and were then asked to play it back using the four keys on the computer keyboard. Each key press evoked a specific tone. Thus, participants heard both the target and their own responses, which was designed to allow them to learn the key-to-pitch associations. Melodies were five-note sequences in which the same four notes were rearranged to create different sequences. In each melody, three pitches (C, D, G) were used once and a fourth (E) was used twice, (e.g., D E C E G). There were 45 different melodies in total, and each block of training contained 15 melodies. Each Block lasted approximately 5 minutes, for a total of 15 minutes of training.

Following the three blocks of training, participants were retested on the Pitch Matching task (Post-training). The inhibitory effect of 1 Hz rTMS is usually thought to last for a maximum of 20-30 minutes (Fitzgerald et al., 2006). To be sure that the effect had dissipated before retesting on the Pitch Matching task, we allowed 40 minutes to elapse. To do this we introduced a short break after the end of the last training block. This break was on average 20 minutes and participants were simply asked to relax, and not to listen to music, or type on a computer keyboard.

Finally, in order to assess the flexibility of the pitch-to-key associations, we tested participants on one block of transfer in which they listened to and reproduced a single novel melody repeated 15 times (Transfer).

All pitches and melodies were presented through headphones, and were created with the "GarageBand" music editing software (GarageBand 6.0.4, Apple Inc. 2011) using a synthesized piano timbre. Each of the four tones lasted 600 msec. Keypress responses and RTs (Reaction Times) were recorded by the computer.

7.2.3 Transcranical magnetic stimulation

Each participant underwent inhibitory 1 Hz TMS stimulation over the dPMC or over V1. Stimulation intensity was 90% of individual active motor threshold (aMT). The mean stimulation intensity over dPMC was 38.60% of maximum stimulator output and 41.28% over V1, with no significant difference between the two areas (p = .15). For each of the sites, 900 pulses were applied at a frequency of 1 Hz (train duration 15 min). The site for dPMC stimulation was located 1 cm medial and 2.5 cm anterior at the same laterality as the motor ‘hot-spot’ (Ortu et al., 2009) defined as the site where the largest MEPs could be evoked in the relaxed first dorsal interosseus (FDI) muscle (see Figure 3). The site for V1 stimulation was localized as the point lying 1.5 cm superior to the inion on the midline (Cattaneo et al., 2014; Heinen et al., 2005; Pascual-Leone & Walsh, 2001). We checked in each participant whether stimulation over the defined dPMC evoked any MEPs and moved the coil 0.5 cm anterior in four subjects where this was the case (Bestmann et al., 2005). TMS was applied through a 70 mm figure-of-eight coil, using a Super Rapid Biphasic Stimulator (Magstim, Whitland, UK) with the handle pointing 45° postero-laterally away from the midline for both M1 and the dPMC. For the V1 control site the coil was placed with the handle pointing upward, parallel to the inion-nasion line. A TMS neuronavigation system (Brainsight, Rogue Research Inc., Canada) was used to ensure a constant coil position during the 15 minutes of stimulation. The aMT was determined according to standard procedure during slight tonic contraction of the FDI muscle (20% of maximal force), using the software based ‘adaptive method’ developed by Awiszus (Awiszus, 2011) (Motor Threshold Assessment Tool (MTAT, version 2.0: http://www.clinicalresearcher.org/software). An MEP $\geq$ 200 µV peak-to-peak amplitude was fed back to the software as valid response. EMG recordings were obtained from the right FDI muscle, with conventional surface electrodes in a belly-tendon montage. Signals were amplified, bandpass filtered (1 Hz – 2 kHz) and sampled at a rate of 10 kHz.

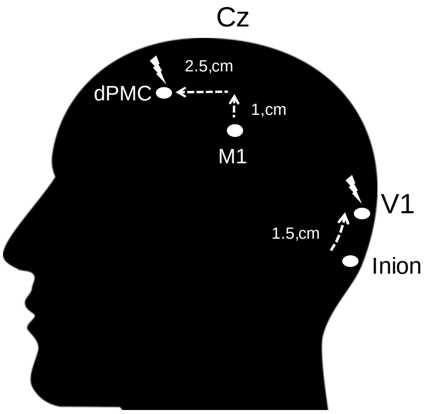

Figure 3. Diagram of the stimulated site. The dPMC was located 1 cm medial and 2.5 cm anterior at the same laterality as the motor ‘hot-spot’ (M1), defined as the site where the largest MEPs could be evoked in the relaxed first dorsal interosseus (FDI) muscle. The site for V1 stimulation was localized as the point lying 1.5 cm superior to the inion on the midline.

7.3 RESULTS

Analysis were performed on both reaction times and accuracy. No significant effect of TMS was found for the reaction times. Thus, we report here only the analysis on the accuracy scores.

7.3.1 Experiment 1

Pitch Matching Task. A 2x2 repeated-measures ANOVA with Session (Pre- and post- melody playback training) as the within-subjects variable and TMS location (dPMC and V1) as the between-subjects variable was carried out on the percentage of correctly played pitches. The ANOVA revealed a significant interaction between Session and TMS location (F(1,22) = 6.08, *p*=.02, $\eta_p^2 = .21$) (see Figure 4a). Post-hoc comparisons (Bonferroni-Holmes correction) revealed that the V1 group showed a significant improvement in pitch matching performance between the first and the second session, (*t*(11) = 3.55, *p* = .02), but the dPMC group did not (*t*(11) = .32, *p* = .75). Importantly, there were also no significant differences between the V1 and dPMC groups at pre-test (*p* = .66).

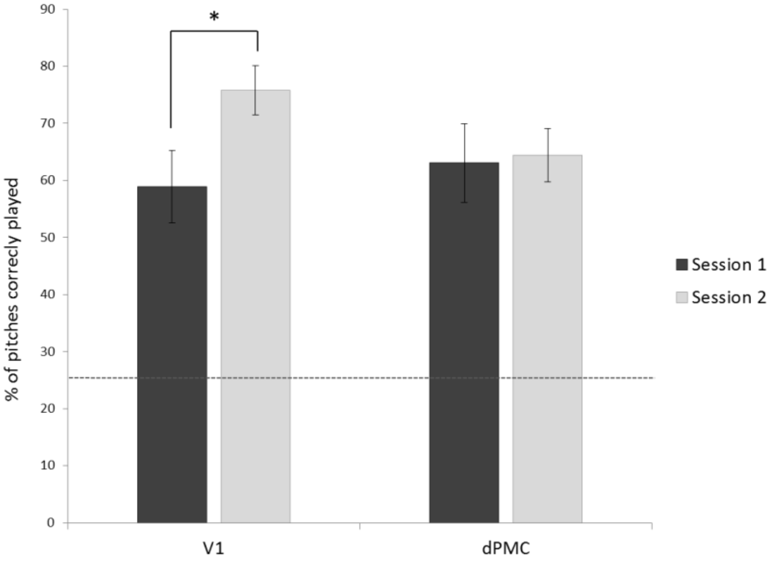

Figure 4a. Results of the Pitch matching task of Experiment 1 (ordered mapping). Columns represent average data with standard error bars. The y-axis represents percent correct scores plotted across the TMS site (dPMC and V1). As indicated by the asterisk, TMS over the dPMC significantly reduced participants' accuracy compared to the V1 stimulation. Horizontal dashed line indicates the level of chance (25%)

Training. A 4x2 repeated-measures ANOVA with Session (Baseline, Block 1, Block 2, Block 3) as the within-subjects variable and TMS location (dPMC and V1) as the between-subjects variable was carried out on the percentage of correctly played pitches. The analysis revealed a significant main effect of Session ($F(3,66) = 5.25$, $p<.01$, $\eta_p^2 = .19$). Neither the main effect of TMS ($p = .88$), nor the interaction between Session and TMS ($p > .48$) reached significance. Post-hoc comparisons with Bonferroni correction showed that for both groups performance improved significantly between Blocks 1 and 3 ($p=.05$) and Blocks 2 and 3 ($p=.002$) (see fig. 4b). Since we expected a linear improvement across the four Blocks we also conducted a linear trend analysis, which revealed a significant linear effect of Session ($F(1,22) = 4.69$, $p=.04$, $\eta_p^2 = .18$), but no interaction between Session and TMS location ($F(1,22) = .78$, $p=.39$, $\eta_p^2 = .03$).

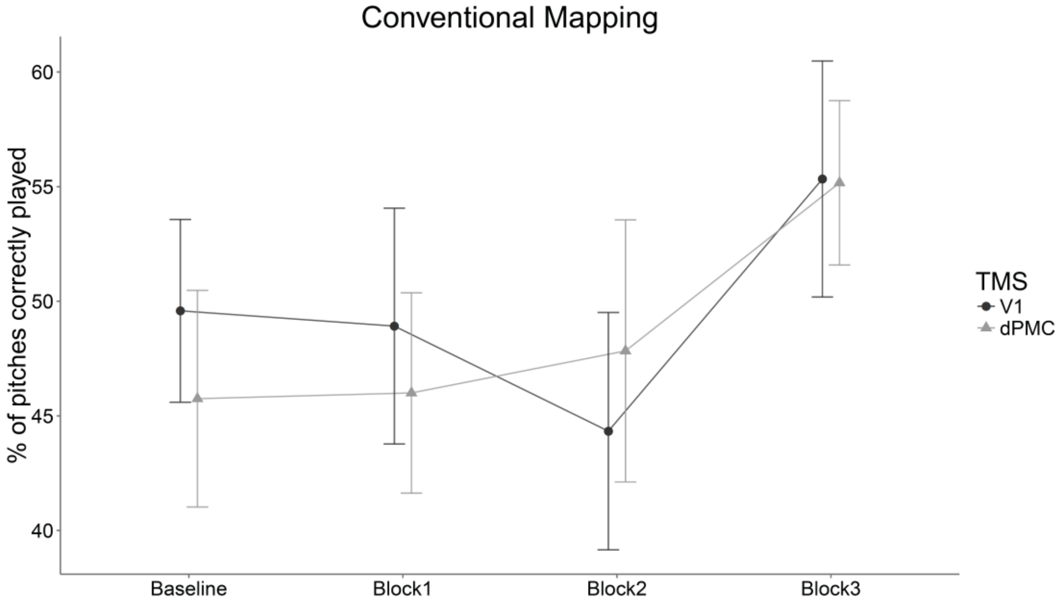

Figure 4b. Results of the training phase of the study of Experiment 1. Graph represents the percentage of pitches correctly played plotted across blocks of trials (Baseline, Block 1, Block 2 and Block 3) for the dPMC and the V1 groups. Error bars represent ±1 SEM.

Transfer task. To test the effect of TMS on transfer to learning a new melody, we compared the percentage of correctly played pitches on the Transfer block between the V1 and the dPMC group using an independent samples t-test. The analysis showed no significant difference between the two groups ($t(22) = .18$, $p = .86$). In addition, we used a second measure of learning, the number of 100%-correct trials. This analysis also showed no significant difference between groups ($t(22) = .00$, $p = 1.0$).

7.3.2 Experiment 2

Pitch Matching Task. The outcome of this task was overall much more difficult than Experiment 1, as expected. This is confirmed by comparing the pre-training performance of both groups in Experiment 1 and Experiment 2, showing a significance difference between the two mappings, ($t(48) = 6.04$, $p < .01$). Moreover, in Experiment 2 performance pre-training did not differ from chance for either the premotor group ($t(12) < 1$, $p = .56$) and the V1 group ($t(12) < 1$, $p = .66$), whereas it did in Experiment 1 (premotor group ($t(11) = 5.51$, $p < .001$); V1 group ($t(11) = 5.37$, $p < .001$)). A 2x2 repeated-measures ANOVA with Session (Session 1, before training and Session 2, after training) as the within-subjects variable and TMS location (dPMC and V1) as between-subjects variable was carried out on the percentage of correctly played pitches. The ANOVA revealed a significant interaction between Session and TMS location ($F(1,22) = 4.98$, $p=.03$, $\eta_p^2 = .17$). Post-hoc comparisons (Bonferroni-Holmes correction) revealed a significant improvement between the first and the second session for the V1 group ($t(12) = 4.55$, $p = .004$), but not for the dPMC group ($t(12) = 1.84$, $p = .16$). Importantly, there were no differences between groups at Pre-test ($t(24) = .19$, $p = .85$), and the V1 group out-performed the dPMC group at Post-test ($t(24) = 2.87$, $p = .02$) (see Figure 5a).

