AN INVESTIGATION OF SPATIAL REPRESENTATIONS OF PITCH IN INDIVIDUALS WITH CONGENITAL AMUSIA

By

Xuejing Lu

B.Sc., M.Ed.

A thesis submitted in fulfillment of the requirements for the degree of

Doctor of Philosophy

March 18th 2016

Department of Psychology

ARC Centre of Excellence in Cognition and its Disorders

Macquarie University

Sydney, Australia

Contents

Abstract

Congenital amusia, a disorder that affects individuals' musical abilities, has been widely investigated over the past dozen years. Although considerable progress has been made in understanding the psychophysical characteristics and the neural mechanisms of congenital amusia, most is based on the evidence coming from studies that focus primarily on the auditory modality. In this thesis, I explore the relationship between music perception and visuospatial features in individuals with congenital amusia. Chapter 1 reviews existing theory and evidence concerning spatial representations of music, and Chapter 2 provides the background knowledge about congenital amusia and outlines the experimental work that is included in the thesis. In Chapter 3, I describe two experiments on spatial representations of pitch in amusia using a stimulus-response compatibility paradigm. The results confirm that amusic individuals represent pitch with a vertical dimension. Chapter 4 describes new empirical evidence that the perceived magnitude of supra-threshold pitch changes is equivalent for amusic and non-amusic individuals, contrary to what one might expect. In Chapter 5, I investigate whether the impairment observed in contour processing exhibited by amusic individuals is caused by a failure in pitch change direction identification, or whether it arises from poor pitch memory. To disentangle these two stages of contour processing, a novel Self-paced Audio-visual Contour Task was devised. Compared with non-amusic participants, amusics exhibited reduced sensitivity to audio-visual incongruence of the direction of change in pitch and vertical space. This reduced sensitivity reflects a failure to consciously access a spatial representation of pitch. In Chapter 6, I used EEG to measure the extent to which visuospatial information influences amusic individual's judgement of pitch change direction. The behavioural and ERP results show that the influence of visual information is significantly greater for amusic individuals than for non-amusic individuals, suggesting that amusics focused strongly on visual information to complete the auditory task. In Chapter 7, I summarise these findings and argue that amusic individuals have an unstable spatial representation of pitch, which has cascade effects that lead to high-level musical impairments. However, they also draw upon spatial information from the visual modality to clarify or reinforce their

auditory representations. As such, visuospatial cues have potential to help amusic individuals understand and perceive music more precisely.

Statement of Candidate

I, Xuejing Lu, certify that the material presented in this thesis is original. It has not previously been submitted in whole or as part of requirements for a degree to any other university or institution other than Macquarie University. Any help and assistance that I have received in my research and preparation of the thesis itself has been appropriately acknowledged.

I also certify that the sources of information used are indicated in this thesis. The research presented in this thesis has been approved by the Macquarie University Ethics Committee (Ref: 5201200890).

Signature:

Xuejing Lu (Student ID: 42814413)

March 2016

Acknowledgements

First and foremost, special appreciation goes to my principal supervisor Bill Thompson and associate supervisor Blake Johnson for your constant support and encouragement. Thank you for your time, guidance and wisdom in shaping this work. Your insightful ideas and advice on research are invaluable. It has been a great honour to be one of your students.

I would like to thank all past and present members of the *Music, Sound and Performance Lab*, with special thanks to Yanan Sun, Tam Ho, and Lena Quinto. Thank you for your scientific knowledge and many discussions and suggestions. Also, a very big thank you goes to Anna Fiveash for your help with proofreading.

I am very grateful to Lauren Stewart and Daniela Sammler for your comments on this work. I would also like to thank Genevieve McArthur, Nicholas Badcock, and Peter de Lissa from the *ERP lab*, and friendly staff at Macquarie University, especially Avril Moss, Chandera Rajah, Jane Teh, and Ruth Taulia from Department of Psychology, and Faulty IT support team.

Thank you to all my good friends here and overseas for being there when I need you. Special thanks go to Meihua, Nancy, Fei, Terrence, David, Liqin, Hui, Thembi, Ben, Valerie, and Leo for being generous with your time and helping me deal with difficulties.

Finally, thank you to my lovely Mum and Dad for all your love and support. Words are not enough to express my gratitude. Your unconditional love has always been the most beautiful thing in my life.

Thesis by Publication

The format of this thesis is that of a "thesis by publication". Chapter 1 (manuscript 1) has been written with the intention of contributing to the *Oxford Handbook on Music and the Body*. Chapters 3, 5 and 6 (manuscripts 2 to 4) have been written for independent publications in peer-reviewed journals. As a result, the format of in-text citation and reference may vary from chapter to chapter. In addition, tables and figures have been inserted within each chapter. Before each chapter, a brief introduction is provided, explaining how this chapter contributes to this thesis and how it is connected to other chapters. When applicable, my contribution on co-authored papers is specified.

x

Chapter 1: Introduction

This chapter (manuscript 1) reviews previous research showing that musical pitch is represented spatially, and forms the basis for an invited chapter to be published in the *Oxford Handbook on Music and the Body*, edited by Y. Kim and S. Gilman. The manuscript was co-authored by Bill Thompson, and I was the primary contributor. In particular, I was responsible for the literature review and preparation of the first draft. Bill Thompson provided guidance and advice on the literature search and theoretical framework, and made critical comments, suggestions, and revisions on the manuscript.

This chapter was prepared as:

Lu, X., & Thompson, W. F. (in press). Spatial representations common to music and bodily experience. In Y. Kim, & S. Gilman (Ed.), *The Oxford Handbook on Music and the Body*. New York: Oxford University Press.

SPATIAL REPRESENTATIONS COMMON TO MUSIC AND BODILY EXPERIENCE

Xuejing Lu and William Forde Thompson

Macquarie University

ABSTRACT: This chapter discusses evidence that musical pitch is conceived and represented spatially, and that bodily experience provides a rich source of conceptualising music metaphorically. It also describes how bodily gestures may be combined with perceptual representations of music, focusing on music-related movements of performers, such as facial expressions and gestures. Such expressive bodily movements help to shape listeners' perception of music structure and link perception to action. Furthermore, it describes the function of spatial representations of music, and discusses evidence that musical expertise affects the stability and reliability of these spatial representations. Finally, a cognitive-motor framework for understanding spatial representations of music is proposed, which makes predictions about how this representation is manifested, differentially relied upon, and sometimes disrupted in individuals with varying levels of expertise.

KEYWORDS: musical pitch, spatial representation, bodily movement, facial expression, gesture

CROSS-MODAL ASSOCIATIONS

To perceive the external environment accurately and efficiently, the human brain draws upon multiple sources of sensory information—such as vision, audition, touch, and olfaction—and integrates these inputs to form a coherent multimodal percept (Ernst and Bulthoff 2004). Because objects in the world are perceived multimodally, strong cross-modal associations are formed, especially between vision and audition,

which process most of the information that we perceive in everyday life. From an evolutionary standpoint, cross-modal associations are essential to survival. For example, when hearing a sound with a high frequency in a potentially dangerous environment, people tend to look up rather than look down to locate the sound source. In this case, pitch is a powerful localisation cue because it is perceived regardless of the position of one's body or head. Audition can provide a rapid but imprecise cue about the location of objects, whereas vision can be used for more precise localisation. As a result, people often use pitch to evaluate the likely spatial height of an object, and then draw upon the vision system for more precise localisation.

Pitch is not only associated with spatial height. Indeed, there is a body of evidence showing that pitch is associated with other visuo-spatial features such as shape, thickness, and brightness. High-pitched sounds usually correspond to sharp-edged, thin, or light-coloured objects, whereas low-pitched sounds are associated with round, thick, or dim-coloured objects. In addition to pitch, loudness that reflects the intensity of a sound is also associated with spatial location, brightness, and size (Lewkowicz and Turkewitz 1980; Wagner et al. 1981; Marks 1982, 1987, 1989; Melara 1989; Eitan, Schupak, and Marks 2008; Walker et al. 2010; Dolscheid et al. 2013; Jeschonek, Pauen, and Babocsai 2013).

Importantly, all of these visual associations are related to "space," either in a literal sense (the location of a sound-emitting object in real three-dimensional space) or in a metaphorical sense (an association that does not reflect the actual spatial location of the sound-emitting object). Eitan and Granot (2006) discuss how physical and metaphorical uses of spatial concepts may be applied to several features of music. As examples, changes in pitch, loudness, or attack rate can be associated with visuo-spatial changes in verticality, distance, or velocity. In other words, a number of different spatial concepts can be used to make sense of music.

This chapter focuses on musical pitch and its association with space and bodily experience. Pitch is one of the most fundamental elements of music, and the tendency to map pitch onto space is one of the most robust cross-modal associations (for a review, see Spence 2011). Nonetheless, the implications drawn from discussions on musical pitch are applicable to other music attributes.

SPATIAL REPRESENTATIONS OF PITCH

Musical pitch is usually discussed, conceived, and represented spatially. The most obvious example is the Western music notation system, in which pitches are mapped to vertical locations: notes corresponding to higher pitches are presented at higher spatial positions while notes corresponding to lower pitches are presented at lower spatial positions on the staff. Another example is the layout of the piano keyboard, in which pitches are mapped to horizontal locations: high-pitched notes are on the right and low-pitched notes are on the left.

Formal theoretic and informal descriptions of pitch also employ spatial terms. One early geometric model depicts musical pitch as a helix with two psychological dimensions, referred to as pitch height and chroma, respectively. In language, most Western speakers use spatial metaphors such as "up-down," "ascending-descending," and "rising-falling" to describe pitch height and movement. However, spatial height is not the only metaphor for this attribute of sound. For Turkish, Farsi, and Zapotec speakers, pitch is referred to as "thin" or "thick." In Kpelle, a language spoken in Liberia, speakers use "large" or "heavy" to describe low-pitched sound and "small" or "light" to describe high-pitched sound (Stone 1981; Shayan, Ozturk, and Sicoli 2011; Dolscheid et al. 2013). These alternative characterisations of pitch challenge the notion that pitch is universally represented in terms of spatial height. Nonetheless, the association between pitch and spatial height is encoded in most languages of the world, and all conceptions of pitch appear to reflect some connection between pitch and space, suggesting that pitch is represented in an embodied manner across cultures.

Pitch-space associations have been extensively investigated since the late 1800s. Why are tones characterised as "high" or "low" in most languages? Stumpf (1883, as cited in Pratt 1930) argued that the association between pitch and vertical location might be accounted for by the frequent use of spatial metaphors derived from the experiences elicited by tones and other stimuli. For instance, low-pitched tones tend to give people impressions of objects that are massive, gloomy, or heavy. Creatures that emit low-pitched vocalisations are usually found at a lower location (i.e., close to the ground). Conversely, high-pitched tones tend to give people impressions of objects that are small, bright, and light. In turn, creatures that emit high-pitched vocalisations tend to be found at higher locations (i.e., up in the sky).

Pratt (1930) made the first empirical attempt to discover the mechanisms underlying pitch-space associations. He presented subjects with tones differing in pitch and asked them to locate the position of incoming tones on a scale running from the floor to the ceiling. Every subject placed tones in a top to bottom order corresponding to their pitch height from high to low. Based on this observation, Pratt reasoned that high- and low-pitched tones are phenomenologically higher and lower in space, respectively, suggesting a psychological bias in the perception of pitch. At the time, it was unknown whether the connection between pitch and space was learned or hard-wired, but it was assumed that the association helps to facilitate sensory processing. For instance, when a visual object at a relatively high position is simultaneously presented with a high-pitched sound rather than a low-pitched sound, both the visual object and the sound are more rapidly and accurately detected (Bernstein and Edelstein 1971; Melara and O'Brien 1987; Ben-Artzi and Marks 1995).

Other evidence showing pitch-space associations comes from studies adopting the stimulus-response compatibility (SRC) paradigm. The SRC effect can be observed in binary-choice reaction time tasks when one manipulates the assignment of the response and stimulus. The basic logic of the SRC paradigm is as follows: participants are asked to classify a stimulus as rapidly and accurately as possible using response keys, whereby the spatial location of the response key for a correct classification is manipulated such that it is either congruent or incongruent with the implied (or metaphoric) spatial location of the stimulus. Subjects should respond faster and more accurately when the position of the (correct) response key is congruent with the implied position of the stimulus (Kornblum, Hasbroucq, and Osman 1990; Proctor and Vu 2006).

Adopting this paradigm, two research teams independently demonstrated an association between the pitch height of a tone and the spatial location of the response keys. This is known as the Spatial-Musical/Pitch Association of Response Code (SMARC/SPARC) effect (Rusconi et al. 2006; Lidji et al. 2007). Both teams found that subjects performed better when the pitch of the tone (higher or lower) was congruent with the position of the correct response key (higher or lower on the device) than when they were incongruent with each other. This effect occurred for both musically trained and untrained participants and hence cannot be explained as a

learned association between pitch and actions required to produce tones on instruments such as a piano. Indeed, when response keys were aligned horizontally (i.e., left-right rather than top-bottom, as on a piano), the advantage for the combinations of high-pitched tone with right response key, and low-pitched tone with left response key, was only found in subjects with musical training but not in those who were musically naïve (Rusconi et al. 2006; Lidji et al. 2007).

SPATIAL METAPHORS OF PITCH ON THE BASIS OF BODILY EXPERIENCE

In an effort to understand the world, humans form mental representations through bodily and sensory experiences, and their mental representations reflect the nature of the world itself. As the world is three-dimensional, spatial coding is an efficient way to represent and interact with it. Spatial representations also appear to be universal and can be expressed using both language metaphors and bodily gestures (Lakoff and Johnson 1980; Parkinson et al. 2012; Nunez and Cooperrider 2013).

The examination of spatial metaphors can shed light on mental processes, as they mirror cross-modal experiences. Lakoff and Johnson (1980) argued that we conceptualise the external world through our own bodily experience in the physical environment, and bodily experiences are reflected in the metaphors we use. Orientational metaphors arise from spatial experience and assign non-spatial concepts a (metaphoric) spatial orientation. As an example, we often use the expression "wake up" to awaken somebody. This expression is formed on the basis that our body tends to be in an upright position when we are awake—standing, sitting or walking, with the ground under our feet and the sky above our head. Therefore, the concept "wake" is oriented in the direction "up," leading to such an expression.

Similarly, we use spatial metaphors to describe the pitch of a tone and its movement on the basis of our experiences of environmental sounds that vary in average fundamental frequency. For example, bird song typically has a high average fundamental frequency and arises from up in trees or in the sky; large animals that produce low-pitched sounds tend to be heavy and emit vocalisations at ground level.

When we are singing high-pitched tones, the larynx rises, and when we are singing low-pitch tones, it descends. Such bodily experiences link pitch height to spatial position, which can lead to orientation metaphors of pitch. Analogously, musicians use movement metaphors to describe tempo changes in music such as "accelerando" (gradually accelerating) and "ritardando" (being slower), because they "feel" these movements in music. In short, we use bodily experiences as our most trusted way of representing, understanding, and interacting with the external world. As a consequence, our conceptual system and the metaphors we use are shaped and constrained by our bodily experiences. These embodied metaphors, in turn, allow us to map concrete experiences in the sensorimotor domain to subjective experience.

MUSIC-RELATED GESTURE AND ITS AFFORDANCES

Music performance is always accompanied by bodily movements. Such movements include actions required to produce sounds on musical instruments, ancillary bodily gestures and facial expressions, and physical responses to perceiving music. The analysis of such movements can provide considerable insight into mental representations of music.

In this chapter, the term "music-related movements" refers to movements required to produce sounds (i.e., music-producing gestures such as piano key press actions), and ancillary movements that often accompany music performance (e.g., emotive facial expressions, gestures that illustrate or reinforce a structural element) or music perception (e.g., dancing or tapping along with music). For performers, physical actions are required to make sounds. For example, percussionists must raise a mallet into the air before striking a drum. However, bodily movements are also a non-verbal means of communicating with audience members and with co-performers, and may function to express an emotional interpretation of the music, to reinforce metric and phrase boundaries, and to interact with and entertain audiences (Davidson and Correia 2002). For those perceiving the sounds, seeing performers' bodily movements can enrich their music experience. For instance, facial expressions shown by performers affect perceived interval size, dissonance, and sung syllables (Thompson, Graham and Russo 2005; Quinto et al. 2010). In addition, audiences are likely to feel more engaged if their response to the music includes bodily movement and facial expressions, such as smiling.

Although a number of researchers have examined music-related gestures, this area of research is still in its infancy. In contrast, there is a large body of literature concerned with language-related gestures. Using gestures during talking is common in all cultural and linguistic backgrounds. Language-related gestures can even be observed in young children before their development of language, as well as in congenitally blind speakers who have never seen gestures and have no model for gesturing (Iverson and Goldin-Meadow 1998). Iverson and Thal (1998) define gestures as actions "produced with the intention to communicate . . . [they] typically involve fingers, hands, and arms (e.g., in pointing or showing) and also may employ facial features (e.g., lip smacking for cookies) or even the entire body (e.g., bouncing up and down for horse)" (59-60). Broadly speaking, gestures can be understood to consist of all facial expressions and body movements that have the potential contribute to communication, whether such movements are intentional or not.

Iverson and Thal classify two primary categories of gesture: deictic and representational. Deictic gestures are used to indicate an object or event by directing listeners' attention (e.g., pointing or reaching for a toy). Representational gestures are used to indicate an object-related semantic content by signifying some feature of the referent (e.g., cupped hand to mouth to represent "drinking," and swinging arms rapidly to express "running"). Representational gestures can also be used in a more abstract or symbolic way (e.g., shrugging shoulders for "helpless").

Based on the categories of non-verbal behaviour described by Ekman and Friesen (1969), McNeill (1992) provides another category—beat gestures—which typically do not convey information about concrete objects or semantic content, but are used rhythmically with speech to emphasize the significant or accented points in the accompanying speech. McNeil further divides representational gestures into three subcategories: iconic, metaphoric, and emblems. Iconic gestures indicate object-related semantic content by mirroring a specific feature of the referent, whereas metaphoric gestures represent abstract features of an object or event. Emblems are stereotypical gestures that are shared within a culture or a group of people, and can usually be translated into a word such as "hello," "ok," or "peace." For example, when greeting in Hindu culture, people usually bow slightly and put their palms together with fingers pointing upwards and thumbs close to the chest (i.e., the gesture of Namaste).

According to Darwin (1871), music and language have a common evolutionary origin as a "musical protolanguage" that was used in courtship and territoriality. If so, then one might expect gestures used in music to have similar communicative function as language-related gestures. For instance, conductors usually use deictic gestures with hands, arms, face, and head to direct instrument players' or performers' attention, thereby unifying them, setting the tempo, and so on. It is also common for performers to use gestures as a way of signaling musical changes in an iconic or metaphoric way. For example, singers often move their head or eyebrows in systematic ways, or display upward hand gestures when there is an increase in pitch or intensity. Such gestures contain secondary information about changes in music, supplementing and reinforcing the acoustic channel, and hence providing a safety net for breakdowns in acoustic communication.

Thompson, Graham, and Russo (2005) point out that music-related gestures are often used to signal "the timing of musical events, focusing listeners' attention to (or away from) critical acoustic information at specific moments in time" (203). In their view, "musical events" can be any significant melodic, harmonic, or rhythmic event based on performers' "understanding of segmentation (points of closure), intervallic information (whether a melodic interval is large or small), and points of expectancy fulfillment or violation" (204). In other words, performers can use bodily movements to convey their intentions and share their experience with listeners.

Although some music-evoked movements appear to be universal (e.g., tapping of feet to music), the same piece of music can induce different bodily responses from individual to individual. It is also common to observe differences in movements that people make to a single piece of music for the first and the second time. These variations in bodily responses to music may reflect unique interactions of individuals with music, in which music-related movements are made depending upon several factors, such as individuals' background, their mood at a given time, and so on.

Godøy (2010) discusses the idea of "gestural affordances of musical sound" by adopting the notion of "affordance" from ecological psychology, which was first introduced by James J. Gibson in the 1970s. Gibson (1979) suggested that "the affordances of the environment are what it offers the animal, what it provides or furnishes either for good or ill" (127). Applying this to the gestural affordances of musical sounds, Godøy (2010) suggests that "music has very rich gestural affordance"

and "gestures shape our perception of the music" (104). In his view, "musical sound is a transducer of source-information" (106), which is fulfilled by the potential gestures that music affords. People acquire knowledge of music-producing gestures through experience, such as seeing music performances. Once the association between musical sounds and gestures is established, people can mentally recreate these music-producing gestures even when they listen to music in the absence of visual images. By combining these mental images with the human nature of imitating others, listeners readily move to musical sounds on the basis of the gestures that they have seen or experienced previously.

Most features of musical sounds can be understood in relation to their gestural affordances. Therefore, Godøy considers these musical features as "trajectory shapes in time and space" (113). For instance, gestural responses such as moving hands up or down, left or right, near or far, and so on are elicited in response to the musical changes. These spatial gestures, in turn, shape music perception of listeners as well as performers, if we consider performers as the listeners of their own performance. From this perspective, music perception is not about perceiving music itself but also concerns how one interacts with its affordances. To support of this idea, empirical evidence for the influence of bodily movements, such as hand gestures and facial expressions, on perceptual representations of music is reviewed, and then the connection between music perception and action planning is discussed.

THE INFLUENCES OF MUSIC-RELATED GESTURES ON THE REPRESENTATIONS OF MUSIC

A number of psychological studies on bodily movements and music experience have been conducted in the past few decades. In most of these studies, researchers record performers' bodily movements using video recorders or motion capture cameras and then conduct movement analysis of the gestures or facial expressions. Captured movements are also sometimes used as stimuli in perceptual experiments in which participants are asked to rate, judge, or reproduce these movements.

At the perceptual level, it has been well documented that the presence of visual information may enhance or alter one's auditory perception (e.g., McGurk and MacDonald 1976; van Wassenhove, Grant, and Poeppel 2005). Thus, music-related

movements introduced intentionally by performers have potential to enhance the listener's music experience. Listeners with varied musical background can judge the size of sung intervals in the absence of acoustic information based on performers' facial expressions and head movements alone (Thompson, Russo, and Livingstone 2010). More generally, music-related movements greatly influence listeners' perception of musical structure. For instance, Connell, Cai and Holler (2013) investigated how visuo-spatial movement biases pitch perception. They first presented a group of subjects with a synthesized note, and then showed them a video where an actress sang a target note and moved her hand either upward or downward at the same time. The task was to judge whether the first and the second note were the same or different in terms of the pitch height while ignoring the visual information. If the subject's answer was "different," the participant was asked whether the second note was higher or lower. Their results indicated an obvious bias on perceived pitch height caused by concurrent spatial gestures, such that downward gestures made notes seem lower in pitch than they really were, and upward gestures made notes seem higher in pitch.

The same research team conducted another study to examine the influence of gesture on perceived tone duration (Cai, Connell, and Holler 2013). This time, researchers asked a group of subjects to view a video of an actress singing a musical note while making a long- or short-distance gesture. After watching the video, subjects were asked to reproduce the duration of the sung note by holding down the spacebar on a computer keyboard. As expected, the perceived duration of the sung note was longer when it was accompanied by a long-distance gesture than when it was accompanied by a short-distance gesture.

In a related study, Schutz and Kubovy (2006) presented subjects with a video showing a world-renowned percussionist performing a single percussive note. They observed that the visual stroke type, that is, the gesture employed for producing a long note or a short note, affected subjects' perceived tone duration, even when they were informed that the visual and auditory information might be mismatched. More importantly, such influences of bodily movements on perceived music structure were not dependent on the level of musical training (Schutz 2007; Schutz and Lipscomb 2007).

A LINK BETWEEN MUSIC PERCEPTION AND ACTION PLANNING

William James (1890) stated that "every representation of a movement awakens in some degree the actual movement which is its object" (2: 526), suggesting a direct perception-action link. More recently, this idea was developed into a formal model known as "common coding theory" (Prinz 1984, 1997; Hommel et al. 2001). Common coding theory holds that perception and action modulate each other because they share a common representational coding system. Hommel and colleagues (2001) emphasise that "the stimulus representations underlying perception and the action representations underlying action planning are coded and stored not separately, but together in a common representational medium" (849). Hommel (2011) further explains that "the human brain seems to code all sorts of spatial aspects of a stimulus event and there is evidence that all these codes can interact with the spatial representation of an action directed to, or at least triggered by a given stimulus" (193). In other words, using common coding allows people to combine information from different sensory sources into an amodal spatial representation, and also to transfer the mental representation to a behavioural level. From this perspective, the common coding theory also reflects an embodied view of cognition, where the subjective experience and physical or biological mechanisms are connected as tightly as possible (Leman 2008).

In the music domain, the perception-action link is well supported by evidence from behavioural studies, such as the priming effect of perception on action that we mentioned above (the SMARC/SPARC effect). On the other hand, both active movements (adults bouncing themselves in synchrony with the experimenter) and passive movements (babies being bounced in the experimenter's arms) elicit a biased hearing of an ambiguous rhythmic pattern (Phillips-Silver and Trainor 2005, 2007, 2008). In addition, Ammirante and Thompson (2012) also found that motor plans could be incorporated into non-musicians' perception of melodic motion. In this study, when subjects were asked to tap continuously to trigger successive tones in a melody, their finger movement was systematically affected by the auditory feedback that they were instructed to ignore. These findings suggest that actions such as bodily movements are not merely the outcome of high-level cognitive functions. Instead, they have great potential to exert a reciprocal influence on perception.

Further evidence of a perception-action link comes from neuroimaging studies (Grezes and Decety 2001; Lotze et al. 2003, Popescu, Otsuka, and Ioannides 2004; Bangert et al. 2006; Baumann et al. 2007; for a review, see Zatorre et al. 2007). Such studies indicate that listening to music without playing it, or even just imagining a music performance, can lead to activation in motor areas of the brain in both musicians and non-musicians. Conversely, when playing music, even with no auditory feedback, the right primary auditory cortex can be activated.

THE ROLE OF MUSICAL SKILL

Based on the current empirical evidence summarised above, it is clear that people with different levels of expertise in music make gestures in response to music, and the influences of gestures on music perception do not seem to depend upon musical training. On the other hand, musical training may increase the stability of spatial representations of music, thereby affecting music perception. Indeed, in addition to a vertical representation of pitch, a horizontal representation of pitch tends to occur in musicians but not in non-musicians (Rusconi et al. 2006; Lidji et al. 2007; Stewart et al. 2012; Küssner et al. 2014; Lega et al. 2014; Taylor and Witt 2014). A possible explanation is that musical training allows people to establish new associations by linking musical sounds and corresponding actions via a shared spatial representation. In general, the association between pitch and vertical space is more straightforward than the association between pitch and horizontal space for most people with no musical training on the basis of their daily experience. Therefore, when faced with situations that require associating pitch with horizontal space, musically untrained individuals may need extra effort to mentally project horizontal space to vertical space. In contrast, musicians, such as pianists, are exposed to associative learning processes due to the immense amount of practice accompanying their training, where they can progressively learn the horizontal representation of pitch. In particular, pianists move towards the left of the keyboard to correspond to decreasing pitch and to the right of the keyboard to correspond to increasing pitch. As a consequence, the horizontal spatial representation of musical pitch is strongly established without involving extra mental spatial manipulations.

On the other end of spectrum of musical ability, people with congenital amusia show unstable vertical spatial representations of pitch when compared with normal

listeners in the SRC task, suggesting that spatial conceptions of musical pitch might be disrupted and cause difficulties in associating pitch height with spatial action among these people (Douglas and Bilkey 2007; Lu et al. 2014). Systematic investigations into congenital amusia started only a dozen years ago (Ayotte, Peretz, and Hyde 2002; Foxton et al. 2004; Hyde and Peretz 2004). Based on the current understanding of this disorder, some researchers believe amusia to be a disorder of pitch awareness caused by a disconnection between auditory perceptual and high-level processing that involves conscious awareness (Loui et al. 2008; Loui, Alsop, and Schlaug 2009; Peretz et al. 2009; Hyde, Zatorre, and Peretz 2010; Lu et al. 2015; Zendel et al. 2015). In other words, the musical input fails to be cognitively represented for further processing. The unstable representation of musical pitch can, of course, lead to other musical impairments at higher levels, such as melodic contour processing, timbre perception, and pitch memory (Gosselin, Jolicoeur, and Peretz 2009; Tillmann, Schulze, and Foxton 2009; Williamson and Stewart 2010; Marin, Gingras, and Stewart 2012; Albouy et al. 2013).

Drawing from findings of musician and amusia studies, the possibility that spatial representations of pitch normally play a valuable function for listeners and performers, bridging music perception and actions and reinforcing the ability to encode and retain auditory information, is considered.

A COGNITIVE-MOTOR FRAMEWORK

This chapter reviewed theories related to music perception and music-related gestures from embodied accounts and provided an overview of recent evidence in support of these ideas. In essence, embodied music cognition stresses the human body as a biological transducer that decodes music and translates it into intended expressive movements. Music perception, which is shaped by bodily experience, encourages action tendencies due to the rich gestural affordances of music (Leman 2008; Godøy 2010). More recently, combining the new evidence showing the functional role of the motor system in auditory perception, Maes and colleagues (2014) provide a theoretical framework to emphasise that action can modulate music perception reciprocally. This framework draws on the common coding theory, which suggests a representational overlap between perception and action planning.

Here, the focus is on spatial representation, and a cognitive-motor framework is proposed to capture its important role in linking music perception and bodily movements. Given the fact that the human body is the only medium between the mental world and the physical world, it makes sense to assume that spatial representation is an efficient and universal way to exchange information between these two worlds. Both concrete and abstract information can be represented in a kind of virtual space. Indeed, spatial representation of music is not specific to pitch. Other attributes of music—loudness, timbral brightness, tempo, rhythm, note duration, temporal order, and melodic accent—can also be represented spatially. The shared spatial representation allows people to link musical sounds to the corresponding gestures by means of progressively associative learning, and also allows people to execute actions in response to music based on these learned associations. Based on the evidence from studies on musical expertise and congenital amusia, this framework emphasises that the degree of stability of spatial representations in music relies upon the levels of music skill. Musical training can reinforce associations between music sounds and corresponding actions. Furthermore, the representation of musical pitch may be disrupted in individuals with congenital amusia. The unstable representations may lead to impairments at both low and high levels of music processing.

SUMMARY

Music is not purely acoustic. Rather, it is multisensory expressed and processed activity that involves both perception and motion systems. Indeed, the affordances of musical sound allow the body to interact with it. In other words, bodily gestures can be incorporated into music perception and shaped by it. This chapter proposed that spatial representations on the basis of bodily experience can account for the connections between musical sounds and music-related movements. Furthermore, these representations might be reinforced in musicians but disrupted in individuals with musical impairments.

References

Albouy, Philippe, Jeremie Mattout, Romain Bouet, Emmanuel Maby, Gaetan Sanchez, Pierre-Emmanuel Aguera, Sebastien Daligault, Claude Delpuech, Olivier Bertrand, Anne Caclin, and Barbara Tillmann. 2013. "Impaired Pitch Perception and Memory in Congenital Amusia: The Deficit Starts in the Auditory Cortex." *Brain* 136: 1639–1661.

Ammirante, Paolo, and William F. Thompson. 2012. "Continuation Tapping to Triggered Melodies: Motor Resonance Effects of Melodic Motion." *Experimental Brain Research* 216: 51–60.

Ayotte, Julie, Isabelle Peretz, and Krista L. Hyde. 2002. "Congenital Amusia: A Group Study of Adults Afflicted with a Music-Specific Disorder." *Brain* 125: 238–251.

Bangert, Marc, Thomas Peschel, Gottfried Schlaug, Michael Rotte, Dieter Drescher, Hermann Hinrichs, Hans-Jochen Heinze, and Eckart Altenmuller. 2006. "Shared Networks for Auditory and Motor Processing in Professional Pianists: Evidence from fMRI Conjunction." *Neuroimage* 30: 917–926.

Baumann, Simon, Susan Koeneke, Conny F. Schmidt, Martin Meyer, Kai Lutz, and Lutz Jancke. 2007. "A Network for Audio-Motor Coordination in Skilled Pianists and Non-musicians." *Brain Research* 1161: 65–78.

Ben-Artzi, Elisheva, and Lawrence E. Marks. 1995. "Visual-Auditory Interaction in Speeded Classification: Role of Stimulus Difference." *Perception & Psychophysics* 57: 1151–1162.

Bernstein, Ira H., and Barry A. Edelstein. 1971. "Effects of Some Variations in Auditory Input Upon Visual Choice Reaction Time." *Journal of Experimental Psychology* 87: 241–247.

Cai, Zhenguang G., Louise Connell, and Judith Holler. 2013. "Time Does Not Flow without Language: Spatial Distance Affects Temporal Duration Regardless of Movement or Direction." *Psychonomic Bulletin & Review* 20: 973–980.

Connell, Louise, Zhenguang G. Cai, and Judith Holler. 2013. "Do You See What I'm Singing? Visuospatial Movement Biases Pitch Perception." *Brain and Cognition* 81: 124–130.

Darwin, Charles. 1871. *The Descent of Man, and Selection in Relation to Sex.* London: John Murray.

Davidson, Jane W., and Jorge S. Correia. 2002. "Body movement." In *The Science and Psychology of Music Performance: Creative Strategies for Teaching and Learning*, edited by Richard Parncutt and Gary McPherson, 237–251. New York: Oxford University Press.

Dolscheid, Sarah, Shakila Shayan, Asifa Majid, and Daniel Casasanto. 2013. "The Thickness of Musical Pitch: Psychophysical Evidence for Linguistic Relativity." *Psychological Science* 24: 613–621.

Douglas, Katie M., and David K. Bilkey. 2007. " Amusia Is Associated with Deficits in Spatial Processing." *Nature Neuroscience* 10: 915–921.

Eitan, Zohar, Assi Schupak, and Lawrence E. Marks. 2008. "Louder Is Higher: Cross-Modal Interaction of Loudness Change and Vertical Motion in Speeded Classification." Paper presented at the 10th International Conference on Music Perception and Cognition, Adelaide, Australia, August 25-29.