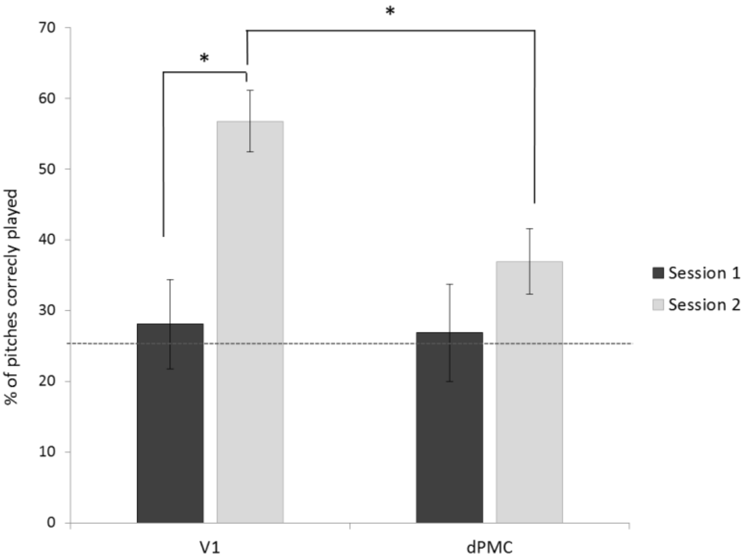

Figure 5a. Results of the Pitch matching task of Experiment 2 (scrambled mapping). Columns represent average data with standard error bars. The y-axis represents percent correct scores plotted across the TMS site (dPMC and V1). As indicated by the asterisk, TMS over the dPMC significantly reduced participants' accuracy compared to the V1 stimulation. Horizontal dashed line indicates the level of chance (25%).

Training. A 4x2 repeated-measures ANOVA with Session (Baseline, Block 1, Block 2, Block 3) as within-subjects variable and TMS (dPMC and V1) as between-subjects variable was carried out on the percentage of pitch correctly played. Analysis revealed a significant main effect of Session, (F(3,72) = 18.30, $p < .01$, $\eta_p^2 = .43$). Post-hoc comparisons (Bonferroni-Holmes correction) showed that both the Baseline and the Block 1 significantly differed when compared to both the Block 2 and Block 3 ($p < .001$). The main effect of TMS was also significant (F(1,24) = 8.58, $p < .01$, $\eta_p^2 = .26$), indicating significantly higher scores for the V1 compared to the dPMC group. Particularly, post-hoc comparisons (Bonferroni-Holmes correction) revealed that the V1 group and the dPMC group differed significantly for Block 2 (t(24)= 3.29, p=.01) and Block 3 (t(24)= 2.64, p = .04). The groups did not differ at Baseline or Block 1 (p= .09) (see fig. 5b). There was no significant interaction between Session and TMS location (F(3,72) = 2.12, $p = .10$, η_p^2 = .08).

Using a linear trend analysis confirmed a significant linear effect of Session (F(1,24) = 39.15, $p < .01$, $\eta_p^2 = .62$), but also revealed a significant interaction between

Session and TMS location (F(1,24) = 4.35, *p* = .04, η_p^2 = .15), indicating that the learning rate differed across groups. Specifically, this was due to a stronger linear effect for the V1 group (*p* < .001), compared to the dPMC group (*p* =. 01) (See also Figure 5b). Indeed, considering the two groups separately, post hoc analysis (Bonferroni correction) revealed that the V1 group showed a significant improvement between the baseline and Block 2 (*p* < .001) and Block 3 (*p* = .002) and between Block 1 and both the Block 2 (*p* = .01) and Block 3 (*p* = .01). Conversely, the dPMC group only showed a significant improvement when comparing Baseline to Block 2 (*p* = .03).

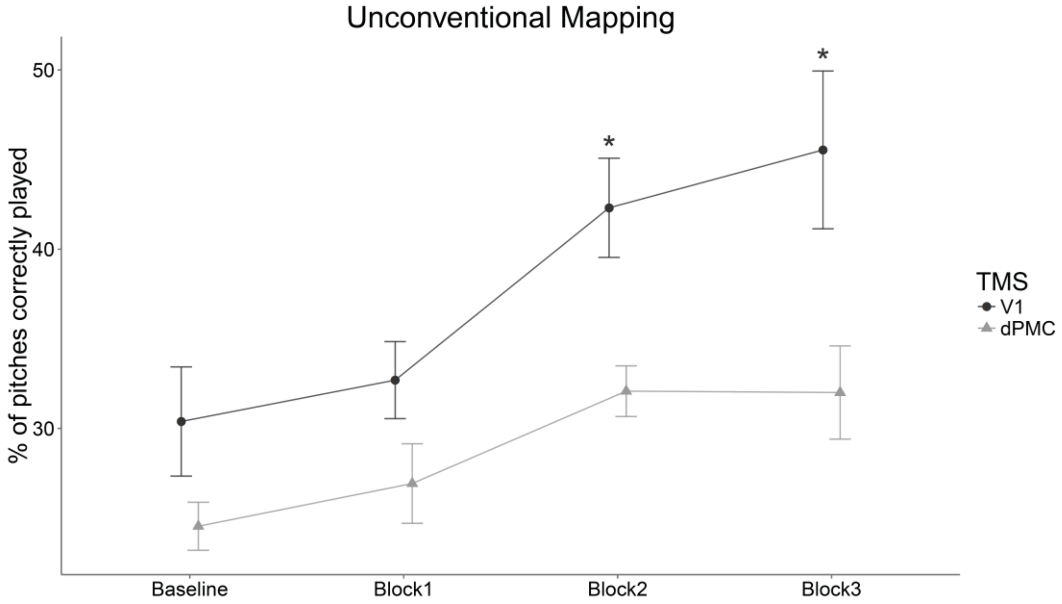

Figure 5b. Results of the training phase of the study of Experiment 2. Graph represents the percentage of pitches correctly played plotted across blocks of trials (Baseline, Block 1, Block 2 and Block 3) for the dPMC and the V1 groups. As indicated by asterisks, TMS over the dPMC significantly reduced participants' accuracy in Block 2 and 3 compared to the V1 stimulation. Error bars represent ±1 SEM.

Transfer task. The transfer task was analysed using an independent t-test on the percentage of pitches correctly played between the V1 and the dPMC group. Analysis showed no significant difference between the two groups (*t*(24)= 1.39, *p* = .17). Notably, when we considered the number of 100% correct melodies, the analysis revealed a near significant difference between the two groups (*t*(24) = 1.93, *p* = .06), with the V1 group out-performing the dPMC group.

7.4 DISCUSSION

The current results provide some of the first direct causal evidence in humans that dPMC is involved in the learning and expression of auditory-motor associations. Inhibitory 1 Hz rTMS over dPMC impaired participants' ability to learn the association between a pitch and a keypress in two independent samples, and this effect was greatest when they were required to learn an unconventional, novel association. These findings are consistent with the hypothesized role of the dPMC in encoding sensory-motor associations, particularly when they are complex or abstract. It is also consistent with findings of previous neuroimaging studies showing that the dPMC is part of a network of regions engaged during learning of auditory-motor associations in the context of music. Finally, our finding that learning of the conventional low to- high/left-to-right mapping was less impaired by rTMS indicates that some auditory-motor associations may be learned implicitly in the absence of explicit musical training.

Evidence from animals and humans suggests that dPMC is important for learning and expression of abstract or higher-order sensory-motor associations (Hoshi & Tanji, 2006; Petrides, 2005; Cisek & Kalaska, 2004; Picard & Strick, 2001). Globally, evidence from electrophysiological studies in animals has shown that neurons in the PMC respond to auditory and visual stimuli that are linked to known actions (Keysers et al., 2003). As described in the Introduction, the premotor cortex can be subdivided into dorsal and ventral subregions (Hoshi & Tanji, 2006; Barbas & Pandya, 1987), both of which are part of the dorsal auditory stream that links auditory and motor representations. These regions have been found to be active in both musical and speech contexts, when interactions between auditory and motor systems are critical (Zatorre et al., 2007; Warren et al., 2005; Hickok and Poeppel, 2007). In the visual domain it has been proposed that the vPMC and dPMC are involved in direct and indirect visuo-motor transformations, respectively (Hoshi & Tanji, 2006, 2007). Specifically, vPMC seems to be critical anytime there is a direct mapping from sensory information into the motor system (Zatorre et al., 2007; Hoshi & Tanji, 2007). In the grasping movement, vPMC neurons are responsible of processing the shape of an object, and selective lesions of the vPMC in the macaque monkey impair hand shaping, leaving sensory processing undamaged (Fogassi et al., 2001). In contrast, the dPMC is thought to be involved in indirect or higher-order sensorimotor integration (Zatorre et al., 2007; Hoshi & Tanji, 2007) including the selection of movements that are conditionally linked to a sensory stimulus (Petrides, 2005; Amiez et al., 2012; Halsband &

Freund, 1990; Passingham, 1989; Wise et al., 1996). This conclusion comes largely from studies in the visual system, where it has been demonstrated in monkeys that the inactivation of the dPMC, but not of the vPMC, impairs conditional motor behaviors (Kurata & Hoffman, 1994). Taken together, there is compelling evidence that both the learning and the performance of arbitrary sensorimotor mappings in conditional associative tasks depends on a complex neural network that includes the dPMC (Amiez et al., 2012; Fogassi et al., 2001; Wise et al., 1996; Grafton et al., 1998; Germain & Lamarre, 1993, Grol et al., 2006; Petrides, 1985). This functional property of the dPMC perfectly fits with the results of the Pitch matching task in both Experiment 1 and Experiment 2, where a conditional motor response driven by a sensory input is required because participants must decide which movement to select among four competing alternatives.

Interestingly, we also observed that rTMS over dPMC reduced the expression of learning at recall, and slowed learning when the auditory-to-motor mapping was novel, but not when it was conventional. The finding that learning was impaired for the novel (unconventional) mapping only, appears to be consistent with the role of the dPMC in learning new associations, especially when those associations are more abstract and/or complex. Indeed, the manipulation of the pitch-key mapping in the second experiment allowed us to test the hypothesis that dPMC engagement would be more important when learning a more complex or unfamiliar auditory-motor association: scrambling the natural pitch-to-key mapping forced participants to explicitly learn the associations rather than relying on a pre-existing mapping from low to high. Previous studies in both the visual (Amiez et al., 2012) and the auditory (Chen et al., 2008) domain reported a correlation between activation of the dPMC and complexity of the task. Amiez et al. (Amiez et al., 2012), analyzing the brain activation during the learning and the execution of conditional visuo-motor responses, reported that the dPMC is the only area modulated by the number of visuo-motor mappings to be acquired, i.e. the dPMC is sensitive to the difficulty of the task. Combined with these findings, our results show that dPMC is critical in learning conditional sensorimotor associations in the auditory, as well as the visual modality. Thus, we cannot exclude that the greater interference of the TMS observed in Experiment 2 is correlated with the greater involvement of the dPMC once an unconventional (more difficult) mapping is required. The results of the transfer task seem to go in the same direction: when we considered the number of 100% correct melodies, results revealed a trend for the control group to perform better compared to the premotor group in Experiment 2, but not in Experiment 1. These results of the transfer task partially support the role of the dPMC in acquiring new auditory-motor associations. Notably, the main difference between the transfer task and the training blocks lies in the fact that during training the sequences were novel for every trial, whereas in the transfer task the auditory-motor sequence was always the same. Based on the concept that vPMC is more important for direct associations, we could also hypothesize that inhibiting this region via rTMS could have a greater impact on the transfer task, compared to the dPMC stimulation. Another interesting result of these experiments is that rTMS interfered with learning of the novel, but not the standard left-to-right/low-to-high mapping. This is consistent with evidence showing that some musical associations can be learned without explicit training, supporting the idea that some abilities can be acquired just by being exposed to music (Bigand & Poulin-Charronnat, 2006). Our results are also consistent with studies that demonstrated the so-called SMARC effect (Spatial Musical Association of Response Codes). They showed that higher pitches facilitate (in terms of speed and accuracy) up\right motor responses, and low pitches facilitate down\left motor responses, even when tones were irrelevant to the task (Rusconi et al., 2006; Lidji et al., 2007; Cho et al., 2012). The SMARC effect would explain the fact that pitch-mapping performance for Experiment 1 was above chance even before training, but for Experiment 2 it was not. Moreover, hearing tones with a conventional mapping facilitates sequence learning in the serial reaction-time task (Hoffmann et al., 2001; Stöcker et al., 2003). Similarly, Keller & Koch (2008) demonstrated faster action planning when the mapping between keys and tones was compatible than when it was incompatible. We have proposed that rTMS interfered with learning of the key-to-tone mapping in Experiment 2 because it was unconventional and novel. In addition, the key-to-tone mapping was not ordered in a spatially sequential manner, i.e., adjacent keys did not correspond to adjacent tones. Thus, future experiments could test an unconventional mapping that is still sequentially ordered, i.e. left-to-right, high-to-low. Further, it is also possible that based on the functional dissociation between dorsal and ventral premotor cortex, the conventional low-to-high mapping may be encoded in the vPMC because more direct and implicit, rather than in the dPMC. Thus future studies could compare the effects of rTMS over both ventral and dorsal regions.

Dorsal premotor cortex is part of a network of regions previously shown in neuroimaging studies to be engaged during learning of auditory-motor associations in the context of music (Zatorre et al., 2007; Herholz et al., 2015; Bermudez & Zatorre, 2005; Lahav et al., 2007). In particular, the study on which our paradigm is based demonstrated that left dPMC was engaged during melody learning, and that activity in this region was related to performance (Chen et al., 2012). The left dPMC was chosen as site of interest based on previous studies on auditory-motor learning, which identified more significant changes in the left as compared to the right dPMC after auditory-motor training (Chen et al., 2012; Herholz et al., 2015). In the future, it would be interesting to compare the effects of left and right dorsal premotor cortex stimulation. The role of the dPMC in auditory-motor integration is likely based on its pattern of connectivity. Indeed, neuroanatomical studies in non-human primates show that the dorsal and ventral PMC are directly connected to both the posterior temporal gyrus and the primary motor cortex (M1), which makes them a critical node in connecting and integrating auditory and motor information (Chen et al., 2009, 2012; Brown et al., 2013; Herholz et al., 2015; Dum & Strick, 2005). Human studies using diffusion tensor imaging (DTI) to assess language pathways indicate that there are similar connections between the Superior Temporal Gyrus (STG) and the PMC via a dorsal route along the arcuate and superior longitudinal fasciculi, although there is controversy about the precise organization of these fibers (Saur et al., 2008; Glasser & Rilling, 2008; Thiebaut de Schotten et al., 2012; Maffei et al., 2015). Nonetheless, premotor cortices are one link in a complex network of brain regions, which includes cerebellum, posterior auditory and inferior parietal cortices. Studies in both animals and humans have already demonstrated the involvement of parietal, sensorimotor and premotor cortices in the control of movement when the integration of spatial, sensory and motor information is required (Hikosaka et al., 2002; Parsons et al., 2005). By analogy with the functional division between ventral and dorsal stream proposed in the visual system, different models suggested that the dorsal auditory stream would be responsible for preparing motor responses from incoming auditory information and in the localization of sounds in space (Warren et al., 2005; Hickok & Poeppel, 2000, 2007). This concept of the dorsal stream not only unifies its function between vision and audition, but also theorizes its role in auditory-motor integration, critical for both music and speech. Finally, previous work shows that the rTMS protocol used in this experiment is effective in interfering with cortical excitability over the dPMC (Giovannelli et al., 2014; Pollok et al., 2008). Moreover, localizing the stimulation site relative to the M1 hot-spot (2.5 cm anterior and 1 cm medial to M1) (Ortu et al., 2009), allowed us to take into account inter-individual differences in the functional architecture of the brain. However, this approach may not provide the same accuracy and precision as fMRI–guided stimulation (Sparing et al., 2008; Feredoes et al., 2007; Sack et al., 2009). In the future, the use of fMRI-guided TMS localization could increase stimulation precision, reducing the variability across subjects, and potentially strengthen some of the current results.

8.

General conclusions

In this dissertation, I investigated how pitch dimension interacts with spatial and motor processing. Overall, our results demonstrated that the frequency of a sound is able to affect both motor planning and spatial representations, with this effect being reinforced and modulated by musical expertise.

In particular, Study 1 and Study 2 investigated the influence of musical expertise and early visual experience on the SMARC effect, which refers to the to the tendency to associate bottom/left responses with low-pitched tones, and top/right responses with high-pitched tones (see Rusconi et al. 2006; Lidji et al., 2007). Study 1 revealed that sensorimotor experience due to musical training affects the horizontal representation of pitch. This suggests that instrumental expertise plays a crucial role in the way our cognitive system maps pitch onto a mental representation of space, at least when spatial coordinates are not already associated with auditory frequency through the statistics of natural auditory scenes. Previous studies already demonstrated the effect of musical expertise on the representation of tones (Rusconi et al., 2006; Lidji et al., 2007). Nonetheless, the current findings take a step forward, suggesting the importance of considering the specific instrument played (i.e the specific sensorimotor experiences). Study 2 suggested that the lack of a normal visual experience does not prevent the development of a music mental (vertical) line. Indeed, our data show that blind participants associated tones along a spatial metal music line, with low tones preferentially associated with bottom responses and high tones with top responses, much like sighted participants. These findings extend previous studies on blind individuals trying to confine the role of vision in the conceptualization of abstract concepts, like numbers or temporal and serial order (Bottini et al., 2015; Bottini et al., 2016; Crollen et al., 2013; Pasqualotto et al., 2014; Castronovo & Seron, 2007; Cattaneo et al., 2010; Rinaldi et al., 2015). Furthermore, although previous studies have shown that sound frequency is represented in a spatial format and that the “music mental line” can affect bimanual motor responses (Rusconi et al., 2006; Lidji et al., 2007; Nishimura and Yokosawa, 2009; Cho et al., 2012), Study 3 provides the first evidence that pitch height influences the allocation of spatial attention cross-modally in tactile and visual peri-personal space in musicians, but not in non-musicians. This third study corroborated Study 1, suggesting the importance of considering musical experience in the association between pitch and space processing.

Taken together, the first three studies confirm previous evidence claiming that the vertical spatial association of tones would be more robust and independent from experience compared to the horizontal one (Rusconi et al., 2006; Lidji et al., 2007; Pitteri et al., 2015; Weis et al., 2016a). Accordingly, we did not find any effect of both musical practice, timbre and early visual experience in the vertical response setting. In turn, we corroborated the evidence showing an impact of musical expertise on horizontal plane (Rusconi et al., 2006; Lidji et al., 2007), especially in the case of piano players (probably for their familiarity with the piano keyboard). Furthermore, we extended these observations, demonstrating an effect of the mental music line also in the external space selectively for musicians.

Study 4 provided the first evidence that pitch information affects the response latencies and movement kinematics of grasping movements, as well as of symbolic manual gestures. The current findings support the hypothesis of an influence of auditory pitch on grasping movements in the context of size processing. Indeed, the observed effect of auditory pitch on initiation times is in agreement with the stimulus-response compatibility effect so far reported in pitch-size perceptual processing (Gallace & Spence, 2006). The present results show that this compatibility effect extends to more complex motor processing. More critically, grasping kinematic was also influenced by auditory pitch. Indeed, auditory pitch modulated the grip and hand aperture independently from the object size and prior to any interaction with it. Moreover, this effect is robust and automatic, as confirmed by its presence in Experiment 5, where the pitch dimension was no more explicitly processed. This means that auditory pitch was *per se* informative about size in motor planning. The present study offers support to the view that pitch-size natural correspondences might also bias our actions in the environment, supporting the tight link between perception and action. Such link seems also to be strengthened in case of high level of musical practice, which probably promotes a more automatic processing of pitch dimension, as already suggested by previous studies (Koelsch et al., 1999; Rusconi et al., 2006; Lidji et al., 2007). Furthermore, this strong sound-action association in expert musicians seems to be at least partially mediated by their body experience with instrumental practice. Indeed, our results indicated that musicians are faster with high tones with the right hand, compared to the left one, effect that was probably mediated by their experience with the piano keyboard, where high tones are usually played using the right hand.

From a theoretical point of view, results of the first four studies might be interpreted in light of the ATOM theory (A Theory of Magnitude) model, according to which space, time and quantity would operate on a similar magnitude representation with the aim to learn about the environment for acting on it (Walsh, 2003). Interestingly, the model has been recently proposed to include metathetic (i.e., concerned with qualitative variation, Stevens, 1957) dimensions as well, such as auditory pitch (see Bottini & Casasanto, 2013, for a discussion). Furthermore, the finding that the lack of a normal visual experience does not prevent the development of a music mental (vertical) line that the mapping of pitch onto vertical space may be grounded in the everyday bodily experience. For instance, everyone experiences that body resonates differently depending on pitch range (i.e. the resonance of the chest with low pitch versus the head resonance with high pitch). Likewise, the proprioception experience of our larynx position, which produces higher tones and lower tones, rising and lowering respectively, might explain the cross-modal correspondence between pitch and elevation (see Shayan et al., 2011). This interpretation would be in line with the embodied cognition (Lakoff & Johnson, 1999; Fauconnier & Turner, 2002; Gibbs, 2003) which claims in favor of a close link between cognition and bodily experience. Recent investigation tends to explain the origin of cross-modal associations as a reflection of cross-modal correspondences on natural statistics of the environment, including simple co-occurrence of events (i.e. bigger/smaller objects producing lower/higher sounds, respectively) (Spence, 2011; Parise et al., 2014). In line with this interpretation, visual information and audio-visual regularities of the external world would be crucial in mediating cross-modal associations. With our studies we cannot exclude neither corroborate this hypothesis. Nonetheless, we demonstrated that even early blind individuals mentally represent tones in a vertical music line, thus suggesting that vision is not strictly critical in establishing such mapping. We can here speculate that in case of a normal visual development, people mostly rely on the visual system for experiencing statistical regularities of the environment. However, in case of lack of vision, people are still able to learn cross-modal associations between pitch and elevation through other senses, especially thanks to body experience. Congruently with this explanation, early blind individuals are able to generate spatial images (Noordzij, Zuidhoek & Postma, 2006, 2007; Vecchi, 2001) on the bases of tactile, motor, linguistic and also auditory information (Cattaneo et al., 2008; Kaski, 2002). This interpretation would be in line with a generalized spatial representation of pitch shared across different modalities in agreement with others spatial representation of magnitude (see for instance Lourenco & Longo, 2010). Furthermore, a large amount of studies showed that learning, practice and experience have a measurable effect in brain structure and function, altering both the anatomy and the way in which the brain processes information (related to the learned domain or to different domains). In this framework, intense musical practice involves the interaction of several modalities and high cognitive functions leading to behavioral, structural and functional changes (see Herholz & Zatorre, 2012 for a review). Demonstrating consistent differences between musicians and non-musicians, this dissertation also supports and extends this literature. Moreover, by showing an effect of sensori-motor experience due to musical training (Study 1, Study 3, Study 5, Experiment 6) our results support the idea that musical mind is highly embodied, where action and perception are closely interwoven.

From a more practical point of view, by unveiling the impact of auditory pitch in motor control, our study has reasonable implications for both neurorehabilitation of motor disorders and the development of virtual-reality interfaces. For instance, recent studies have shown a benefit of auditory contact cues on the planning and control of grasping movements (Zahariev & Mackenzie, 2003; 2008). Accordingly, auditory pitch might be exploited to facilitate the fulfillment of actions requiring size processing, such as grasping movements, or provide augmented feedback for actions performed in immersive virtual-reality. Moreover, disclosing the effect of pitch on mental representation and the impact of musical expertise, this work contributes to substantiate scientific bases for the use of music in clinical setting. For example, previous studies already demonstrated the benefic effect of the mental music line in neglect patients. Ishihara et al., (2013) demonstrated that auditory cues effectively modulated the direction of attentional bias in neglect patients, in line with previous studies showing transient beneficial spatial effect of auditory cueing (Robertson et al., 1998; Golay et al., 2005). Congruently with this studies, recent findings showed the possibility to reduce spatial neglect through instrument playing (Bernardi et al., 2015; Bodak et al., 2014), but selectively when notes were played in the canonical and natural way from low to high (i.e. following the music mental line).

Study 5 and study 6 focused more on brain regions mediating auditory-motor associations. By means of Transcranical Magnetic Stimulation we aimed to better understand the role of cerebellum and dorsal premotor cortex in auditory perceptual processing and in integrating sounds and movement. Specifically, in study 5 we found that the (right) cerebellum plays a causal role in pitch processing, extending previous studies that showed that discrimination of pitch and melody elicits activation of cerebellar regions (Gaab et al., 2003; Griffiths et al., 1999; Parsons et al., 2009). More generally, the current results provide new evidence in favor of a strong link between the perceptual and the motor systems. Indeed, cerebellum activation has been reported in motor sequencing and motor control tasks (Lewis and Miall, 2003; Doyon, Penhune & Ungerleider, 2003; Penhune and Doyon, 2005; Garraux et al., 2005; Sakai et al., 2004), in the acquisition and integration of sensory information (Bower, 1997) and in purely auditory perceptual task (Petacchi et al., 2005). Future research is needed to better clarify the role of cerebellum in other aspects of auditory processing, such as rhythm or complex melody recognition and discrimination. Moreover, level of expertise in determining the involvement of the cerebellum in auditory functions deserves consideration. In fact, prior studies showed greater engagement of the cerebellum in rhythm perception and synchronization in musicians compared to non-musicians (Chen et al., 2008; Grahn and Brett, 2007). Musicians have been found to detect pitch changes and rhythmic irregularities faster and more accurately than non-musicians (Tervaniemi et al., 2005; Micheyl et al., 2006; Rammsayer and Altenmüller, 2006), an ability that may also depend on different cerebellar involvement. The relation between level of expertise and cerebellar involvement in perceptual functions is an important topic to which brain stimulation may significantly contribute. Finally, our results are important in a clinical perspective helping understanding the impact of cerebellar lesions on sensory and cognitive functions.