Eitan, Zohar, and Roni Y. Granot. 2006. "How Music Moves: Musical Parameters and Listeners' Images of Motion." *Music Perception* 23: 221–248.

Ekman, Paul, and Wallace V. Friesen. 1969. "The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding." *Semiotica* 1: 49–98.

Ernst, Marc O., and Heinrich H. Bulthoff. 2004. "Merging the Senses into a Robust Percept." *Trends in Cogntive Sciences* 8: 162–169.

Foxton, Jessica M., Jennifer L. Dean, Rosemary Gee, Isabelle Peretz, and Timothy D. Griffiths. 2004. "Characterization of Deficits in Pitch Perception Underlying 'Tone Deafness.'" *Brain* 127: 801–810.

Gibson, James J. 1979. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.

Godøy, Rolf I. 2010. "Gestural Affordances of Musical Sound." In *Musical Gestures: Sound, Movement, and Meaning*, edited by Rolf I. Godøy and Marc Leman, 103–125. New York: Routledge.

Gosselin, Nathalie, Pierre Jolicoeur, and Isabelle Peretz. 2009. "Impaired Memory for Pitch in Congenital Amusia." *Annals of the New York Academy of Sciences* 1169: 270–272.

Grezes, Julie, and Jean Decety. 2001. "Functional Anatomy of Execution, Mental Simulation, Observation, and Verb Generation of Actions: A Meta-Analysis." *Human Brain Mapping* 12: 1–19.

Hommel, Bernhard. 2011. "The Simon Effect as Tool and Heuristic." *Acta Psychologica* 136: 189–202.

Hommel, Bernhard, Jochen Müsseler, Gisa Aschersleben, and Wolfgang Prinz. 2001. "The Theory of Event Coding (TEC): A Framework for Perception and Action Planning." *Behavioral and Brain Sciences* 24: 849–937.

Hyde, Krista L., and Isabelle Peretz. 2004. "Brains That Are out of Tune but in Time." *Psychological Science* 15: 356–360.

Hyde, Krista. L., Robert J. Zatorre, and Isabelle Peretz. 2010. " Functional MRI Evidence of an Abnormal Neural Network for Pitch Processing in Congenital Amusia." *Cerebral Cortex* 21: 292–299.

Iverson, Jana M., and Susan Goldin-Meadow. 1998. "Why People Gesture When They Speak." *Nature* 396: 228.

Iverson, Jana M., and Donna J. Thal. 1998. "Communicative Transitions: There's More to the Hand Than Meets the Eye." In *Transitions in Prelinguistic Communication: Preintentional to Intentional and Presymbolic to Symbolic*, edited by Amy M. Wetherby, Steven F. Warren and Joe Reichle, 59–86. Baltimore: Paul H. Brookes.

James, William. 1890. *The Principles of Psychology*, vol. 2. New York: Holt.

Jeschonek, Susanna, Sabina Pauen, and Lysett Babocsai. 2013. "Cross-Modal Mapping of Visual and Acoustic Displays in Infants: The Effect of Dynamic and Static Components." *European Journal of Developmental Psychology* 10: 337–358.

Kornblum, Sylvan, Thierry Hasbroucq, and Allen Osman. 1990. "Dimensional Overlap: Cognitive Basis for Stimulus-Response Compatibility – a Model and Taxonomy." *Psychological Review* 97: 253–270.

Küssner, Mats B., Dan Tidhar, Helen M. Prior, and Daniel Leech-Wilkinson. 2014. "Musicians Are More Consistent: Gestural Cross-Modal Mappings of Pitch, Loudness and Tempo in Real-Time." *Frontiers in Psychology* 5: 789.

Lakoff, George, and Mark Johnson. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.

Lega, Carlotta, Zaira Cattaneo, Lotfi B. Merabet, Tomaso Vecchi, and Slivia Cucchi. 2014. "The Effect of Musical Expertise on the Representation of Space." *Frontiers in Human Neuroscience* 8: 250.

Leman, Marc. 2008. *Embodied Music Cognition and Mediation Technology*. Cambridge: MIT Press.

Lewkowicz, David J., and Gerald Turkewitz. 1980. "Cross-Modal Equivalence in Early Infancy: Auditory–Visual Intensity Matching." *Developmental Psychology* 16: 597–607.

Lidji, Pascale, Régine Kolinsky, Aliette Lochy, and José Morais. 2007. "Spatial Associations for Musical Stimuli: A Piano in the Head?" *Journal of Experimental Psychology: Human Perception and Performance* 33: 1189–1207.

Lotze, M., G. Scheler, H. R. Tan, C. Braun, and N. Birbaumer. 2003. "The Musician's Brain: Functional Imaging of Amateurs and Professionals During Performance and Imagery." *Neuroimage* 20: 1817–1829.

Loui, Psyche, David Alsop, and Gottfried Schlaug. 2009. "Tone Deafness: A New Disconnection Syndrome?." *The Journal of Neuroscience* 29: 10215–10220.

Loui, Psyche, Frank H. Guenther, Christoph Mathys, and Gottfried Schlaug. 2008. "Action-Perception Mismatch in Tone-Deafness." *Current Biology* 18: R331–332.

Lu, Xuejing, Pin Zhang, Yanan Sun, and William F. Thompson. 2014. "Spatial Representations of Pitch among Individuals with Congenital Amusia." Paper presented at the 13th International Conference on Music Perception and Cognition, Seoul, South Korea. August 4–8.

Lu, Xuejing, Hao Tam Ho, Fang Liu, Daxing Wu, and William F. Thompson. 2015. "Intonation Processing Deficits of Emotional Words among Mandarin Chinese Speakers with Congenital Amusia: An Erp Study." *Frontiers in Psychology* 6: 385.

Maes, Pieter-Jan, Marc Leman, Caroline Palmer, and Marcelo M. Wanderley. 2014. "Action-Based Effects on Music Perception." *Frontiers in Psychology* 4: 1008.

Marin, Manuela M., Bruno Gingras, and Lauren Stewart. 2012. "Perception of Musical Timbre in Congenital Amusia: Categorization, Discrimination and Short-Term Memory." *Neuropsychologia* 50: 367–378.

Marks, Lawrence E. 1982. "Bright Sneezes and Dark Coughs, Loud Sunlight and Soft Moonlight." *Journal of Experimental Psychology: Human Perception and Performance* 8: 177–193.

Marks, Lawrence E. 1987. "On Cross-Modal Similarity: Auditory-Visual Interactions in Speeded Discrimination." *Journal of Experimental Psychology: Human Perception and Performance* 13: 384–394.

Marks, Lawrence E. 1989. "On Cross-Modal Similarity: The Perceptual Structure of Pitch, Loudness, and Brightness." *Journal of Experimental Psychology: Human Perception and Performance* 15: 586–602.

McGurk, Harry, and John MacDonald. 1976. "Hearing Lips and Seeing Voices." *Nature* 264: 746–748.

McNeill, David. 1992. *Hand and Mind: What Gestures Reveal About Thought*. Chicago: University of Chicago Press.

Melara, Robert D. 1989. "Dimensional Interaction between Color and Pitch." *Journal of Experimental Psychology: Human Perception and Performance* 15: 69–79.

Melara, Robert D., and Thomas P. O'Brien. 1987. "Interaction between Synesthetically Corresponding Dimensions." *Journal of Experimental Psychology: General* 116: 323–336.

Nunez, Rafael, and Kensy Cooperrider. 2013. "The Tangle of Space and Time in Human Cognition." *Trends in Cognitive Sciences* 17: 220–229.

Parkinson, Carolyn, Peter Jes Kohler, Beau Sievers, and Thalia Wheatley. 2012. "Associations between Auditory Pitch and Visual Elevation Do Not Depend on Language: Evidence from a Remote Population." *Perception* 41: 854–861.

Peretz, Isabelle, Elvira Brattico, Miika Jarvenpaa, and Mari Tervaniemi. 2009. "The Amusic Brain: In Tune, out of Key, and Unaware." *Brain* 132: 1277–1286.

Phillips-Silver, Jessica, and Laurel. J. Trainor. 2005. "Feeling the Beat: Movement Influences Infant Rhythm Perception." *Science* 308: 1430.

Phillips-Silver, Jessica, and Laurel J. Trainor. 2007. "Hearing What the Body Feels: Auditory Encoding of Rhythmic Movement." *Cognition* 105: 533–546.

Phillips-Silver, Jessica, and Laurel J. Trainor. 2008. "Vestibular Influence on Auditory Metrical Interpretation." *Brain and Cognition* 67: 94–102.

Popescu, Mihai, Asuka Otsuka, and Andreas A. Ioannides. 2004. "Dynamics of Brain Activity in Motor and Frontal Cortical Areas During Music Listening: A Magnetoencephalographic Study." *Neuroimage* 21: 1622–1638.

Pratt, Carroll C. 1930. "The Spatial Character of High and Low Tones." *Journal of Experimental Psychology* 13: 278–285.

Prinz, Wolfgang. 1984. "Modes of Linkage between Perception and Action." In *Cognition and Motor Processes*, edited by Wolfgang Prinz and Andries F. Sanders, 185–193. Berlin: Springer.

Prinz, Wolfgang. 1997. "Perception and Action Planning." *European Journal of Cognitive Psychology* 9: 129–154.

Proctor, Robert W., and Kim-Phuong L. Vu. 2006. *Stimulus-Response Compatibility Principles: Data, Theory, and Application*. Boca Raton, FL: CRC Press.

Quinto, Lena, William F. Thompson, Frank A. Russo, and Sandra E. Trehub. 2010. "A Comparison of the McGurk Effect for Spoken and Sung Syllables." *Attention, Perception, & Psychophysics* 72: 1450–1454.

Rusconi, Elena, Bonnie Kwan, Bruno L. Giordano, Carlo Umilta, and Brian Butterworth. 2006. "Spatial Representation of Pitch Height: The SMARC Effect." *Cognition* 99: 113–129.

Schutz, Michael. 2007. "Do We Hear Silent Gestures?" Paper presented at the 153rd Meeting of the Acoustical Society of America, Salt Lake City, Utah, June 4–8.

Schutz, Michael, and Michael Kubovy. 2006. "Seeing Music: The Percussionists' Strategic Use of Gesture in Live Performance." Paper presented at the 2nd International Conference on Music and Gesture, Royal Northern College of Music, Manchester, UK, July 20–23.

Schutz, Michael, and Scott Lipscomb. 2007. "Hearing Gestures, Seeing Music: Vision Influences Perceived Tone Duration." *Perception* 36: 888–897.

Shayan, Shakila, Ozge Ozturk, and Mark A. Sicoli. 2011. "The Thickness of Pitch: Crossmodal Metaphors in Farsi, Turkish, and Zapotec." *The Senses and Society* 6: 96–105.

Spence, Charles. 2011. "Crossmodal Correspondences: A Tutorial Review." *Attention, Perception & Psychophysics* 73: 971–995.

Stewart, Lauren, Rinus G. Verdonschot, Patrick Nasralla, and Jennifer Lanipekun. 2012. "Action-Perception Coupling in Pianists: Learned Mappings or Spatial Musical Association of Response Codes (SMARC) Effect?" *The Quarterly Journal of Experimental Psychology* 66: 37–50.

Stone, Ruth M. 1981. "Toward a Kpelle Conceptualization of Music Performance." *Journal of American Folklore* 94: 188–206.

Stumpf, Carl. 1883. *Tonpsychologie*, vol. 1. Lepizig: S. Hirzel.

Taylor, J. Eric T., and Jessica K. Witt. 2014. "Listening to Music Primes Space: Pianists, but Not Novices, Simulate Heard Actions." *Psychoogical Research* 79: 175–182.

Thompson, William F., Philip W. Graham, and Frank A. Russo. 2005. "Seeing Music Performance: Visual Influences on Perception and Experience." *Semiotica* 156:203–227.

Thompson, William F., Frank A. Russo, and Steven R. Livingstone. 2010. "Facial Expressions of Singers Influence Perceived Pitch Relations." *Psychonomic Bulletin & Review* 17: 317–322.

Tillmann, Barbara, Katrin Schulze, and Jessica M. Foxton. 2009. "Congenital Amusia: A Short-Term Memory Deficit for Non-Verbal, but Not Verbal Sounds." *Brain and Cognition* 71: 259–64.

van Wassenhove, Virginie, Ken W. Grant, and David Poeppel. 2005. "Visual Speech Speeds up the Neural Processing of Auditory Speech." *Proceedings of the National Academy of Sciences of the United States of America* 102: 1181–1186.

Wagner, Sheldon, Ellen Winner, Dante Cicchetti, and Howard Gardner. 1981. ""Metaphorical" Mapping in Human Infants." *Child Development* 52: 728–731.

Walker, Peter, J. Gavin Bremner, Uschi Mason, Jo Spring, Karen Mattock, Alan Slater, and Scott P. Johnson. 2010. "Preverbal Infants' Sensitivity to Synaesthetic Cross-Modality Correspondences." *Psychological Science* 21: 21–25.

Williamson, Victoria J., and Lauren Stewart. 2010. "Memory for Pitch in Congenital Amusia: Beyond a Fine-Grained Pitch Discrimination Problem." *Memory* 18: 657–669.

Zatorre, Robert J., Joyce L. Chen, and Virginia B. Penhune. 2007. "When the Brain Plays Music: Auditory-Motor Interactions in Music Perception and Production." *Nature Reviews Neuroscience* 8: 547–558.

Zendel, Benjamin R., Marie-Elaine Lagrois, Nicolas Robitaille, and Isabelle Peretz. 2015. "Attending to Pitch Information Inhibits Processing of Pitch Information: The Curious Case of Amusia." The *Journal of Neuroscience* 35: 3815–3824.

Chapter 2: Outline of Experimental Work

As reviewed in Chapter 1, the spatial representation of pitch plays an important role in auditory processing, reinforcing one's ability to encode and retain auditory information. Therefore, it makes sense to believe that an unstable representation of pitch may lead to impairments at low and high levels of musical pitch processing. Although investigations into people who suffer from musical handicaps (either congenital or acquired after brain injury or stroke) have been occurring since the late 19th century (Ayotte, Peretz, & Hyde, 2002; Geschwind, 1984; Grant-Allen, 1879; Kalmus & Fry, 1980; Peretz et al., 2002; Piccirilli, 2000; see Stewart, von Kriegstein, Warren, & Griffiths, 2006, for a review), the role of spatial representations for pitch processing in these populations is poorly understood.

To explore whether impaired auditory processing is associated with impaired pitch-space mapping, and how spatial information affects pitch perception, this thesis draws upon the data from individuals with congenital amusia, a musical disorder that mainly affects one's pitch perception (Ayotte et al., 2002; Foxton, Dean, Gee, Peretz, & Griffiths, 2004; Hyde & Peretz, 2004). Although there are historical anecdotes about people born with deficits in musical abilities, Monica is the first systematically documented case of congenital amusia (Peretz et al., 2002). Monica is well educated and has neither neurological nor hearing problems. However, she lacks basic musical abilities, such as melodic discrimination and pitch change detection, despite normal language skills. This music-specific disorder was corroborated by an investigation and characterisation of a group of amusic individuals (Ayotte et al., 2002), who also showed impairments in music memory and music production, in addition to deficits in fine-grained pitch perception. In 2003, Peretz and colleagues published the *Montreal Battery of Evaluation of Amusia* (MBEA; Peretz, Champod, & Hyde, 2003), a tool to identify individuals with congenital amusia. The MBEA comprises six subtests (Scale, Contour, Interval, Rhythm, Meter, and Memory) that assess each of the musical processing components described by Peretz and Coltheart (2003). Although the utility of the MBEA to effectively identify amusic individuals has been questioned in recent years (Henry & McAuley, 2010, 2013; Pfeifer & Hamann, 2015), the MBEA is the most widely used tool in the research field of congenital

amusia. Depending on the goal of the study, individuals are defined as amusics if their average scores for three melodic subtests (Scale, Contour, and Interval), or all six subtests are at or below two standard deviations of the mean obtained in the general population (e.g., Liu, Patel, Fourcin, & Stewart, 2010; Nan, Sun, & Peretz, 2010).

Following these initial studies, extensive research on individuals with congenital amusia has been conducted, reporting that amusics have difficulties in melodic contour processing (Albouy et al., 2013), timbre perception (Marin, Gingras, & Stewart, 2012), emotional prosody processing (Thompson, Marin, & Stewart, 2012), pitch memory (Gosselin, Jolicoeur, & Peretz, 2009; Tillmann, Schulze, & Foxton, 2009; Williamson & Stewart, 2010), pitch production (Dalla Bella, Berkowska, & Sowinski, 2011; Dalla Bella, Giguère, & Peretz, 2009), and spatial processing (Douglas & Bilkey, 2007; but see, Tillmann et al., 2010; Williamson, Cocchini, & Stewart, 2011). However, amusic individuals only show mild or no impairment in linguistic pitch processing (Ayotte et al., 2002; Hutchins, Gosselin, & Peretz, 2010; Jiang, Hamm, Lim, Kirk, & Yang, 2010; Liu et al., 2012; Liu et al., 2010; Nan et al., 2010; Nguyen, Tillmann, Gosselin, & Peretz, 2009; see Peretz, 2013 for a review). Furthermore, they are usually sensitive to musical emotions (Gosselin, Paquette, & Peretz, 2015), and show some understanding of musical structure (Omigie, Pearce, & Stewart, 2012; Omigie & Stewart, 2011; Tillmann, Gosselin, Bigand, & Peretz, 2012; but see Jiang, Lim, Wang, & Hamm, 2013). These findings suggest that amusia is a very specific disorder mainly affecting music abilities.

Although considerable progress has been made over the past dozen years in understanding the cognitive and neural mechanisms of congenital amusia, the cause of this impairment is still under debate. To date, it has been suggested that congenital amusia arises from difficulties in fine-grained pitch perception (Ayotte et al., 2002; Vuvan, Nunes-Silva, & Peretz, 2015), pitch awareness (Loui, Alsop, & Schlaug, 2009; Loui, Guenther, Mathys, & Schlaug, 2008; Lu, Ho, Liu, Wu, & Thompson, 2015; Moreau, Jolicoeur, & Peretz, 2013; Peretz, Brattico, Jarvenpaa, & Tervaniemi, 2009; Zendel, Lagrois, Robitaille, & Peretz, 2015), abnormal brainstem response (Lehmann, Skoe, Moreau, Peretz, & Kraus, 2015; but see, Liu, Maggu, Lau, & Wong, 2015), and impaired short-term memory (Gosselin et al., 2009; Tillmann et al., 2009; Williamson & Stewart, 2010; see Tillmann, Leveque, Fornoni, Albouy, & Caclin, 2016, for a recent review). Neurologically, congenital amusia may be associated with

decreased white matter concentration in the inferior frontal gyrus (Hyde, Zatorre, Griffiths, Lerch, & Peretz, 2006; Mandell, Schulze, & Schlaug, 2007), decreased volume of the arcuate fasciculus (Loui, Alsop, & Schlaug, 2009; Loui & Schlaug, 2009), and reduced functional connectivity between the inferior frontal gyrus and the auditory cortex (Albouy et al., 2013; Hyde, Zatorre, & Peretz, 2011).

Considering existing evidence summarised in Chapter 1, I argue that congenital amusia is associated with a failure to form stable mental representations of pitch. To test this hypothesis, this thesis explores the spatial representation of pitch in congenital amusia. To screen individuals with congenital amusia, over 500 participants were administrated the three melodic subtests of the MBEA (because this thesis focuses on pitch processing) at the Music, Sound, and Performance lab during 2012-2015. The majority of participants were university students recruited from the Psychology Participant Pool at Macquarie University. Individuals whose average scores were below the cut-off scores (i.e., 2 SD below the population mean across three subtests) were defined as amusics and invited to participate in further studies, and individuals whose average scores within the normal range were considered as potential control participants. Amusics and controls were matched at the group level in terms of age, gender, handedness, years of education, years of musical training, and hours of daily music listening.

Specifically, I first examined whether amusic and non-amusic people represent pitch in the same way. Chapter 3 describes experimental work that compares the mental representation of pitch in individuals with and without congenital amusia using a stimulus-response compatibility paradigm. To rule out the possibility that group differences arise because the perceived size of pitch changes is not equivalent in amusic and non-amusic participants, a control experiment is reported in Chapter 4.

Second, I focused on distinguishing perceptual and memory stages of contour processing. Although amusic individuals usually exhibit deficits in processing contours of pitch sequences (Foxton et al., 2004; Jiang et al., 2010), it is still unclear whether the impairments result from a failure in perceiving pitch properly, or in retaining the pitch information in memory. To disentangle these two stages of contour processing, a novel Self-paced Audiovisual Contour Task (SACT) that places minimal demands on pitch memory is introduced in Chapter 5, and experimental work drawing on this task is described.

Third, Chapter 6 describes an ERP study that investigates the extent to which visuospatial information affects pitch perception in individuals with congenital amusia. This study was motivated by survey evidence that amusic participants tend to rely heavily on facial expressions and gestures when interpreting the moods and feelings of people with whom they are speaking (Thompson et al., 2012).

Chapter 7 summarises the main findings of this thesis, discusses the theoretical implications of the research, and outlines potential directions for future research. To anticipate this discussion, I conclude that there is strong evidence for the hypothesis that individual with amusia fail to form a stable mental representation of pitch, and this failure arises at a perceptual level of processing.

References

Albouy, P., Mattout, J., Bouet, R., Maby, E., Sanchez, G., Aguera, P. E., . . . Tillmann, B. (2013). Impaired pitch perception and memory in congenital amusia: The deficit starts in the auditory cortex. *Brain, 136*(5), 1639-1661. doi: 10.1093/brain/awt082

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

Dalla Bella, S., Berkowska, M., & Sowinski, J. (2011). Disorders of pitch production in tone deafness. *Frontiers in Psychology, 2,* 164. doi: 10.3389/fpsyg.2011.00164

Dalla Bella, S., Giguère, J.-F., & Peretz, I. (2009). Singing in congenital amusia. *The Journal of the Acoustical Society of America, 126*(1), 414. doi: 10.1121/1.3132504

Douglas, K. M., & Bilkey, D. K. (2007). Amusia is associated with deficits in spatial processing. *Nature Neuroscience, 10*(7), 915-921. doi: 10.1038/nn1925

Foxton, J. M., Dean, J. L., Gee, R., Peretz, I., & Griffiths, T. D. (2004). Characterization of deficits in pitch perception underlying 'tone deafness'. *Brain, 127*(4), 801-810. doi: 10.1093/brain/awh105

Geschwind, N. (1984). The brain of a learning-disabled individual. *Annuals of Dyslexia, 34*, 319-327.

Gosselin, N., Jolicoeur, P., & Peretz, I. (2009). Impaired memory for pitch in congenital amusia. *Annals of the New York Academy of Sciences, 1169*, 270-272. doi: 10.1111/j.1749-6632.2009.04762.x

Gosselin, N., Paquette, S., & Peretz, I. (2015). Sensitivity to musical emotions in congenital amusia. *Cortex, 71*, 171-182. doi: 10.1016/j.cortex.2015.06.022

Grant-Allen. (1879). Note-deafness. *Mind, 3*(10), 157-167.

Hutchins, S., Gosselin, N., & Peretz, I. (2010). Identification of changes along a continuum of speech intonation is impaired in congenital amusia. *Frontiers in Psychology, 1*, 236. doi: 10.3389/fpsyg.2010.00236

Henry, M. J., & McAuley, J. D. (2013). Failure to apply signal detection theory to the Montreal Battery of Evaluation of Amusia may misdiagnose amusia. *Music Perception, 30*(5), 480-496. doi:10.1525/mp.2013.30.5.480

Henry, M. J., & McAuley, J. D. (2010). On the Prevalence of Congenital Amusia. *Music Perception, 27*(5), 413-418. doi:10.1525/mp.2010.27.5.413

Hyde, K. L., & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science, 15*(5), 356-360. doi: 10.1111/j.0956-7976.2004.00683.x

Hyde, K. L., Zatorre, R. J., Griffiths, T. D., Lerch, J. P., & Peretz, I. (2006). Morphometry of the amusic brain: A two-site study. *Brain, 129*(10), 2562-2570. doi: 10.1093/brain/awl204

Hyde, K. L., Zatorre, R. J., & Peretz, I. (2011). Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. *Cerebral Cortex, 21*(2), 292-299. doi: 10.1093/cercor/bhq094

Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. (2010). Processing melodic contour and speech intonation in congenital amusics with Mandarin Chinese. *Neuropsychologia, 48*(9), 2630-2639. doi: *10.1016/j.neuropsychologia.2010.05.009*

Jiang, C., Lim, V. K., Wang, H., & Hamm, J. P. (2013). Difficulties with pitch discrimination influences pitch memory performance: Evidence from congenital amusia. *PLoS One, 8*(10), e79216. doi: 10.1371/journal.pone.0079216.t001

Kalmus, H., & Fry, D. B. (1980). On tune deafness (dysmelodia) : Frequency, development, genetics and musical background. *Annals of Human Genetics, 43*(4), 369-382. doi: 10.1111/j.1469-1809.1980.tb01571.x

Lehmann, A., Skoe, E., Moreau, P., Peretz, I., & Kraus, N. (2015). Impairments in musical abilities reflected in the auditory brainstem: evidence from congenital amusia. *The European Journal of Neuroscience, 42*(1), 1644-1650. doi: 10.1111/ejn.12931

Liu, F., Jiang, C., Thompson, W. F., Xu, Y., Yang, Y., & Stewart, L. (2012). The mechanism of speech processing in congenital amusia: Evidence from Mandarin speakers. *PLoS One, 7*(2), e30374. doi: 10.1371/journal.pone.0030374

Liu, F., Maggu, A. R., Lau, J. C. Y., & Wong, P. C. M. (2015). Brainstem encoding of speech and musical stimuli in congenital amusia: evidence from Cantonese speakers. *Frontiers in Humam Neuroscience, 8*. doi: 10.3389/fnhum.2014.01029

Liu, F., Patel, A. D., Fourcin, A., & Stewart, L. (2010). Intonation processing in congenital amusia: Discrimination, identification and imitation. *Brain, 133*(6), 1682-1693. doi: 10.1093/brain/awq089

Loui, P., Alsop, D., & Schlaug, G. (2009). Tone deafness: a new disconnection syndrome? *Journal of Neuroscience, 29*(33), 10215-10220. doi: 10.1523/jneurosci.1701-09.2009

Loui, P., Guenther, F. H., Mathys, C., & Schlaug, G. (2008). Action-perception mismatch in tone-deafness. *Current Biology, 18*(8), R331-332. doi: 10.1016/j.cub.2008.02.045

Loui, P., & Schlaug, G. (2009). Investigating musical disorders with diffusion tensor imaging. *Annals of the New York Academy of Sciences, 1169*(1), 121-125. doi: 10.1111/j.1749-6632.2009.04781.x

Lu, X., Ho, H. T., Liu, F., Wu, D., & Thompson, W. F. (2015). Intonation processing deficits of emotional words among Mandarin Chinese speakers with congenital amusia: an ERP study. *Frontiers in Psychology, 6*, 385. doi: 10.3389/fpsyg.2015.00385

Mandell, J., Schulze, K., & Schlaug, G. (2007). Congenital amusia: An auditory-motor feedback disorder? *Restorative Neurology and Neuroscience, 25*(3), 323-334.

Marin, M. M., Gingras, B., & Stewart, L. (2012). Perception of musical timbre in congenital amusia: Categorization, discrimination and short-term memory. *Neuropsychologia, 50*(3), 367-378. doi: 10.1016/j.neuropsychologia.2011.12.006

Moreau, P., Jolicoeur, P., & Peretz, I. (2013). Pitch discrimination without awareness in congenital amusia: Evidence from event-related potentials. *Brain and Cognition, 81*(3), 337-344. doi: 10.1016/j.bandc.2013.01.004

Nan, Y., Sun, Y., & Peretz, I. (2010). Congenital amusia in speakers of a tone language: Association with lexical tone agnosia. *Brain, 133*(9), 2635-2642. doi: 10.1093/brain/awq178

Nguyen, S., Tillmann, B., Gosselin, N., & Peretz, I. (2009). Tonal language processing in congenital amusia. *Annals of the New York Academy of Sciences, 1169*(1), 490-493. doi: 10.1111/j.1749-6632.2009.04855.x

Omigie, D., Pearce, M. T., & Stewart, L. (2012). Tracking of pitch probabilities in congenital amusia. *Neuropsychologia, 50*(7), 1483-1493. doi: 10.1016/j.neuropsychologia.2012.02.034

Omigie, D., & Stewart, L. (2011). Preserved statistical learning of tonal and linguistic material in congenital amusia. *Frontiers in Psychology, 2*, 109. doi: 10.3389/fpsyg.2011.00109

Peretz, I. (2013). The biological foundations of music: Insights from congenital amusia. In D. Deustch (Ed.), *The psychology of music* (pp. 551-564). San Diego: Academic Press.

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., & Jutras, B. (2002). Congenital amusia: A disorder of fine-grained pitch discrimination. *Neuron, 33*(2), 185-191. doi: 10.1016/S0896-6273(01)00580-3

Peretz, I., Brattico, E., Jarvenpaa, M., & Tervaniemi, M. (2009). The amusic brain: in tune, out of key, and unaware. *Brain, 132*(5), 1277-1286. doi: 10.1093/brain/awp055

Peretz, I., Champod, A. S., & Hyde, K. L. (2003). Varieties of musical disorders: The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences, 999*(1), 58-75. doi:10.1196/annals.1284.006

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6*(7), 688-691. doi:10.1038/nn1083

Pfeifer, J., & Hamann, S. (2015). Revising the diagnosis of congenital amusia with the Montreal Battery of Evaluation of Amusia. *Frontiers in Human Neuroscience, 9*, 161. doi:10.3389/fnhum.2015.00161

Piccirilli, M. (2000). Modularity of music: Evidence from a case of pure amusia. *Journal of Neurology, Neurosurgery & Psychiatry, 69*(4), 541-545. doi: 10.1136/jnnp.69.4.541

Stewart, L., von Kriegstein, K., Warren, J. D., & Griffiths, T. D. (2006). Music and the brain: Disorders of musical listening. *Brain, 129*, 2533-2553. doi: 10.1093/brain/awl171

Thompson, W. F., Marin, M. M., & Stewart, L. (2012). Reduced sensitivity to emotional prosody in congenital amusia rekindles the musical protolanguage hypothesis. *Proceedings of the National Academy of Sciences, 109*(46), 19027-19032. doi: 10.1073/pnas.1210344109

Tillmann, B., Gosselin, N., Bigand, E., & Peretz, I. (2012). Priming paradigm reveals harmonic structure processing in congenital amusia. *Cortex, 48*(8), 1073-1078. doi: 10.1016/j.cortex.2012.01.001

Tillmann, B., Jolicoeur, P., Ishihara, M., Gosselin, N., Bertrand, O., Rossetti, Y., & Peretz, I. (2010). The amusic brain: Lost in music, but not in space. *PLoS One, 5*(4), e10173. doi: 10.1371/journal.pone.0010173.g001

Tillmann, B., Leveque, Y., Fornoni, L., Albouy, P., & Caclin, A. (2016). Impaired short-term memory for pitch in congenital amusia. *Brain Research, 1640*, 251-263. doi: 10.1016/j.brainres.2015.10.035

Tillmann, B., Schulze, K., & Foxton, J. (2009). Congenital amusia: A short-term memory deficit for non-verbal, but not verbal sounds. *Brain and Cognition, 71*(3), 259-264. doi: 10.1016/j.bandc.2009.08.003

Vuvan, D. T., Nunes-Silva, M., & Peretz, I. (2015). Meta-analytic evidence for the non-modularity of pitch processing in congenital amusia. *Cortex, 69*, 186-200. doi: 10.1016/j.cortex.2015.05.002

Williamson, V. J., Cocchini, G., & Stewart, L. (2011). The relationship between pitch and space in congenital amusia. *Brain and Cognition, 76*(1), 70-76. doi: 10.1016/j.bandc.2011.02.016

Williamson, V. J., & Stewart, L. (2010). Memory for pitch in congenital amusia: Beyond a fine-grained pitch discrimination problem. *Memory, 18*(6), 657-669. doi: 10.1080/09658211.2010.501339

Zendel, B. R., Lagrois, M. E., Robitaille, N., & Peretz, I. (2015). Attending to pitch information inhibits processing of pitch information: The curious case of amusia. *The Journal of Neuroscience, 35*(9), 3815-3824. doi: 10.1523/JNEUROSCI.3766-14.2015

Chapter 3: Spatial Representation of Pitch

As discussed in Chapter 1 (manuscript 1), spatial representation of pitch plays a central role in auditory processing. However, it is still unclear whether the impairments in auditory processing seen in congenital amusia are associated with impaired pitch-space mapping. In this chapter, two experiments using a stimulus-response capability (SRC) paradigm are described to address the intriguing association between pitch and spatial height in individuals with and without congenital amusia. First, I used the SRC to assess the spatial representation of pitch in individuals with congenital amusia (Experiment 1). Second, I evaluated whether the SRC effect exhibited by amusic participants reflects a linguistic association (Experiment 2).

This manuscript (manuscript 2) was co-authored by Yanan Sun and Bill Thompson. My contribution to the manuscript was roughly 80%. I designed the experiments, collected the data, analysed the results, and wrote the first draft of the article. Yanan Sun and Bill Thompson provided helpful comments and feedback regarding experimental design, data interpretation, and manuscript preparation.

This chapter was prepared as:

Lu, X., Sun, Y., & Thompson, W. F (in press). An investigation of spatial representation of pitch in individuals with congenital amusia. *The Quarterly Journal of Experimental Psychology.* doi: 10.1080/17470218.2016.1213870

AN INVESTIGATION OF SPATIAL REPRESENTATION OF PITCH IN INDIVIDUALS WITH CONGENITAL AMUSIA

Xuejing Lu[1,2], Yanan Sun[2,3], William Forde Thompson[1,2]

[1]Department of Psychology, Macquarie University, Sydney, NSW, Australia
[2]ARC Centre of Excellence in Cognition and its Disorders, Sydney, NSW, Australia
[3]Department of Cognitive Science, Macquarie University, Sydney, NSW, Australia

Running head: pitch representation in amusics

Correspondence should be addressed to Xuejing Lu, Department of Psychology, Macquarie University, NSW 2109, Australia.