Study 6 is the first to demonstrate a causal role for the dPMC in learning and implementation of auditory-motor associations. Our findings show that inhibitory rTMS over dPMC impairs the ability to learn and apply auditory-motor associations, and that this effect is greater when a novel association must be explicitly acquired. The present results contribute to a better understanding of the role of dPMC in auditory-motor integration, suggesting its critical role in learning the mapping between an action and its associated sound, a key function allowing us to speak and to play music. Nicely, this last study also created a link with the first part of the thesis. Indeed, results suggested a stronger contribution of the dorsal premotor cortex in case of auditory-motor associations learning following the mental music line, compared to when sounds were scrambled in an unconventional mapping. In line with this, Amiez et al. (Amiez et al., 2012) analyzed the brain activation during the learning and the execution of conditional visuo-motor responses and reported that the dPMC is the only area modulated by the number of visuo-motor mappings to be acquired, i.e. the dPMC is sensitive to the difficulty of the task. Finally, in this study we proposed that rTMS interfered with learning of the key-to-tone mapping in Experiment 2 because it was unconventional and novel. In addition, the key-to-tone mapping was not ordered in a spatially sequential manner, i.e., adjacent keys did not correspond to adjacent tones. Thus, it might be interesting for future experiments to test an unconventional mapping that is still sequentially ordered, i.e. left-to-right, high-to-low. Furthermore, it is also possible that based on the functional dissociation between dorsal and ventral premotor cortex, the conventional low-to-high mapping may be encoded in the vPMC because more direct and implicit, rather than in the dPMC. Thus, future studies could compare the effects of rTMS over both ventral and dorsal regions.

In summary, the findings discussed so far suggest that pitch dimension strictly interact with our mental representation of the external world and with our body and motor system. Overall, the contribution of the present work is to provide converging data about the role of pitch processing in promoting an efficient interaction with the world around us. Although vision is the prevalent sense that we use to know about the environment in case of normal perceptual systems development, the current data suggest that pitch is informative for motor planning and consequently for the way we interact with objects. This interaction is modulated by life experiences, especially by musical expertise, where the sound-action association is particularly developed. It would be interesting for future studies to shed light on the neural underpinning of cross-modal correspondences. One hypothesis suggests that the inferior parietal lobule (IPL) acts as a generalized magnitude system for action, reflecting the processing of space, time and quantity to be used in sensorimotor transformation (Walsh, 2003). Thus, it represents a good candidate to be a brain area extremely important for cross-modal associations. Interestingly, recent neuroimaging studies have demonstrated an activation of the intraparietal sulcus during both pitch processing (Foster & Zatorre, 2010) and size processing (Pinel et al., 2004). Furthermore, there are evidence suggesting the involvement of a fronto-parietal network in the audiovisual integration guiding action (Rizzolatti and Sinigaglia, 2010). Nonetheless, this is a new field of research (see Bien et al., 2012; Sadaghiani et al., 2009; Spence & Parise, 2012), where lot of questions are still unsolved.

In this context, this thesis project contributes to our knowledge on the neural basis underlying pitch processing and integration between auditory and motor information. Given the complexity of such processes, it might be of wide interest to investigate whether and how other brain regions are also involved. An important challenge for future research is to further identify the neural underpinning underlying associative learning processes. For instance, with respect to the premotor system it would be interesting to better understand the dissociation between ventral and dorsal premotor cortex in auditory-motor association learning. Moreover, we have yet to understand how premotor cortex interacts with other brain areas to form the network that is crucial for both music and speech.

References

Ackermann, H., Mathiak, K., & Riecker, A. (2007). The contribution of the cerebellum to speech production and speech perception: clinical and functional imaging data. *The Cerebellum*, *6*(3), 202-213.

Ackermann, H. (2008). Cerebellar contributions to speech production and speech perception: psycholinguistic and neurobiological perspectives. *Trends in neurosciences*, *31*(6), 265-272.

Alluri, V., Toiviainen, P., Jääskeläinen, I. P., Glerean, E., Sams, M., & Brattico, E. (2012). Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *Neuroimage*, *59*(4), 3677-3689.

Andres, M., Davare, M., Pesenti, M., Olivier, E., & Seron, X. (2004). Number magnitude and grip aperture interaction. *Neuroreport*, *15*, 2773–2777.

Amiez, C., Hadj-Bouziane, F., & Petrides, M. (2012). Response selection versus feedback analysis in conditional visuo-motor learning. *Neuroimage*, *59*(4), 3723-3735.

Antović, M., Bennett, A., & Turner, M. (2013). Running in circles or moving along lines: Conceptualization of musical elements in sighted and blind children. *Musicae scientiae*, *17*(2), 229-245.

Akhtar, N., & Enns, J. T. (1989). Relations between convert orienting and filtering in the development of visual attention. *Journal of Experimental Child Psychology*, *48*, 315-334.

Arasanz, C. P., Staines, W. R., Roy, E. A., & Schweizer, T. A. (2012). The cerebellum and its role in word generation: a cTBS study. *Cortex*, *48*(6), 718-724.

Argyropoulos, G. P., & Muggleton, N. G. (2013). Effects of cerebellar stimulation on processing semantic associations. *The Cerebellum*, *12*(1), 83-96.

Awiszus, F. (2011). Fast estimation of transcranial magnetic stimulation motor threshold: is it safe?. *Brain stimulation*, *4*(1), 58-59.

Baek, M. J., Lee, B. H., Kwon, J. C., Park, J. M., Kang, S.J., Chin J., et al. (2002). Influence of final search direction on tactile line bisection in normal subjects. *Neurology* 58, 1833-1838.

Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., ... & Altenmüller, E. (2006). Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *Neuroimage*, *30*(3), 917-926.

Barbas, H., & Pandya, D. N. (1987). Architecture and frontal cortical connections of the premotor cortex (area 6) in the rhesus monkey. *Journal of comparative neurology*, *256*(2), 211-228.

Baumann, S., Koeneke, S., Schmidt, C. F., Meyer, M., Lutz, K., & Jancke, L. (2007). A network for audio–motor coordination in skilled pianists and non-musicians. *Brain Res*. 1161, 65–78.

Baumann, S., Koeneke, S., Schmidt, C. F., Meyer, M., Lutz, K., & Jancke, L. (2007). A network for audio–motor coordination in skilled pianists and non-musicians. *Brain research*, *1161*, 65-78.

Baumann, O., & Mattingley, J. B. (2010). Scaling of neural responses to visual and auditory motion in the human cerebellum. *The Journal of Neuroscience*, *30*(12), 4489-4495.

Belin, P., McAdams, S., Smith, B., Savel, S., Thivard, L., Samson, S., & Samson, Y. (1998). The functional anatomy of sound intensity discrimination. *The Journal of Neuroscience*, *18*(16), 6388-6394.

Belin, P., McAdams, S., Thivard, L., Smith, B., Savel, S., Zilbovicius, M., ... & Samson, Y. (2002). The neuroanatomical substrate of sound duration discrimination. *Neuropsychologia*, *40*(12), 1956-1964.

Ben-Artzi, E., & Marks, L. E. (1995). Visual-auditory interaction in speeded classification: Role of stimulus difference. *Perception & Psychophysics*, *57*(8), 1151-1162.

Ben-Yehudah, G., Guediche, S., & Fiez, J. A. (2007). Cerebellar contributions to verbal working memory: beyond cognitive theory. *The Cerebellum*, *6*(3), 193-201.

Berlucchi, G., Tassinari, G., & Aglioti, S. (1994). Callosal pathways for simple visuomotor control in man. Rendiconti Lincei, 5(2), 191–201.

Bermudez, P., & Zatorre, R. J. (2005). Conditional associative memory for musical stimuli in nonmusicians: implications for absolute pitch. *The Journal of neuroscience*, *25*(34), 7718-7723.

Bernardis, P., & Gentilucci, M. (2006). Speech and gesture share the same communication system. *Neuropsychologia*, *44*, 178–190.

Bernardi, N. F., Cioffi, M. C., Ronchi, R., Maravita, A., Bricolo, E., Zigiotto, L., ... & Vallar, G. (2015). Improving left spatial neglect through music scale playing. *Journal of neuropsychology*.

Bestmann, S., Baudewig, J., Siebner, H. R., Rothwell, J. C., & Frahm, J. (2005). BOLD MRI responses to repetitive TMS over human dorsal premotor cortex. *Neuroimage*, *28*(1), 22-29.

Bien, N., ten Oever, S., Goebel, R., & Sack, A. T. (2012). The sound of size: Crossmodal binding in pitch-size synesthesia: A combined TMS, EEG and psychophysics study. *NeuroImage, 59,* 663–672.

Bigand, E., & Poulin-Charronnat, B. (2006). Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. *Cognition*, *100*(1), 100-130.

Blakemore, S. J., Frith, C. D., & Wolpert, D. M. (2001). The cerebellum is involved in predicting the sensory consequences of action. *Neuroreport*, *12*(9), 1879-1884.

Bodak, R., Malhotra, P., Bernardi, N. F., Cocchini, G., & Stewart, L. (2014). Reducing chronic visuo-spatial neglect following right hemisphere stroke through instrument playing. *Frontiers in human neuroscience*, *8*, 413.

Bonato, M., Priftis, K., Marenzi, R., & Zorzi, M. (2008). Modulation of hemispatial neglect by directional and numerical cues in the line bisection task. *Neuropsychologia* 46, 426-433.

Booth, J. R., Wood, L., Lu, D., Houk, J. C., & Bitan, T. (2007). The role of the basal ganglia and cerebellum in language processing. *Brain research*, *1133*, 136-144.

Bottini, R., & Casasanto, D. (2013). Space and time in the child's mind: metaphoric or ATOMic? *Frontiers in Psychol*ogy, *4*, 803.

Bottini, R., Crepaldi, D., Casasanto, D., Crollen, V., & Collignon, O. (2015). Space and time in the sighted and blind. *Cognition*, *141*, 67-72.

Bottini, R., Mattioni, S., & Collignon, O. (2016). Early blindness alters the spatial organization of verbal working memory. *Cortex*, *83*, 271-279.

Bower, J. M. (1997). Control of sensory data acquisition. *International review of neurobiology*, *41*, 489-513.

Braaten, R. (1993). Synesthetic correspondence between visual location and auditory pitch in infants. In *34th Annual Meeting of the Psychonomic Society*.

Brown, S., Martinez, M. J., & Parsons, L. M. (2004). Passive music listening spontaneously engages limbic and paralimbic systems. *Neuroreport*, *15*(13), 2033-2037.

Brown, R. M., Chen, J. L., Hollinger, A., Penhune, V. B., Palmer, C., & Zatorre, R. J. (2013). Repetition suppression in auditory–motor regions to pitch and temporal structure in music. *Journal of Cognitive Neuroscience*, *25*(2), 313-328.

Brown, R. M., Zatorre, R. J., & Penhune, V. B. (2015). Expert music performance: cognitive, neural, and developmental bases. *Progress in brain research, 217*, 57-86.

Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*, 1831–1840.

Callan, D. E., Tsytsarev, V., Hanakawa, T., Callan, A. M., Katsuhara, M., Fukuyama, H., & Turner, R. (2006). Song and speech: brain regions involved with perception and covert production. *Neuroimage*, *31*(3), 1327-1342.

Callan, D. E., Kawato, M., Parsons, L., & Turner, R. (2007). Speech and song: the role of the cerebellum. *The Cerebellum*, *6*(4), 321-327.

Casasanto, D., & Bottini, R. (2014). Mirror-reading can reverse the flow of time. Journal of Experimental Psychology: General, 143(2), 473–479.

Castiello, U., Giordano, B. L., Begliomini, C., Ansuini, C., & Grassi, M. (2010). When ears drive hands: the influence of contact sound on reaching to grasp. *PLoS ONE*, *5*, e12240.

Castiello, U., Zucco, G. M., Parma, V., Ansuini, C., & Tirindelli, R. (2006). Cross–modal interactions between olfaction and vision when grasping. *Chemical Senses*, *31*, 665–671.

Castronovo, J., & Seron, X. (2007). Semantic numerical representation in blind subjects: The role of vision in the spatial format of the mental number line. *The Quarterly journal of experimental psychology*, *60*(1), 101-119.

Cattaneo, Z., Vecchi, T., Cornoldi, C., Mammarella, I., Bonino, D., Ricciardi, E., & Pietrini, P. (2008). Imagery and spatial processes in blindness and visual impairment. *Neuroscience & Biobehavioral Reviews*, *32*(8), 1346-1360.

Cattaneo, Z., Fantino, M., Tinti, C., Silvanto, J., & Vecchi, T. (2010). Crossmodal interaction between the mental number line and peripersonal haptic space representation in sighted and blind individuals. *Attention, Perception, & Psychophysics*, *72*(4), 885-890.

Cattaneo, Z., Fantino, M., Tinti, C., Pascual-Leone, A., Silvanto, J., & Vecchi, T. (2011). Spatial biases in peripersonal space in sighted and blind individuals revealed by a haptic line bisection paradigm. *J. Exp. Psychol. Hum. Percept. Perform.* 37(4), 1110-1121.