E-mail address: xuejing.lu@mq.edu.au

ABSTRACT

Spatial representation of pitch plays a central role in auditory processing. However, it is unknown whether impaired auditory processing is associated with impaired pitch-space mapping. Experiment 1 examined spatial representation of pitch in individuals with congenital amusia using a stimulus-response compatibility (SRC) task. For amusic and non-amusic participants, pitch classification was faster and more accurate when correct responses involved a physical action that was spatially congruent with the pitch height of the stimulus than when it was incongruent. However, this spatial representation of pitch was not stable in amusic individuals, revealed by slower response times when compared with control individuals. One explanation is that the SRC effect in amusics reflects a linguistic association, requiring additional time to link pitch height and spatial location. To test this possibility, Experiment 2 employed a colour-classification task. Participants judged colour while ignoring a concurrent pitch by pressing one of two response keys positioned vertically to be congruent or incongruent with the pitch. The association between pitch and space was found in both groups, with comparable response times in the two groups, suggesting amusic individuals are only slower to respond to tasks involving explicit judgments of pitch.

KEYWORDS

Congenital amusia; Spatial representation; Pitch perception; Stimulus-response compatibility effect

A rich body of evidence supports the idea that musical pitch is conceived and represented spatially. Pitches are talked about using spatial metaphors such as "*high* pitch", "*descending* scale", and "pitch *contour*". Furthermore, pitches are mapped vertically in the Western musical notation system, where high-pitched notes are represented at higher positions, and low-pitched notes are represented at lower positions on the staff. With regards to empirical evidence, two research teams have independently demonstrated an association between the pitch of a tone and the spatial location of corresponding response keys using the stimulus-response compatibility (SRC) task (Lidji, Kolinsky, Lochy, & Morais, 2007; Rusconi, Kwan, Giordano, Umilta, & Butterworth, 2006). The basic logic of the SRC paradigm is as follows: If the response features overlap with stimulus features, subjects should respond more quickly and more accurately when the overlapped features between response and stimulus are compatible with each other than when they are not (Kornblum, Hasbroucq, & Osman, 1990; Proctor & Vu, 2006). Using this paradigm, Rusconi et al. (2006) and Lidji et al. (2007) found that both musician and non-musician subjects performed better when pitch height was congruent with the vertically spatial position of response keys (i.e., higher tone with top response key and lower tone with bottom response key) than when they were incongruent with each other (i.e., higher tone with bottom response key and lower tone with top response key). However, the strength of such association depends on musical expertise (Küssner, Tidhar, Prior, & Leech-Wilkinson, 2014; Lega, Cattaneo, Merabet, Vecchi, & Cucchi, 2014; Stewart, Verdonschot, Nasralla, & Lanipekun, 2013; Taylor & Witt, 2015), suggesting that spatial representations of musical pitch are more consistent and robust in musicians than in non-musicians.

Furthermore, an ambiguous visual stimulus is perceived as moving upward when presented with a rising tone and moving downward when presented with a falling tone, whereas no movement is perceived when the same stimulus is presented with white noise (Maeda, Kanai, & Shimojo, 2004; see also Eitan & Granot, 2006). This finding suggests that auditory stimuli with no spatial information can alter human visual motion perception. Manipulations of pitch height have even been shown to affect infants' attention to the spatial location of visual stimuli (Jeschonek, Pauen, & Babocsai, 2013; Wagner, Winner, Cicchetti, & Mardner, 1981; Walker et al., 2010). Conversely, visual stimuli or gestures that indicate spatial height also

affect the estimation of pitch height (Casasanto, Phillips, & Boroditsky, 2003; Connell, Cai, & Holler, 2013).

These findings suggest that a spatial representation of pitch plays a central role in auditory processing. As such, an impaired pitch-space mapping might be associated with impaired auditory processing, decreasing sensitivity to aspects of pitch structure such as melodic contour and pitch interval size. Indeed, a developmental failure to form cognitive representations of musical pitch might be associated with congenital amusia (Patel, 2003), a music-specific disorder that affects one's fine-grained pitch perception (Ayotte, Peretz, & Hyde, 2002; Foxton, Dean, Gee, Peretz, & Griffiths, 2004; Hyde & Peretz, 2004), melodic contour processing (Albouy et al., 2013), timbre perception (Marin, Gingras, & Stewart, 2012), emotional prosody processing (Thompson, Marin, & Stewart, 2012), pitch memory (Gosselin, Jolicoeur, & Peretz, 2009; Tillmann, Schulze, & Foxton, 2009; Williamson & Stewart, 2010), and pitch production (Dalla Bella, Berkowska, & Sowinski, 2011; Dalla Bella, Giguère, & Peretz, 2009).

Douglas and Bilkey (2007) used the SRC task from Rusconi et al. (2006) to examine the possibility that musical pitch is represented spatially in individuals with congenital amusia (Griffiths et al., 1997; Krumhansl, 1990). Participants were presented with pairs of tones and were asked to identify whether the second tone was higher or lower in relation to the first tone by pressing one of two response keys aligned vertically. They found that spatial congruency (e.g., indicating a higher pitch by pressing a response key that was physically higher) had a greater effect on response speed and accuracy for the control group than for the amusic group, resulting in a significant interaction between group and congruency. However, Stewart and Walsh (2007) pointed out that the auditory stimuli in that study were ambiguous, in that they included small intervals between the first and second tones that were below the pitch threshold of amusics. Thus, the impairment observed in the amusic group might reflect a deficit in pitch discrimination rather than an unstable spatial representation of pitch. Furthermore, only one subtest of the Montreal Battery of Evaluation of Amusia (MBEA; Peretz, Champod, & Hyde, 2003) was used to diagnose congenital amusia, raising additional questions about the validity of the diagnosis (Ayotte et al., 2002). Finally, other cognitive abilities, such as intelligence and working memory, were not taken into account. Such abilities are associated with

speed of information processing and should be matched or statistically controlled in the two groups (Fry & Hale, 1996, 2000; Sheppard & Vernon, 2008).

In addition to the absence of an SRC effect in amusics, Douglas and Bilkey (2007) reported that amusic participants exhibited impairment on a mental rotation task when compared with matched controls, suggesting that amusia is associated with impaired spatial imagery. However, the latter finding could not be replicated (Tillmann et al., 2010; Williamson, Cocchini, & Stewart, 2011) and appears unrelated to pitch representations. Taken together, these findings suggest that amusia is not associated with a general impairment in spatial processing, but rather with deficits that are restricted to the spatial representation of pitch.

In the present investigation, we employed the SRC paradigm to re-examine pitch-space associations in individuals with congenital amusia, while controlling for a range of potentially confounding factors. If reduced performance on the SRC paradigm by amusic individuals were restricted to pitch changes that are below pitch-discrimination thresholds, then it would not be possible to conclude that they have an unstable spatial representation of pitch, given that performance deficits could be readily explained by an inability to detect pitch changes. However, if amusic individuals genuinely had an unstable or distorted spatial representation of pitch, then group difference should be observed in SRC task even when differences in pitch from trial to trial are above the pitch-discrimination threshold for all participants, and when other cognitive abilities are matched between groups.

Experiment 1: Pitch classification task

Method

Participants

Sixteen individuals with congenital amusia and 16 matched controls participated in Experiment 1. Three melodic subtests of the MBEA were administered for diagnosis because we were interested in individuals with deficits in pitch processing. In each subtest, participants were presented with pairs of melodies and were asked to judge whether they were the "same" or "different" (for more details, see Peretz et al., 2003). Participants whose composite scores of three melodic subtests were equal to or lower than 65 out of 90 points—that is, 72% correct—were considered as congenital

amusics (Liu, Patel, Fourcin, & Stewart, 2010). All participants had normal or corrected-to-normal visual acuity. Their hearing thresholds were at 20 dB or less for the pure tones in the 0.25-8-kHz range in at least one ear. The amusic and control groups were matched in age, handedness, gender, years of education, years of musical training, hours of music listening, non-verbal IQ, and memory for digits. However, amusic participants performed significantly worse on the three melodic subtests of the MBEA when compared with control participants (see details in Table 1). All participants were informed of the experiment procedures and consented to take part in the experiment before testing, which was approved by the Macquarie University Ethics Committee.

Table 1. Participants' characteristics, percentage correct on the melodic subtests of the MBEA, and independent-sample *t* test results between two groups in Experiment 1.

Contents	Amusic	Control	t (2-tailed)
Age (years)	23.83 ± 4.91	22.06 ± 7.04	0.83
Handedness[a]	13R/1L/2A	13R/1L/2A	–
Gender[b]	7F/9M	7F/9M	–
Years of education	15.44 ± 2.45	14.00 ± 1.85	1.87
Years of musical training	0.38 ± 1.09	1.58 ± 3.23	1.41
Hours of music listening daily	2.31 ± 2.94	2.06 ± 1.21	0.31
Non-verbal IQ[c]	103.44 ± 10.89	109.88 ± 11.91	1.60
Memory for digits[d]	6.88 ± 1.09	7.5 ± 1.37	1.43
MBEA (average)	66.88 ± 6.03	86.04 ± 5.91	9.08 ***
Scale	71.25 ± 10.32	91.46 ± 5.83	6.82 ***
Contour	65.21 ± 10.40	85.42 ± 7.88	6.20 ***
Interval	64.17 ± 10.29	81.25 ± 9.10	4.97 ***

Note: MBEA = Montreal Battery of Evaluation of Amusia. Values are mean ± SD unless otherwise indicated.

[a]According to the Edinburgh Inventory (Oldfield, 1971) (R = right-handed; L = left-handed; A = Ambidextrous). [b]F: female; M: male. [c]Tested with the Matrices subset from Kaufmann Brief Intelligence Test (KBIT-2). [d]Tested with Memory for Digits subtest from Comprehension Test of Phonological Processing (CTOPP-2). The maximum number of correctly recalled items was recorded and represents the capacity of the short-term auditory memory store. ***: $p < .001$.

Stimuli

A low-pitched (D4, 293.66 Hz) and a high-pitched (F6, 1396.9 Hz) pure tone were used as stimuli, whereby the pitch distance between the two tones (27 semitones) was well above participants' threshold for pitch discrimination. Tones were generated using Audacity (Version 2.0.3) with constant amplitude and ramps on both onset and offset. Both frequencies had two durations—50 ms and 1000 ms. This manipulation was used to ensure that the spatial representation of pitch is independent of tone duration. The interaction of pitch height and response location has previously been shown to be highly significant in the vertical alignment but only a trend toward significance in the horizontal alignment in individuals with no musical training (Lidji et al., 2007; Rusconi et al., 2006), suggesting that the horizontal representation of pitch relates to the level of expertise in music. Therefore, only the vertical alignment was employed in the present study. The response pad (RB-530, Cedrus Corporation, San Pedro, U.S.A.) was placed upright with the surface off vertical by approximately 30°. The distance between the top and bottom response keys was 88 mm, and there was a "HOME" key located in the middle of the box. All stimuli were presented using SuperLab 4.5 (Cedrus Corporation, San Pedro, USA) at a comfortable hearing level via noise-cancelling headphones (Sennheiser PXC 350).

Procedure

Participants were tested individually in a quiet room. They were familiarized with the auditory stimuli before commencing the experiment, and they were asked to rest their index finger of their preferred hand over the "HOME" key in advance, and after each response. On each trial, a fixation-cross appeared for 500 ms in the centre of the screen, at which point either the high-pitched or the low-pitched tone was presented. Participants were asked to identify whether the tone they heard was high or low by pressing one of two response keys as accurately and quickly as possible. Critically, the spatial position of the correct response key was either congruent or incongruent with the pitch height, depending on the assignment of the response keys. Specifically, a congruent trial refers to trials in which a high-pitched sound aligns with a top-response key or a low-pitched sound aligns with a bottom-response key, whereas an incongruent trial refers to trials with reversed mappings—that is, a high-pitched

sound aligns with a bottom- response key or a low-pitched sound aligns with a top-response key. Congruent and incongruent trials and tones with different durations were presented in four different blocks. The presentation order of blocks was counterbalanced within each group. Participants were informed of the correct response keys before each block, after which they completed four blocks of 96 trials, where each stimulus was presented 48 times. No more than four trials with the same stimulus (i.e., high or low pitch) were presented in a row. A blank screen was presented for 500 ms between trials. Before each block, eight practice trials with feedback were presented.

Results

The primary dependent measures were accuracy, indicated by error rate (ER), and speed, taken as the mean response times (RT; i.e., the interval between the stimulus onset and the key-press) for correct trials only. Trials with reaction times shorter than 100 ms, reflecting anticipatory or accidentally correct responses, and reaction times longer than 1500 ms, were classified as outliers and excluded from analysis. With these criteria, 0.37% of the correct trials in total were discarded in Experiment 1.

Repeated-measures analysis of variance (ANOVA) on RT and ER was conducted with the between-subjects factor of group (amusic and control) and within-subject factors of congruency (congruent and incongruent) and duration (short and long). There was a main effect of congruency on RT, $F(1, 30) = 24.51$, $MSE = 3756.86$, $p < .001$, $\eta_P^2 = .45$, revealing a 54-ms advantage in response times for congruent (amusic: $M = 532$ ms, $SE = 22$ ms; control: $M = 453$ ms, $SE = 22$ ms) over incongruent trials (amusic: $M = 589$ ms, $SE = 30$ ms; control: $M = 503$ ms, $SE = 30$ ms). Similarly, a significant main effect of congruency on ER was found in both groups, $F(1, 30) = 11.74$, $MSE = 0.001$, $p < .01$, $\eta_P^2 = .28$, with lower error rates for congruent trials (amusic: $M = 1.2\%$, $SE = 0.4\%$; control: $M = 1.8\%$, $SE = 0.4\%$) than incongruent trials (amusic: $M = 3.9\%$, $SE = 0.8\%$; control: $M = 2.6\%$, $SE = 0.8\%$). This tendency towards faster responses to congruent trials was found in both groups, reflected in no interaction between congruency and group on RT, $F(1, 30) = 0.12$, $MSE = 3756.86$, $p = .73$, $\eta_P^2 < .01$, and on ER, $F(1, 30) = 3.57$, $MSE = 0.001$, $p = .07$, $\eta_P^2 < .10$, respectively. However, control participants responded faster than amusic

participants across congruent and incongruent conditions, as revealed by a significant main effect of group, $F(1, 30) = 5.29$, $MSE = 40,744.01$, $p < .05$, $\eta_P^2 = .15$. The faster responses by the control group were not at the cost of accuracy, however, as the group difference was not significant on ER, $F(1, 30) = 0.19$, $MSE = 0.002$, $p = .67$, $\eta_P^2 < .01$. Although responses to the tones with short durations were slightly longer than those to tones with long durations, the effect of duration on RT was not statistically reliable, $F(1, 30) = 3.27$, $MSE = 2128.32$, $p = .08$, $\eta_P^2 = .10$. Therefore, all following analyses were collapsed across duration. As shown in Figure 1, five out of 32 participants showed a reversed SRC effect (3 amusic and 2 control participants). No other significant effects on RT or ER were found (all $p > .05$).
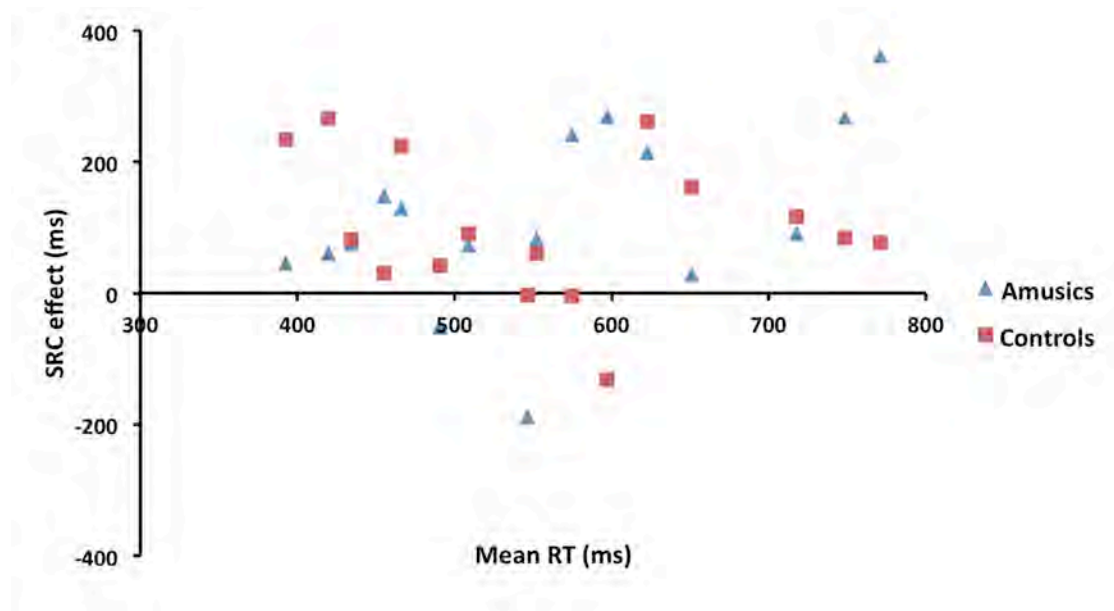
Figure 1. Stimulus-response compatibility (SRC) effect (the difference in response time, RT, between incongruent and congruent trials) collapsed across durations for each participant as a function of mean RT (Experiment 1).

When response time performance is assessed in terms of average speed of responding, information such as trial-to-trial variability is lost. In other words, the temporal dynamics of the SRC effect may be varied between the amusic and the control group, although the magnitude of the SRC effect shown by the two groups was comparable. Therefore, we conducted a distribution analysis on response times to investigate the temporal dynamics of the SRC effect (De Jong, Liang, & Lauber, 1994; Keller & Koch, 2006; Proctor, Miles, & Baroni, 2011; Proctor, Yamaguchi,

Dutt, & Gonzalez, 2013; Ratcliff, 1985). To do this, each participant's RT distribution was computed separately for congruent and incongruent trials, and then the average RT in each quintile bin was calculated. The SRC effect for each bin was obtained as the difference in mean RT for congruent and incongruent trials. The data were submitted to repeated measures ANOVA with the within-subject factor bin, and between-subjects factor group. As shown in Figure 2, a main effect of bin was found, $F(1.31, 39.17) = 33.99$, $MSE = 4983,12$, $p < .001$, $\eta_P^2 = .53$, where the magnitude of the SRC effect increased by 106 ms from the first RT bin (amusic = 8 ms; control = 23 ms) to the last bin (amusic = 148 ms; control = 95 ms). This analysis also revealed a significant interaction between group and bin, $F(1.31, 39.17) = 3.91$, $MSE = 4983.12$, $p < .05$, $\eta_P^2 = .12$, with no main effect of Group, $F(1, 30) = 0.11$, $MSE = 18,856.53$, $p = .74$, $\eta_P^2 < .01$.
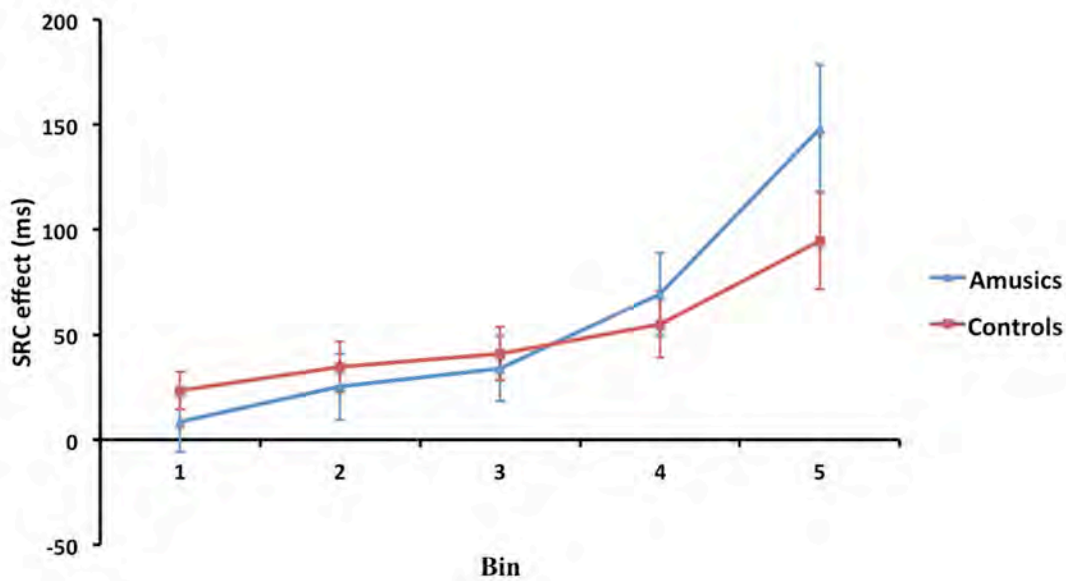
Figure 2. The binned difference response time data on stimulus-response compatibility (SRC) effect (the difference in response time between incongruent and congruent conditions) for each group (Experiment 1). Error bars represent ±1 standard error.

Discussion

Experiment 1 examined the vertical representation of pitch in individuals with or without congenital amusia using a pitch classification task. Like non-amusic participants, amusic participants represented pitch with a vertical dimension, which

generated a faster response time and fewer errors when responses to pitches were spatially congruent than when responses to pitches were spatially incongruent. Moreover, the magnitude and the temporal dynamics of the SRC effect were comparable in the two groups. These findings indicate that amusic participants are able to implement a vertical representation of pitch. However, we found that the response times across different congruency conditions were slower for amusic participants than for control participants, even though the distance between the high and low pitch was extremely large and well above amusics' threshold for pitch discrimination. This suggests the association between pitch and vertical space in congenital amusics may not be as stable as that in control individuals. On the other hand, it has been suggested that language can strengthen pre-existing associations between pitch and space (Dolscheid, Shayan, Majid, & Casasanto, 2013; Eitan, 2013). Thus, one possibility is that amusic individuals rely on semantic labels to link the pitch with spatial location, and this reliance consumes additional processing time. It is also possible that individuals with deficits in pitch perception take more time to respond to any tasks that involve auditory stimuli. To test these two possibilities, we measured the pitch-space association using an implicit SRC task, in which either no auditory stimuli were presented or pitch was irrelevant to the task.

Experiment 2: Color classification task

In Experiment 2, a color classification task was administered to disentangled semantic labels from pitch-space association. It has been reported that the SRC effect can be observed in normal listeners when they are asked to identify colours by pressing either top or bottom response keys while ignoring an accompanying pitch that is either congruent or incongruent with the spatial layout of the response keys (Keller & Koch, 2006). SRC effects based on information from irrelevant stimulus dimensions can be interpreted in the same manner as the Simon effect (Simon & Rudell, 1967; Simon & Small, 1969; see Hommel, 2011, for a review). If amusics rely on explicit semantic labels such as *"high"* or *"low"* to map pitch onto vertical positions, then they should not exhibit an association between pitch and space in a colour classification task.

Method

Participants

Sixteen amusics and 16 matched controls (12 amusics and 6 controls were from Experiment 1) participated in Experiment 2. The procedure of recruitment was the same as that in Experiment 1 (see Table 2 for details).

Table 2. Participant's characteristics, percentage correct on the melodic subtests of the MBEA, and independent-sample *t* test results between two groups in Experiment 2.

Contents	Amusic	Control	t (2-tailed)
Age (years)	23.12 ± 5.28	22.16 ± 8.16	0.39
Handedness	15R/1A	15R/1A	–
Years of education	14.81 ± 2.56	13.69 ± 2.02	1.38
Years of musical training	0.38 ± 1.09)	0.53 ± 1.14	0.40
Hours of music listening daily	2.52 ± 2.94	1.75 ± 1.35	0.95
Non-verbal IQ	103.38 ± 10.91	108.75 ± 15.56	1.13
Memory for digits	6.56 ± 1.21	7.38 ± 1.20	1.90
MBEA (average)	66.67 ± 5.24	83.47 ± 3.49	10.68 ***
Scale	71.88 ± 10.82	90.42 ± 6.65	5.84 ***
Contour	64.79 ± 8.25	83.95 ± 7.02	7.08 ***
Interval	63.33 ± 10.26	76.04 ± 7.52	4.00 ***

Note: MBEA = Montreal Battery of Evaluation of Amusia; R = right-handed; A = ambidextrous. Values are mean ± SD unless otherwise indicated. ***: $p < .001$.

Stimuli

The stimuli and apparatus were the same as those in Experiment 1, except that the duration of auditory stimuli used in Experiment 2 was 500 ms. Additionally, a red and a blue patch (90 × 90 pixels) served as task-related visual stimuli.

Procedure

The procedure was similar to that in Experiment 1. However, a colour patch was presented in the centre of the screen simultaneously with a tone in each trial. Instead

of being asked to identify whether a tone was high or low, participants were required to identify whether the colour patch was red or blue while ignoring the tone that accompanied the patch. As in Experiment 1, we were interested in the effects of congruency between the unattended auditory stimulus and the response location. However, unlike the fully blocked design in Experiment 1, congruent and incongruent trials were interspersed within blocks in the current experiment.

There were two combinations of visual stimulus and response location—that is, red-top/blue-bottom and blue-top/red-bottom. Half of the participants were randomly assigned to one combination and the other half were assigned to the other combination. In each combination, there were 96 trials divided equally into two blocks of presentations, such that half of the trials within each block had a congruent pitch-space mapping and the other half had an incongruent mapping. We also included a visual-only block of 48 trials presented at the beginning of the task that contained visual but no auditory stimuli to measure whether the two groups showed differences to colour alone without auditory stimuli.

Results

A total of 0.19% of correct trials with a RT outside the range of 100-1500 ms were classified as outliers and excluded from analysis. Two-way repeated measures ANOVA on RT and ER was conducted, with congruency as the within-subject factor and group as the between-subjects factor. Across the two groups, participants responded 12 ms faster when the correct response location was congruent with the pitch of the unattended tone than when it was incongruent, as suggested by a significant main effect of congruency, $F(1, 30) = 6.37$, $MSE = 349.47$, $p < .05$, $\eta_P^2 = .18$ (congruent: $M = 527$ ms, $SE = 13$ ms; incongruent: $M = 538$ ms, $SE = 13$ ms). There was no significant interaction between congruency and group, $F(1, 30) = 2.51$, $MSE = 349.47$, $p = .12$, $\eta_P^2 = .08$, indicating that the effect of congruency was similar in the two groups. Although the overall response time was again slightly longer in the amusic group than in the control group, the group difference was not statistically significant, $F(1, 30) = 3.03$, $MSE = 10,314.71$, $p = .09$, $\eta_P^2 = .09$. An independent t test on the visual-only block showed no difference between amusics ($M = 518$ ms, $SD = 68$ ms) and controls ($M = 504$ ms, $SD = 58$ ms), $t(30) = 0.62$, $p = .54$. As shown

in Figure 3, nine out of 32 participants showed a reversed SRC effect (5 amusic and 4 control participants). No other significant effects on RT or ER were found (all *p* > .05).
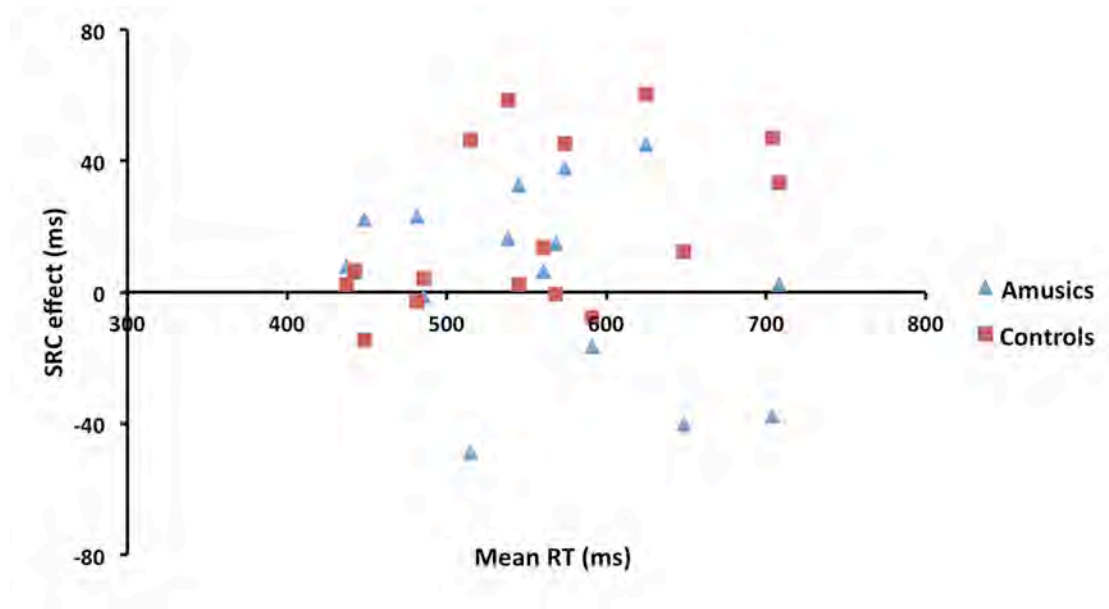
Figure 3. Stimulus-response compatibility (SRC) effect (the difference in response time, RT, between incongruent and congruent trials) for each participant as a function of mean RT (Experiment 2).

Discussion

Experiment 2 investigated whether semantic labelling underlies the pitch-space association in individuals with congenital amusia. As with controls, amusic participants in our sample exhibited a pitch-space congruency effect even though the sounded pitch was irrelevant to the task, which suggests that the SRC effect observed in the amusic group is unlikely due to semantic overlap. Furthermore, the two groups showed comparable reaction times in general, regardless of whether auditory stimuli were presented. This finding indicates that the slower response times shown by amusic participants in Experiment 1 might relate to explicit pitch processing.

General discussion

The present study examined whether the association between pitch and vertical space that is observed in Western listeners is also observed in people with congenital

amusia, and whether the mechanism that accounts for this association differs in the two populations. Although the origin of the association between pitch and space is currently unknown, it might be a consequence of internalizing naturally occurring audio-visual correlations in the environment (Miller, 1986; Pratt, 1930; Proctor & Vu, 2002). Contrary to findings reported by Douglas and Bilkey (2007), both amusic and non-amusic individuals in our sample showed faster responses and lower error rates for congruent than for incongruent mappings between pitch and key-press positions. It is possible that the discrepancy between the current findings and those of by Douglas and Bilkey (2007) relate to differences in task demands (pitch classification task vs. pitch comparison task) and/or inclusion criteria of congenital amusia (three pitch-related subtests vs. contour subtest alone). The present findings confirm that a pitch-space association does exist in amusic individuals, extending reports that an SRC effect can be observed in both musically trained and untrained individuals (Lidji et al., 2007; Rusconi et al., 2006). The temporal dynamics of the SRC effect also appears to be similar in amusic and control groups. In Experiment 1, the SRC effect increased across the RT distribution, indicating an indirect activation of stimulus-response association rather than a direct activation of the corresponding response (Vu & Proctor, 2008).

However, response times were slower for amusic than for control participants in Experiment 1, even when the pitch distance between high and low tones was well above amusics' threshold for pitch discrimination. What is the source of this delay in response times? Stage models suggest that the time taken to make a response can be estimated by the time required to accumulate evidence for a decision, plus time taken for non-decision processes, such as event detection and motor execution (Donkin, Averell, Brown, & Heathcote, 2009). It is unlikely that the slower response times can be explained by a difference in motor execution times, given that the two groups showed no difference in response times to the colour patches displayed in Experiment 2. The data from Experiment 2 further indicate that the slower response times for amusics are not due to a reliance on semantic labels to link pitch with spatial location, which consumes additional processing time, as the amusic group also showed an association between pitch height and response location, even when semantic labels were irrelevant to the task and were asked to be ignored. Instead, our data suggest that amusia is associated with slower or unstable processing at the stage

of pitch height classification, and not at other stages of processing. Moreover, given that the two groups were matched on gender, education, IQ, and working memory, it is unlikely that the increased RT can be explained by overall differences in cognitive function.

As such, we interpret the slower response times shown by amusics to reflect a slower access to spatial representation of pitch, as an outcome of an unstable representation of pitch, exacerbated by low confidence for tasks involving pitch processing. This interpretation is consistent with the proposal that amusic individuals have difficulty encoding pitch information (Patel, 2003; Tillmann, Leveque, Fornoni, Albouy, & Caclin, 2016). Low confidence on pitch-related tasks is also a prediction of the "pitch awareness deficit" hypothesis, which holds that amusic individuals are unable to access pitch information consciously (e.g., Hyde, Zatorre, & Peretz, 2011; Peretz, Brattico, Jarvenpaa, & Tervaniemi, 2009; Tillmann, Albouy, Caclin, & Bigand, 2014). More specifically, some researchers believe that amusic individuals can represent pitch information, but that the information does not reach conscious awareness, thereby resulting in poor performance and low confidence in tasks that require explicit judgements (Loui, Alsop, & Schlaug, 2009; Loui, Guenther, Mathys, & Schlaug, 2008; Lu, Ho, Liu, Wu, & Thompson, 2015; Peretz et al., 2009; Zendel, Lagrois, Robitaille, & Peretz, 2015). In support of this interpretation, Hutchins and Peretz (2013) reported a group of amusics with poor pitch perception as indicated by their performance on pitch discrimination and identification tasks, but with intact vocal production skills (also see Hutchins, Gosselin, & Peretz, 2010; Loui et al., 2008). The preserved production skills indicate that these amusic individuals show mild or no impairment in production or imitation of pitches, where the conscious processing of pitch information is not required.

Furthermore, a number of neurological investigations have observed no dysfunction in amusics at the stage of brainstem encoding of pitch (Liu, Maggu, Lau, & Wong, 2015), or at other stages of early auditory processing (Moreau, Jolicoeur, & Peretz, 2013, 2009; Peretz et al., 2009; Peretz, Brattico, & Tervaniemi, 2005). These findings suggest that congenital amusia arises from deficits at relatively late stages of processing involving conscious awareness, which may be correlated with reduced functional connectivity between the inferior frontal gyrus and the auditory cortex (Albouy et al., 2013; Hyde et al., 2011) or the reduced volume or absence of the right

arcuate fasciculus in amusics (Loui et al., 2009; but see Chen et al., 2015). Such neurological differences might also help to explain why amusic participants are slower to respond to tasks involving explicit judgements of pitch, and why they seem to have an unstable or distorted spatial representation of pitch.

Disclosure statement

No potential conflict of interest was reported by the authors.

REFERENCES

Albouy, P., Mattout, J., Bouet, R., Maby, E., Sanchez, G., Aguera, P. E., … Tillmann, B. (2013). Impaired pitch perception and memory in congenital amusia: The deficit starts in the auditory cortex. *Brain, 136*, 1639–1661.

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

Casasanto, D., Phillips, W., & Boroditsky, L. (2003). *Do we think about music in terms of space: Metaphoric representation of musical pitch.* Paper presented at the 25th annual conference of the cognitive science society, Boston, MA.

Chen, J. L., Kumar, S., Williamson, V. J., Scholz, J., Griffiths, T. D., & Stewart, L. (2015). Detection of the arcuate fasciculus in congenital amusia depends on the tractography algorithm. *Frontiers in Psychology, 6*, 9. doi: 10.3389/fpsyg.2015.00009

Connell, L., Cai, Z. G., & Holler, J. (2013). Do you see what I'm singing? Visuospatial movement biases pitch perception. *Brain and Cognition, 81*, 124–130.

Dalla Bella, S., Berkowska, M., & Sowinski, J. (2011). Disorders of pitch production in tone deafness. *Frontiers in Psychology, 2*, 164. doi:10.3389/fpsyg.2011.00164

Dalla Bella, S., Giguère, J.-F., & Peretz, I. (2009). Singing in congenital amusia. *The Journal of the Acoustical Society of America*, *126*, 414–424.

De Jong, R., Liang, C. C., & Lauber, E. (1994). Conditional and unconditional automaticity: A dual-process model of effects of spatial stimulus-response correspondence. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 731–750.

Dolscheid, S., Shayan, S., Majid, A., & Casasanto, D. (2013). The thickness of musical pitch: Psychophysical evidence for linguistic relativity. *Psychological Science, 24*, 613–621.

Donkin, C., Averell, L., Brown, S., & Heathcote, A. (2009). Getting more from accuracy and response time data: Methods for fitting the linear ballistic accumulator. *Behavior Research Methods, 41*, 1095–1110.

Douglas, K. M., & Bilkey, D. K. (2007). Amusia is associated with deficits in spatial processing. *Nature Neuroscience, 10*, 915–921.

Eitan, Z. (2013). Musical objects, cross-domain correspondences, and cultural choice: commentary on "Cross-cultural representations of musical shape" by George Athanasopoulos and Nikki Moran. *Empirical Musicology Review, 8,* 204–207.

Eitan, Z., & Granot, R. Y. (2006). How music moves. *Music Perception, 23*, 221–248.

Foxton, J. M., Dean, J. L., Gee, R., Peretz, I., & Griffiths, T. D. (2004). Characterization of deficits in pitch perception underlying 'tone deafness'. *Brain, 127*, 801–810.

Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science, 7*, 237–241.

Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology, 54*, 1–34.

Gosselin, N., Jolicoeur, P., & Peretz, I. (2009). Impaired memory for pitch in congenital amusia. *Annals of the New York Academy of Sciences*, *1169*, 270–272.

Griffiths, T. D., Rees, A., Witton, C., Cross, P. M., Shakir, R. A., & Green, G. G. R. (1997). Spatial and temporal auditory processing deficits following right hemisphere infarction. A psychophysical study. *Brain, 120*, 785–794.

Hommel, B. (2011). The Simon effect as tool and heuristic. *Acta Psychologica, 136,* 189–202.

Hutchins, S., Gosselin, N., & Peretz, I. (2010). Identification of changes along a continuum of speech intonation is impaired in congenital amusia. *Frontiers in Psychology, 1*, 236. doi: 10.3389/fpsyg.2010.00236

Hutchins, S., & Peretz, I. (2013). Vocal pitch shift in congenital amusia (pitch deafness). *Brain and Language, 125*, 106–117.

Hyde, K. L., & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science, 15*, 356–360.

Hyde, K. L., Zatorre, R. J., & Peretz, I. (2011). Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. *Cerebral Cortex, 21*, 292–299.

Jeschonek, S., Pauen, S., & Babocsai, L. (2013). Cross-modal mapping of visual and acoustic displays in infants: The effect of dynamic and static components. *European Journal of Developmental Psychology, 10*, 337–358.

Keller, P. E., & Koch, I. (2006). Exogenous and endogenous response priming with auditory stimuli. *Advances in Cognitive Psychology, 2*, 269–276.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility - A model and taxonomy. *Psychological Review, 97*, 253–270.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

Küssner, M. B., Tidhar, D., Prior, H. M., & Leech-Wilkinson, D. (2014). Musicians are more consistent: Gestural cross-modal mappings of pitch, loudness and tempo in real-time. *Frontiers in Psychology, 5*, 789. doi: 10.3389/fpsyg.2014.00789

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. doi: 10.3389/fnhum.2014.00250

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*, 1189–1207.

Liu, F., Maggu, A. R., Lau, J. C. Y., & Wong, P. C. M. (2015). Brainstem encoding of speech and musical stimuli in congenital amusia: Evidence from Cantonese speakers. *Frontiers in Human Neuroscience, 8*, 1029. doi: 10.3389/fnhum.2014.01029

Liu, F., Patel, A. D., Fourcin, A., Stewart, L. (2010). Intonation processing in congenital amusia: Discrimination, identification and imitation. *Brain, 133*, 1682–1693.

Loui, P., Alsop, D., & Schlaug, G. (2009). Tone deafness: A new disconnection syndrome? *Journal of Neuroscience, 29*, 10215–10220.

Loui, P., Guenther, F. H., Mathys, C., & Schlaug, G. (2008). Action-perception mismatch in tone-deafness. *Current Biology, 18*, R331–R332.

Lu, X., Ho, H. T., Liu, F., Wu, D., & Thompson, W. F. (2015). Intonation processing deficits of emotional words among Mandarin Chinese speakers with congenital amusia: an ERP study. *Frontiers in Psychology, 6*, 385. doi: 10.3389/fpsyg.2015.00385

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

Marin, M. M., Gingras, B., & Stewart, L. (2012). Perception of musical timbre in congenital amusia: Categorization, discrimination and short-term memory. *Neuropsychologia, 50*, 367–378.

Miller, J. (1986). Timecourse of coactivation in bimodal divided attention. *Perception & Psychophysics, 40*, 331–343.

Moreau, P., Jolicœur, P., & Peretz, I. (2009). Automatic brain responses to pitch changes in congenital amusia. *Annals of the New York Academy of Sciences, 1169*, 191–194.

Moreau, P., Jolicœur, P., & Peretz, I. (2013). Pitch discrimination without awareness in congenital amusia: Evidence from event-related potentials. *Brain and Cognition, 81*, 337–344.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia, 9*, 97–113.

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

Peretz, I., Brattico, E., Jarvenpaa, M., & Tervaniemi, M. (2009). The amusic brain: In tune, out of key, and unaware. *Brain, 132*, 1277–1286.

Peretz, I., Brattico, E., & Tervaniemi, M. (2005). Abnormal electrical brain responses to pitch in congenital amusia. *Annals of Neurology, 58*, 478–482.

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders: The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences, 999*, 58–75.

Pratt, C. C. (1930). The spatial character of high and low tones. *Journal of Experimental Psychology, 13*, 278–285.

Proctor, R. W., Miles, J. D., & Baroni, G. (2011). Reaction time distribution analysis of spatial correspondence effects. *Psychonomic Bulletin & Review, 18*, 242–266. Proctor, R. W., & Vu, K.-P. L. (2002). Eliminating, magnifying, and reversing spatial compatibility effects with mixed location-relevant and irrelevant trials. In W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action: Attention and performance* (pp. 443-473). Oxford, UK: Oxford University Press.

Proctor, R. W., & Vu, K.-P. L. (2006). *Stimulus-response compatibility principles: Data, theory, and application*. Boca Raton: CRC Press.

Proctor, R. W., Yamaguchi, M., Dutt, V., & Gonzalez, C. (2013). Dissociation of S-R compatibility and Simon effects with mixed tasks and mappings. *Journal of Experimental Psychology: Human Perception and Performance, 39*, 593–609.

Ratcliff, R. (1985). Theoretical interpretations of the speed and accuracy of positive and negative responses. *Psychological Review, 92*, 212–225.

Rusconi, E., Kwan, B., Giordano, B., Umilta, C., & Butterworth, B. (2006). Spatial representation of pitch height: The SMARC effect. *Cognition, 99*, 113–129.

Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information-processing: A review of 50 years of research. *Personality and Individual Differences, 44*, 535–551.

Simon, J. R., & Rudell, A. P. (1967). Auditory S–R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology, 51,* 300–304.

Simon, J. R., & Small, A. M. Jr. (1969). Processing auditory information: Interference from an irrelevant cue. *Journal of Applied Psychology, 53*, 433–435.

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*, 37–50.

Stewart, L., & Walsh, V. (2007). Music perception: Sounds lost in space. *Current Biology, 17*, R892–R893.

Taylor, J. E., & Witt, J. K. (2014). Listening to music primes space: Pianists, but not novices, simulate heard actions. *Psychological Research, 79*, 175–182.

Thompson, W. F., Marin, M. M., & Stewart, L. (2012). Reduced sensitivity to emotional prosody in congenital amusia rekindles the musical protolanguage hypothesis. *Proceedings of the National Academy of Sciences, 109*, 19027–19032.

Tillmann, B., Albouy, P., Caclin, A., & Bigand, E. (2014). Musical familiarity in congenital amusia: evidence from a gating paradigm. *Cortex, 59*, 84–94.

Tillmann, B., Jolicoeur, P., Ishihara, M., Gosselin, N., Bertrand, O., Rossetti, Y., & Peretz, I. (2010). The amusic brain: Lost in music, but not in space. *PLoS One, 5,* e10173.

Tillmann, B., Leveque, Y., Fornoni, L., Albouy, P., & Caclin, A. (2016). Impaired short-term memory for pitch in congenital amusia. *Brain Research, 1640*, 251–263.

Tillmann, B., Schulze, K., & Foxton, J. (2009). Congenital amusia: A short-term memory deficit for non-verbal, but not verbal sounds. *Brain and Cognition, 71*, 259–264.

Vu, K.-P. L., & Proctor, R. W. (2008). Age differences in response selection for pure and mixed stimulus-response mappings and tasks. *Acta Psychologica, 129*, 49–60.

Wagner, S., Winner, E., Cicchetti, D., & Mardner, H. (1981). "Metaphorical" mapping in human infants. *Child Development, 52*, 728–731.

Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science, 21*, 21–25.

Williamson, V. J., Cocchini, G., & Stewart, L. (2011). The relationship between pitch and space in congenital amusia. *Brain and Cognition, 76*, 70–76.

Williamson, V. J., & Stewart, L. (2010). Memory for pitch in congenital amusia: Beyond a fine-grained pitch discrimination problem. *Memory, 18*, 657–669.

Zendel, B. R., Lagrois, M. E., Robitaille, N., & Peretz, I. (2015). Attending to pitch information inhibits processing of pitch information: The curious case of amusia. *The Journal of Neuroscience, 35*, 3815–3824.

Chapter 4: Control Experiment

The two experiments described in Chapter 3 (manuscript 2), demonstrate that individuals with congenital amusia exhibit unstable spatial representations of pitch when tasks involved explicit judgments of pitch, as revealed by slower response times in the pitch task but not in the colour classification task. However, the slower response times may simply reflect a general difference between amusic and non-amusic individuals in the perceived magnitude of pitch changes, even when the interval is well above the threshold of discrimination for both groups. To rule out this possibility, this chapter describes an experiment on whether the perceived magnitude of supra-threshold pitch changes differs significantly for amusic and non-amusic individuals.

Comparison of magnitude estimations of pitch change in amusic and non-amusic individuals

In comparison to *frequency* - an objective physical quantity - *pitch* is a subjective psychoacoustic attribute of sound that cannot be directly measured. For listeners with normal hearing, the perception of pitch is approximately logarithmic with respect to frequency. In other words, category labels that describe the perceived intervals between the equivalent log frequencies are the same. For instance, the perceived interval size between 220 Hz (A_3) and 440 Hz (A_4) is the same as that between 196 Hz (G_3) and 392 Hz (G_4). However, the psychophysical relation between pitch and frequency may not be logarithmic (Stevens, Volkmann, & Newman, 1937). The estimates of interval size can be affected by level of music skill (Russo & Thompson, 2005).

Given that individuals with congenital amusia have difficulty in pitch perception (Ayotte, Peretz, & Hyde, 2002; Foxton, Dean, Gee, Peretz, & Griffiths, 2004), it is possible that the perceived magnitude of the interval separating two pitches is not equivalent for amusic and non-amusic individuals. For example, pitch differences that are close to the discrimination threshold of amusic individuals – in that they are virtually indistinguishable – are likely to be perceived as "very small" by amusic individuals but not by non-amusic individuals who can readily distinguish the pitches (see also, Jiang, Lim, Wang, & Hamm, 2013). On the other hand, it is unknown whether larger intervals that are well above the threshold of discrimination for both groups are still perceived as "smaller" by amusic than control participants. If there is a general difference in the perceived size of pitch intervals for amusic and non-amusic individuals, then the group differences reported in Chapter 3 may reflect this overall difference in perceived interval size, as the judgements of tones closer in pitch yields longer reaction times than more distant tones (Elkin & Leuthold, 2011). In particular, "perceived interval size" could predict responses to pitch change direction equally for amusic and non-amusic participants, which would eliminate the need to posit anomalies in spatial representations of pitch by amusic individuals.

To examine whether the perceived magnitude of supra-threshold pitch changes is equivalent for amusic and non-amusic individuals, participants were presented with pairs of tones and asked to estimate the interval size between these two tones.

Method

Participants

Congenital amusia was diagnosed using three melodic subtests (Scale, Contour, and Interval) of the Montreal Battery of Evaluation of Amusia (MBEA; Peretz, Champod, & Hyde, 2003). Individuals with composite scores equal or less than 65 out of 90 points (i.e., 72% correct) were considered as congenital amusics. With this criterion, eight individuals with congenital amusia and their matched control participants took part in this experiment. None reported any auditory, neurological, or psychiatric disorders. The amusic and control group were matched in terms of age, gender, years of education, years of musical training, and hours of musical listening (see Table 1). The study was approved by the Macquarie University Ethics Committee, and written informed consent for participation was obtained from all participants prior to testing.

Table 1. Participants' characteristics, mean $\pm$ SD percentage correct on the melodic subtests of MBEA, and independent-sample t-test results between two groups. DF refers to the degrees of freedom and is corrected if the equal variances assumption is violated. *: $p < .05$; ***: $p < .001$.

	Amusics	**Controls**	**DF**	**t**	**p (2-tailed)**
Age	24.21 ± 5.32	25.33 ± 4.96	14	0.44	.67
Gender	4F/4M	4F/4M	-	-	-
Year of education	15.50 ± 1.60	16.00 ± 3.16	10.38	0.40	.70
Years of musical training	0.19 ± 0.53	0.25 ± 0.53	14	0.24	.82
Hours of music listening daily	1.93 ± 1.94	1.75 ± 0.89	14	- 0.24	.82
MBEA (percentage correct)					
Scale	0.71 ± 0.05	0.93 ± 0.05	14	8.25 ***	$< .001$
Contour	0.67 ± 0.11	0.83 ± 0.12	14	2.87 *	.01
Interval	0.65 ± 0.05	0.88 ± 0.06	14	8.48 ***	$< .001$
Total	0.68 ± 0.05	0.88 ± 0.06	14	7.06 ***	$< .001$

Stimuli

The fundamental frequency of each tone used in this experiment ranged from 103.83 Hz ($G^{\#}_2$) to 659.26 Hz (E_5), and was separated by one semitone from its nearest neighbour, yielding a total of 33 tones. All tones were generated using the computer software GarageBand (Version 6.0.4; Apple Inc., U.S.A.) with flute timbre. Tones were 1000 ms in duration, including a 50 ms ramp onset and offset. Each tone was assigned a number within the range of 10 to 170 stepped by 5; that is, the smallest number (10) was assigned to the lowest pitch ($G^{\#}_2$) and the largest number (170) was assigned to the highest pitch (E_5).

The set of tones was then used to create 13 melodic intervals (two-tone sequences) – 1 to 12, and 27 semitones. These intervals were selected because they had been used in other experiments of the present thesis (see Chapters 3 and 5). Specifically, $G^{\#}_2$, $G^{\#}_3$, C_3, C_4, E_3 and E_4 served as the starting note of six chromatic octaves. The melodic intervals of 1 to 12 semitones were created within each chromatic octave, and the melodic intervals of 27 semitones were created using $G^{\#}_2$-B_4, A_2-C_5, $A^{\#}_2$-$C^{\#}_5$, B_2-D_5, C_3-$D^{\#}_5$, and $C^{\#}_3$-E_5. Therefore, there were six two-tone sequences for each melodic interval, yielding 78 melodic intervals in total.

Procedure

Participants were tested in a quiet room. At the beginning of each trial, participants were instructed to fixate on the centre of the display. The fixation point appeared for 1000 ms, followed by the first tone with its assigned number appearing on the display for 1000 ms. After an inter-stimulus interval of 1000 ms the second tone was presented for 1000 ms. Participants were instructed that the number "0" represents the lowest pitch they can imagine, and were asked to input a number to best describe the pitch of the second tone they heard in relation to the given number of the first tone. The inter-trial interval was 500 ms.

All melodic intervals were presented twice – in ascending and descending pitch direction. Therefore, each participant completed a total of 156 experimental trials. All trials were scrambled and presented randomly in three equal blocks for each participant. Ten random trials were presented prior to the formal experiment to

familiarise participants with the task. All sounds were presented at a comfortable level (50 dB SPL) via headphones.

Results and Discussion

The perceived amount of pitch change was evaluated using the difference score between the estimated number of the second tone and the given number of the first tone. The average of difference scores for each interval was then calculated (i.e., mean magnitude estimation of interval size). A repeated-measure ANOVA was conducted with the between-subject factor of Group and the within-subject factor of Interval. Greenhouse-Geisser adjustments were used in light of observed violations of sphericity, and the corrected degrees of freedom was reported, where appropriate. A main effect of Interval was found, $F(1.86, 26.07) = 99.14$, $p < .001$, $\eta_p^2 = .88$. As shown in Fig. 1, both groups showed a trend for increased estimation of interval size as the actual interval size became larger. Neither the group difference, $F(1, 14) = 1.72$, $p = .21$, $\eta_p^2 = .11$, nor interaction between Group and Interval was significant, $F(1.86, 26.07) = 0.48$, $p = .61$, $\eta_p^2 = .03$. The results reveal that the perceived amount of pitch change is equivalent between amusic and non-amusic individuals.
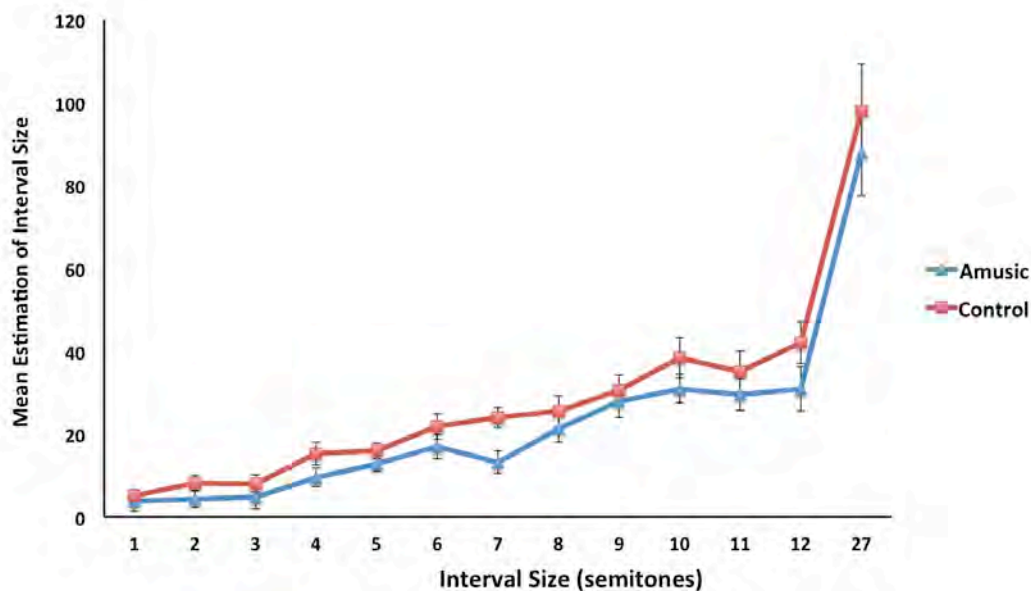
Fig.1 – Mean estimation of interval size by amusic and non-amusic participants of the intervals in semitones.

Conclusion

The present experiment investigated whether the psychological distance between two given pitches differs between amusic and control participants by asking them to provide magnitude estimations of the interval between two tones presented sequentially. The results indicated that amusic individuals perceive pitch changes equivalently to control individuals when intervals are above their threshold for pitch discrimination, suggesting that amusic individuals have similar perceptions of pitch interval size as long as the separation between two tones exceeds their pitch discrimination thresholds. The finding suggests that the impairments in music shown by individuals with congenital amusia cannot be accounted for by a low-level perceptual problem.

References

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

Elkin, J., & Leuthold, H. (2011). The representation of pitch in auditory imagery: Evidence from S-R compatibility and distance effects. *Journal of Cognitive Psychology,* 23(1), 76-91. doi: 10.1080/20445911.2011.455251

Foxton, J. M., Dean, J. L., Gee, R., Peretz, I., & Griffiths, T. D. (2004). Characterization of deficits in pitch perception underlying 'tone deafness'. *Brain, 127*(4), 801-810. doi: 10.1093/brain/awh105

Jiang, C., Lim, V. K., Wang, H., & Hamm, J. P. (2013). Difficulties with pitch discrimination influences pitch memory performance: Evidence from congenital amusia. *PLoS One, 8*(10), e79216. doi: 10.1371/journal.pone.0079216.t001

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders: The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences, 999*(1), 58-75. doi: 10.1196/annals.1284.006

Russo, F. A., & Thompson, W. F. (2005). The subjective size of melodic intervals over a two-octave range. *Psychonomic Bulletin & Review, 12*(6), 1068-1075. doi: 10.3758/BF03206445

Stevens, S. S., Volkmann, J., & Newman, E. B. (1937). A scale for the measurement of the psychological magnitude pitch. *The Journal of the Acoustical Society of America, 8*(3), 185-190. doi: 10.1121/1.1915893

Chapter 5: Pitch Contour Processing

Chapter 3 (manuscript 2) revealed that individuals with congenital amusia exhibit unstable spatial representations of pitch when tasks involved explicit judgments of pitch. However, the reason for this unstable representation of pitch is difficult to ascertain because the task that is used to diagnose this disorder – the MBEA – requires both intact perceptual processing and intact short-term memory. For example, it could be that amusic participants did poorly because they could not perceive the pitches properly, and were therefore unable to represent them. However, it is also possible that amusic participants perceived the pitches accurately, but then could not retain them in short-term memory long enough to perform the task at the levels achieved by the control participants. This latter possibility is supported by evidence from Williamson and Stewart (2010) (see also, Gosselin et al., 2009; Tillmann et al., 2009), showing that many amusic individuals do indeed show deficits in short-term memory for pitch. To disentangle these two possibilities, a new task that places minimal demands on pitch memory was devised – called the Self-paced Audio-visual Contour Task (SACT). Using this task, I examined contour processing in both the pitch dimension (Experiment 1) and in two non-pitch dimensions (Experiment 2).

The manuscript that describes this study (manuscript 3) was co-authored by Yanan Sun, Tam Ho, and Bill Thompson. My contribution to the manuscript was roughly 70%, including experimental design, data collection, analysis, and preparation of the first draft. The pitch task was jointly designed by Bill Thompson, Tam Ho and myself. Yanan Sun provided assistance in participant recruitment. All co-authors made critical comments on the manuscript and helped with interpretation of the data.

This chapter was prepared as:

Lu, X., Sun, Y., Ho, H. T., & Thompson, W. F (under review). Pitch contour impairment in congenital amusia: New insights from the Self-paced Audio-visual Contour Task (SACT).

PITCH CONTOUR IMPAIRMENT IN CONGENITAL AMUSIA: NEW INSIGHTS FROM THE SELF-PACED AUDIO-VISUAL CONTOUR TASK (SACT)

Xuejing Lu [a,b,]*, Yanan Sun [b,c], Hao Tam Ho [a], William Forde Thompson [a,b]

[a] Department of Psychology, Macquarie University, Sydney, NSW, Australia
[b] ARC Centre of Excellence in Cognition and its Disorders, NSW, Australia
[c] Department of Cognitive Science, Macquarie University, Sydney, NSW, Australia

* Corresponding author at: Department of Psychology, Macquarie University, NSW 2109, Australia.

E-mail address: xuejing.lu@mq.edu.au

Abstract

Individuals with congenital amusia usually exhibit impairments in melodic contour processing when asked to compare pairs of melodies that may or may not be identical to one another. However, it is still unclear whether the impairment observed in contour processing is mainly caused by a failure in pitch change direction identification, or results from poor pitch memory. To disentangle these two stages of contour processing, the present study examined the perceptual stage of contour processing independently from the storage stage. In Experiment 1, we employed a Self-paced Audio-visual Contour Task (SACT) that placed minimal or no burden on memory. Participants controlled the pace of an auditory contour that was simultaneously accompanied by a visual contour, and judged whether the two contours were congruent. Compared with control participants, amusics showed reduced sensitivity to audio-visual congruency, even when the intervals between consecutive tones in each sequence were well above their threshold for pitch discrimination. This finding suggests that impaired contour processing in congenital amusia arises from a perceptual stage of processing, but cannot be explained as a problem of pitch discrimination. Such impairment may reflect a failure to consciously access knowledge of pitch direction, but it could also arise from a more general deficit in cross-modal mapping. To exclude the latter possibility, and to confirm that the impairment is restricted to the pitch dimension, we administered the cross-modal mapping task using two other auditory dimensions: loudness and timbral brightness in Experiment 2. These dimensions were manipulated by adjusting the intensity and spectral centroid of sounds while fixing their fundamental frequencies. Results showed that amusics performed as well as controls when the auditory contour varied in loudness or timbral brightness, indicating that the impairment observed in Experiment 1 was not driven by a general impairment in cross-modal mapping. The results indicate that impaired contour processing in congenital amusia is a perceptual problem that is restricted to the pitch dimension, and occurs between pitches that are above discrimination thresholds.

Keywords

congenital amusia, cross-modal mapping, auditory dimension, contour processing, the perceptual stage

1. Introduction

Congenital amusia is a disorder of music perception that has been the subject of considerable research and theory (Ayotte, Peretz, & Hyde, 2002; see Peretz, 2013 for a review). Although much remains to be understood about this rare disorder, it is generally agreed that amusic individuals exhibit difficulties in fine-grained pitch perception (see Vuvan, Nunes-Silva, & Peretz, 2015 for meta-analysis results), as revealed by elevated thresholds for pitch change detection and discrimination at a group level, when compared with normal listeners (Foxton, Dean, Gee, Peretz, & Griffiths, 2004; Hyde & Peretz, 2004; Jiang, Lim, Wang, & Hamm, 2013; Tillmann, Schulze, & Foxton, 2009). Despite this difficulty, people with amusia are still able to name and recognise voices and environmental sounds, and they can identify and discriminate speech intonation that involves coarse changes in pitch (Ayotte et al., 2002; Patel, Foxton, & Griffiths, 2005). These findings have led to the proposal that the core deficit of congenital amusia is impaired low-level fine-grained pitch processing (Peretz & Hyde, 2003).

Curiously, when amusics were presented with the prosodic aspect of spoken sentences in the absence of linguistic content (i.e., non-speech analogues in which semantic content was removed while preserving pitch contour), they were significantly impaired as discriminating speech prosody relative to their own performance on spoken sentences (Patel et al., 2005) and compared to performance by control participants (Ayotte et al., 2002). Given that non-speech analogues share the same pitch pattern as spoken sentences, this dissociation challenged the hypothesis that the core deficit of congenital amusia is restricted to impaired fine-grained pitch processing. To explain the finding, Patel et al. (2005) proposed the "*Melodic Contour Deafness Hypothesis*", which suggests that the principle deficit of congenital amusia lies at a higher level of processing and does not arise from a low-level impairment in pitch.

Research generally supports Patel et al. (2005) suggestion that individuals with congenital amusia exhibit difficulties at a relatively high level of processing, such as melodic contour processing (Foxton et al., 2004; Jiang, Hamm, Lim, Kirk, & Yang, 2010), although the source of this difficulty is still under debate. Melodic contour refers to the rising and falling pattern of intervals within a melody, and indicates whether adjacent notes are higher or lower than one another (Jones, 1976). Unlike

intervals, which represent the exact difference between two pitches, contour only represents a sequence of the direction of pitch changes (but see, Prince, Schmuckler, & Thompson, 2009a). On hearing a piece of music, people remember little about the absolute pitches or the precise pitch intervals (i.e., interval processing), but are good at remembering the pattern of upward and downward shifts between notes (i.e., contour processing) (Attneave & Olson, 1971; Dowling, 1978; Dowling & Fujitani, 1971; Edworthy, 1985; Hébert & Peretz, 1997). The ability to identify pitch change direction is considered to be one of the fundamental outcomes of melodic contour processing, as efficient contour processing requires accurate perception of pitch changes. Another fundamental outcome is the ability to retain a succession of pitch changes in memory, which is crucial to melody recognition (Dowling, 1978; Dowling & Fujitani, 1971; Mikumo, 1992; Trehub, Bull, & Thorpe, 1984).

From this perspective, there are two plausible explanations for the impaired melodic contour processing observed in congenital amusia. One possibility is that amusic individuals have a reduced sensitivity to the direction of pitch movement at a perceptual stage of processing (Liu, Patel, Fourcin, & Stewart, 2010). This reduced sensitivity, in turn, leads to an unstable and distorted mental representation of musical pitch. Alternatively, the disorder may be caused by a failure to store and retain pitch information in memory (Albouy, Schulze, Caclin, & Tillmann, 2013; Gosselin, Jolicoeur, & Peretz, 2009; Tillmann et al., 2009). Early psychophysical models of pitch memory suggest that pitch is mentally represented in the form of a helix, having both linear (height) and circular (proximity) dimensions, that is, pitch height and pitch chroma. The helical model accounts for the observation that pitches are perceived to be similar if they are (a) proximal in fundamental frequency and (b) separated by an octave (i.e., they have the same chroma). Later models have incorporated psychological relationships between pitches based on their tonal function within a musical context (Krumhansl, 1979; Krumhansl & Kessler, 1982).

A major challenge of investigating pitch deficits in congenital amusia is that many musical tasks require both pitch perception and pitch memory. For example, performance on tasks that involve a comparison between two successive events can be explained by deficits at a perceptual stage, a memory stage (one must hold the first event in memory in order to compare it to the second event), or both. Thus, to disentangle perceptual and storages stages of processing, we developed a Self-paced

Audio-visual Contour Task (SACT), which was designed to place minimal or no burden on short-term memory. As will be described in the *Methods* section, melodic and visual contours were presented simultaneously at a rate that was self-paced by each participant. Participants were asked to judge whether melodic and visual contours were congruent with one another (i.e., on-line matching).

The decision to combine melodic and visual contours was motivated by two lines of evidence for a cross-modal connection between audio and visual domains. One line of evidence comes from cross-modal dimensional interactions at the psychophysical level. For instance, sounds with high frequency (i.e., pitch) or intensity (i.e., loudness) are associated with a higher spatial location (i.e., height) and bright lights or colours (i.e., visual brightness), whereas those with low frequency or intensity are associated with a lower spatial location and dim lights or colours (Eitan, Schupak, & Marks, 2008; Lewkowicz & Turkewitz, 1980; Lidji, Kolinsky, Lochy, & Morais, 2007; Marks, 1982, 1987, 1989; Marks, Hammeal, Bornstein, & Smith, 1987; Melara, 1989; Rusconi, Kwan, Giordano, Umilta, & Butterworth, 2006). Another line of evidence comes from the similarity of auditory and visual contours. For instance, untrained listeners can create visual representations of melodic contour (Prince, Schmuckler, & Thompson, 2009b) and make fairly accurate drawings of the contours of melodies they have just heard (Davies & Roberts, 1975). These findings suggest that there is an intermodal association between melodic and visual contour (Mikumo, 1997).

As well as enabling us to analyse explicit judgements of audio-visual congruence, the SACT also allowed us to examine implicit contour processing. It has been suggested that the brainstem and auditory cortex function normally in amusia (Liu, Maggu, Lau, & Wong, 2015; Moreau, Jolicœur, & Peretz, 2009; Peretz, Brattico, Jarvenpaa, & Tervaniemi, 2009; Peretz, Brattico, & Tervaniemi, 2005; but see, Lehmann, Skoe, Moreau, Peretz, & Kraus, 2015), and hence amusic individuals typically exhibit preserved implicit knowledge of tonal structure (Albouy, Schulze, et al., 2013; Omigie, Pearce, & Stewart, 2012; Tillmann, Gosselin, Bigand, & Peretz, 2012). These findings suggest that congenital amusia may result from a disconnection between auditory perceptual system and higher processes that involves conscious awareness (Loui, Alsop, & Schlaug, 2009; Loui, Guenther, Mathys, & Schlaug, 2008; Lu, Ho, Liu, Wu, & Thompson, 2015; Peretz et al., 2009; Tillmann,

Albouy, Caclin, & Bigand, 2014; Zendel, Lagrois, Robitaille, & Peretz, 2015). In the self-paced paradigm employed, the pace of the melodic and visual contour was controlled by self-pacing, and participants indicated explicitly at the end of each trial whether the melodic and visual contours were the same or different. If individuals with congenital amusia can identify the pitch change direction implicitly, then a slower self-paced rate might be expected in response to incongruent audio-visual contours than in response to congruent ones, even if the incongruent trials were not noticed explicitly (i.e., incongruent trials with incorrect explicit responses).