Cattaneo, Z., Fantino, M., Mancini, F., Mattioli, F., & Vallar, G. (2012a). Listening to numbers affects visual and haptic bisection in healthy individuals and neglect patients. *Neuropsychologia* 50, 913-925.

Cattaneo, Z., Lega, C., Vecchi, T., & Vallar, G. (2012b). Listening to white noise counteracts visual and haptic pseudoneglect. *Perception* 41(11), 1395-98.

Cattaneo, Z., Lega, C., Boehringer, J., Gallucci, M., Girelli, L., Carbon, C. C. (2013). Happiness takes you right: the effect of emotional stimuli on line bisection. *Cogn. Emot.* 28(2), 325-44.

Cattaneo, Z., Renzi, C., Casali, S., Silvanto, J., Vecchi, T., Papagno, C., & D'Angelo, E. (2014). Cerebellar vermis plays a causal role in visual motion discrimination. *Cortex*, *58*, 272-280.

Chandrasekaran, B., & Kraus, N. (2010). The scalp- recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology*, *47*(2), 236-246.

Chen, R., Classen, J., Gerloff, C., Celnik, P., Wassermann, E. M., Hallett, M., & Cohen, L. G. (1997). Depression of motor cortex excitability by low- frequency transcranial magnetic stimulation. *Neurology*, *48*(5), 1398-1403.

Chen, J. L., Zatorre, R. J., & Penhune, V. B. (2006). Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. *Neuroimage*, *32*(4), 1771-1781.

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

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2009). The role of auditory and premotor cortex in sensorimotor transformations. *Annals of the New York Academy of Sciences*, *1169*(1), 15-34.

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.

Chiou, R., & Rich, A. N. (2012). Cross–modality correspondence between pitch and spatial location modulates attentional orienting. *Perception*, *41*, 339–353.

Cho, Y. S., & Proctor, R. W. (2003). Stimulus and response representations underlying orthogonal stimulus-response compatibility effects. *Psychonomic Bulletin & Review*, *10*(1), 45-73.

Cho, Y. S., Bae, G. Y., & Proctor, R. W. (2012). Referential coding contributes to the horizontal SMARC effect. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 726.

Cisek, P., & Kalaska, J. F. (2004). Neural correlates of mental rehearsal in dorsal premotor cortex. *Nature*, *431*(7011), 993-996.

Cohen Kadosh, R., Henik, A., & Walsh, V. (2009). Synaesthesia: Learned or lost? Developmental Science, 12, 484–491.

Collignon, O., Charbonneau, G., Lassonde, M., & Lepore, F. (2009). Early visual deprivation alters multisensory processing in peripersonal space. *Neuropsychologia*, *47*(14), 3236-3243.

Coward, S. W., & Stevens, C. J. (2004). Extracting meaning from sound: Nomic mappings, everyday listening, and perceiving object size from frequency. *Psychological Record*, *54*, 349–364.

Crisinel, A.-S., & Spence, C. (2010). A sweet sound? Exploring implicit associations between basic tastes and pitch. Perception, 39, 417– 425.

Crisinel, A. S., & Spence, C. (2012). A fruity note: crossmodal associations between odors and musical notes. *Chemical Senses*, *37*(2), 151-158.

Crollen, V., Dormal, G., Seron, X., Lepore, F., & Collignon, O. (2013). Embodied numbers: The role of vision in the development of number–space interactions. *cortex*, *49*(1), 276-283.

Crummer, G. C., Walton, J. P., Wayman, J. W., Hantz, E. C., & Frisina, R. D. (1994). Neural processing of musical timbre by musicians, nonmusicians, and musicians possessing absolute pitch. *The Journal of the Acoustical Society of America*, *95*(5), 2720-2727.

D’Angelo, E., & Casali, S. (2013). Seeking a unified framework for cerebellar function and dysfunction: from circuit operations to cognition. *Frontiers in neural circuits*, *6*, 116.

D’Ausilio, A., Altenmüller, E., Olivetti Belardinelli, M., & Lotze, M. (2006). Cross- modal plasticity of the motor cortex while listening to a rehearsed musical piece. *Eur. J. Neurosci.* 24, 955–958.

Davare, M., Andres, M., Cosnard, G., Thonnard, J. L., & Olivier, E. (2006). Dissociating the role of ventral and dorsal premotor cortex in precision grasping. *The Journal of neuroscience*, *26*(8), 2260-2268.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol.* 122, 371-396.

Dehaene-Lambertz, G. (2000). Cerebral specialization for speech and non-speech stimuli in infants. *Journal of Cognitive Neuroscience*, *12*(3), 449-460.de Schotten, M. T., Dell’Acqua, F., Valabregue, R., & Catani, M. (2012). Monkey to human comparative anatomy of the frontal lobe association tracts. *Cortex*, *48*(1), 82-96.

Deike, S., Gaschler-Markefski, B., Brechmann, A., & Scheich, H. (2004). Auditory stream segregation relying on timbre involves left auditory cortex. *Neuroreport*, *15*(9), 1511-1514.

Del Olmo, M. F., Cheeran, B., Koch, G., & Rothwell, J. C. (2007). Role of the cerebellum in externally paced rhythmic finger movements. *Journal of Neurophysiology*, *98*(1), 145-152.

Deroy, O., & Spence, C. (2013). Are we all born synaesthetic? Examining the neonatal synaesthesia hypothesis. *Neuroscience & Biobehavioral Reviews*, *37*(7), 1240-1253.

Deroy, O., Fasiello, I., Hayward, V., & Auvray, M. (2016). Differentiated Audio-Tactile Correspondences in Sighted and Blind Individuals.

Dolscheid, S., Shayan, S., Majid, A., & Casasanto, D. (2013). The Thickness of Musical Pitch Psychophysical Evidence for Linguistic Relativity. *Psychological Science*, *24*(5), 613-621.

Dolscheid, S., Hunnius, S., Casasanto, D., & Majid, A. (2014). Prelinguistic infants are sensitive to space-pitch associations found across cultures. *Psychological Science*, *25*(6), 1256-61.

Doyon, J., Penhune, V., & Ungerleider, L. G. (2003). Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia*, *41*(3), 252-262.

Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005a). Action-effect coupling in pianists. *Psychological Research, 69*(4), 233-241.

Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005b). When hearing turns into playing: Movement induction by auditory stimuli in pianists. *The Quarterly Journal of Experimental Psychology Section A*, *58*(8), 1376-1389.

Drost, U. C., Rieger, M., & Prinz, W. (2007). Instrument specificity in experienced musicians. *The Quarterly Journal of Experimental Psychology*,*60*(4), 527-533.

Dum, R. P., & Strick, P. L. (2005). Frontal lobe inputs to the digit representations of the motor areas on the lateral surface of the hemisphere. *The Journal of neuroscience*, *25*(6), 1375-1386.

Eitan, Z., Ornoy, E., & Granot, R. Y. (2012). Listening in the dark: Congenital and early blindness and cross-domain mappings in music. *Psychomusicology: Music, Mind, and Brain*, *22*(1), 33.

Elliott, D., Roy, E. A., Goodman, D., Carson, R. G., Chua, R., & Maraj, B. K. (1993). Asymmetries in the preparation and control of manual aiming movements. Canadian Journal of Experimental Psychology, 47(3), 570-589.

Elliott, D., & Chua, R. (1996). Manual Asymmetries in Motor Performance. (D. Elliott, & R. Chua, Eds.) CRC Press.

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

Ernst, M. O. (2006). A Bayesian view on multimodal cue integration. In G. Knoblich, I. M. Thornton, M. Grosjean, & M. Shiffrar (Eds.), *Human body perception from the inside out* (pp. 105–131). Oxford, UK: Oxford University Press.

Evans, K. K., & Treisman, A. (2009). Natural cross-modal mappings between visual and auditory features. *Journal of vision*, *10*(1), 6-6.

Fauconnier, Gilles and Turner, Mark. 2002. The Way We Think: Conceptual Blending and the Mind's Hidden Complexities. New York: Basic Books.

Feredoes, E., Tononi, G., & Postle, B. R. (2007). The neural bases of the short-term storage of verbal information are anatomically variable across individuals. *The Journal of Neuroscience*, *27*(41), 11003-11008.

Fias, W., Brysbaert, M., Geypens, F., d'Ydewalle, G. (1996). The importance of magnitude information in numerical processing: evidence from the SNARC effect. *Mathematical Cogn.* 2, 95-110.

Fitzgerald, P. B., Fountain, S., & Daskalakis, Z. J. (2006). A comprehensive review of the effects of rTMS on motor cortical excitability and inhibition. *Clinical Neurophysiology*, *117*(12), 2584-2596.

Flindall, J. W., Doan, J. B., & Gonzalez, C. L. (2014). Manual asymmetries in the kinematics of a reach-to-grasp action. *Laterality: Asymmetries of Body, Brain and Cognition*, *19*(4), 489-507.

Fogassi, L., Gallese, V., Buccino, G., Craighero, L., Fadiga, L., & Rizzolatti, G. (2001). Cortical mechanism for the visual guidance of hand grasping movements in the monkey. *Brain*, *124*(3), 571-586.

Foster, N. E., & Zatorre, R. J. (2010). A role for the intraparietal sulcus in transforming musical pitch information. *Cerebral Cortex*, *20*, 1350–1359.

Frings, M., Maschke, M., Gerwig, M., Diener, H. C., & Timmann, D. (2006). Acquisition of simple auditory and visual sequences in cerebellar patients. *The Cerebellum*, *5*(3), 206-211.

Fuhrman, O., & Boroditsky, L. (2010). Cross-cultural differences in mental representations of time: Evidence from an implicit non-linguistic task. Cognitive Science, 34, 1430–1451.

Gaab, N., Gaser, C., Zaehle, T., Jancke, L., & Schlaug, G. (2003). Functional anatomy of pitch memory—an fMRI study with sparse temporal sampling. *Neuroimage*, *19*(4), 1417-1426.

Gallace, A., & Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size. *Perception & Psychophysics*, *68*, 1191–1203.

Gallese†, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive neuropsychology*, *22*(3-4), 455-479.

Garraux, G., Hallett, M., & Talagala, S. L. (2005). CASL fMRI of subcortico-cortical perfusion changes during memory-guided finger sequences. *Neuroimage*, *25*(1), 122-132.

Gentilucci, M. (2003). Grasp observation influences speech production. *European Journal of Neuroscience*, *17*, 179–184.

Gentilucci, M., Benuzzi, F., Gangitano, M., & Grimaldi, S. (2001). Grasp with hand and mouth: a kinematic study on healthy subjects. *Journal of Neurophysiology*, *86*, 1685–1699.

Gentilucci, M., & Corballis, M. C. (2006). From manual gesture to speech: A gradual transition. *Neuroscience & Biobehavioral Reviews*, *30*, 949–960.

Gentilucci, M., Santunione, P., Roy, A. C., & Stefanini, S. (2004). Execution and observation of bringing a fruit to the mouth affect syllable pronunciation. *European Journal of Neuroscience*, *19*, 190–202.

GeRMain, L., & Lamarre, Y. (1993). Neuronal activity in the motor and premotor cortices before and after learning the associations between auditory stimuli and motor responses. *Brain research*, *611*(1), 175-179.

Gevers, W., Reynvoet, B., & Fias, W. (2003). The mental representation of ordinal sequences is spatially organised. *Cognition* 87, B87-B95.

Gibbs, R. W. (2003). Embodied experience and linguistic meaning. *Brain and language*, *84*(1), 1-15.

Gilbert, A. N., Martin, R., & Kemp, S. E. (1996). Cross-modal correspondence between vision and olfaction: the color of smells. *The American journal of psychology*, 335-351.

Giovannelli, F., Innocenti, I., Rossi, S., Borgheresi, A., Ragazzoni, A., Zaccara, G., ... & Cincotta, M. (2014). Role of the dorsal premotor cortex in rhythmic auditory–motor entrainment: a perturbational approach by rTMS. *Cerebral Cortex*, *24*(4), 1009-1016.

Glasser, M. F., & Rilling, J. K. (2008). DTI tractography of the human brain's language pathways. *Cerebral cortex*, *18*(11), 2471-2482.

Goffaux, V., Hault, B., Michel, C., Vuong, Q. C., & Rossion, B. (2005). The respective role of low and high spatial frequencies in supporting configural and featural processing of faces. *Perception*, *34*, 77-86.

Golay, L., Hauert, C. A., Greber, C., Schnider, A., & Ptak, R. (2005). Dynamic modulation of visual detection by auditory cues in spatial neglect. *Neuropsychologia*, *43*(9), 1258-1265.

Goodale, M. A., & Humphrey, G. K. (1998). The objects of action and perception. *Cognition*, *67*, 181–207.

Grafton, S. T., Fagg, A. H., & Arbib, M. A. (1998). Dorsal premotor cortex and conditional movement selection: a PET functional mapping study. *Journal of Neurophysiology*, *79*(2), 1092-1097.