The SACT also allowed us to examine the capacity of individuals with congenital amusia to process contour in other auditory dimensions. Although the concept of melodic contour has traditionally been applied to melodies consisting of a sequence of tones that vary in pitch, it can also be perceived and used for recognition of familiar melodies in dimensions other than pitch (Graves, Micheyl, & Oxenham, 2014; McDermott, Lehr, & Oxenham, 2008), for example with timbral brightness (a dimension that relates to a sound's spectral centroid) and loudness (a dimension that relates to a sound's intensity). Brightness is one of the most salient dimensions of timbre (McAdams, Winsberg, Donnadieu, De Soete, & Krimphoff, 1995), and reflects the spectral profile of the sound. The perception of brightness correlates with the centre of mass of the frequency spectrum. Sounds with more energy in the high-frequency range of the spectrum are perceived as brighter, whereas sounds with more energy in the low-frequency range are perceived as duller, even when they have the same fundamental frequency (F0). In other words, brightness can be varied independently of the F0 (i.e., pitch height). Loudness, on the other hand, is a non-spectral dimension, and reflects the intensity of a sound. To the best of our knowledge, no study has tested whether individuals with congenital amusia show impairments in contour processing in either spectrally related or spectrally unrelated dimensions rather than in just the pitch dimension. If the contour processing impairment were restricted to the pitch dimension, then individuals with congenital amusia should be able to discriminate patterns of change in other dimensions. However, if problems in contour processing extend to other auditory dimensions, it would imply that individuals with congenital amusia have a more abstract level of pattern perception impairment.

The aim of this study was to address three research questions: a) Does the impairment in melodic contour processing in congenital amusia occur at a perceptual

or a short-term memory stage of processing? b) Does the impairment occur at an implicit level of processing, and explicit level of processing, or both, and c) Are the impairments observed among individuals with congenital amusia restricted to the pitch dimension, or do they extend to non-pitch dimensions?

To address the first question, we examined melodic contour processing by experimentally separating perceptual and storage stages of processing in Experiment 1. Height-varying visual stimuli served as visual contours in this experiment, as the link between pitch and elevation is one of the most robust cross-modal correspondences (Spence, 2011). If amusics' impairment in contour processing arises because of limits in auditory short-term memory, then they should be able to perceive up-down patterns within a melody as easily as controls in this on-line matching task, given that the SACT places minimal or no load on memory. On the other hand, if amusic individuals were unable to extract the direction of changes between adjacent notes, they should fail to detect any incongruence between melodic and visual contours. To address the second question, we compared judgements of audio-visual congruence with the rate of self-pacing for congruent and incongruent trials, thereby shedding light on explicit and implicit level of processing, respectively. To address the third question, we experimentally manipulated the spectral centroid and intensity of the auditory stimuli in order to create auditory brightness and loudness contours. This paradigm was adopted in Experiment 2 to enable us to investigate the cross-modal mapping capacity of individuals with congenital amusia for auditory dimensions other than pitch.

2.   Experiment 1

2.1. Methods

2.1.1. Participants

The Montreal Battery of Evaluation of Amusia (MBEA; Peretz, Champod, & Hyde, 2003) has been widely used for diagnosing amusia over the past decade. Three melodic subtests (Scale, Contour, and Interval) of the MBEA were administered, as we were specifically interested in amusic individuals with deficits in pitch processing. For each of the melodic subtests, listeners were presented with pairs of melodies and

asked to judge whether they were the "same" or "different". Participants in the amusic group were 14 individuals with composite scores at or below 65 out of 90 scores (72% correct). Another 14 participants with composite scores at or above 72 out of 90 scores (80% correct) comprised the control group. All participants reported normal hearing and normal or corrected-to-normal vision. None reported any auditory, neurological, or psychiatric disorder. As shown in Table 1, although amusics performed significantly worse than controls in three melodic tests ($ps <$ 0.001), the two groups were matched in age, gender, years of education, years of musical training and hours of music listening ($ps > .1$). Written informed consent for participation was obtained from all participants prior to the study. This experimental protocol was approved by Macquarie University Ethics Committee.

Table 1 – Participants' characteristics, correct rates (mean ± SD) on the melodic subtests of MBEA, and independent-sample t-test results between two groups in Experiment 1. DF refers to the degrees of freedom and was corrected if the equal variances assumption was violated. ***: $p < .001$.

	Amusics	Controls	DF	t	p (2-tailed)
Age	23.22 ± 6.02	25. 46 ± 8.39	26	-0.81	.42
Gender	9F/5M	9F/5M	-	-	-
Years of education	14.21 ± 1.97	14.57 ± 2.31	26	-0.44	.66
Years of musical training	0.36 ± 0.63	1.00 ± 1.40	18.10	-1.57	.14
Hours of music listening daily	2.52 ± 2.96	2.75 ± 2.48	26	-0.23	.82
MBEA (percentage correct)					
Scale	0.76 ± 0.07	0.92 ± 0.06	26	-6.08 ***	< .001
Contour	0.68 ± 0.08	0.89 ± 0.05	26	-8.08 ***	< .001
Interval	0.61 ± 0.08	0.86 ± 0.07	26	-9.11 ***	< .001
Composite	0.69 ± 0.04	0.89 ± 0.04	26	-13.24 ***	< .001

2.1.2. Pre-tests

2.1.2.1.   Stimuli and procedure

To further explore the relationship between pitch perception and melodic contour processing, we measured the thresholds for pitch discrimination and pitch

direction identification using a transformed "3-down 1-up"[1] staircase design with an expected accuracy rate of 0.79 (Levitt, 1971), implemented in the PSYCHOACOUSTICS toolbox (Soranzo & Grassi, 2014) in MATLAB (Version 2013b; Mathworks, U.S.A.). In the pitch discrimination task, participants were presented with a sequence of three steady-state pure tones of 500 ms separated by 500 ms of silence. Two tones were identical in frequency and the third tone was either higher or lower than the others. Participants were required to judge which of the three tones was the oddball by pressing one of three response keys referring to the first, second, or third tone. In the pitch direction identification task, only two tones were presented. Participants were asked to identify which one was higher in frequency by pressing one of two response keys. The procedure was terminated after 12 reversals and the threshold was then calculated as the arithmetic mean excursion value of the pitch intervals for the last eight reversals. Feedback was provided after each trial.

2.1.2.2.   Results

Independent-samples t-tests were conducted on log-transformed thresholds between two groups, as the raw data (see Table 2) were not distributed normally (One-sample Kolmogorov-Smirnov test, $ps < .001$). In line with previous studies (Foxton et al., 2004; Hyde & Peretz, 2004; Liu et al., 2010; Tillmann et al., 2009), amusics' thresholds for pitch discrimination and pitch direction identification were significantly higher than controls' at the group level, $t(19.34) = 2.25$, $p = .03$ and $t(18.37) = 2.07$, $p = .05$, respectively.

Table 2 – Thresholds in cents (Mean ± SD) for pitch discrimination and direction identification of each group.

Task	Amusics	Controls
Pitch discrimination	133.62 ± 165.93	18.98 ± 15.24
Pitch direction identification	104.28 ± 160.81	13.16 ± 13.53

---

[1] In the present study, three consecutive correct responses were required for a reduction in pitch excursion of the target while a single incorrect response resulted in an increase in pitch excursion, defined as a reversal.

2.1.3. Stimuli

As illustrated in Fig. 1, auditory stimuli were seven-note tonal sequences. Each auditory stimulus was accompanied by a visual presentation of a sequence of dots – one for each note of the auditory sequence. The fourth, fifth or sixth tone of each auditory sequence served as a target tone, in which an audio-visual incongruence could occur. The chance of incongruence occurrence was equal at each position. Moreover, the size of the interval separating the preceding and target tones was systematically varied from one to twelve semitones stepped by one semitone, where incongruence should be most obvious for large interval sizes. There were six trials
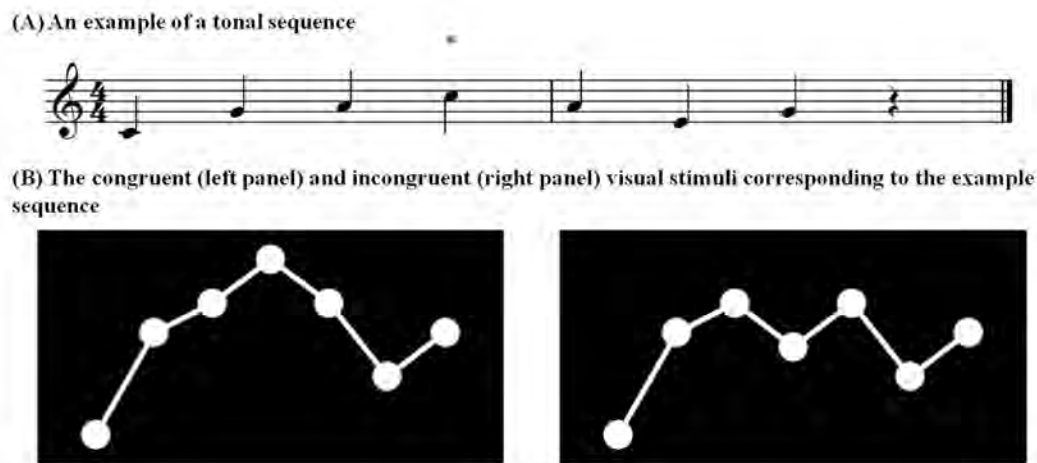
Fig. 1 – Illustration of the (A) auditory and (B) visual stimuli. * indicates the note that is either congruent (left panel) or incongruent (right panel) with the visual representation. In the incongruent condition, the change in the vertical position of dots is inconsistent with the change in direction of pitches in the accompanying melody.

for each interval size condition – half of them with upward changes and the rest with downward changes. All sequences were tonally structured and constructed using tones from the C major scale. The tonality of each tone sequence was measured using the "key-finding algorithm" (Krumhansl, 1990) implemented in the MATLAB MIDI toolbox (Eerola & Toiviainen, 2004), where the maximum positive correlation provides an indication of the most strongly established key. As expected, the maximum correlation for each tone sequence was with the C major key, $r(10) = .76$ (range from .45-.92), $ts > 3.75$, $ps < .05$. Furthermore, to ensure that results were not affected by the tonal stability of notes, we compared the degree of fit (i.e., stability

value) of the target notes and notes immediately preceding the target notes with respect to the local key context (C major in this case) for each interval size and change direction. The stability of the target tones did not differ from that of the immediately preceding tones in any condition (*ps* > .05). Following these rules, 72 short sequences were constructed. To ensure results could be generalized beyond a C major key context, all tones in one third of sequences were shifted up and in another one third were shifted down by four semitones. All tones were generated using the computer software GarageBand (Version 6.0.4; Apple Inc., U.S.A.) with flute timbre, and a duration of 500 ms.

Visual stimuli consisted of seven white dots (50-pixel diameter; screen resolution: 1980 × 1024 pixels) that varied in both vertical and horizontal spatial location and that were displayed on a black background and connected by short lines. Each dot corresponded to a single note in the tonal sequence, with pitch height represented as a location on the vertical dimension, and temporal position represented as a spatial location on the horizontal dimension. The first dot was fixed on the left vertical centre of the display, and the rest were presented successively from left to right simultaneously with the presentation of each tone in the auditory sequence. Two visual stimulus sets were constructed such that melodic and visual contours were either congruent or incongruent with one another.

2.1.4. Procedure

Participants were tested in a quiet and dimly lit room. Each trial started with a fixation on the left vertical centre of the display for 500 ms, at which time the first tone and dot were presented simultaneously. Participants were required to push the spacebar on the computer keyboard at a comfortable pace to trigger the following tones and dots, one after another. After the presentation of the whole sequence, participants were asked to make a non-speeded judgement on whether the melodic and visual contours were congruent with one another ("yes" or "no") by pressing one of two response keys. Assignment of the two response keys for congruent and incongruent trials was counterbalanced across participants. For the congruent trials, melodic and visual contours always matched; for incongruent trials, the dot in the visual contour went in the opposite direction to the direction implied by the pitch change. Twelve practice trials were presented prior to the formal experiment. Feedback was provided during the practice trials but not the experimental trials.

Instead, participants were asked to rate their confidence level for their judgement on a five-point scale (1 = "no confidence at all"; 5 = "complete confidence"). The congruent and incongruent trials were scrambled independently for each participant, and assigned to six equal blocks. Participants were encouraged to take a short break after each block to minimize fatigue effects. Auditory stimuli were delivered via noise-cancelling headphones (Sennheiser PXC 350) at a comfortable hearing level. The experiment was programmed and presented in SuperLab 4.5 (Cedrus Corporation, U.S.A.).

2.2. Results

Participants' judgements of audio-visual congruence and self-pacing times for each tone (i.e., inter-stimulus interval, ISI) were recorded for subsequent analysis. For analysis purposes, we classified the interval sizes into three categories: small (1-4 semitones), medium (5-8 semitones) and large (9-12 semitones). This procedure allowed us to reveal broad effects of interval size on the detection of audio-visual incongruence[2].

### 2.2.1. Judgements of audio-visual congruence (pitch dimension)

The audio-visual judgements were evaluated using d-prime (d') (Macmillan & Creelman, 2005). A response of "no" on an incongruent trial was defined as a "hit" and on a congruent trial was defined as a "false-alarm". A repeated-measures analysis of variance (ANOVA) was conducted with the between-subject factor of Group (amusic and control) and the within-subject factor of Interval size (small, medium, and large). Greenhouse-Geisser adjustments were used in light of observed violations of sphericity, and the Bonferroni correction was used in post-hoc analysis, where appropriate.

As shown in Fig. 2, amusic individuals performed worse than control participants in all interval size conditions, as revealed by a significant main effect of Group, $F(1, 26) = 11.73$, $p < .01$, $\eta_p^2 = .31$. A main effect of Interval size was also found, $F(1.59, 41.22) = 72.20$, $p < .001$, $\eta_p^2 = .70$, with increased performance associated with larger interval sizes. This effect was similar for both groups, as there

---

[2] The analysis of all 12 interval sizes also yielded significant main effects of group and interval size, $F(1, 30) = 12.05$, $p < .005$, $\eta_p^2 = .32$, and $F(1, 30) = 29.16$, $p < .001$, $\eta_p^2 = .53$, respectively.

was no significant interaction between Interval size and Group, $F(1.59, 41.22) = 0.86$, $p = .41$, $\eta_p^2 = .03$.
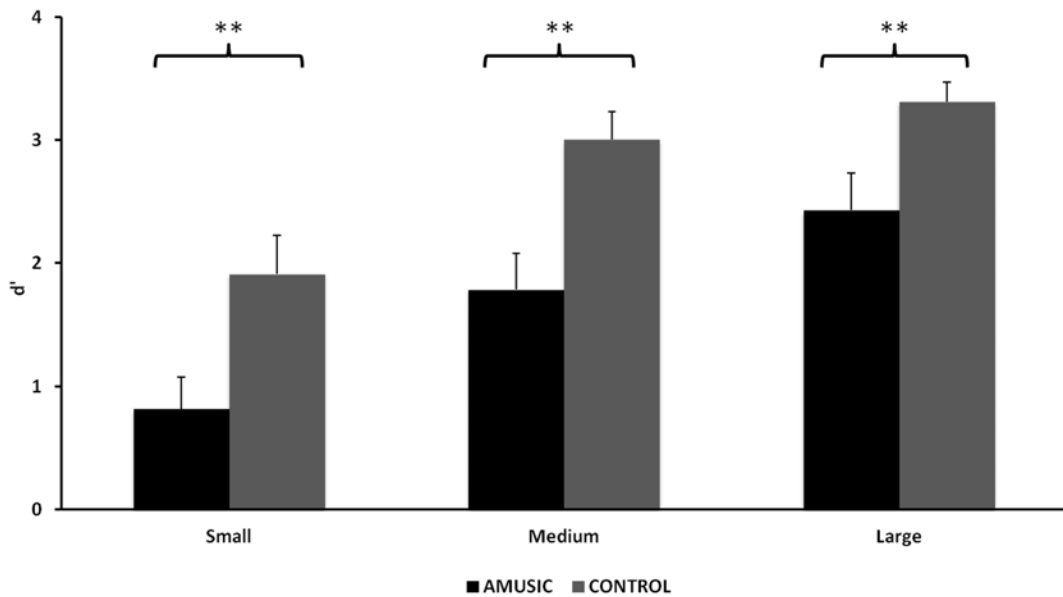
Fig. 2 – Task performance in Experiment 1 evaluated by d' on each interval size condition for amusic (black bars) and control (grey bars) groups. Error bars represent +1 SE. **: $p < .01$.

2.2.2. Self-pacing time differences on congruent and incongruent audio-visual contours

To examine the implicit processing of audio-visual congruence, we computed self-pacing times for the target note in congruent and incongruent trials separately. Self-pacing time was calculated as the duration from the onset of the target note to the onset of the post-target note (i.e., ISI). The difference in the ISI for incongruent and congruent trials was then calculated. Implicit knowledge of audio-visual congruence was inferred if the mean ISI for incongruent trials is longer than the mean ISI for congruent trials. That is, implicit detection of audio-visual incongruence should result in a longer processing time at the target note of incongruent than congruent trials, even if participants are not consciously aware of it. Thus, a positive value of self-pacing time difference on trials with incorrect explicit responses indicates an implicit processing of audio-visual incongruence. One-sample t-tests (1-tailed) were conducted for trials in which incongruences were not explicitly detected. The mean ISI difference on the target note ($M = 174.49$ ms; $SD = 319.33$ ms) was significantly larger than zero, $t(13) = 2.04$, $p < .05$, indicating that

individuals with congenital amusia were detecting audio-visual incongruences implicitly to some degree. It should be noted that we did not conduct the same analysis for the control group, as control participants made almost no explicit errors in their judgements of audio-visual congruence. As such, there were too few incorrect trials in the control group to enable an analysis of implicit detection of incongruences.

We next conducted one-sample t-tests (2-tailed) on the difference in ISI values for post-target notes of trials with correct responses, as an index of the confidence with which an audio-visual incongruence was detected. High confidence in detecting incongruence should result in rapid progression to the next event, given that the target event (i.e., incongruence) has already occurred. Low confidence in detecting incongruence should result in slower progression to the next event, given that the participant would be uncertain whether the target event had actually occurred, and hence would attend more carefully to the subsequent stimuli. The self-pacing time on post-target notes was defined as the mean duration of each key press after the target note appeared. A negative value indicates acceleration in incongruent trials after the target event was detected. The results showed that the key pressing was accelerated after audio-visual incongruence was detected in the control group ($M = $-170.37 ms; $SD = $ 166.85 ms), $t(13) = 3.82$, $p < .01$, whereas acceleration was not observed in the amusic group ($M = $-103.93 ms; $SD = $ 294.30 ms), $t(13) = 1.32$, $p = .21$.

2.2.3. Confidence ratings (pitch dimension)

A repeated-measures ANOVA was conducted on the confidence ratings for trials with a correct response, with the factors of Group, Interval size, and Congruence. As shown in Table 3, individuals with congenital amusia ($M = 3.78$; $SD = 0.58$) reported lower confidence levels than control participants ($M = 4.23$; $SD = 0.54$), $p < .05$. Furthermore, the main effect of Interval size was significant, $p < .001$. A post-hoc test suggested participants' confidence level increased with the size of the pitch interval leading to the target note (small: $M = 3.75$, $SE = 0.12$; medium: $M = 3.98$, $SE = 0.11$; large: $M = 4.18$, $SE = 0.10$; $ps < .001$). An interaction between Interval size and Congruence was also found. In the small-change condition, confidence ratings were higher for congruent trials than for incongruent trials ($p < .001$), whereas the pattern was reversed in the large-change condition ($p < .05$). There was no difference in the medium interval size condition ($p = .70$). Further paired-sample t-tests also

showed that confidence ratings for the three interval size conditions were different for incongruent trials ($ps < .001$) but not congruent ones ($ps > .1$).

Table 3 – The ANOVA results of main effects and two-way interactions on confidence rating scores in Experiment 1, as none of higher-order interactions was significant. Effect size was estimated using partial eta-squared (η_p^2), and DF refers to the degrees of freedom. *: $p < .05$; ***: $p < .001$.

Effect	DFn	DFd	F	p	η_p^2
Group	1	26	5.03 *	.03	.16
Interval size	2	52	44.71 ***	< .001	.63
Congruence	1	26	0.53	.48	.02
Interval size × Group	2	52	1.70	.19	.06
Congruence × Group	1	26	3.15	.09	.11
Interval size × Congruence	2	52	32.77 ***	< .001	.56

2.2.4. Correlations between pitch perception and congruence judgements

To investigate the relationship between pitch perception and audio-visual congruence judgements, we computed the correlations (1-tailed) between the two log-transformed pitch thresholds (identification and discrimination) and d' values in amusic and control groups separately as well as across groups. As shown in Table 4, the task performance negatively correlated with log-transformed pitch thresholds for both discrimination and identification ($ps < .01$). The two groups showed the same pattern, suggesting task performance was affected by pitch perception, even though most of the change intervals were above participant's pitch thresholds.

Table 4 – Correlation coefficients between pitch thresholds and congruence judgements of audio-visual contours. *: $p < .05$; *: $p < .01$; ***: $p < .001$.

	Amusics	Controls	All participants
Pitch identification	-.56 * ($p = .02$)	-.74 ** ($p = .001$)	-.68 *** ($p < .001$)
Pitch discrimination	-.51 * ($p = .03$)	-.47 * ($p = .05$)	-.57 ** ($p = .001$)

2.3. Discussion

Experiment 1 revealed that individuals with congenital amusia, in comparison with non-amusic listeners, showed reduced sensitivity to the incongruence between melodic and visual contours. The elevated pitch thresholds observed in the amusic group might account for the difficulties in judging audio-visual congruence, in that significant correlations between pitch perception measures and participants' congruence judgements were found. However, it is possible that pitch perception measures and congruence judgements are related to each other by their connection to a common high-level factor, such as attention. Combined with the finding that amusics exhibited preserved implicit knowledge of audio-visual incongruence to some degree (revealed by longer self-pacing times for incongruent than for congruent targets), even when they were not aware of the incongruences explicitly, the findings suggest that impaired contour processing arises at a perceptual stage that involves explicit processing of pitch change direction. The dissociation between the explicit task performance and implicit processing suggests a failure of conscious access to pitch information in congenital amusia within the perceptual stage.

On the other hand, it is also possible that a more general impairment in cross-modal mapping can account for the finding that amusics have difficulties in detecting incongruence of audio-visual contours, given that cross-modal mapping is needed in order to compare stimuli from two sensory modalities. To test the latter possibility and examine whether the impaired contour processing extends beyond the pitch dimension, Experiment 2 employed the same paradigm as that of Experiment 1 to investigate contour processing in auditory dimensions other than pitch: either loudness or timbral brightness.

3.   Experiment 2

3.1. Methods

3.1.1. Participants

Sixteen individuals with congenital amusia (four were recruited from Experiment 1) and 16 matched controls took part in Experiment 2. A summary of the participants' characteristics and results of the MBEA for each group is shown in

Table 5.

Table 5 – Participants' characteristics, correct rates (mean ± SD) on the melodic subtests of MBEA, and independent-sample t-test results between two groups in Experiment 2. DF refers to the degrees of freedom and was corrected if the equal variances assumption was violated. ***: $p < .001$.

	Amusics	Controls	DF	t	p (2-tailed)
Age	23.34 ± 5.42	23.37 ± 8.73	30	- 0.01	.99
Gender	6F/10M	6F/10M	-	-	-
Years of education	14.38 ± 2.03	13.88 ± 2.33	30	0.65	.52
Years of musical training	0.44 ± 1.09	0.47 ± 0.96	30	-0.07	.94
Hours of music listening daily	2.57 ± 2.15	1.84 ± 0.94	30	1.24	.22
MBEA (percentage correct)					
Scale	0.69 ± 0.09	0.92 ± 0.06	30	-8.45 ***	< .001
Contour	0.65 ± 0.08	0.85 ± 0.09	30	-6.29 ***	< .001
Interval	0.66 ± 0.08	0.79 ± 0.08	30	-4.56 ***	< .001
Composite	0.67 ± 0.04	0.85 ± 0.05	30	-12.25 ***	< .001

3.1.2. Stimuli

Auditory stimuli were constructed to vary in either timbral brightness or loudness, but not in pitch. Auditory brightness-varying stimuli were digitally synthesized and manipulated by shifting the spectral centroid while keeping the F0 fixed using MATLAB, following the strategy described by Russo and Thompson (2005) and Warrier and Zatorre (2002). The energy in the dull timbre was weighted in lower partials, and the energy in the bright timbre was weighted in the higher partials. In the present experiment, F3 (174.62 Hz), G3 (196 Hz), A3 (220 Hz) or B3 (246.94 Hz) was selected to serve as the F0. As illustrated in Fig. 3, five levels of timbral brightness (very dull, somewhat dull, intermediate, somewhat bright, and very bright) were generated for each frequency by varying the intensities of ten partial harmonics while keeping F0 at a fixed intensity level that was always higher than that of any other partial. This strategy ensured that tones differed only in brightness but not on the intensity of the F0. The spectral centroid changes from the intermediate timbre to the very dull, somewhat dull, somewhat bright, and very

bright timbre were approximately 330, 114, 107, and 278 cents, respectively. To be comparable to the pitch-varying sequence used in Experiment 1, timbral brightness-varying sequences consisted of seven notes, and the target note occurred at the fourth, fifth or sixth position at an equal rate. There were 24 timbral brightness-varying sequences in total and the notes in each sequence shared the same F0 but varied in spectral centroid. The degree of change in brightness between the preceding and target note was manipulated experimentally: for half of the trials, small changes in timbre were introduced (a change by one or two levels) and the rest were large changes (a change by three or four levels). Every sequence started with the tone in the intermediate timbre and was presented binaurally at 65 dB SPL through headphones.
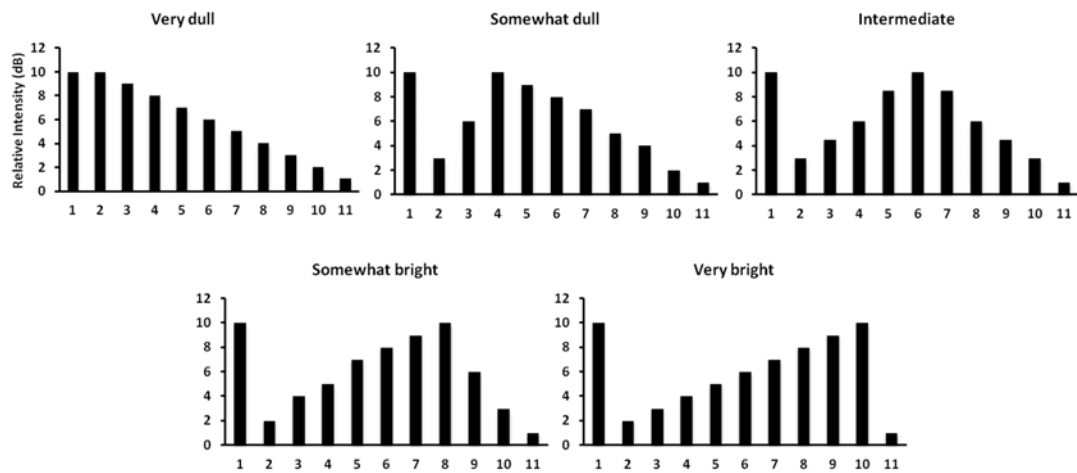
Fig. 3 – Spectra of five levels of brightness used in Experiment 2. The x-axis and y-axis refer to the harmonic components and the relative intensity, respectively.

The loudness-varying stimuli were generated using the intermediate timbre tones from the timbral brightness-varying stimuli, and then manipulated by altering the amplitude to increase or reduce intensity by 5 or 10 dB. Loudness-varying sequences that shared the same pitch and timbral brightness but varied in intensity were then created following similar procedures with timbral brightness-varying ones, where the degree of change in loudness between the preceding and target note was either small or large.

To get a better understanding of whether the impairments are specific to spatial height or generalise to other physiological relations, two types of visual stimuli were

employed – height-varying and visual brightness-varying contours. The former stimuli were similar to those used in Experiment 1; the latter stimuli were created by varying visual brightness while maintaining the vertical spatial location of the dots. In congruent trials, the audio and visual information were congruent, where the sound went brighter or louder aurally as the dot moved up or went brighter visually, or the sound went duller or softer aurally as the dot moved down or went dimmer visually. Incongruent trials were characterized by the reversed mapping.

3.1.3. Procedure

The paradigm and procedure were the same as that of Experiment 1. The experiment consisted of four audio-visual conditions: spatial height and timbral brightness; spatial height and loudness; visual brightness and timbral brightness; and visual brightness and loudness. Within each condition there were three blocks of 16 trials, yielding 48 trials per condition. Assignment of the two response keys to the congruent and incongruent responses and the orders of four conditions were counterbalanced across participants.
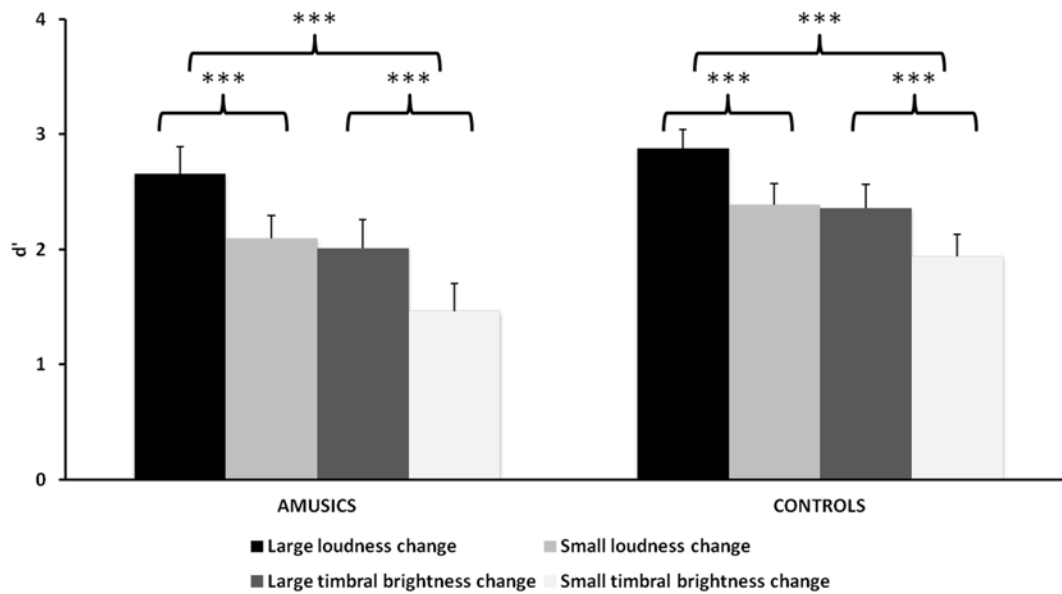
3.2. Results

3.2.1. Judgements of audio-visual congruence (timbral brightness and loudness dimensions)

A repeated-measures ANOVA on d' was conducted with the factors of Group, Audio change (timbral brightness and loudness), Visual change (height and visual brightness), and Change size (small and large). The mean performance in the amusic group ($M = 2.06$, $SE = 0.16$) was slightly lower than the mean performance in the control group ($M = 2.39$, $SE = 0.16$); however, the difference between the two groups was not significant, $p = .16$ (see Table 6 and Fig. 4). All participants did better on loudness-varying conditions (visual height: $M = 2.52$, $SE = 0.12$; visual brightness: $M = 2.49$, $SE = 0.13$) in comparison to brightness-varying conditions (visual height: $M = 2.03$, $SE = 0.15$; visual brightness: $M = 1.85$, $SE = 0.14$), as revealed by the main effect of Audio change, $p < .001$. Similar to Experiment 1, individuals both with and without congenital amusia showed better performance on trials with large changes ($M = 2.48$, $SE = 0.12$) than on trials with small changes ($M = 1.97$, $SE = 0.12$), $p < .001$.

Table 6 – The ANOVA results of main effects and two-way interactions on judgements of audio-visual congruence evaluated by d' score in Experiment 2, as none of higher-order interactions was significant. Effect size was estimated using partial eta-squared (η_p^2), and DF refers to the degrees of freedom. ***: $p < .001$.

Effect	DFn	DFd	F	p	η_p^2
Group	1	30	2.11	.16	.07
Audio	1	30	34.79 ***	< .001	.54
Visual	1	30	2.42	.13	.08
Change size	1	30	36.20 ***	< .001	.55
Audio × Group	1	30	0.63	.43	.02
Visual × Group	1	30	< 0.001	.99	< .001
Change size × Group	1	30	0.34	.56	.01
Audio × Visual	1	30	1.22	.28	.04
Audio × Change size	1	30	0.11	.74	< .005
Visual × Change size	1	30	0.32	.58	.01

Fig. 4 – Task performance in Experiment 2 evaluated by d' on each condition for amusic and control groups. Error bars represent +1 SE. ***: $p < .001$.

3.2.2. Confidence ratings (timbral brightness and loudness dimensions)

A repeated-measures ANOVA was conducted with the factors of Group, Audio change, Visual change, Change size, and Congruence. As shown in Table 7, there was no group difference in self-reported confidence level (Amusic: $M = 4.11$, $SE = 0.13$; Control: $M = 3.88$, $SE = 0.13$; $p = .24$). Furthermore, participants reported higher confidence levels on loudness-varying conditions ($M = 4.15$, $SE = 0.10$) when compared with brightness-varying conditions ($M = 3.83$, $SE = 0.11$), $p < .001$.

Table 7 – The ANOVA results of main effects and two-way interactions on confidence rating scores in Experiment 2, as none of higher-order interactions was significant. Effect size was estimated using partial eta-squared (η_p^2), and DF refers to the degrees of freedom. *: $p < .05$; ***: $p < .001$.