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of cognitive neuroscience*, *19*(5), 893-906.

Grassi, M., Pastore, M., & Lemaitre, G. (2013). Looking at the world with your ears: How do we get the size of an object from its sound? *Acta Psychologica*, *143*, 96–104.

Grassi, M. (2005). Do we hear size or sound: Balls dropped on plates. *Perception & Psychophysics*, *67*, 274–284.

Griffiths, T. D., Johnsrude, I., Dean, J. L., & Green, G. G. (1999). A common neural substrate for the analysis of pitch and duration pattern in segmented sound?. *Neuroreport*, *10*(18), 3825-3830.

Grol, M. J., de Lange, F. P., Verstraten, F. A., Passingham, R. E., & Toni, I. (2006). Cerebral changes during performance of overlearned arbitrary visuomotor associations. *The Journal of Neuroscience*, *26*(1), 117-125.

Grube, M., Lee, K. H., Griffiths, T. D., Barker, A. T., & Woodruff, P. W. (2010). Transcranial magnetic theta-burst stimulation of the human cerebellum distinguishes absolute, duration-based from relative, beat-based perception of subsecond time intervals. *Frontiers in psychology*, *1*, 171.

Halpern, A. R., Zatorre, R. J., Bouffard, M., & Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia*, *42*(9), 1281-1292.

Halsband, U., & Freund, H. J. (1990). Premotor cortex and conditional motor learning in man. *Brain*, *113*(1), 207-222.

Hardwick, R. M., Lesage, E., & Miall, R. C. (2014). Cerebellar transcranial magnetic stimulation: the role of coil geometry and tissue depth. *Brain stimulation*, *7*(5), 643-649.

Haueisen, J., & Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience*, *13*(6), 786-792.

He, Y., Zang, Y., Jiang, T., Liang, M., & Gong, G. (2004, September). Detecting functional connectivity of the cerebellum using low frequency fluctuations (LFFs). In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 907-915). Springer Berlin Heidelberg.

Heinen, K., Jolij, J., & Lamme, V. A. (2005). Figure–ground segregation requires two distinct periods of activity in V1: a transcranial magnetic stimulation study. *Neuroreport*, *16*(13), 1483-1487.

Herholz, S. C., & Zatorre, R. J. (2012). Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron*, *76*(3), 486-502.

Herholz, S. C., Coffey, E. B., Pantev, C., & Zatorre, R. J. (2015). Dissociation of Neural Networks for Predisposition and for Training-Related Plasticity in Auditory-Motor Learning. *Cerebral Cortex*, bhv138.

Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in cognitive sciences*, *4*(4), 131-138.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, *92*(1), 67-99.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393-402.

Hikosaka, O., Nakamura, K., Sakai, K., & Nakahara, H. (2002). Central mechanisms of motor skill learning. *Current opinion in neurobiology*, *12*(2), 217-222.

Hoffmann, J., Sebald, A., & Stöcker, C. (2001). Irrelevant response effects improve serial learning in serial reaction time tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(2), 470.

Holcomb, H. H., Medoff, D. R., Caudill, P. J., Zhao, Z., Lahti, A. C., Dannals, R. F., & Tamminga, C. A. (1998). Cerebral blood flow relationships associated with a difficult tone recognition task in trained normal volunteers. *Cerebral Cortex*, *8*(6), 534-542.

Hommel, B. (2009). Action control according to TEC (theory of event coding). *Psychological Research PRPF*, *73*(4), 512-526.

Hoshi, E., & Tanji, J. (2006). Differential involvement of neurons in the dorsal and ventral premotor cortex during processing of visual signals for action planning. *Journal of neurophysiology*, *95*(6), 3596-3616.

Hoshi, E., & Tanji, J. (2007). Distinctions between dorsal and ventral premotor areas: anatomical connectivity and functional properties. *Current opinion in neurobiology*, *17*(2), 234-242.

Iordanescu, L., Grabowecky, M., & Suzuki, S. (2013). Action enhances auditory but not visual temporal sensitivity. *Psychon. Bull. Rev.* 20, 108–114.

Ishihara, M., Revol, P., Jacquin-Courtois, S., Mayet, R., Rode, G., Boisson, D., ... & Rossetti, Y. (2013). Tonal cues modulate line bisection performance: preliminary evidence for a new rehabilitation prospect?. *Frontiers in psychology*, *4*, 704.

Ivry, R. B., & Keele, S. W. (1989). Timing functions of the cerebellum. *Journal of cognitive neuroscience*, *1*(2), 136-152.

Ivry, R. B., & Diener, H. C. (1991). Impaired velocity perception in patients with lesions of the cerebellum. *Journal of Cognitive Neuroscience*, *3*(4), 355-366.

Ivry, R. B., & Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends in cognitive sciences*, *12*(7), 273-280.

Jansen, A., Flöel, A., Van Randenborgh, J., Konrad, C., Rotte, M., Förster, A. F., ... & Knecht, S. (2005). Crossed cerebro–cerebellar language dominance. *Human brain mapping*, *24*(3), 165-172.

Jeannerod, M., Arbib, M. A., Rizzolatti, G., & Sakata, H. (1995). Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends in Neurosciences*, *18*, 314–320.

Jewell, G., & McCourt, M. E. (2000). Pseudoneglect: a review and meta-analysis of performance factors in line bisection tasks. *Neuropsychologia*,*38*(1), 93-110.

Jokisch, D., Troje, N. F., Koch, B., Schwarz, M., & Daum, I. (2005). Differential involvement of the cerebellum in biological and coherent motion perception. *European Journal of Neuroscience*, *21*(12), 3439-3446.

Kadosh, R. C., Henik, A., & Walsh, V. (2007). Small is bright and big is dark in synaesthesia. Current Biology, 17, R834–R835.

Kaski, D. (2002). Revision: Is visual perception a requisite for visual imagery? Perception, 31(6), 717–731.

Kawashima, R., Okuda, J., Umetsu, A., Sugiura, M., Inoue, K., Suzuki, K., ... & Yanagawa, I. (2000). Human cerebellum plays an important role in memory-timed finger movement: an fMRI study. *Journal of neurophysiology*, *83*(2), 1079-1087.

Keysers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L., & Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental brain research*, *153*(4), 628-636.

Keller, P. E., & Koch, I. (2006). The planning and execution of short auditory sequences. *Psychonomic Bulletin & Review, 13*(4), 711-716.

Keller, P. E., & Koch, I. (2008). Action planning in sequential skills: Relations to music performance. *The Quarterly Journal of Experimental Psychology, 61*(2), 275-291.

Kishon-Rabin, L., Amir, O., Vexler, Y., & Zaltz, Y. (2001). Pitch discrimination: Are professional musicians better than non-musicians?. *Journal of basic and clinical physiology and pharmacology*, *12*(2), 125-144.

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

Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science, 297,* 846–848.

Kurata, K., & Hoffman, D. S. (1994). Differential effects of muscimol microinjection into dorsal and ventral aspects of the premotor cortex of monkeys. *Journal of neurophysiology*, *71*(3), 1151-1164.

Laeng, B., Buchtel, H. A., & Butter, C. M. (1996). Tactile rod bisection: Hemispheric activation and sex differences. *Neuropsychologia* 34, 1115-1121.

Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *The journal of neuroscience, 27*(2), 308-314.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. Basic books.

Lee, T. M., Liu, H. L., Hung, K. N., Pu, J., Ng, Y. B., Mak, A. K., ... & Chan, C. C. (2005). The cerebellum's involvement in the judgment of spatial orientation: a functional magnetic resonance imaging study. *Neuropsychologia*, *43*(13), 1870-1877.

Lega, C., Cattaneo, Z., Merabet, L. B., Vecchi, T., & Cucchi, S. (2014). The effect of musical expertise on the representation of space. *Frontiers in Human Neuroscience, 8,* 250.

Lewis, P. A., & Miall, R. C. (2003). Brain activation patterns during measurement of sub- and supra-second intervals. *Neuropsychologia*, *41*(12), 1583-1592.

Lidji, P., Kolinsky, R., Lochy, A., & Morais, J. (2007). Spatial associations for musical stimuli: a piano in the head?. *Journal of Experimental Psychology: Human Perception and Performance, 33*(5), 1189-207.

Lindemann, O., Abolafia, J. M., Girardi, G., & Bekkering, H. (2007). Getting a grip on numbers: numerical magnitude priming in object grasping. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 1400–1409.

Lotze, M., Scheler, G., Tan, H. R., Braun, C., & Birbaumer, N. (2003). The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *Neuroimage*, *20*(3), 1817-1829.

Linkovski, O., Akiva-Kabiri, L., Gertner, L., & Henik, A. (2012). Is it for real? Evaluating authenticity of musical pitch-space synesthesia. *Cognitive Processing*, *13*(1), 247-251.

Lourenco, S. F., & Longo, M. R. (2010). General magnitude representation in human infants. *Psychological Science*.

Ludwig, V. U., Adachi, I., & Matzuzawa, T. (2011). Visuoauditory mappings between high luminance and high pitch are shared by chimpanzees (*Pan troglodytes*) and humans. *Proceedings of the National Academy of Sciences, 108,* 20661–20665.

Maeda, F., Kanai, R., & Shimojo, S. (2004). Changing pitch induced visual motion illusion. *Current Biology*, *14*(23), R990-R991.

Maes, P. J., & Leman, M. (2013). The influence of body movements on children’s perception of music with an ambiguous expressive character. *PloS one*, *8*(1), e54682.

Maes, P. J., Leman, M., Palmer, C., & Wanderley, M. (2014). Action-based effects on music perception. *Frontiers in psychology*, *4*, 1008.

Maffei, C., Soria, G., Prats-Galino, A., Catani, M. (2015). Imaging white-matter pathways of the auditory system with diffusion imaging tractography. *The Human Auditory System: Fundamental Organization and Clinical Disorders, 129,* 277

Manto, M., Bower, J. M., Conforto, A. B., Delgado-García, J. M., da Guarda, S. N. F., Gerwig, M., ... & Molinari, M. (2012). Consensus paper: roles of the cerebellum in motor control—the diversity of ideas on cerebellar involvement in movement. *The Cerebellum*, *11*(2), 457-487.

Marks, L. (1978). The unity of the senses: Interrelations among the modalities. New York: Academic Press.

Marks, L. E. (2004). Cross–modal interactions in speeded classification. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of mutisensory processes* (pp. 85–105). Cambridge, MA: MIT Press.

Martino, G., & Marks, L. E. (1999). Perceptual and linguistic interactions in speeded classification: Tests of the semantic coding hypothesis. *Perception*,*28*(7), 903-923.

Martino, G., & Marks, L. E. (2000). Cross-modal interaction between vision and touch: The role of synesthetic correspondence. Perception, 29, 745–754.

Martino, G., & Marks, L. E. (2001). Synesthesia: Strong and weak. Current Directions in Psychological Science, 10, 61–65.

McCourt, M. E., & Olafson, C. (1997). Cognitive and perceptual influences on visual line bisection: Psychophysical and chronometric analyses of pseudoneglect. *Neuropsychologia*, *35*(3), 369-380.

McCourt, M. E., Garlinghouse, M., & Reuter-Lorenz, P. A. (2005). Unilateral visual cueing and asymmetric line geometry share a common attentional origin in the modulation of pseudoneglect. *Cortex*, *41*(4), 499-511.

Meehan, S. K., Zabukovec, J. R., Dao, E., Cheung, K. L., Linsdell, M. A., & Boyd, L. A. (2013). One hertz repetitive transcranial magnetic stimulation over dorsal premotor cortex enhances offline motor memory consolidation for sequence- specific implicit learning. *European Journal of Neuroscience*, *38*(7), 3071-3079.

Mendonça, C., Campos, G., Dias, P., & Santos, J. A. (2013). Learning auditory space: Generalization and long-term effects. *PloS one*, *8*(10), e77900.

Mennemeier, M., Wertman, E., Heilman, K. M. (1992). Neglect of near peripersonal space: evidence for multidirectional attentional systems in humans. *Brain*, 115:37±50.

Menon, V., Levitin, D. J., Smith, B. K., Lembke, A., Krasnow, B. D., Glazer, D., ... & McAdams, S. (2002). Neural correlates of timbre change in harmonic sounds. *Neuroimage*, *17*(4), 1742-1754.

Miall, R. C. (1997). Sequences of sensory predictions. *Behavioral and Brain Sciences*, *20*(02), 258-259.

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing research, 219*(1), 36-47.

Mieschke, P. E., Elliott, D., Helsen, W. F., Carson, R. G., & Coull, J. A. (2001). Manual asymmetries in the preparation and control of goal-directed movements. Brain and Cognition, 45, 129-140.

Mikumo, M. (1998). Encoding strategies for pitch information. *Japanese Psychological Monographs,* 27.