Effect	DFn	DFd	F	p	η_p^2
Group	1	30	1.47	.24	.05
Audio	1	30	16.69 ***	< .001	.36
Visual	1	30	1.07	.31	.03
Change size	1	30	84.95 ***	< .001	.74
Congruence	1	30	4.29 *	.05	.13
Audio × Group	1	30	0.55	.46	.02
Visual × Group	1	30	0.25	.62	.01
Change size × Group	1	30	0.30	.86	< .005
Congruence × Group	1	30	1.30	.26	.04
Audio × Visual	1	30	0.09	.77	< .005
Audio × Change size	1	30	0.04	.84	.001
Audio × Congruence	1	30	3.30	.08	.10
Visual × Change size	1	30	1.54	.23	.05
Visual × Congruence	1	30	2.26	.14	.07
Change size × Congruence	1	30	54.73 ***	< .001	.65

The main effects of Change size and Congruence were found, $ps < .05$. There was also an interaction between these two factors, $p < .05$. Further paired-sample t-tests showed that the congruence effect only remained significant in the small-change condition (congruent: $M = 3.94$, $SE = 0.11$; incongruent: $M = 4.30$, $SE = 0.10$; p

< .001) but not in large-change condition (congruent: M = 3.87, SE = 0.11; incongruent: M = 3.86, SE = 0.11; p = .92). In other words, the ratings were higher on incongruent trials than congruent ones in the small-change condition, whereas this pattern was not observed in the large-change condition.

3.3. Discussion

Experiment 2 showed no evidence that the impairment of melodic contour processing extends to other auditory dimensions (either timbral brightness or loudness in this case). In other words, individuals with congenital amusia have no difficulty in cross-modal matching, as long as the task does not require explicit pitch processing.

4.   General discussion

In this investigation, we developed and employed an on-line self-paced audio-visual contour task (i.e., SACT), in which minimal or no load was placed on short-term memory. The novel task allowed us to largely disentangle perceptual and storage stages of melodic contour processing among individuals with and without congenital amusia. Although amusic participants showed reduced explicit sensitivity to the audio-visual incongruence when compared with controls, they preserved some implicit knowledge of the on-line melodic and visual contour congruence. These findings suggest that congenital amusia is associated with impaired melodic contour processing at an explicit perceptual stage. We next examined contour processing in other auditory dimensions (i.e., timbral brightness and loudness) using same paradigm. No deficits in contour processing for non-pitch dimensions of sound were observed, suggesting that the reduced sensitivity found in Experiment 1 cannot be explained by either a general deficit in auditory contour processing or a deficit in cross-modal mapping.

Previous studies have examined melodic contour processing in amusia using a discrimination paradigm, in which participants were presented with pairs of melodies and were then required to judge whether they were the same or different (e.g., Albouy, Mattout, et al., 2013; Foxton et al., 2004; Jiang et al., 2010; Peretz et al., 2003). However, such an experimental design cannot exclude potential influences

from the limited capacity of auditory short-term memory (Gosselin et al., 2009; Tillmann et al., 2009; Williamson & Stewart, 2010). Thus, in Experiment 1 we used the SACT that places minimal or no load on memory, allowing us to focus on the perceptual stage of contour processing.

We found that the judgements of audio-visual congruence were significantly worse in amusic participants when compared with control participants, which indicates that amusic individuals have difficulty explicitly detecting cross-modal incongruence at a perceptual stage. On the other hand, our analysis of the incorrect responses in the amusic group indicated that the self-pacing time on target notes were significantly different for congruent and incongruent trials, as revealed by a longer ISI for the targets of incongruent trials in comparison to those of congruent trials. Given that the incongruent audio-visual contour required a longer processing time, our results suggest that amusic participants had some implicit ability to distinguish the targets of incongruent trials from those of congruent trials, even when they failed to detect such incongruences explicitly.

The dissociation between implicit knowledge and explicit judgements on audio-visual congruence in the amusic group is consistent with the hypothesis that congenital amusia is a disorder of pitch awareness, caused by a disconnection between auditory perceptual and high-level processing that involves conscious awareness (e.g., Peretz et al., 2009; Hyde, Zatorre, & Peretz, 2011, Tillmann et al., 2014). This hypothesis predicts that amusic individuals can represent pitch change directions, but that the information does not reach conscious awareness, thereby resulting in poor performance in tasks that require explicit judgements (Loui et al., 2009; Loui et al., 2008; Lu et al., 2015; Peretz et al., 2009; Zendel et al., 2015).

This interpretation is supported by evidence from both neuroimaging and behavioural studies. First, amusic individuals show normal auditory N1 ERP components in response to pitch changes (Moreau et al., 2009; Peretz et al., 2005). Given that the neural generators have been localized within the auditory cortex (Näätänen & Picton, 1987), a normal N1 in amusics reflects intact auditory processing at an early stage. Second, Loui et al. (2008) demonstrated that individuals with congenital amusia could reproduce the direction of a pitch change even when they were at chance in reporting whether the change went up or down. In line with this study, Hutchins, Gosselin, and Peretz (2010) reported that when amusic

participants were asked to sing out loud, their produced responses were highly correlated with the target pitches with a systematic downward shift in pitch height. These findings suggest amusic individuals show mild or no impairment in production or imitation of pitches, where the conscious processing of pitch information is not required. Therefore, Loui, Kroog, Zuk, Winner, and Schlaug (2011) have proposed in their later work that the pitch perception problems in amusia may be linked to an inability to process pitch information consciously, or a lack of "pitch awareness".

Dehaene, Changeux, Naccache, Sackur, and Sergent (2006) distinguished "accessibility" from "access," whereby some attended stimuli have the potential to gain access to conscious awareness (accessibility), but they are nonetheless not consciously accessed (access). In the case of congenital amusia, it is possible that the auditory stimuli are processed after they arrive in the brain, and the pitch information extracted from the stimuli is then accessible for further processing. However, the reduced functional connectivity between the inferior frontal gyrus and the auditory cortex (Albouy, Mattout, et al., 2013; Hyde et al., 2011) and the reduced volume or absence of right arcuate fasciculus in amusia (Loui et al., 2009; but see, Chen et al., 2015) may cause a failure to access the information consciously.

Another prediction of the "pitch awareness" hypothesis is that individuals with amusia might show a bias toward reporting low confidence compared to control individuals in tasks involving auditory processing in pitch compared to other dimensions, as they are more likely to be wary of their judgements due to their inability to access the pitch information consciously. It was therefore no surprise that amusic participants reported lower confidence levels than controls even on their correct judgements in the pitch-related task (Experiment 1) but not in the pitch-unrelated tasks (Experiment 2). The comparable self-reported confidence levels in Experiment 2 suggest that individuals with congenital amusia do not underestimate their judgements in general. In other words, amusics are quite confident in their judgements concerning auditory input, as long as the task does not require pitch processing. The prediction is also supported by our findings that only control participants showed acceleration after the incongruent targets appeared in Experiment 1, whereas amusic participants took all the chances to gather information for the final decision.

Taken together, the findings of Experiments 1 and 2 suggest that the impairment

of contour processing in amusics is restricted to the pitch dimension. These results are consistent with a recent study suggesting that temporal, spectral, and intensity perception is intact in the amusic auditory system (Cousineau, Oxenham, & Peretz, 2013, 2015). The comparable performance in Experiment 2 between the two groups further suggests that the impaired task performance observed in Experiment 1 was not caused by a general deficit in cross-modal mapping. In other words, amusic individuals have a normal ability to compare auditory and visual stimuli with respect to their psychological relations (i.e., auditory loudness and timbral brightness correspond to spatial height and visual brightness).

What is the underlying mechanism of contour processing in congenital amusia? Based on the observations that amusics showed intact cross-modal mapping in non-pitch auditory dimensions, our results suggest that the impairments in congenital amusia are restricted to melodic contour processing involving pitch perception, and that this impairment can be traced down to the perceptual stage of auditory processing, when pitch information fails to reach conscious awareness. However, it should be highlighted that based on our findings we are not arguing that deficits at the perceptual stage can account for all impairments of melodic contour processing reported in previous studies (Foxton et al., 2004; Jiang et al., 2010). As we did not examine the storage stage of auditory processing independently by excluding the potential influences of pitch perception, it is possible that amusia is also associated with auditory short-term memory deficits (Tillmann et al., 2009; Williamson & Stewart, 2010). It is also possible that auditory short-term memory deficits are not caused by impaired pitch perception, but are concurrent with it. Although individuals with congenital amusia can represent pitch movement implicitly to some degree within the perceptual stage, they fail to bring it to conscious awareness. This limited access leads to unstable representations of pitch, thereby causing impaired task performance and low confidence levels.

Conflict of interest

The authors declare no conflict of interest.

References

Albouy, P., Mattout, J., Bouet, R., Maby, E., Sanchez, G., Aguera, P. E., et al. (2013). Impaired pitch perception and memory in congenital amusia: The deficit starts in the auditory cortex. *Brain, 136*(5), 1639-1661.

Albouy, P., Schulze, K., Caclin, A., & Tillmann, B. (2013). Does tonality boost short-term memory in congenital amusia? *Brain Research, 1537*(6), 224-232.

Attneave, F., & Olson, R. K. (1971). Pitch as a medium: A new approach to psychophysical scaling. *The American Journal of Psychology, 84*(2), 147-166.

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia: a group study of adults afflicted with a music-specific disorder. *Brain, 125*(2), 238-251.

Chen, J. L., Kumar, S., Williamson, V. J., Scholz, J., Griffiths, T. D., & Stewart, L. (2015). Detection of the arcuate fasciculus in congenital amusia depends on the tractography algorithm. *Frontier in Psychology, 6*, 9.

Cousineau, M., Oxenham, A. J., & Peretz, I. (2013). The role of peripheral spectro-temporal coding in congenital amusia. *The Jouranl of the Acoustical Society of America, 133*(5), 3381.

Cousineau, M., Oxenham, A. J., & Peretz, I. (2015). Congenital amusia: a cognitive disorder limited to resolved harmonics and with no peripheral basis. *Neuropsychologia, 66*, 293-301.

Davies, A. D. M., & Roberts, E. (1975). Poor pitch singing: a survey of its incidence in school children. *Psychology of Music, 3*(2), 24-36.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences, 10*(5), 204-211.

Dowling, W. J. (1978). Scale and contour: two components of a theory of memory for melodies. *Psychological Review, 85*(4), 341-354.

Dowling, W. J., & Fujitani, D. S. (1971). Contour, interval, and pitch recognition in memory for melodies. *The Journal of the Acoustical Society of America, 49*(2), 524-531.

Edworthy, J. (1985). Interval and contour in melody processing. *Music Perception, 2*(3), 375-388.

Eerola, T., & Toiviainen, P. (2004). MIDI Toolbox: MATLAB Tools for Music Research. Retrieved from

http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/miditoolbox/

Eitan, Z., Schupak, A., & Marks, L. E. (2008). *Louder is higher: cross-modal interaction of loudness change and vertical motion in speeded classification.* Proceedings of the 10th international conference on music perception and cognition, Adelaide, Australia.

Foxton, J. M., Dean, J. L., Gee, R., Peretz, I., & Griffiths, T. D. (2004). Characterization of deficits in pitch perception underlying `tone deafness'. *Brain, 127*(4), 801-810.

Gosselin, N., Jolicoeur, P., & Peretz, I. (2009). Impaired memory for pitch in congenital amusia. *Annals of the New York Academy of Sciences, 1169*(1), 270-272.

Graves, J. E., Micheyl, C., & Oxenham, A. J. (2014). Expectations for melodic contours transcend pitch. *Journal of Experimental Psychology: Human Perception and Performance, 40*(6), 2338-2347.

Hébert, S., & Peretz, I. (1997). Recognition of music in long-term memory: are melodic and temporal patterns equal partners? *Memory & Cognition, 25*(4), 518-533.

Hutchins, S., Gosselin, N., & Peretz, I. (2010). Identification of changes along a continuum of speech intonation is impaired in congenital amusia. *Frontiers in Psychology, 1*, 236.

Hyde, K. L., & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science, 15*(5), 356-360.

Hyde, K. L., Zatorre, R. J., & Peretz, I. (2011). Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. *Cerebral Cortex, 21*(2), 292-299.

Jiang, C., Hamm, J. P., Lim, V. K., Kirk, I. J., & Yang, Y. (2010). Processing melodic contour and speech intonation in congenital amusics with Mandarin Chinese. *Neuropsychologia, 48*(9), 2630-2639.

Jiang, C., Lim, V. K., Wang, H., & Hamm, J. P. (2013). Difficulties with pitch discrimination influences pitch memory performance: evidence from congenital amusia. *PLoS One, 8*(10), e79216. doi: 10.1371/journal.pone.0079216.t001

Jones, M. R. (1976). Levels of structure in the reconstruction of temporal and spatial serial patterns. *Journal of Experimental Psychology, 2*(4), 475-488.

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology, 11*(3), 346-374.

Krumhansl, C. L. (1990). *Cognitive Foundations of Musical Pitch*. Oxford, UK: Oxford University Press.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review, 89*(4), 334-368.

Lehmann, A., Skoe, E., Moreau, P., Peretz, I., & Kraus, N. (2015). Impairments in musical abilities reflected in the auditory brainstem: Evidence from congenital amusia. *European Journal of Neurosciene, 42*(1), 1644-1650.

Levitt, H. (1971). Transformed up–down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*, 467-477.

Lewkowicz, D. J., & Turkewitz, G. (1980). Cross-modal equivalence in early infancy: auditory–visual intensity matching. *Developmental psychology, 16*(6), 597-607.

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), 1189-1207.

Liu, F., Maggu, A. R., Lau, J. C. Y., & Wong, P. C. M. (2015). Brainstem encoding of speech and musical stimuli in congenital amusia: evidence from Cantonese speakers. *Frontiers in Human Neuroscience, 8*, 1029.

Liu, F., Patel, A. D., Fourcin, A., & Stewart, L. (2010). Intonation processing in congenital amusia: discrimination, identification and imitation. *Brain, 133*(6), 1682-1693.

Loui, P., Alsop, D., & Schlaug, G. (2009). Tone deafness: a new disconnection syndrome? *The Journal of Neuroscience, 29*(33), 10215-10220.

Loui, P., Guenther, F. H., Mathys, C., & Schlaug, G. (2008). Action-perception mismatch in tone-deafness. *Current Biology, 18*(8), R331-R332.

Loui, P., Kroog, K., Zuk, J., Winner, E., & Schlaug, G. (2011). Relating pitch awareness to phonemic awareness in children: implications for tone-deafness and dyslexia. *Frontier in Pyscholgoy, 2*, 111.

Lu, X., Ho, H. T., Liu, F., Wu, D., & Thompson, W. F. (2015). Intonation processing deficits of emotional words among Mandarin Chinese speakers with congenital amusia: an ERP study. *Frontiers in Psychology, 6*, 385.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide*. Mahwah, NJ: Erlbaum.

Marks, L. E. (1982). Bright sneezes and dark coughs, loud sunlight and soft moonlight. *Journal of Experimental Psychology: Human Perception and Performance, 8*(2), 177-193.

Marks, L. E. (1987). On cross-modal similarity: auditory-visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance, 13*(3), 384-394.

Marks, L. E. (1989). On cross-modal similarity: The perceptual structure of pitch, loudness, and brightness. *Journal of Experimental Psychology: Human Perception and Performance, 15*(3), 586-602.

Marks, L. E., Hammeal, R. J., Bornstein, M. H., & Smith, L. B. (1987). Perceiving similarity and comprehending metaphor. *Monographs of the Society for Research in Child Development, 52*(1), i-100.

McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: common dimensions, specificities, and latent subject classes. *Psychological research, 58*(3), 177-192.

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2008). Is relative pitch specific to pitch? *Psychological Science, 19*(12), 1263-1271.

Melara, R. D. (1989). Dimensional interaction between color and pitch. *Journal of Experimental Psychology: Human Perception and Performance, 15*(1), 69-79.

Mikumo, M. (1992). Encoding strategies for tonal and atonal melodies. *Music Perception, 10*(1), 73-81.

Mikumo, M. (1997). Multi-encoding for pitch infornation of tone sequences. *Japanese Psychological Research, 39*(4), 300-311.

Moreau, P., Jolicœur, P., & Peretz, I. (2009). Automatic brain responses to pitch changes in congenital amusia. *Annals of the New York Academy of Sciences, 1169*(1), 191-194.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology, 24*(4), 375-425.

Omigie, D., Pearce, M. T., & Stewart, L. (2012). Tracking of pitch probabilities in congenital amusia. *Neuropsychologia, 50*(7), 1483-1493.

Patel, A. D., Foxton, J. M., & Griffiths, T. D. (2005). Musically tone-deaf individuals have difficulty discriminating intonation contours extracted from speech. *Brain and Cognition, 59*(3), 310-313.

Peretz, I. (2013). The biological foundations of music: insights from congenital amusia. In D. Deustch (Ed.), *The psychology of music*. San Diego: Acdemic Press.

Peretz, I., Brattico, E., Jarvenpaa, M., & Tervaniemi, M. (2009). The amusic brain: in tune, out of key, and unaware. *Brain, 132*(5), 1277-1286.

Peretz, I., Brattico, E., & Tervaniemi, M. (2005). Abnormal electrical brain responses to pitch in congenital amusia. *Annals of Neurology, 58*(3), 478-482.

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders: the Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences, 999*(1), 58-75.

Peretz, I., & Hyde, K. L. (2003). What is specific to music processing? insights from congenital amusia. *Trends in Cognitive Sciences, 7*(8), 362-367.

Prince, J. B., Schmuckler, M. A., & Thompson, W. F. (2009a). The effect of task and pitch structure on pitch-time interactions in music. *Memory & Cognition, 37*(3), 368-381.

Prince, J. B., Schmuckler, M. A., & Thompson, W. F. (2009b). Cross-model melodic contour similarity. *Canadian Acoustics, 37*(1), 35-49.

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

Russo, F. A., and Thompson, W. F. (2005). An interval size illusion: The influence of timbre on the perceived size of melodic intervals. *Perception & Psychophysics, 67*, 559-568.

Soranzo, A., & Grassi, M. (2014). PSYCHOACOUSTICS: a comprehensive MATLAB toolbox for auditory testing. *Frontier in Psychology, 5*, 712.

Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Attention, Perception, & Psychophysics, 73*(4), 971-995.

Tillmann, B., Albouy, P., Caclin, A., & Bigand, E. (2014). Musical familiarity in congenital amusia: evidence from a gating paradigm. *Cortex, 59*, 84-94.

Tillmann, B., Gosselin, N., Bigand, E., & Peretz, I. (2012). Priming paradigm reveals harmonic structure processing in congenital amusia. *Cortex, 48*(8), 1073-1078.

Tillmann, B., Schulze, K., & Foxton, J. (2009). Congenital amusia: a short-term memory deficit for non-verbal, but not verbal sounds. *Brain and Cognition, 71*(3), 259-264.

Trehub, S. E., Bull, D., & Thorpe, L. A. (1984). Infants' perception of melodies: The role of melodic contour. *Child development, 55*(3), 821-830.

Vuvan, D. T., Nunes-Silva, M., & Peretz, I. (2015). Meta-analytic evidence for the non-modularity of pitch processing in congenital amusia. *Cortex, 69*, 186-200.

Warrier, C. M., and Zatorre, R. J. (2002). Influence of tonal context and timbral variation on perception of pitch. *Perception & Psychophysics, 64*, 198-207.

Williamson, V. J., & Stewart, L. (2010). Memory for pitch in congenital amusia: beyond a fine-grained pitch discrimination problem. *Memory, 18*(6), 657-669.

Zendel, B. R., Lagrois, M. E., Robitaille, N., & Peretz, I. (2015). Attending to pitch information inhibits processing of pitch information: The curious case of amusia. *The Journal of Neuroscience, 35*(9), 3815-3824.

Chapter 6: Influences of Visuospatial Information on Pitch Processing

Chapter 5 (manuscript 3) examined contour processing in individuals with congenital amusia using a novel cross-modal task. The results revealed that the impairments in contour processing exhibited by amusic participants occur at a perceptual stage and are restricted to the pitch dimension. In addition to the difficulties in representing pitch change direction, the potential influences of concurrent visual stimuli presented with auditory stimuli may also partly account for the impairments observed in Chapter 5. Indeed, people with congenital amusia report that they often rely on facial expressions and gestures when interpreting the moods and feelings of people with whom they are speaking due to their deficits in auditory processing (Thompson et al. 2012). To examine the extent to which visuospatial information affects amusics' pitch perception, an ERP study was conducted. In this study, I examined amusics' brain responses to pitches when presented with unattended and attended visual stimuli.

The manuscript that describes this study (manuscript 4) was co-authored by Tam Ho, Yanan Sun, Blake Johnson, and Bill Thompson, and I was the main contributor. My contribution included experimental design, data collection, data analysis, and preparation of the first draft of the manuscript. Bill Thompson and Tam Ho provided helpful comments and feedback in experimental design and interpretation of the data. Yanan Sun provided assistance in participant recruitment and experiment programming. Blake Johnson made critical comments regarding the writing of the manuscript.

This chapter has been published as:

Lu, X., Ho, H. T., Sun, Y., Johnson, B. W., & Thompson, W. F (2016). The influence of visual information on auditory processing in individuals with congenital amusia: An ERP study. *NeuroImage, 135,* 142–151. doi: 10.1016/j.neuroimage.2016.04.043

THE INFLUENCE OF VISUAL INFORMATION ON AUDITORY PROCESSING IN INDIVIDUALS WITH CONGENITAL AMUSIA: AN ERP STUDY

Xuejing Lu [a,b,*], Hao T. Ho [a], Yanan Sun [b,c], Blake W. Johnson [b,c], William F. Thompson [a,b]

[a] Department of Psychology, Macquarie University, Sydney, NSW, Australia
[b] ARC Centre of Excellence in Cognition and its Disorders, NSW, Australia
[c] Department of Cognitive Science, Macquarie University, Sydney, NSW, Australia

* Corresponding author at: Department of Psychology, Macquarie University, NSW 2109, Australia
E-mail address: xuejing.lu@mq.edu.au

ABSTRACT

While most normal hearing individuals can readily use prosodic information in spoken language to interpret the moods and feelings of conversational partners, people with congenital amusia report that they often rely more on facial expressions and gestures, a strategy that may compensate for deficits in auditory processing. In this investigation, we used EEG to examine the extent to which individuals with congenital amusia draw upon visual information when making auditory or audio-visual judgments. Event-related potentials (ERP) were elicited by a change in pitch (up or down) between two sequential tones paired with a change in spatial position (up or down) between two visually presented dots. The change in dot position was either congruent or incongruent with the change in pitch. Participants were asked to judge (1) the direction of pitch change while ignoring the visual information (AV implicit task), and (2) whether the auditory and visual changes were congruent (AV explicit task). In the AV implicit task, amusic participants performed significantly worse in the incongruent condition than control participants. ERPs showed an enhanced N2-P3 response to incongruent AV pairings for control participants, but not for amusic participants. However, when participants were explicitly directed to detect AV congruency, *both* groups exhibited enhanced N2-P3 responses to incongruent AV pairings. These findings indicate that amusics are capable of extracting information from both modalities in an AV task, but are biased to rely on visual information when it is available, presumably because they have learned that auditory information is unreliable. We conclude that amusic individuals implicitly draw upon visual information when judging auditory information, even though they have the capacity to explicitly recognize conflicts between these two sensory channels.

1. Introduction

Congenital amusia is a disorder of musical abilities including pitch perception (Ayotte et al. 2002; Foxton et al. 2004; Hyde and Peretz, 2004; Peretz 2013; Peretz et al. 2002; Peretz and Hyde 2003), melodic contour processing (Albouy et al. 2013), timbre perception (Marin et al. 2012) and pitch memory (Gosselin et al. 2009; Tillmann et al. 2009; Williamson and Stewart 2010). Experimental findings indicate that congenital amusia can also affect aspects of speech perception, such as the recognition of linguistic prosody. Interestingly, however, this deficit occurs mainly when speech is stripped of semantic information: affected individuals rarely show difficulties with natural speech and other complex sounds (e.g., environmental sounds) in everyday life (e.g., Ayotte et al. 2002; Liu et al. 2010; Patel et al. 2005). One possible explanation for this discrepancy is that amusics make use of other available cues, such as semantic or visual information, to compensate for their auditory deficits. For example, Thompson et al. (2012) found that individuals with amusia show reduced sensitivity to prosody of spoken sentences conveying a happy, sad, tender or irritated emotional state. When queried, the same individuals reported that they often rely on facial expressions and gestures to interpret the moods and feelings of people with whom they interact. Thus, the reduced sensitivity to emotional prosody may not pose a problem to amusic individuals in real life because they can fall back on cues delivered through the visual modality.

A number of researchers have noted the tendency to compensate for unreliable sensory information by drawing upon information from intact sensory systems (e.g., Lessard et al. 1998; Massaro and Light 2004; Neville 1990; Rauschecker 1995). For instance, individuals with hearing loss benefit substantially from visible cues, and hence engage in lip-reading for optimal speech recognition (Grant et al. 1998; Middelweerd and Plomp 1987; Massaro and Cohen 1999). Visual information arising from the face and lips helps to compensate for hearing impairment, leading to more efficient auditory processing (i.e., faster and more accurate recognition of speech). Neurologically, the dependency on visual cues by individuals with a handicap in auditory processing, such as cochlear implant recipients, results in increased brain activation within the visual cortex in response to meaningful sounds (e.g., words and environmental sounds), even in the absence of visual stimulation (Giraud et al. 2000; Giraud et al. 2001; Giraud and Truy 2002; Zatorre 2001). Even

for people with normal hearing, speech perception improves considerably when a speaker's lip movements are visible (Besle et al. 2004). As such, it is likely that amusic individuals also draw heavily from visual information during auditory tasks, in order to compensate for their deficits in pitch perception.

Albouy et al. (2015) provided indirect evidence that individuals with congenital amusia use a compensatory strategy. They showed that auditory perception in amusics benefits from the presence of informative visual information. Amusic and control participants were asked to detect a deviant pitch in a 5-tone sequence, which was simultaneously presented with or without visual stimuli – 5 dots appearing consecutively from the left to the right along a horizontal line but at the same vertical position. Even though the visual stimuli provided no task-relevant (pitch) information about the deviant tone (it did not vary in height, for example), the mere presence of the dots facilitated amusics' ability to detect the deviant. The authors argued that the presence of dots provided an additional temporal signal that "prepared" the processing of the concurrent auditory stimuli, thereby enhancing detection of deviant pitches (Jones 1976; Nickerson 1973). However, it is also possible that the visual stimuli improved performance by increasing the confidence levels of participants with amusia. If amusics tend to use visual information in daily life to compensate for their auditory difficulties, then they may have found it reassuring to have a visual accompaniment while judging the auditory signal. In other words, because amusics believe visual information is potentially valuable, its presence may increase their confidence in decision making, which may benefit performance.

To explore this issue further, we required amusics and normal hearing controls to judge the direction (up or down) of pitch change in two consecutive tones. Pitch change stimuli were presented by themselves (auditory only, henceforth AO) or concurrently with visual cues on a computer screen (audio-visual, henceforth AV). In the AV condition, a sequence of two dots appeared on the screen, the second of which appeared either above or below the first one. Within the AV condition, the pitch change of the two tones was either congruent (second tone higher pitch and second dot higher location, or second tone lower pitch and second dot lower location) or incongruent (second tone higher pitch and second dot lower location, or vice versa) with the spatial change of the two dots. If the visual enhancement effect is entirely

task independent, we predicted that amusics should show the same increment in performance for both AV conditions relative to AO condition, because the temporal cues and visual information in the AV-congruent and AV-incongruent displays are identical. However if there is an element of task-dependence in the enhancement effect, then amusics should show a larger increment in performance in the AV-congruent task relative to the AV-incongruent task. Notably, the intervals involved in Albouy et al.'s study (i.e., 12.5, 25 and 50 cents) were below or close to amusics' pitch threshold for pitch change detection (Foxton et al. 2004; Hyde and Peretz 2004; Tillmann et al. 2009). Therefore, we cannot tell whether facilitation by visual information is restricted to ambiguous contexts involving small intervals, or whether it can be observed even for pitch intervals that are well above the pitch threshold for amusic individuals. Typically, people combine and integrate multiple sources of sensory information by weighting them in proportion to their reliability (e.g., Alais and Burr 2004; Ernst and Banks 2002). Therefore we also examined performance in detecting both near threshold and well-above threshold pitch changes.

The use of congruent and incongruent AV pairings in the present study provided an interesting opportunity to obtain additional information about the neural mechanisms underlying visual enhancement of auditory information in amusics. A number of EEG studies have reported that multisensory stimuli that are incongruent in spatial location or temporal synchrony elicit a N2 component, a negative polarity event-related potential (ERP) component with a latency of about 200 ms after the multisensory stimulus onset (Forster and Pavone 2008; Lindstrom et al. 2012). This component is strongly associated with conflict detection at both a response level and a stimulus-representation level (Yeung et al. 2004), and its amplitude is typically larger following incongruent than congruent stimuli (Nieuwenhuis et al. 2003). We reasoned that if amusic individuals tend to ignore AV conflicts because of an over reliance on visual information, then the N2 response to incongruent stimuli may be attenuated for amusic individuals but not control participants. However, an inability to detect the AV conflict in amusics (i.e., failure to discriminate incongruent from congruent pairings) would also account for an absence of N2 effect. To assess whether amusics can explicitly detect AV conflicts, we included a task in which participants were instructed to pay attention to both auditory and visual information, and indicate whether the auditory and visual changes were congruent with one

another. In addition, the combination of auditory and visual stimuli allowed us to examine the N1 response to multisensory stimuli, which reflects a relatively earlier stage of perceptual processing compared to the N2 component (Luck 2014). Amusic individuals tend to show normal N1 responses to pitch changes despite their impairments in fine-grained pitch perception (Moreau et al. 2009; Peretz et al. 2005; but see Albouy et al. 2013); but it is unknown whether amusics exhibit typical N1 responses to multisensory stimuli.

2. Materials and method

2.1 Participants

Sixteen individuals with congenital amusia and 16 control participants took part in the present study. To screen for individuals with deficits in pitch processing, three melodic subtests (Scale, Contour, and Interval) of the Montreal Battery of Evaluation of Amusia (MBEA; Peretz et al. 2003) were administered. Participants

Table 1

Participants' characteristics, mean ± SD percent correct on the melodic subtests of MBEA, and independent-sample t-test results between the amusic and control group. Amusic participants performed significantly worse than control participants in the three melodic subtests of MBEA (all $p < 0.001$). However, the two groups were matched in other aspects (all $p > 0.40$). DF refers to the degrees of freedom and is corrected if the equal variances assumption is violated.

	Amusics	Controls	DF	t	p (2-tailed)
Age	21.28 ± 3.89	20.61 ± 2.82	30	0.56	0.58
Gender	8F/8M	8F/8M	–	–	–
Year of education	14.38 ± 1.89	14.13 ± 1.78	30	0.38	0.70
Years of musical training	0.53 ± 1.02	0.72 ± 1.22	30	-0.47	0.64
Hours of music listening daily	2.55 ± 2.55	1.97 ± 0.85	18.25	0.87	0.40
MBEA (percent correct)					
Scale	64.58 ± 10.17	92.50 ± 6.50	25.49	-9.25 ***	< 0.001
Contour	67.29 ± 11.00	86.88 ± 7.93	30	-5.79 ***	< 0.001
Interval	64.38 ± 10.66	85.00 ± 5.02	21.34	-7.00 ***	< 0.001
Total	65.42 ± 4.85	88.13 ± 4. 60	30	-13.59 ***	< 0.001

*** $p < 0.001$.

were diagnosed as congenital amusics when their composite scores (based on the melodic subtests) were equal or less than 65 out of 90 points, that is, 72% correct (Liu et al. 2010). All participants were right-handed and had normal hearing and normal or corrected-to-normal vision. None reported any auditory, neurological, or psychiatric disorder. The amusic and control group were matched in terms of age, gender, and education (see Table 1). The study was approved by the Macquarie University Ethics Committee, and written informed consent for participation was obtained from all participants prior to testing.

2.2 Stimuli

Auditory and visual stimuli were delivered separately in two unimodal tasks and then combined with each other in two bimodal tasks (see *Tasks and conditions* below). The auditory stimuli comprised 800 ms tones, which had a flute timbre and were generated with the computer software Garageband (Version 6.0.4, Apple Inc., USA). The visual stimuli were white dots (60 × 60 pixels, screen resolution: 1980 × 1024 pixels) presented for 800 ms at the center of a black computer background. Each trial contained two stimuli. We refer to the first and second stimulus in a pair as the *standard* and *target*, respectively. The auditory standard could have one of the following tones in C major scale – C (261.63 Hz), D (293.66 Hz), E (329.63 Hz), F (349.23 Hz), and G (392 Hz). The pitch of the auditory target was shifted upward or downward by 3 and 4 semitones (small interval) or 8 and 9 semitones (large interval) with respect to the standard. These intervals were selected as amusic individuals show a relatively higher threshold for pitch direction discrimination (Liu et al. 2010). The visual standard always appeared at the center of the computer screen, while the visual target could be shifted 300 pixels upward or downward along with the vertical line in relation to the standard. In the bimodal tasks, auditory and visual stimuli were presented simultaneously, and auditory and visual target could be congruent or incongruent in terms of the direction of their shift (e.g., *congruent*: both upward; *incongruent*: auditory upward, but visual downward).

2.3 Tasks and conditions

Participants performed four different tasks – two unimodal and two bimodal – in four blocks of trials (for more details, see *Procedure* below). In the AO task, participants were asked to judge whether the auditory target was shifted upward or

downward with respect to the standard. In a similar manner, they had to indicate whether the visual target was shifted upward or downward with respect to the standard in the visual only (henceforth, VO) task. In the AV implicit task, participants performed the same task as in the AO task but were instructed to ignore the visual information. Finally, in the AV explicit task, participants were required to determine whether the visual change in direction was congruent with the change in pitch. Participants performed the four tasks in the following fixed order: (1) AO, (2) AV implicit, (3) AV explicit, and (4) VO.

2.4 Procedure

As illustrated in Fig. 1, each trial began with a fixation-cross that appeared for 500-800 ms randomly in the center of the screen. Subsequently, standard and target stimuli (each of 800 ms duration) were presented consecutively with a jittered inter-stimulus interval (ISI) of 300-500 ms. To minimize the possibility of interference at the response stage, participants made their responses by pressing one of two response keys with no time limit. The inter-trial interval (ITI) was also jittered
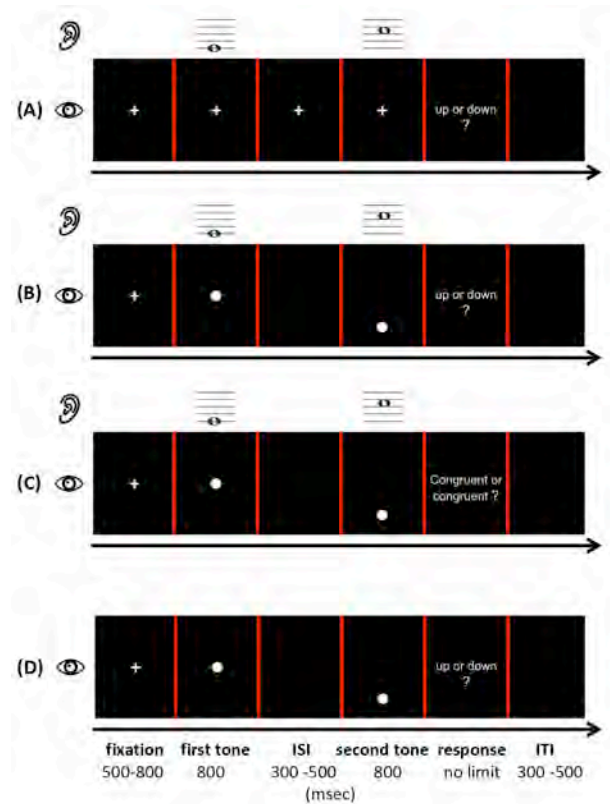
Fig. 1. Schematic illustration of the stimuli employed for four experimental tasks: (A) AO task, (B) AV implicit task, (C) AV explicit task, and (D) VO task.

between 300-500 ms. Participants underwent 240 trials in total in each AV task, that is, 60 trials per interval (small and large) and congruency (congruent and incongruent). The total number of trials was 120 (60 trials per interval) for the AO task, and 60 for the VO task. All trials were presented in randomized order. Participants were given a break after every 60 trials. They completed ten practice trials prior to each block. Feedback was provided in the practice but not in the actual experiment. Participants were seated approximately 50 cm from the computer screen in an electrically shielded and sound-attenuated room with dimmed light. All sounds were presented at a comfortable level via headphones (Sennheiser HD 280).

2.5 EEG Recording

The EEG was recorded with a sampling rate of 1000 Hz from 32 Ag-AgCl electrodes placed according to the extended International 10-20 electrode system and mounted to a head-cap (EASYCAP GmbH, Germany). The left mastoid electrode served as reference electrode during the recording. Four additional electrodes were employed to monitor horizontal and vertical eye movements. Electrode impedances were below 5 kΩ. Using a SynAmps 2.0 RT amplifier (Compumedics Neuroscan, USA), the EEG was filtered online with an analogue band-pass filter (0.05-200 Hz).

2.6 EEG Processing

Offline processing of the EEG was performed in MATLAB (R2013b; Mathwork, USA) using the EEGLAB toolbox (Delorme and Makeig 2004). The raw data was first downsampled to 500 Hz and re-referenced to the average of the left and right mastoids. It was then segmented into one-second epochs extending from 200 ms before to 800 ms after the onset of the target. Trials with incorrect response[1] or in which the potential exceeded ±150 µV were excluded from further data processing. The mean of each epoch was removed (see Groppe et al. 2009 for details) before an independent component analysis (ICA) was performed using the *runica* algorithm. We used an ICA-based method to identify and reject trials with unusual activity (i.e., rejection by kurtosis; see Delorme et al. 2001 for details). In addition, we employed the EEGLAB plugin *ADJUST* to automatically reject ocular ICs (Mognon et al.

---

[1] The analyses on all trials (including the trials with incorrect response) yielded the same results, except that the interaction between Group and Congruency was non-significant within the N1 time window for the AV explicit task (p = 0.15). Please see full results in the Supplementary Table 5.

2011). After IC rejection, the EEG data was low-pass filtered with a Windowed Sinc FIR Filter (Widmann and Schröger 2012) and a cut-off frequency at 30 Hz (Blackman window; filter order: 276). Epochs were subsequently averaged per participant, condition and task and baseline corrected by subtracting a pre-stimulus interval of 200 ms.

2.7 ERP Analysis

The main analysis focused on the comparison of the ERP components of interest involved in congruent and incongruent visual information processing between amusic and control groups in both implicit and explicit AV tasks. Based on visual inspection of the grand averages and previous studies, two pronounced ERP deflections within the following time windows were selected: 100-180 ms and 260-380 ms after the stimulus onset, encompassing the N1 component and N2-P3 complex respectively (e.g., Folstein and Van Petten 2008; Talsma et al. 2007). The N2 ERP overlapped with a larger and longer duration P3. Consequently, although the N2 is a negative-going ERP, peak amplitudes in both congruent and incongruent conditions were positive. Note that even in such circumstances, the experimental contrast (e.g., subtracting the ERP in the congruent condition from that in the incongruent condition) yields a negative difference N2 "component" (Luck, 2014). Additionally, we identified two negative ERP deflections in the AO and VO tasks corresponding to auditory and visual N1 ERPs, peaking within a post-stimulus time window of 90-150 ms (Näätänen and Picton 1987) and 120-180 ms (Luck, 2014; Vogel and Luck, 2000), respectively. A positive ERP deflection within 200-380 ms after stimulus onset was found in both AO and VO tasks, although both groups showed a negative-going deflection in response to targets with small interval (see Fig. 3) within this time window in the AO task. Finally, for each participant, condition and task, we computed the mean amplitudes within these pre-defined time windows. Furthermore, given the topographical distributions of ERP components of interest, the analyses were conducted at the central sites (CP3, CPZ, CP4, C3, CZ, and C4) for the AO and two AV tasks, and at the centroparietal sites (CP3, CPZ, CP4, P3, PZ, and P4) for the VO task. Scalp topographies were assessed using all scalp electrodes.

2.8 Statistical analysis

For the statistical analysis, task performance on unimodal tasks and the AV implicit task was evaluated by mean percent correct (PC). However, we computed d-prime (d') (Macmillan and Creelman 2005) for the AV explicit task to exclude the potential response biases reported in previous amusia studies (e.g., Henry and McAuley 2013; Pfeifer and Hamann 2015). Repeated-measures ANOVA for the AV implicit task was conducted, with the between-subject factor Group (amusic and control) and the within-subject factors Interval (large and small) and Congruency (congruent and incongruent). The factor Congruency was excluded from the analysis of the AV explicit and AO tasks, and independent two-sample t-tests were conducted for the VO task. Similar analyses were conducted on the mean amplitudes of ERPs for each task and time window. Below we report and discuss in detail only significant results of interest (see Supplementary Tables 1-6 for full results).

3. Results

3.1 Behavioural results

We first compared the task performance of the amusic and control group in the two unimodal tasks. For the AO task, the results revealed a significant interaction between Group and Interval, $F(1, 30) = 14.78$, $p = 0.001$, $\eta_p^2 = 0.33$. Further analyses showed that amusic participants performed as well as controls when the interval was large (amusics: M = 0.95, SE = 0.01; controls: M = 0.96, SE = 0.01), $F(1, 30) = 0.03$, $p = 0.87$, $\eta_p^2 < 0.01$. However, their performance significantly deteriorated compared to that of controls when the interval was small (amusics: M = 0.80, SE = 0.02; controls: M = 0.89, SE = 0.02), $F(1, 30) = 7.71$, $p < 0.01$, $\eta_p^2 = 0.21$. In addition, both amusic and control groups performed better when interval size was large, $F(1, 30) = 89.31$, $p < 0.001$, $\eta_p^2 = 0.75$, and $F(1, 30) = 16.11$, $p < 0.001$, $\eta_p^2 = 0.35$, respectively. For the VO task, both groups performed exceedingly well (> 99%); the statistical results confirmed that there was no significant group difference, $t(30) = 1.75$, $p = 0.10$.

Next we examined how participants' performance was affected by the presence of visual information in the AV tasks. When participants were instructed to ignore visual information, as in the AV implicit task, we found all main effects and two-way interactions were significant (all $p < 0.01$). Although there was a trend, the three-way

interaction of Group, Congruency and Interval did not reach significance, $F(1, 30) = 3.49$, $p = 0.07$, $\eta_p^2 = 0.10$. Further examination of the interaction between Group and Congruency revealed that the performance of amusics was significantly worse than controls in small (amusics: $M = 0.72$, $SE = 0.03$; controls: $M = 0.87$, $SE = 0.03$) and large interval conditions (amusics: $M = 0.92$, $SE = 0.01$; controls: $M = 0.96$, $SE = 0.02$), $F(1, 30) = 12.63$, $p = 0.001$, $\eta_p^2 = 0.30$, when the visual change in direction was incongruent with the pitch change. However, amusics (small: $M = 0.88$, $SE = 0.02$; large: $M = 0.97$, $SE = 0.01$) performed just as well as controls (small: $M = 0.93$, $SE = 0.02$; large: $M = 0.98$, $SE = 0.01$), $F(1, 30) = 3.64$, $p = 0.07$, $\eta_p^2 = 0.11$, when visual and auditory information was congruent. In addition, all participants were affected by unattended visual information to some degree, revealed by the significant congruency effect in both amusic, $F(1, 30) = 58.39$, $p < 0.001$, $\eta_p^2 = 0.66$, and control groups, $F(1, 30) = 9.71$, $p < 0.01$, $\eta_p^2 = 0.24$.
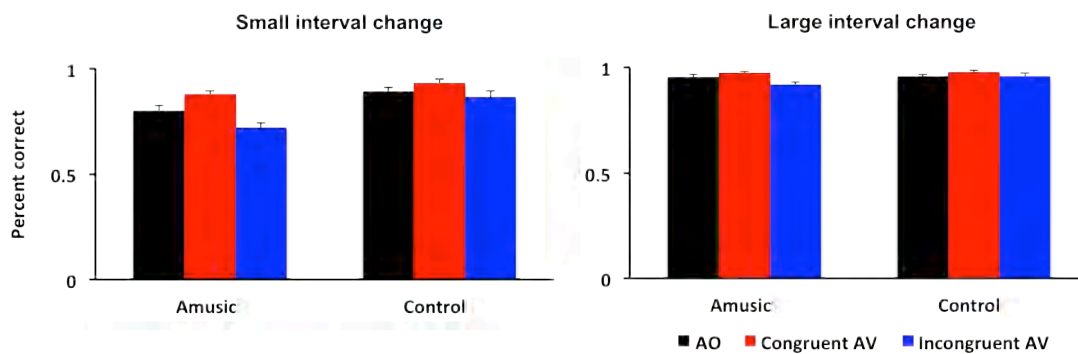
Fig. 2. Behavioural performance on the AO (black bar) and AV implicit (congruent AV pairings: red bar; incongruent AV pairings: blue bar) tasks evaluated by percent correct. Error bars represent 1 SE.

To further explore the extent to which unattended visual information affects one's pitch change direction identification, additional comparisons (separate paired-sample t-tests) between AO and implicit AV tasks were conducted for each group (see Fig. 2). The mean difference is considered significant at a significance level of 0.0125, due to the adjustment for multiple comparisons. When congruent visual information was presented, task performance of both amusics and controls was significantly improved in small interval condition (all $p < 0.006$), but not in large interval condition due to a ceiling effect (all $p > 0.03$). More interestingly, amusics' performance was significantly worse in the presence of incongruent visual

information, regardless of the interval size (all $p < 0.005$). By contrast, controls' performance was unaffected by incongruent visual information (all $p > 0.14$).

In the AV explicit task, participants judged the congruency of the visual and auditory information. Unsurprisingly, participants' sensitivity to AV incongruence increased with the increase of interval size between standard and target stimuli, revealed by a significant interval effect, $F(1, 30) = 70.28$, $p < 0.001$, $\eta_p^2 = 0.70$. Moreover, a significant group difference, $F(1, 30) = 5.07$, $p = 0.03$, $\eta_p^2 = 0.15$, and Interval $\times$ Group interaction were found, $F(1, 30) = 4.61$, $p = 0.04$, $\eta_p^2 = 0.13$. Post-hoc analysis revealed that amusics performed significantly worse in comparison to controls in small interval condition (amusics: $M = 1.77$, $SE = 0.26$; controls: $M = 2.75$; $SE = 0.26$), $F(1, 30) = 7.33$, $p < 0.05$, $\eta_p^2 = 0.20$, but not in large interval condition (amusics: $M = 3.23$, $SE = 0.21$; controls: $M = 3.61$; $SE = 0.21$), $F(1, 30) = 1.16$, $p = 0.22$, $\eta_p^2 = 0.05$.
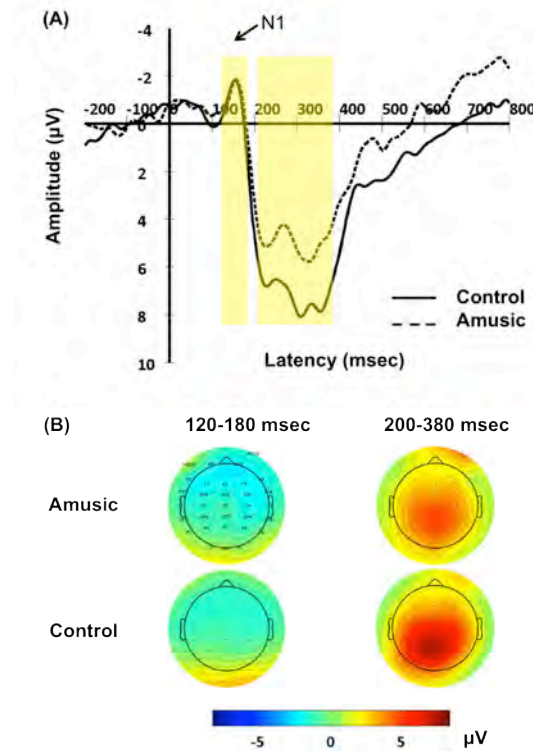
3.2 ERP results
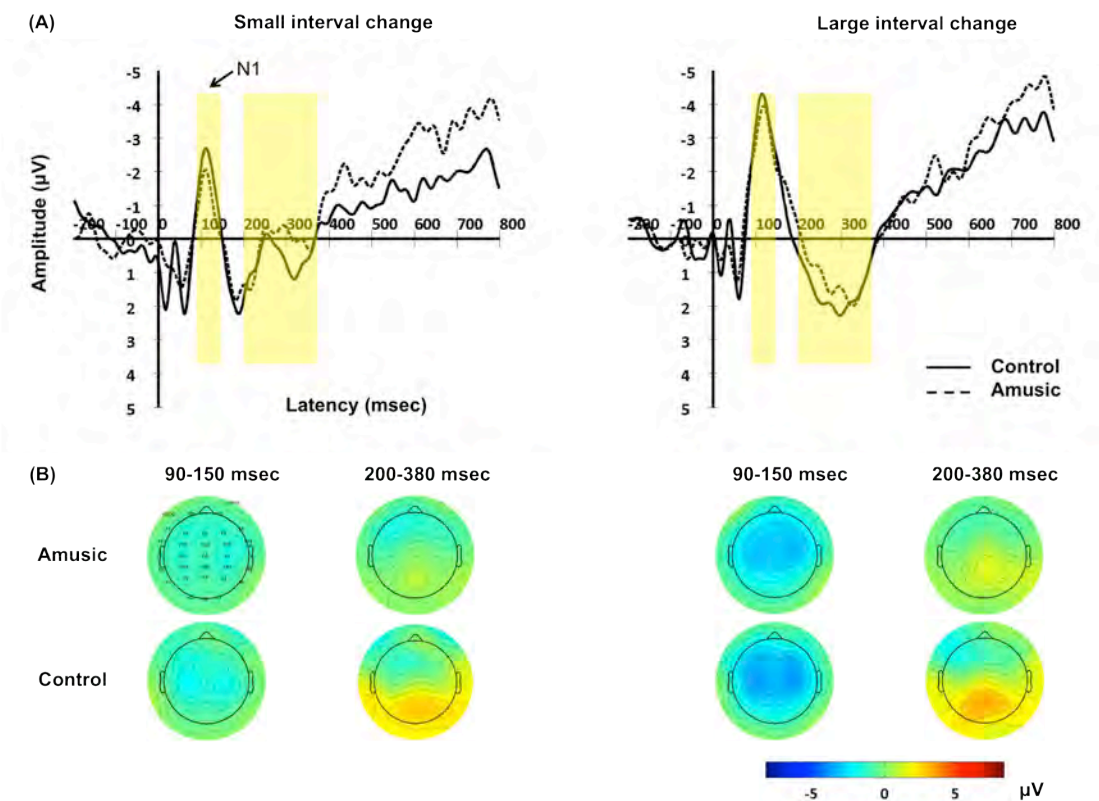
Fig. 3. ERP results of the AO task. (A) Grand-averaged ERPs of amusics (dash line) and controls (solid line) at electrode CZ. No group difference was observed in two time windows of interest (highlighted by yellow). (B) Group average topographic maps for each time window.

Amusic and control participants showed no significant differences in the AO and VO task within any of the selected time windows (all $p > 0.05$, see Fig. 3 and 4).

In line with behavioral results of the AO task, the amplitude of N1 component was larger for the large than small pitch change, $F(1, 30) = 8.14$, $p < 0.01$, $\eta_p^2 = 0.21$. No other significant effect was observed in unimodal tasks. Group differences were only found in the two bimodal tasks.
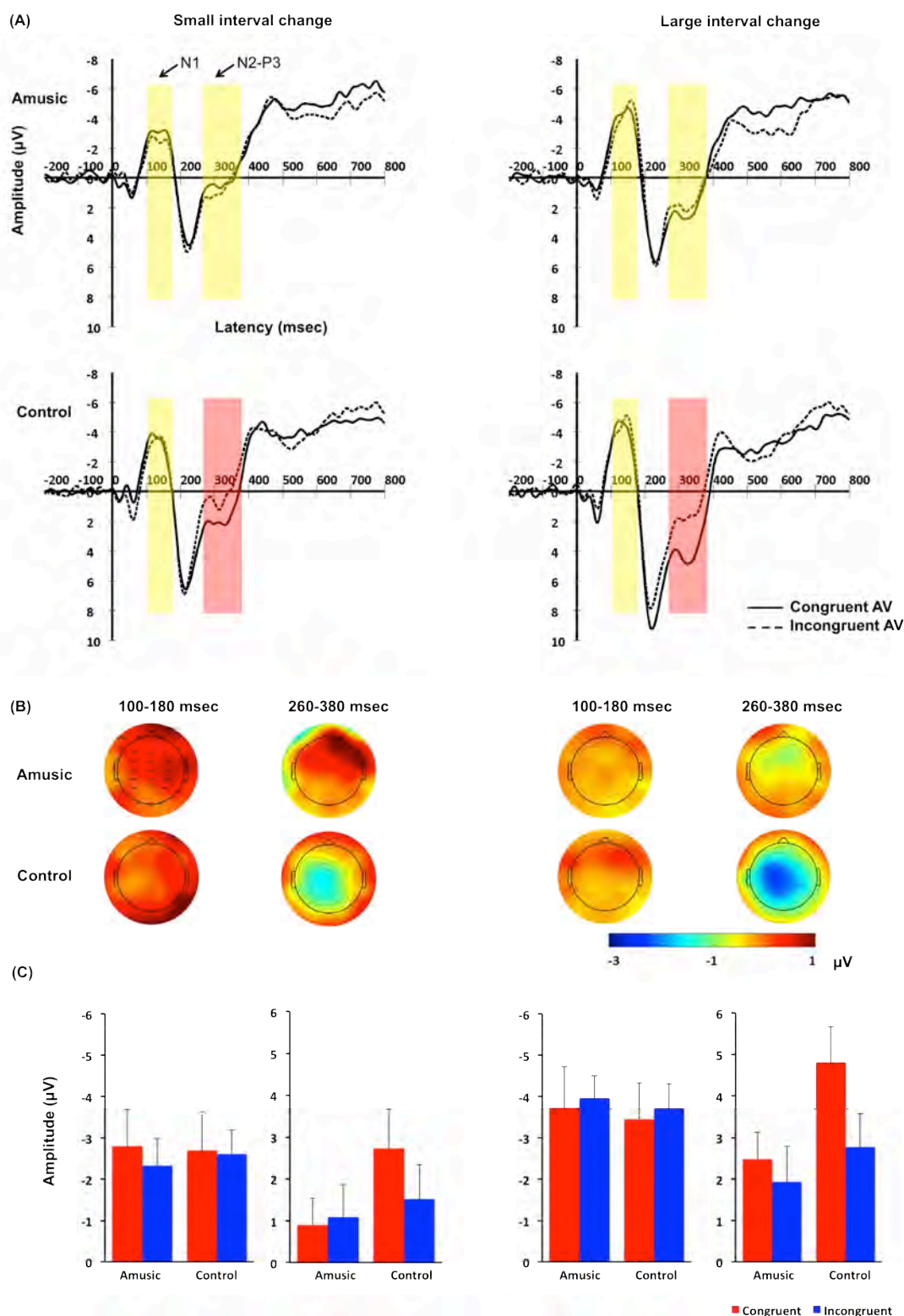

Fig. 4. ERP results of the VO task. (A) Grand-averaged ERPs of amusics and controls at electrode CPZ. No significant group difference was observed in either time window of interest (highlighted by yellow). (B) Group average topographic maps for each time window.

For the implicit task (see Fig. 5), both groups exhibited larger amplitude response to AV pairings with large interval change within the N1 and N2-P3 time windows, as revealed by the main effect of Interval with no other interactions, $F(1, 30) = 13.91$, $p = 0.001$, $\eta_p^2 = 0.32$, and $F(1, 30) = 48.36$, $p < 0.001$, $\eta_p^2 = 0.62$, respectively. More importantly, the statistical results yielded a significant interaction between Congruency and Group in the N2-P3 time window, $F(1, 30) = 4.75$, $p < 0.05$, $\eta_p^2 = 0.14$. For the control group, the negativity of the N2-P3 complex was greater for incongruent trials (M = 2.15 μV, SE = 0.83 μV) than for congruent trials (M = 3.77 μV, SE = 0.70 μV), $F(1, 30) = 12.02$, $p < 0.01$, $\eta_p^2 = 0.29$. For the amusic group, the enhanced negativity of the N2-P3 complex for incongruent trials was not found

(incongruent: M = 1.51 µV, SE = 0.83 µV; congruent: M = 1.69 µV, SE = 0.70 µV), $F(1, 30) = 0.15$, $p = 0.70$, $\eta_p^2 < 0.01$. No other significant effect was observed in the AV implicit task.

By contrast, in the AV explicit task, both control and amusic groups showed enhanced N2-P3 complex response to incongruent AV pairings (controls: M = 2.16 µV, SE = 0.84 µV; amusics: M = 1.00 µV, SE = 0.84 µV) relative to congruent AV pairings (controls: M = 3.61 µV, SE = 0.79 µV; amusics: M = 2.50 µV, SE = 0.79 µV) AV pairings, as the main effect of Congruency indicated, $F(1, 30) = 29.08$, $p < 0.001$, $\eta^2 = 0.49$. Furthermore, the interaction between Congruency and Group was found to be non-significant, $F(1, 30) = 0.01$, $p = 0.93$, $\eta_p^2 < 0.001$. Within the N1 time window, however, there was a significant interaction between Group and Congruency in the AV explicit task, $F(1, 30) = 5.84$, $p < 0.05$, $\eta_p^2 = 0.16$, with no three-way interaction involving Interval, $F(1, 30) = 0.19$, $p = 0.67$, $\eta_p^2 < 0.01$. As can be seen in Fig. 6, the mean N1 amplitude elicited by incongruent AV pairs is larger than that elicited by congruent AV pairings in the amusic group (congruent: M = -2.29 µV, SE = 0.76 µV; incongruent: M = -3.20 µV, SE = 0.72 µV), $F(1, 30) = 8.87$, $p < 0.01$, $\eta_p^2 = 0.23$. No such difference was observed for the control group (congruent: M = -2.56 µV, SE = 0.76 µV; incongruent: M = -2.43 µV, SE = 0.72 µV), $F(1, 30) = 0.19$, $p = 0.67$, $\eta_p^2 < 0.01$. It should be noted that the analyses on all trials (including trials with incorrect response) showed a different picture, as the interaction between Group and Congruency within the N1 time window was no longer significant, $F(1, 30) = 2.14$, $p = 0.15$, $\eta_p^2 = 0.07$. In addition, an enhanced negativity of the N1 and N2-P3 complex was elicited by large pitch changes across participants, suggested by the main effect of Interval, $F(1, 30) = 11.19$, $p < 0.01$, $\eta_p^2 = 0.27$, and $F(1, 30) = 30.40$, $p < 0.001$, $\eta_p^2 = 0.50$, respectively, which was in line with the results of the AV implicit task.

Fig. 5. Results of the AV implicit task. (A) Grand-averaged ERPs of amusic and control groups at electrode CZ in response to auditory stimuli with congruent (solid line) and incongruent (dash line) visual stimuli. For each time window, a significant congruency effect is highlighted by red, and a non-significant congruency effect is highlighted by yellow. (B) Topographic maps of difference amplitude by subtracting congruent trials from incongruent trials for each group within each time window of interest. (C) Mean amplitude in response to congruent trials (red bar) and incongruent trials (blue bar) over ROI for each time window. Error bars represent 1 SE.
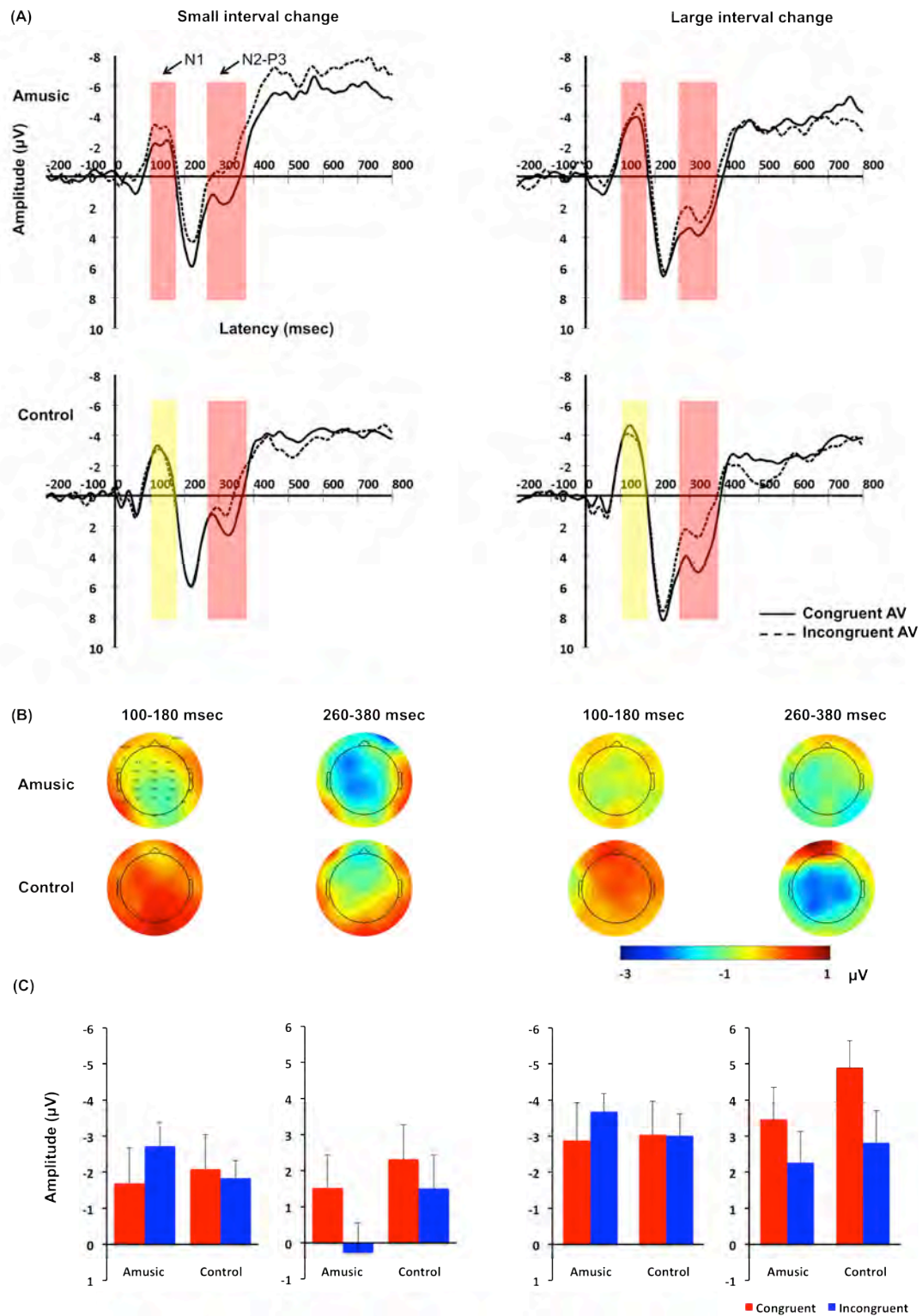
Fig. 6. Results of the AV explicit task. (A) Grand-averaged ERPs of amusic and control groups at electrode CZ in response to auditory stimuli with congruent (solid line) and incongruent (dash line) visual stimuli. For each time window, a significant congruency effect is highlighted by red, and a non-significant congruency effect is highlighted by yellow. (B) Topographic maps of difference amplitude by subtracting congruent trials from incongruent trials for each group within each time window of interest. (C) Mean amplitude in response to congruent trials (red bar) and incongruent trials (blue bar) over ROI for each time window. Error bars represent 1 SE.

4. Discussion

Congenital amusia is characterised by a variety of deficits in musical perception. The present study tested the hypothesis that these auditory processing deficits shown by amusic individuals can be compensated for by visuospatial information presented simultaneously. There were several salient observations from the present results:

(1) amusics performed significantly worse than controls in the AO task, when pitch changes were small;

(2) the two groups performed equally well in the VO task;

(3) in the presence of unattended visual information, the performance of amusics improved significantly when it was congruent with auditory information involving small interval change, but deteriorated significantly when they were incongruent;

(4) the performance of controls was only affected by unattended visual change in direction, when it was congruent with the direction of the small pitch change;

(5) in controls, a robust N2-P3 complex was elicited by incongruent AV pairings, whether implicit or explicit;

(6) in amusics, the N2-P3 complex was elicited by explicit AV pairings but not by implicit AV pairings.

4.1 Behavioural results

Amusic participants displayed difficulties in the AO task when the interval was small (i.e., 3 or 4 semitones), whereas they could identify the direction of the interval just as well as controls when the interval was large (i.e., 8 or 9 semitones). This finding is consistent with previous claims that amusics can discriminate large pitch differences because such differences exceed their threshold for discriminating pitch change direction (Liu et al. 2010; Williamson and Stewart 2010).

Bimodal stimuli are typically detected and recognized more rapidly and accurately than unimodal stimuli by both non-amusic (Besle et al. 2004; Giard and Peronnet 1999; Laurienti et al. 2004; Teder-Sälejärvi et al. 2005; Teder-Sälejärvi et al. 2002) and amusic participants (Albouy et al. 2015). Indeed, when congruent

visual information was supplied in the AV implicit task, amusics' ability to identify the contour of small intervals improved significantly than when auditory stimuli were presented alone. However, our results also showed that amusics' performance deteriorated significantly when the visual information was incongruent with the auditory information. More interestingly, the deterioration in performance shown by amusic participants, especially in large interval condition, did not result from an inability to detect the AV incongruence, as suggested by the results of the AV explicit task. Instead, the finding in the AV implicit task is most likely to reveal visual influences on auditory processing for amusics. The interference may occur at either the perceptual stage, where amusics perceived the changes in pitch direction, or the response stage, where they have to respond. Given that the task required participants to pay attention to the auditory modality while ignoring the visual modality, and allowed a non-speeded response, it is reasonable to conclude that the amusic individuals are readily affected by the unattended visual stimuli at the perceptual stage of auditory processing.

4.2 Neurophysiological results

For the AO task, amusics displayed comparable ERPs in comparison with control participants (see Fig. 3) despite their behaviourally impaired performance for small changes in pitch. This discrepancy between behavioral and neural results for the small interval condition has been observed in other investigations (Peretz et al. 2009; Peretz et al. 2005), and may indicate that early stages of auditory processing (as reflected in ERPs) are intact in amusics, when the task is not demanding (e.g., Moreau et al. 2009; see Albouy et al. 2013 for further discussion). With regard to the VO task, both amusic and control participants performed equally well and showed no difference in terms of ERPs.

For the AV implicit task, control participants exhibited an increase in the negativity of N2-P3 complex to incongruent relative to congruent AV pairings, suggesting a processing of conflict detection even when the visual information was task-irrelevant (Forster and Pavone 2008; Lindstrom et al. 2012; Nieuwenhuis et al. 2003; Yeung et al. 2004). By contrast, amusics failed to show a conflict response in this task, as reflected by an absence of N2-P3 effect. However, the absence of N2-P3 effect cannot be explained by impaired neural circuitry, as amusic participants

exhibited a similar increase in the negativity of N2-P3 complex to incongruent pairings when compared with congruent AV pairings in the AV explicit task, regardless of whether the interval was small or large. Therefore, the absence of N2-P3 effect in amusic individuals observed in the AV implicit task is more likely due to an unconscious neglect of AV conflict. This explanation is further supported by the discrepancy between behavioural and ERP results in the AV implicit task, in which amusics depended on unattended visual information to complete the auditory task, as behavioural results suggested, thereby ignoring AV conflicts, as reflected by the absence of N2-P3 complex.