Miniussi, C., Harris, J. A., & Ruzzoli, M. (2013). Modelling non-invasive brain stimulation in cognitive neuroscience. *Neuroscience & Biobehavioral Reviews*, *37*(8), 1702-1712.

Molinari, M., Chiricozzi, F. R., Clausi, S., Tedesco, A. M., De Lisa, M., & Leggio, M. G. (2008). Cerebellum and detection of sequences, from perception to cognition. *The Cerebellum*, *7*(4), 611-615.

Mondloch, C. J., & Maurer, D. (2004). Do small white balls squeak? Pitch-object correspondences in young children. *Cognitive, Affective, & Behavioral Neuroscience*, *4*(2), 133-136.

Monegato, M., Cattaneo, Z., Pece, A., & Vecchi, T. (2007). Comparing the effects of congenital and late visual impairments on visuospatial mental abilities. *Journal of Visual Impairment & Blindness*, *101*(5), 278.

Morton, E. S. (1977). On the occurrence and significance of motivation–structural rules in some bird and mammal sounds. *American Naturalist, 111*, 855–869.

Müller, D., & Schwarz, W. (2007). Is there an internal association of numbers to hands? The task set influences the nature of the SNARC effect. *Memory & Cognition*, *35*(5), 1151-1161.

Nager, W., Kohlmetz, C., Altenmüller, E., Rodriguez-Fornells, A., & Münte, T. F. (2003). The fate of sounds in conductors' brains: an ERP study. *Cognitive Brain Research*, *17*(1), 83-93.

Nava, E., Grassi, M., & Turati, C. (2015). Audio-visual, visuo-tactile and audio-tactile correspondences in preschoolers. *Multisensory Research*, 22134808-00002493.

Nishimura, A., & Yokosawa, K. (2009). Effects of laterality and pitch height of an auditory accessory stimulus on horizontal response selection: The Simon effect and the SMARC effect. *Psychonomic Bulletin & Review*, *16*(4), 666-670.

Noordzij, M. L., Zuidhoek, S., & Postma, A. (2006). The influence of visual experience on the ability to form spatial mental models based on route and survey descriptions. *Cognition*, *100*(2), 321-342.

Noordzij, M. L., Zuidhoek, S., & Postma, A. (2007). The influence of visual experience on visual and spatial imagery. Perception, 36(1), 101–112

Ohala, J. J. (1983). Cross–language use of pitch: an ethological view. *Phonetica*, *40*, 1–18.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97-113.

Oliver, R., Opavsky, R., Vyslouzil, M., Greenwood, R., & Rothwell, J. C. (2011). The role of the cerebellum in 'real'and 'imaginary'line bisection explored with 1- Hz repetitive transcranial magnetic stimulation. *European Journal of Neuroscience*, *33*(9), 1724-1732.

Oliveri, M., Koch, G., Torriero, S., & Caltagirone, C. (2005). Increased facilitation of the primary motor cortex following 1Hz repetitive transcranial magnetic stimulation of the contralateral cerebellum in normal humans. *Neuroscience letters*, *376*(3), 188-193.

Ortu, E., Ruge, D., Deriu, F., & Rothwell, J. C. (2009). Theta Burst Stimulation over the human primary motor cortex modulates neural processes involved in movement preparation. *Clinical Neurophysiology*, *120*(6), 1195-1203.

Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., & Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport*, *12*(1), 169-174.

Parise, C. V., Knorre, K., & Ernst, M. O. (2014). Natural auditory scene statistics shapes human spatial hearing. *Proceedings of the National Academy of Sciences*, *111*, 6104–6108.

Parise, C. V., & Spence, C. (2008). Synesthetic congruency modulates the temporal ventriloquism effect. *Neuroscience letters*, *442*, 257–261.

Parise, C. V., & Spence, C. (2009). 'When birds of a feather flock together': Synesthetic correspondences modulate audiovisual integration in non–synesthetes. *PLoS ONE*, *4*, e5664.

Parise, C. V., & Spence, C. (2012). Audiovisual crossmodal correspondences and sound symbolism: a study using the implicit association test. *Experimental Brain Research*, *220*, 319–333.

Parise, C. V., & Spence, C. (2013). Audiovisual crossmodal correspondences in the general population. In J. Simner, & E. Hubbard (Eds.), *The Oxford handbook of synesthesia* (pp. 790-815). Oxford, UK: Oxford University Press.

Parise, C. V., Harrar, V., Ernst, M. O., & Spence, C. (2013). Cross-correlation between auditory and visual signals promotes multisensory integration. *Multisensory research*, *26*(3), 307-316.

Parkin, B. L., Ekhtiari, H., & Walsh, V. F. (2015). Non-invasive human brain stimulation in cognitive neuroscience: a primer. *Neuron*, *87*(5), 932-945.

Parkinson, C., Kohler, P. J., Sievers, B., & Wheatley, T. (2012). Associations between auditory pitch and visual elevation do not depend on language: Evidence from a remote population. *Perception*, *41*(7), 854-861.

Parsons, L. M. (2001). Exploring the functional neuroanatomy of music performance, perception, and comprehension. *Annals of the New York Academy of Sciences*, *930*(1), 211-231.

Parsons, M. W., Harrington, D. L., & Rao, S. M. (2005). Distinct neural systems underlie learning visuomotor and spatial representations of motor skills. *Human brain mapping*, *24*(3), 229-247.

Parsons, L. M., Petacchi, A., Schmahmann, J. D., & Bower, J. M. (2009). Pitch discrimination in cerebellar patients: evidence for a sensory deficit. *Brain research*, *1303*, 84-96.

Pascual-Leone, A., Walsh, V., & Rothwell, J. (2000). Transcranial magnetic stimulation in cognitiveneuroscience–virtual lesion, chronometry, and functional connectivity. *Current opinion in neurobiology*, *10*(2), 232-237.

Pascual-Leone, A., & Walsh, V. (2001). Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science*, *292*(5516), 510-512.

Pasqualotto, A., Taya, S., & Proulx, M. J. (2014). Sensory deprivation: visual experience alters the mental number line. *Behavioural brain research*, *261*, 110-113.

Passingham, R. E. (1989). Premotor cortex and the retrieval of movement. *Brain, behavior and evolution*, *33*(2-3), 189-192.

Petacchi, A., Laird, A. R., Fox, P. T., & Bower, J. M. (2005). Cerebellum and auditory function: An ALE meta- analysis of functional neuroimaging studies. *Human brain mapping*, *25*(1), 118-128.

Petacchi, A., Kaernbach, C., Ratnam, R., & Bower, J. M. (2011). Increased activation of the human cerebellum during pitch discrimination: a positron emission tomography (PET) study. *Hearing research*, *282*(1), 35-48.

Patchay, S., Castiello, U., & Haggard, P. (2003). A cross–modal interference effect in grasping objects. *Psychonomic Bulletin & Review*, *10*, 924–931.

Penhune, V. B., & Doyon, J. (2005). Cerebellum and M1 interaction during early learning of timed motor sequences. *Neuroimage*, *26*(3), 801-812.

Peretz, I., & Zatorre, R. J. (2005). Brain organization for music processing. *Annu. Rev. Psychol.*, *56*, 89-114.

Petrides, M. (1985). Deficits in non-spatial conditional associative learning after periarcuate lesions in the monkey. *Behavioural brain research*, *16*(2-3), 95-101.

Petrides, M. (2005). Lateral prefrontal cortex: architectonic and functional organization. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *360*(1456), 781-795.

Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: auditory encoding and rhythmic movement. *Cognition* 105, 533–546.

Picard, N., & Strick, P. L. (2001). Imaging the premotor areas. *Current opinion in neurobiology*, *11*(6), 663-672.

Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, *41*, 983–993.

Pitt, M. A. (1994). Perception of pitch and timbre by musically trained and untrained listeners. *Journal of experimental psychology: human perception and performance*, *20*(5), 976.

Pitteri, M., Marchetti, M., Priftis, K., & Grassi, M. (2015). Naturally together: pitch-height and brightness as coupled factors for eliciting the SMARC effect in non-musicians. *Psychological research*, 1-12.

Plack, C. J. (2005). *The Sense of Hearing.* New Jersey: Lawrence Erlbaum.

Pollok, B., Rothkegel, H., Schnitzler, A., Paulus, W., & Lang, N. (2008). The effect of rTMS over left and right dorsolateral premotor cortex on movement timing of either hand. *European Journal of Neuroscience*, *27*(3), 757-764.

Postma, A., Zuidhoek, S., Noordzij, M. L., & Kappers, A. M. (2007). Differences between early-blind, late-blind, and blindfolded-sighted people in haptic spatial-configuration learning and resulting memory traces. *Perception*, *36*(8), 1253-1265.

Pratt, C. C. (1930). The spatial character of high and low tones. *Journal of Experimental Psychology*, *13*(3), 278.

Proctor, R. W., & Vu, K. P. L. (2002). Mixing location-irrelevant and location-relevant trials: Influence of stimulus mode on spatial compatibility effects. *Memory & Cognition*, *30*(2), 281-293.

Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological bulletin*, *132*(3), 416.

Rammsayer, T., & Altenmüller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Perception: An Interdisciplinary Journal*, *24*(1), 37-48.

Reiterer, S. M., Erb, M., Droll, C. D., Anders, S., Ethofer, T., Grodd, W., & Wildgruber, D. (2005). Impact of task difficulty on lateralization of pitch and duration discrimination. *Neuroreport*, *16*(3), 239-242.

Reiterer, S., Erb, M., Grodd, W., & Wildgruber, D. (2008). Cerebral processing of timbre and loudness: fMRI evidence for a contribution of Broca's area to basic auditory discrimination. *Brain Imaging and Behavior*, *2*(1), 1-10.

Renzi, C., Vecchi, T., D'Angelo, E., Silvanto, J., & Cattaneo, Z. (2014). Phosphene induction by cerebellar transcranial magnetic stimulation. *Clinical Neurophysiology*, *125*(10), 2132-2133.

Repp, B. H., & Knoblich, G. (2007). Action can affect auditory perception. *Psychological Science*, *18*(1), 6-7.

Repp, B. H., and Knoblich, G. (2009). Performed or observed keyboard actions affect pianists' judgements of relative pitch. *Q. J. Exp. Psychol.* 62, 2156–2170.

Rinaldi, L., Vecchi, T., Fantino, M., Merabet, L. B., & Cattaneo, Z. (2015). The effect of hand movements on numerical bisection judgments in early blind and sighted individuals. *Cortex*, *71*, 76-84.

Rinaldi, L., Lega, C., Cattaneo, Z., Girelli, L., & Bernardi, N. F. (2016). Grasping the sound: Auditory pitch influences size processing in motor planning. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(1), 11-22.

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto–frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews Neuroscience*, *11*, 264–274.

Robertson, I. H., Mattingley, J. B., Rorden, C., & Driver, J. (1998). Phasic alerting of neglect patients overcomes their spatial deficit in visual awareness. *Nature*, *395*(6698), 169-172.

Röder, B., Kusmierek, A., Spence, C., & Schicke, T. (2007). Developmental vision determines the reference frame for the multisensory control of action. *Proceedings of the National Academy of Sciences*, *104*(11), 4753-4758.

Roffler, S. K., & Butler, R. A. (1968). Localization of tonal stimuli in the vertical plane. *The Journal of the Acoustical Society of America*, *43*(6), 1260-1266.

Rossi, S., Hallett, M., Rossini, P. M., Pascual-Leone, A., & Safety of TMS Consensus Group. (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical neurophysiology*, *120*(12), 2008-2039.

Roth, M. J., Synofzik, M., & Lindner, A. (2013). The cerebellum optimizes perceptual predictions about external sensory events. *Current Biology*, *23*(10), 930-935.

Rusconi, E., Kwan, B., Giordano, B. L., Umilta', C., & Butterworth, B. (2006). Spatial representation of pitch height: the SMARC effect. *Cognition*, *99*(2), 113-129.

Sack, A. T., Kadosh, R. C., Schuhmann, T., Moerel, M., Walsh, V., & Goebel, R. (2009). Optimizing functional accuracy of TMS in cognitive studies: a comparison of methods. *Journal of Cognitive Neuroscience*, *21*(2), 207-221.

Sadaghiani, S., Maier, J. X., & Noppeney, U. (2009). Natural, metaphoric, and linguistic auditory direction signals have distinct influences on visual motion processing. *The Journal of Neuroscience*, *29*(20), 6490-6499.

Säfström, D., & Edin, B. B. (2006). Acquiring and adapting a novel audiomotor map in human grasping. *Experimental Brain Research*, *173*, 487-497.

Sakai, K., Hikosaka, O., & Nakamura, K. (2004). Emergence of rhythm during motor learning. *Trends in cognitive sciences*, *8*(12), 547-553.

Samson, S. (2003). Neuropsychological studies of musical timbre. *Annals of the New York Academy of Sciences*, *999*(1), 144-151.

Santiago, J., Lupiãnez, J., Pérez, E., & Funes, M. J. (2007). Time, also flies from left to right. Psychonomic Bulletin & Review, 14, 512–516.

Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M. S., ... & Huber, W. (2008). Ventral and dorsal pathways for language. *Proceedings of the national academy of Sciences*, *105*(46), 18035-18040.

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, *41*(3), 341-349.

Schutter, D. J., & van Honk, J. (2009). The cerebellum in emotion regulation: a repetitive transcranial magnetic stimulation study. *The Cerebellum*, *8*(1), 28-34.

Sedda, A., Monaco, S., Bottini, G., & Goodale, M. A. (2011). Integration of visual and auditory information for hand actions: preliminary evidence for the contribution of natural sounds to grasping. *Experimental Brain Research*, *209*, 365–374.

Shayan, S., Ozturk, O., & Sicoli, M. A. (2011). The thickness of pitch: Crossmodal metaphors in Farsi, Turkish, and Zapotec. *The Senses and Society*, *6*(1), 96-105.

Shiffrin, R. M., & Schneider, W. (1984). Automatic and controlled processing revisited.

Shintel, H., Nusbaum, H. C., & Okrent, A. (2006). Analog acoustic expression in speech communication. *Journal of Memory and Language*, *55*, 167–177.

Silvanto, J., & Pascual-Leone, A. (2012). Why the assessment of causality in brain–behavior relations requires brain stimulation. *Journal of cognitive neuroscience*, *24*(4), 775-777.

Simner, J., & Ludwig, V. (2009). What colour does that feel? Cross-modal correspondences from touch to colour. In *Third International Conference of Synaesthesia and Art, Granada, Spain, April*.

Sparing, R., Buelte, D., Meister, I. G., Pauš, T., & Fink, G. R. (2008). Transcranial magnetic stimulation and the challenge of coil placement: a comparison of conventional and stereotaxic neuronavigational strategies. *Human brain mapping*, *29*(1), 82-96.

Spence, C., & Zampini, M. (2006). Auditory contributions to multisensory product perception. Acta Acustica united with Acustica, 92, 1009–1025

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*, 971–995.

Spence, C., & Deroy, O. (2012). Crossmodal correspondences: Innate or learned?. *i-Perception*, *3*(5), 316-318.

Spence, C., & Parise, C. V. (2012). The cognitive neuroscience of crossmodal correspondences. *i-Perception*, *3*(7), 410-412.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, *64*, 153–181.

Stewart, L., Henson, R., Kampe, K., Walsh, V., Turner, R., & Frith, U. (2003). Brain changes after learning to read and play music. *Neuroimage*, *20*(1), 71-83.

Stewart, L., Walsh, V., & Frith, U. (2004). Reading music modifies spatial mapping in pianists. *Perception & psychophysics*, *66*(2), 183-195.

Stewart, L., Verdonschot, R. G., Nasralla, P., & Lanipekun, J. (2013). Action–perception coupling in pianists: Learned mappings or spatial musical association of response codes (SMARC) effect?. *The Quarterly Journal of Experimental Psychology*, *66*(1), 37-50.

Stöcker, C., Sebald, A., & Hoffmann, J. (2003). The influence of response–effect compatibility in a serial reaction time task. *The Quarterly Journal of Experimental Psychology: Section A, 56*(4), 685-703.

Stoodley, C. J. (2012). The cerebellum and cognition: evidence from functional imaging studies. *The Cerebellum*, *11*(2), 352-365.

Strick, P. L., Dum, R. P., & Fiez, J. A. (2009). Cerebellum and nonmotor function. *Annual review of neuroscience*, *32*, 413-434.

Struiksma, M. E., Noordzij, M. L., & Postma, A. (2011). Reference frame preferences in haptics differ for the blind and sighted in the horizontal but not in the vertical plane. *Perception*, *40*(6), 725-738.

Stumpf, K. (1883). Tonpsychologie I [Psychology of the tone]. Leipzig: Hirzel.

Szűcs, D., & Csépe, V. (2005). The parietal distance effect appears in both the congenitally blind and matched sighted controls in an acoustic number comparison task. *Neuroscience letters, 384*(1), 11-16.

Tamietto, M., Latini Corazzini, L., Pia, L., Zettin, M., Gionco, M., Geminiani, G. (2005). Effects of emotional face cueing on line bisection in neglect: a single case study. *Neurocase*. 11(6), 399-404

Taylor, J. E. T., & Witt, J. K. (2015). Listening to music primes space: pianists, but not novices, simulate heard actions. *Psychological Research*, *79*(2), 175-182.

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *The Journal of Neuroscience*, *31*(10), 3805-3812.

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. *Experimental brain research*, *161*(1), 1-10.

Théoret, H., Haque, J., & Pascual-Leone, A. (2001). Increased variability of paced finger tapping accuracy following repetitive magnetic stimulation of the cerebellum in humans. *Neuroscience letters*, *306*(1), 29-32.

Thier, P., Haarmeier, T., Treue, S., & Barash, S. (1999). Absence of a common functional denominator of visual disturbances in cerebellar disease. *Brain*, *122*(11), 2133-2146.

Timmers, R., & Shen, L. (2016). Representation of Pitch in Horizontal Space and Its Dependence on Musical and Instrumental Experience. *Psychomusicology: Music, Mind and Brain, 26*(2).

Tölgyesi, B., & Evers, S. (2014). The impact of cerebellar disorders on musical ability. *Journal of the neurological sciences, 343*(1), 76-81.

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. *Cognitive Theory, 3*, 200-239.

Torriero, S., Oliveri, M., Koch, G., Caltagirone, C., & Petrosini, L. (2004). Interference of left and right cerebellar rTMS with procedural learning. *Journal of Cognitive Neuroscience, 16*(9), 1605-1611.

Trimarchi, P. D., & Luzzatti, C. (2011). Implicit chord processing and motor representation in pianists. *Psychological Research, 75*(2), 122-128.

Ulrich, R., & Maienborn, C. (2010). Left–right coding of past and future in language: The mental timeline during sentence processing. Cognition, 117(2), 126–138.

van Zuijen, T. L., Sussman, E., Winkler, I., Näätänen, R., & Tervaniemi, M. (2005). Auditory organization of sound sequences by a temporal or numerical regularity—a mismatch negativity study comparing musicians and non-musicians. *Cognitive Brain Research, 23*(2), 270-276.

Vecchi, T. (2001). Visuo-spatial processing in congenitally blind people: Is there a gender-related preference? Personality and Individual Differences, 30, 1361–1370.

Vu, K. P. L., Minakata, K., Ngo, M. K. (2013). Influence of auditory and audiovisual stimuli on the right-left prevalence effect. *Psychol. Res.,* DOI 10.1007/s00426-013-0518-4.

Walker, A. R. (1985). Mental imagery and musical concepts: Some evidence from the congenitally blind. *Bulletin of the Council for Research in Music Education*, 229-237.

Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2009). Preverbal Infants' Sensitivity to Synaesthetic Cross-Modality Correspondences. *Psychological Science, 21*(1), 21-5.

Walker, P., & Walker, L. (2012). Size–brightness correspondence: Crosstalk and congruity among dimensions of connotative meaning. Attention, Perception, & Psychophysics, 74, 1226–1240.

Walsh, V., & Cowey, A. (2000). Transcranial magnetic stimulation and cognitive neuroscience. *Nature Reviews Neuroscience, 1*(1), 73-80.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in cognitive sciences*, *7*(11), 483-488.

Warren, J. E., Wise, R. J., & Warren, J. D. (2005). Sounds do-able: auditory–motor transformations and the posterior temporal plane. *Trends in Neurosciences*, *28*(12), 636-643.

Weis, T., Estner, B., & Lachmann, T. (2016a). When speech enhances Spatial Musical Association of Response Codes: Joint spatial associations of pitch and timbre in nonmusicians. *The Quarterly Journal of Experimental Psychology*, *69*(9), 1687-1700.

Weis, T., Estner, B., van Leeuwen, C., Lachmann, T., & Lachmann, T. (2016b). SNARSC meets SPARC: Automaticity and Interdependency in Compatibility Effects. *The Quarterly Journal of Experimental Psychology*, 1-37.

Welch, G. F. (1991). Visual metaphors for sound: A study of mental imagery, language and pitch perception in the congenitally blind. *Canadian Journal of Research in Music Education*, *33*(1), 215-222.

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.

Winter, B., Perlman, M., & Matlock, T. (2013). Using space to talk and gesture about numbers: Evidence from the TV News Archive. *Gesture*, *13*, 377–408.

Wise, S. P., Di Pellegrino, G., & Boussaoud, D. (1996). The premotor cortex and nonstandard sensorimotor mapping. *Canadian journal of physiology and pharmacology*, *74*(4), 469-482.

Wood, G., Nuerk, H. C., & Willmes, K. (2006). Crossed Hands and the Snarc Effect: Afailure to Replicate Dehaene, Bossini and Giraux (1993). *Cortex*, *42*(8), 1069-1079.

Yau, J. M., Olenczak, J. B., Dammann, J. F., & Bensmaia, S. J. (2009). Temporal frequency channels are linked across audition and touch. Current Biology, 19, 561–566.

Zahariev, M. A., & Mackenzie, C. L. (2003). Auditory, graphical and haptic contact cues for a reach, grasp, and place task in an augmented environment. *Proceedings of the 5th International Conference on Multimodal Interfaces*, 273–276.

Zahariev, M. A., & Mackenzie, C. L. (2008). Auditory contact cues improve performance when grasping augmented and virtual objects with a tool. *Experimental Brain Research, 186,* 619–627.

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends in cognitive sciences*, *6*(1), 37-46.

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*(7), 547-558.

Acknowledgments

My best acknowledgment is for my supervisor Zaira Cattaneo, for initiating me at research and supporting me along the entire way. My great thanks for teaching me and giving me the opportunity to follow my passion and personal research interests.

Moreover, I want to acknowledge all the colleagues and collaborators of the different studies presented in this thesis, Tomaso Vecchi, Noemi Ancona, Luisa Girelli, Chiara Ferrari, Micaela Fantino, Lotfi Merabet, Silvia Cucchi, Nicolo’ Bernardi, Egidio D’Angelo. A special thanks to Luca Rinaldi for all the nice scientific discussions, feedbacks, ideas and fun. To Robert Zatorre and Virginia Penhune with whom I have had the honor to work during my Canadian experience: thanks for their guidance, support and constant supervision. My special gratitude to Marianne Stephan: for her help, suggestions, for the encouragement and for being a constant source of motivation. Thanks to all of them: working together enriched me as researcher and as person, improving the quality of my work.

A big thanks to Carlo Toneatto: technician, supporter, partner of coffee and inspiring conversations.

I am grateful to all the volunteers that took part in the different studies. To all musicians who found a moment in their crazy and amazing life to make this thesis a reality. Thanks to all the students who helped me in the data collection.

Thanks to all the magic people I met in Montreal, who made that experience the best I have even done. A special thanks to Ben and Sallamari, Nicolò and Tara, for their kindness, and to Floris, for nicely appearing in my life in unexpected place and unexpected moments. To Lucia, for her contagious happiness. To Reiko, for playing the piano with me, for laughing, discussing, for sharing…to Federica and Daniele and their unconditional support. To Liz (and Cara), for making me feel at home since my first day in Montreal, and for her gift of always seeing the positive aspect of life.

Thanks to the past, the present and the future “Ufficio dei giusti” for being my second family during these four years. To Pietro, Dario, Gragna, Federica, Beatrice, Laura, Valentino, Erica, Susanna, Ilaria, Katia, Paola, Elena, Francesca, for sharing with me the best and the worst moment….this PhD wouldn’t have been possible without them. To all my colleagues for sharing with me this journey (and lot of beers).

To Alberto, Daniele, Giulio, Francesco and Elena, for bringing a bit of rock in my music and in my life.

A double thanks to Giulio, because in different ways he has always been by my side during these four years. For teaching me that with patience and determination we can reach even the most unthinkable goals. I hope this is just the beginning.

Thanks to Chiara, Cecilia, Rossella and Marianna for our friendship, that resists through space and time.

A big thanks to Gio, Sara, Luca, Ilaria, Fede, Davide, Mauro, Paola and Lele, Francesco, Luca and Sara, for making me feel safe, because I know I can always count on them. A special thanks to Eva, for her patience, support and for the precious gift of listening for real and to Claudio, for teaching me to breath, and many many other things. My deepest gratitude to Simone, for giving me the chance to learn and grow with him: he is one of the most beautiful person I know. I am honored and proud to have such great friends in my life.

I would like to express my immeasurable gratitude to my family for their untiring support. To mum and dad, for giving me the opportunity to freely follow my passions. Because I know they are proud of me, which represents the strength to deal with everything in my life. Thanks to grandma and grandpa, because everything started with them. Thanks to my brothers, Edoardo and Tommaso, and to Francesca, for being always by my side.

Thanks because with your presence, inspiration and motivation you helped me to reach this point. I dedicate this thesis to all of you.