Another discrepancy between behavioural and ERP results was found in the AV explicit task, in which amusics showed reduced sensitivity to AV incongruence in small interval condition, while exhibiting a comparable N2-P3 effect (i.e., enhanced N2-P3 complex in response to incongruent than congruent AV pairings), to controls. The failure of amusics to detect AV incongruence with a small interval change may reflect a limited access to pitch information consciously, even though their neural response to incongruent AV pairings differed from that to congruent pairings. This observation is in good agreement with the "pitch awareness hypothesis", which emphasises that pitch information can be registered but not accessed consciously by individuals who suffer from congenital amusia (e.g., Hyde et al. 2011; Loui et al. 2009; Loui et al. 2008; Loui et al. 2011; Lu et al. 2015; Moreau et al. 2013; 2009; Peretz et al. 2009). One prediction of the "pitch awareness hypothesis" is that amusics would show low confidence level in tasks that require explicit pitch processing owing to a lack of pitch awareness (Lu et al. under review). Therefore, it is possible that the discrepancy between behavioural and ERP results observed in the AV explicit task results from a low confidence level in cases with small interval changes.

In line with previous study examining the ERPs to neutral AV stimuli on individuals with normal hearing (Stekelenburg and Vroomen 2007), the N1 component shown by controls in both AV tasks did not depend on the informational congruency between audio and visual stimuli, revealed by comparable N1 amplitude for incongruent and congruent AV pairings. However, we observed significantly larger N1 amplitudes for incongruent than congruent AV pairings in the amusic group but not the control group in the explicit AV task, when the trials with correct

response included only (i.e., when participants detected the AV incongruence successfully). Interestingly, the effect shown by amusics was no longer significant when trials with incorrect response were also included. Since there were few trials with incorrect response, we can only speculate that the enhanced N1 for incongruent AV pairings is necessary for detecting AV conflict for amusics but not controls, resulting from a (strategic) mechanism that they use to compensate for deficient processing of pitch.

4.3 Implications

Based on the observations of the present study, we argue that amusics tend to make use of available visual cues in order to compensate for their difficulties in pitch perception. The tendency by amusics to use visual cues is compatible with the "optimal-integration hypothesis," which suggests that perceivers are more likely to rely on one modality over the other depending on how reliable the information is (Ernst and Banks 2002; Ernst and Bulthoff 2004). For instance, when presented with flashed visual stimuli accompanied by auditory beeps, the perceived number of flashes is influenced by the number of beeps (Shams et al. 2000). In this case, audition dominates over vision, as the auditory modality is more reliable and more precise at determining temporal information.

For amusic individuals, their pitch impairment means that auditory information is not reliable. Therefore, amusic individuals tend to use contextual or facial cues to boost their auditory perception (Albouy et al. 2015; Thompson et al. 2012). Our results show that amusics rely on unattended visual information even when pitch differences are well above their threshold. This finding extends the "optimal-integration hypothesis" to circumstances where there is a long-term bias in weighing visual over auditory information established on the basis of daily experience, generalizing the phenomenon from a particular source of reliability in a sensory modality to the entire modality.

5. Conclusion

In summary, the present study is the first ERP study showing that individuals with congenital amusia rely heavily on unattended visual information when doing auditory task due to their deficits in auditory processing, providing the theoretical

basis for using visual information to improve amusics' auditory perception. Furthermore, the presence of visual information may boost auditory encoding at the perceptual stage in congenital amusia, thereby leading to more accurate neural representations of input stimuli, which yields new insights into the mechanisms of the combination of auditory and visual information in this special population.

Conflict of interest

The authors declare no conflict of interest.

References

Alais, D., Burr, D., 2004. The ventriloquist effect results from near-optimal bimodal integration. Curr. Biol. 14, 257-262.

Albouy, P., Mattout, J., Bouet, R., Maby, E., Sanchez, G., Aguera, P. E., Daligault, S., Delpuech, C., Bertrand, O., Caclin, A., Tillmann, B., 2013. Impaired pitch perception and memory in congenital amusia: the deficit starts in the auditory cortex. Brain 136, 1639-1661.

Albouy, P., Leveque, Y., Hyde, K. L., Bouchet, P., Tillmann, B., Caclin, A., 2015. Boosting pitch encoding with audiovisual interactions in congenital amusia. Neuropsychologia 67, 111-120.

Ayotte, J., Peretz, I., Hyde, K., 2002. Congenital amusia: a group study of adults afflicted with a music-specific disorder. Brain 125, 238-251.

Besle, J., Fort, A., Delpuech, C., Giard, M. H., 2004. Bimodal speech: early suppressive visual effects in human auditory cortex. Eur. J. Neurosci. 20, 2225-2234.

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

Delorme, A., Makeig, S., Sejnowski, T., 2001. Automatic artifact rejection for EEG data using high-order statistics and independent component analysis. Proceedings of the Third International ICA Conference, pp. 9-12.

Ernst, M. O., Banks, M. S., 2002. Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415, 429-433.

Ernst, M. O., Bulthoff, H. H., 2004. Merging the senses into a robust percept. Trends Cogn. Sci. 8, 162-169.

Folstein, J. R., van Petten, C., 2008. Influence of cognitive control and mismatch on the N2 component of the ERP: a review. Psychophysiology 45, 152-170.

Forster, B., Pavone, E. F., 2008. Electrophysiological correlates of crossmodal visual distractor congruency effects: evidence for response conflict. Cogn. Affect. Behav. Neurosci. 8, 65-73.

Foxton, J. M., Dean, J. L., Gee, R., Peretz, I., Griffiths, T. D., 2004. Characterization of deficits in pitch perception underlying "tone deafness". Brain 127, 801-810.

Giard, M. H., Peronnet, F., 1999. Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. J. Cogn. Neurosci. 11, 473-490.

Giraud, A. L., Truy, E., 2002. The contribution of visual areas to speech comprehension: a PET study in cochlear implants patients and normal-hearing subjects. Neuropsychologia 40, 1562-1569.

Giraud, A. L., Lorenzi, C., Ashburner, J., Wable, J., Johnsrude, I., Frackowiak, R., Kleinschmidt, A., 2000. Representation of the temporal envelope of sounds in the human brain. J. Neurophysiol. 84, 1588-1598.

Giraud, A. L., Price, C. J., Graham, J. M., Truy, E., Frackowiak, R. S., 2001. Cross-modal plasticity underpins language recovery after cochlear implantation. Neuron 30, 657-664.

Gosselin, N., Jolicoeur, P., Peretz, I., 2009. Impaired memory for pitch in congenital amusia. Ann. N. Y. Acad. Sci. 1169, 270-272.

Grant, K. W., Walden, B. E., Seitz, P. F., 1998. Auditory-visual speech recognition by hearing-impaired subjects; Consonat recognition, sentence recognition, and auditory-visual integration. J. Acoust. Soc. Am. 103, 2677-2690.

Groppe, D. M., Makeig, S., Kutas, M., 2009. Identifying reliable independent components via split-half comparisons. NeuroImage 45, 1199-1211.

Henry, M. J., McAuley, J. D., 2013. Failure to apply signal detection theory to the Montreal Battery of Evaluation of Amusia may misdiagnose amusia. Music Percept. 30, 480-496.

Hyde, K. L., Peretz, I., 2004. Brains that are out of tune but in time. Psychol. Sci. 15, 356-360.

Hyde, K. L., Zatorre, R. J., Peretz, I., 2011. Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. Cereb. Cortex, 21, 292-299.

Jones, M. R., 1976. Levels of structure in the reconstruction of temporal and spatial serial patterns. J. Exp. Psychol. 2, 475-488.

Laurienti, P. J., Kraft, R. A., Maldjian, J. A., Burdette, J. H., Wallace, M. T., 2004. Semantic congruence is a critical factor in multisensory behavioral performance. Exp. Brain Res. 158, 405-414.

Lessard, N., Pare, M., Lepore, F., Lassonde, M., 1998. Early-blind human subjects localize sound sources better than sighted subjects. Nature, 395, 278-280.

Lindstrom, R., Paavilainen, P., Kujala, T., Tervaniemi, M., 2012. Processing of audiovisual associations in the human brain: dependency on expectations and rule complexity. Front. Psychol. 3, 159.

Liu, F., Patel, A. D., Fourcin, A., Stewart, L., 2010. Intonation processing in congenital amusia: discrimination, identification and imitation. Brain 133, 1682-1693.

Loui, P., Alsop, D., Schlaug, G., 2009. Tone deafness: a new disconnection syndrome? J. Neurosci. 29, 10215-10220.

Loui, P., Guenther, F. H., Mathys, C., Schlaug, G., 2008. Action-perception mismatch in tone-deafness. Curr. Biol. 18, R331-332.

Loui, P., Kroog, K., Zuk, J., Winner, E., Schlaug, G., 2011. Relating pitch awareness to phonemic awareness in children: implications for tone-deafness and dyslexia. Front. Psychol. 2, 111.

Lu, X., Ho, H. T., Liu, F., Wu, D., Thompson, W. F., 2015. Intonation processing deficits of emotional words among Mandarin Chinese speakers with congenital amusia: an ERP study. Front. Psychol. 6, 385.

Luck, S., 2014. An introduction to the event-related potential technique. MIT Press, Cambridge.

Macmillan, N. A., Creelman, C. D., 2005. Detection Theory: A User's Guide. Erlbaum, Mahwah, NJ.

Marin, M. M., Gingras, B., Stewart, L., 2012. Perception of musical timbre in congenital amusia: categorization, discrimination and short-term memory. Neuropsychologia 50, 367-378.

Massaro, D. W., Cohen, M. M., 1999. Speech perception in perceivers with hearing loss. J. Speech Lang. Hear. Res., 42, 21-41.

Massaro, D. W., Light, J., 2004. Using visible speech to train perception and production of speech for individuals with hearing loss. J. Speech Lang. Hear. Res. 47, 304-320.

Middelweerd, M. J., Plomp, R., 1987. The effect of speechreading on the speech-reception threshold of sentences in noise. J. Acoust. Soc. Am. 82, 2145-2147.

Mognon, A., Jovicich, J., Bruzzone, L., Buiatti, M., 2011. ADJUST: an automatic EEG artifact detector based on the joint use of spatial and temporal features. Psychophysiology 48, 229-240.

Moreau, P., Jolicœur, P., Peretz, I., 2009. Automatic brain responses to pitch changes in congenital amusia. Ann. N. Y. Acad. Sci. 1169, 191-194.

Moreau, P., Jolicoeur, P., Peretz, I., 2013. Pitch discrimination without awareness in congenital amusia: Evidence from event-related potentials. Brain Cogn. 81, 337-344.

Näätänen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology 24, 375-425.

Neville, H. J., 1990. Intermodal competition and compenstation in development. Evidence from studies of the visual system in congenitally deaf aduults. Ann. N. Y. Acad. Sci. 608, 71-91.

Nickerson, R. S., 1973. Intersensory facilitation of reaction time: energy summation or preparation enhancement? Psychol. Rev. 80, 489-509.

Nieuwenhuis, S., Yeung, N., van den Wildenberg, W., Ridderinkhof, K. R., 2003. Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. Cogn. Affect. Behav. Neurosci. 3, 17-26.

Patel, A. D., Foxton, J. M., Griffiths, T. D., 2005. Musically tone-deaf individuals have difficulty discriminating intonation contours extracted from speech. Brain Cogn. 59, 310-313.

Peretz, I., 2013. The biological foundations of music: Insights from congenital amusia. In: Deustch, D., (Ed.), The psychology of music. Acdemic Press, San Diego, pp. 551-564.

Peretz, I., Hyde, K. L., 2003. What is specific to music processing? Insights from congenital amusia. Trends Cogn. Sci. 7, 362-367.

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., Jutras, B., 2002. Congenital amusia: a disorder of fine-grained pitch discrimination. Neuron 33, 185-191.

Peretz, I., Champod, A. S., Hyde, K., 2003. Varieties of musical disorders: the Montreal Battery of Evaluation of Amusia. Ann. N. Y. Acad. Sci. 999, 58-75.

Peretz, I., Brattico, E., Tervaniemi, M., 2005. Abnormal electrical brain responses to pitch in congenital amusia. Ann. Neurol. 58, 478-482.

Peretz, I., Brattico, E., Jarvenpaa, M., Tervaniemi, M., 2009. The amusic brain: in tune, out of key, and unaware. Brain 132, 1277-1286.

Pfeifer, J., Hamann, S., 2015. Revising the diagnosis of congenital amusia with the Montreal Battery of Evaluation of Amusia. Front. Hum. Neurosci. 9, 161.

Rauschecker, J. P., 1995. Compensatory plasticity and sensory substitution in the cerebral cortex. Trends Neurosci. 18, 36-43.

Shams, L., Kamitani, Y., Shimojo, S., 2000. Illusions: What you see is what you hear. Nature, 408, 788.

Stekelenburg, J. J., Vroomen, J., 2007. Neural correlates of multisensory integration of ecologically valid audiovisual events. J. Cogn. Neurosci. 19, 1964-1973.

Talsma, D., Doty, T. J., Woldorff, M. G., 2007. Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? Cereb. Cortex 17, 679-690.

Teder-Sälejärvi, W. A., McDonald, J. J., Di Russo, F., Hillyard, S. A., 2002. An analysis of audio-visual crossmodal integration by means of event-related potential (ERP) recordings. Cogn. Brain Res. 14, 106-114.

Teder-Sälejärvi, W. A., Di Russo, F., McDonald, J. J., Hillyard, S. A., 2005. Effects of spatial congruity on audio-visual multimodal integration. J. Cogn. Neurosci. 17, 1396-1409.

Thompson, W. F., Marin, M. M., Stewart, L., 2012. Reduced sensitivity to emotional prosody in congenital amusia rekindles the musical protolanguage hypothesis. Proc. Natl. Acad. Sci. U. S. A. 109, 19027-19032.

Tillmann, B., Schulze, K., Foxton, J., 2009. Congenital amusia: a short-term memory deficit for non-verbal, but not verbal sounds. Brain Cog. 71, 259-264.

Vogel, E. K., Luck, S. J., 2000. The visual N1 component as an index of a discrimination process. Psychophysiology 37, 190-203.

Widmann, A., Schröger, E., 2012. Filter effects and filter artifacts in the analysis of electrophysiological data. Front. Psychol. 3, 233.

Williamson, V. J., Stewart, L., 2010. Memory for pitch in congenital amusia: beyond a fine-grained pitch discrimination problem. Memory 18, 657-669.

Yeung, N., Botvinick, M. M., Cohen, J. D., 2004. The neural basis of error detection: conflict monitoring and the error-related negativity. Psychol. Rev. 111, 931-959.

Zatorre, R.J., 2001. Do you see what I'm saying? Interactions between auditory and visual cortices in cochlear implant users. Neuron 31, 13-14.

Chapter 7: Conclusion

To better understand congenital amusia and to shed light on cognitive mechanisms underlying the spatial representations of music, I investigated the spatial representations of pitch and explored the influences of visuospatial information on auditory processing in amusic individuals. In this Chapter, the experimental results (described in Chapters 3 to 6) are first reviewed. Next, the primary implications of the research, including the mechanisms underlying congenital amusia, are discussed. Finally, I propose a framework for understanding auditory and visual contributions to pitch processing in congenital amusia.

Summary of main experimental findings

Chapter 3 described two experiments that were conducted to address the intriguing association between pitch and spatial height in individuals with congenital amusia. Using a stimulus-response capability (SRC) paradigm, it showed that pitch-space associations exist in amusic individuals, just as they do in individuals with typical music perception capabilities. Such associations were revealed by the finding that pitch classification was faster and more accurate when correct responses involved a physical action that was spatially congruent with the pitch height of the stimulus than when it was incongruent (i.e., the SRC effect). In addition, this association was evident for both amusic and non-amusic participants even when pitch was irrelevant to the task and was supposed to be ignored in a colour classification task. However, amusic participants exhibited slower response times than non-amusic participants in the pitch classification (explicit) task but not in the colour classification (implicit) task, suggesting that spatial representation of pitch for amusic individuals is not as stable as that for non-amusic individuals when the task requires explicit processing of pitch.

To further confirm that the group differences in response times observed in the pitch classification task were not due to a general difference in the perceived size of pitch intervals for amusic and non-amusic individuals, a control experiment examining the perceived magnitude of supra-threshold pitch changes was reported in

Chapter 4. The results indicate that amusic individuals perceived pitch changes equivalently to control individuals when intervals were above their threshold for pitch discrimination, suggesting that amusic individuals have similar perceptions of pitch interval size as long as the separation between two tones exceeds their pitch discrimination thresholds.

In Chapter 5, I attempted to determine whether the impaired contour processing shown by amusic individuals results from poor pitch memory or from a failure in identifying the direction of changes in pitch. One of the major contributions of this study was the development of a novel Self-paced Audio-visual Contour Test (SACT), which allowed me to disentangle the perceptual stage of contour processing from the storage stage by placing minimal demands on memory. Using the SACT, I found that amusic participants still showed reduced sensitivity to the conflict between the pitch and visual contour, suggesting that the deficit in pitch contour processing occurs at the perceptual stage of processing. However, amusic individuals did show some implicit knowledge of pitch change direction, as their response times to incongruent trials (i.e., audio and visual stimuli move in opposite directions) were longer than those to congruent trials (i.e., audio and visual stimuli move in the same direction), even when they were not explicitly aware of the incongruence. In the second experiment, the SACT was employed to examine contour processing in two auditory dimensions other than pitch: timbral brightness and loudness. The results showed that the impaired contour processing observed in the first experiment does not extend to other auditory dimensions, and hence is not associated with a general problem in cross-modal contour mapping.

In Chapter 6, I explored how visuospatial information affects auditory processing in amusic individuals. In particular, I used EEG to evaluate the hypothesis that amusic individuals rely heavily on visuospatial information in order to compensate for their proposed neurophysiological deficits in auditory processing. In this study, participants were presented with pairs of tones that differ in pitch and pairs of dots that differ in spatial position (up or down) simultaneously. The change in dot position was either congruent or incongruent with the change in pitch. As predicted, amusics' performance in judging the direction of pitch change was significantly improved when the visual change in direction was incongruent with the pitch change, while it was deteriorated significantly when auditory and visual

information was incongruent. In contrast, the influence of unattended visual information was less evident for non-amusics. This finding suggests that amusic participants are readily misled by unattended visual information in auditory processing, regardless of whether that visual information is congruent or incongruent with the auditory stimulus. Examination of ERP results showed an absence of the N2-P3 effect (expected when there is an audio-visual incongruence) for amusic participants but not for non-amusic participants when unattended visual stimuli presented in synchrony with auditory stimuli. Given that the N2 component (N2-P3 complex in my case) is strongly associated with conflict detection (Yeung, Botvinick, & Cohen, 2004), and its amplitude is typically larger following incongruent than congruent stimuli (Nieuwenhuis, Yeung, van den Wildenberg, & Ridderinkhof, 2003), the absence of the N2-P3 effect among amusics indicates a failure in detecting audio-visual incongruence. However, when participants were explicitly directed to auditory and visual information at the same time, both groups exhibited the N2-P3 effect, suggesting amusic individuals have no difficulty in extracting information from both audio and visual modalities. Therefore, it is most likely that amusics may have utilised unattended visual information to complete the auditory task, and largely failed to process audio-visual incongruence. In other words, amusic individuals are biased towards relying on visual information when it is available.

Pitch and space association

The SRC effect observed in Chapter 3 revealed that the association between pitch and vertical space exists even for individuals with deficits in pitch perception. In line with a considerable amount of evidence showing that musical pitch processing interacts with visuospatial representation in individuals with and without musical training (Küssner, Tidhar, Prior, & Leech-Wilkinson, 2014; Lega, Cattaneo, Merabet, Vecchi, & Cucchi, 2014; Lidji, Kolinsky, Lochy, & Morais, 2007; Rusconi, Kwan, Giordano, Umilta, & Butterworth, 2006; Stewart, Verdonschot, Nasralla, & Lanipekun, 2012; Taylor & Witt, 2014), this finding suggests that people are able to implement a vertical representation of pitch, regardless of their level of musical skill.

Although Pratt (1930) attributed the association between pitch and height to a psychological bias in the perception of pitch, it is still unknown if the association between pitch and height is innate or not, as he did not test whether pitch can be

mapped to other dimensions and other sensory domains. Nevertheless, two theories have been proposed based on the investigation of listeners with normal music perception abilities. One theory posits that the processing of pitch and space share common representational and processing resources (Connell, Cai, & Holler, 2013), reflecting an innate cross-modal correspondence that is independent of experience or cultural norms (Walker et al., 2010). The other theory emphasises that the association emerges following an abundant experience with pitch-space mapping in daily life (see Spence, 2011, for a review). More specifically, the association may be obtained on the basis of bodily experience (Miller, 1986; Proctor & Vu, 2002). For instance, Stumpf (1883, as cited in Pratt 1930) argued that such an association might be derived from the multisensory experiences elicited by acoustic stimuli in the environment, such as high-pitched sounds often produced by birds up in the sky. This association has been confirmed by Parise, Knorre, and Ernst work (2014), who found that the spatial connotation of pitch corresponds to statistical associations within the natural auditory environment. Furthermore, the association might be strengthened or weakened with language development as certain terms come to be associated with more than one perceptual continuum semantically (Dolscheid, Shayan, Majid, & Casasanto, 2013; Eitan, 2013). Either theory can lead to a prediction that individuals with impaired pitch perception should have impaired spatial representation to some degree. Indeed, data shown in Chapter 3 revealed that amusic individuals have an unstable representation of pitch, as suggested by slower response times to classify the pitch of a tone explicitly when compared with non-amusic participants. Douglas and Bilkey (2007) also argue that amusic participants have a weaker link between spatial representation and pitch discrimination, as the amusic group showed a reduced SRC effect compared with the control group. However, it should be noted that only a trend existed for the comparison between amusic and non-musician participants, when the control group was split into the two subgroups of non-musicians and musicians.

On the other hand, it has been suggested that the judgements of tones closer in pitch yield longer reaction times than more distant tones (Elkin & Leuthold, 2011). Therefore, it is also possible that the slower response to pitch shown by amusic individuals reflects a difference in the perceived level of pitch change compared to non-amusic individuals. That is, a given change in fundamental frequency might be

perceived to be a smaller interval by amusic participants than by control participants. If that is the case, then response times might be predicted by the psychophysical function that relates physical with perceived change in pitch. Although this possibility had not been considered in the original study described in Chapter 3, I subsequently evaluated this hypothesis in Chapter 4, and found that the estimation of interval size (i.e., 27 semitones) was equivalent for a group of amusics and their matched controls. This finding suggests that slower response times shown by amusic participants in the pitch classification task cannot be explained as a difference in the perceived size of pitch changes for amusics and non-amusics. Apparently, once pitch differences exceed the threshold of discrimination, magnitude estimations of the size of such changes are roughly equivalent for amusic and non-amusic individuals.

Furthermore, a tendency towards mapping pitch onto vertical space by amusic and non-amusic participants was also found, when the pitch-space association cannot possibly be mediated by language in the colour classification task. This finding suggests that the association can emerge regardless of linguistic labels. More interestingly, the group difference in response times was no longer significant in the colour classification task, in which explicit processing of pitch was not involved. One explanation is that amusic participants have low confidence in pitch-related tasks, thereby adopting a cautious attitude and taking more time to respond. The low confidence may result from unstable representations of pitch as a consequence of a failure to access pitch information explicitly.

Pitch contour processing

The impairments in contour processing shown by amusic individuals may also result from a failure to access pitch information consciously. Using the SACT, Chapter 5 revealed that amusic participants showed reduced sensitivity to the direction of a change in pitch at the perceptual stage of contour processing, even though they were able to represent pitch movement implicitly to some degree. Furthermore, this perceptual impairment of congenital amusia can neither be explained by the elevated threshold for pitch discrimination, nor by the smaller magnitude estimation of pitch distance, as most intervals used in this study (i.e., 1-12 semitones) were well above amusics' pitch threshold and perceived as equally large by amusic and non-amusic individuals (see Chapter 4). More importantly, it suggests

that the impairment in contour processing is restricted to the pitch dimension of sound, and has not extended to other auditory dimensions such as loudness or timbral brightness. This finding is consistent with previous studies showing that the processing of spectral information and intensity within individual tones is intact in the amusic auditory system (Cousineau, Oxenham, & Peretz, 2013; 2015).

Evidence from Chapters 3 and 5 revealed that the pitch representations of amusic individuals are less stable than that of non-amusic participants. One outcome of unstable representations is that pitch information is prone to rapid decay in short-term memory, thereby causing impairments at higher levels of pitch memory. Indeed, previous studies have shown that amusic individuals have a smaller short-term memory span for pitch that exceeded their pitch perception thresholds (Williamson & Stewart, 2010), and weaker pitch traces (Gosselin, Jolicoeur, & Peretz, 2009; Williamson, McDonald, Deutsch, Griffiths, & Stewart, 2010; see Tillmann, Leveque, Fornoni, Albouy, & Caclin, 2016, for a recent review). In addition, when the pitch distance between two auditory stimuli becomes larger, the performance of pitch memory improves (Albouy et al., 2013; Jiang, Lim, Wang, & Hamm, 2013; Tillmann, Schulze, & Foxton, 2009). Nonetheless, our findings suggest that short-term memory problems exist even for pitch changes that exceed the threshold of discrimination, resulting from unstable representations of pitch regardless of interval size.

Taken together, the impaired contour processing in congenital amusia starts from the perceptual stage, where amusic individuals have some implicit knowledge of pitch information, but fail to bring it to conscious awareness. The limited conscious access to pitch information leads to unstable representations of pitch, which can readily decay and be interfered with.

The influences of visuospatial information

An important prediction in this thesis is that amusic individuals might compensate for their unstable pitch representations by relying on other sources of information when processing pitch. In Chapter 6, I found that amusic individuals relied more heavily than non-amusic individuals on unattended visuospatial information when judging change direction of pitch, even when they could perform as well as non-amusics based on the auditory information alone. Although the ceiling

effect might explain the non-significant group differences between amusic and non-amusic participants in the auditory-only (e.g., large interval change condition) and visual-only tasks, it cannot account for the significant group differences between the two groups observed in the AV tasks. In fact, auditory information is not as reliable as visual information for amusic individuals in their daily experience, so it is reasonable to expect that they draw upon visual information to compensate for their deficits in pitch perception. Combined with previous studies showing that visual information boosts auditory perception for amusics (Albouy, et al., 2015a; Thompson, Marin, & Stewart, 2012), data shown in Chapter 6 provide supporting evidence for the "optimal-integration hypothesis" (Ernst & Banks, 2002; Ernst & Bulthoff, 2004), which suggests that people focus on a certain modality based on which modality provides optimal information.

From a practical view, Chapter 5 provides a theoretical basis for using multimedia devices that include visuospatial information to improve amusics' auditory perception in everyday life. Visuospatial information has potential to influence auditory processing in amusic individuals because it provides a pathway to access spatial representations of pitch, as pitch and space are associated with each other in daily life. To test this possibility, a further study is needed to ensure that visual influences on auditory processing in amusic individuals are negligible when the visual information is non-spatial. Another direction for further research is to examine whether using more realistic visuospatial information, such as head movements or hand gestures, would yield to similar results.

The mechanisms of congenital amusia

What distinguishes amusic individuals from others? As summarised in Chapter 2, considerable progress in understanding the nature of congenital amusia has been made in the past dozen years. Nonetheless, the precise pattern of impairments that give rise to congenital amusia has yet to be fully understood. Data presented in this thesis are in support of the hypothesis that congenital amusia is associated with a lack of "pitch awareness" (e.g., Hyde, Zatorre, & Peretz, 2011; Loui, Alsop, & Schlaug, 2009; Loui, Guenther, Mathys, & Schlaug, 2008; Loui, Kroog, Zuk, Winner, & Schlaug, 2011; Lu, Ho, Liu, Wu, & Thompson, 2015; Moreau, Jolicoeur, & Peretz,

2013; Peretz, Brattico, Jarvenpaa, & Tervaniemi, 2009; Zendel, Lagrois, Robitaille, & Peretz, 2015).

The "pitch awareness hypothesis" emphasises that pitch information can be registered but not accessed consciously by individuals who suffer from congenital amusia. Under this hypothesis, four testable predictions can be made. First, the essential neural circuit to register pitch for congenital amusics should be intact. Indeed, a number of neurological investigations have revealed that mild or no dysfunction was found in amusics at the stage of brainstem encoding of pitch (Liu, Maggu, Lau, & Wong, 2015; but see Lehmann, Skoe, Moreau, Peretz, & Kraus, 2015), nor at early auditory processing stages when tasks were relatively easy, as revealed by a normal auditory N1 component (Moreau et al., 2013; 2009; Peretz et al., 2009; Peretz, Brattico, & Tervaniemi, 2005). In line with those findings, I also observed a normal auditory N1 in amusic participants in Chapter 6, reflecting an intact early stage of auditory processing. Second, congenital amusia results from limited access to pitch information consciously. Therefore, amusic individuals should be able to access pitch information implicitly, given that their neural circuit is intact. This prediction is supported by both behavioural and neurological studies. For instance, Hutchins, Gosselin and Peretz (2010) found that amusic individuals have mild or no impairment in production or imitation of pitches, where the conscious processing of pitch information is not required. In Chapter 5, I also found that amusic individuals showed implicit knowledge of pitch change direction without being aware of it. The limited access to pitch information might be associated with reduced functional connectivity between the inferior frontal gyrus and the auditory cortex (Albouy et al., 2013; Albouy, Mattout, Sanchez, Tillmann, & Caclin, 2015b; Hyde et al., 2011), and reduced volume of the right arcuate fasciculus (Loui & Schlaug, 2009; but see Chen et al., 2015). Third, the limited awareness should be restricted to the pitch dimension. In support of this prediction, no group difference was found in this thesis beyond the tasks involving pitch perception explicitly, in line with previous studies showing that amusic individuals have no difficulty in temporal, spectral or intensity perception (Cousineau, Oxenham, & Peretz, 2013; 2015). Finally, amusic individuals should show low confidence levels in tasks that require explicit pitch perception owing to a lack of pitch awareness. Data from Chapter 5 provide direct evidence that amusic individuals exhibit a lower confidence level in pitch contour

processing, but have a comparable confidence level to non-amusic individuals in loudness and timbral brightness contour processing. This low confidence in pitch perception may account for the slower response times in the pitch classification task described in Chapter 3, and the tendency to use visual information in auditory tasks as discussed in Chapter 6.
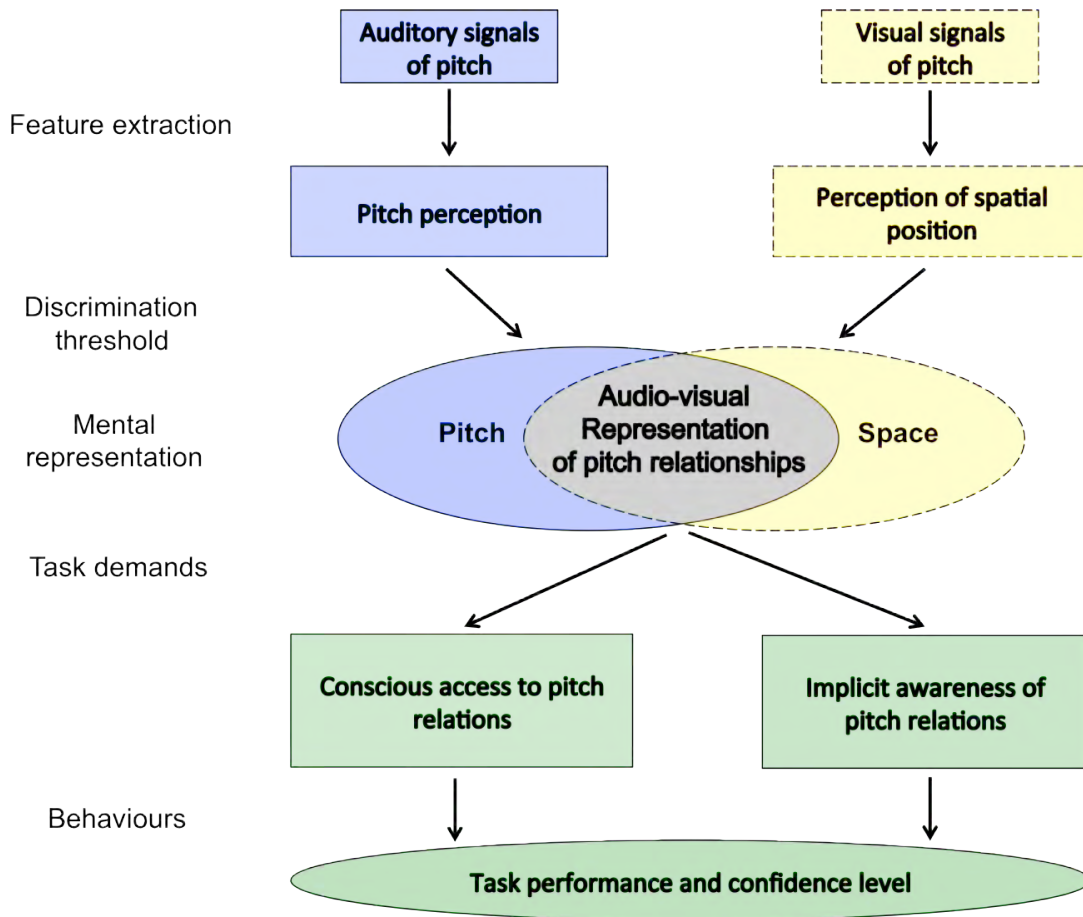
Fig. 1 – A diagram illustrating pitch processing and potential influence of the spatial position of visual information when available in congenital amusia.

The data from this thesis corroborate the "pitch awareness hypothesis". Combined with the findings from previous studies, my findings suggest that congenital amusia is more than a low-level auditory perceptual problem, given that task performance is still impaired when pitch intervals are above the threshold of discrimination. As illustrated in Fig.1, pitch information is encoded by extracting the frequency feature from the auditory stimulation, and spatial position of the visual

stimulation is encoded when it is available. Given that pitch and space are associated, they interact with each other and share a common representation in a way that depends on factors such as clarity of auditory and visual signals, thresholds for encoding differences in pitch and spatial location, long term sensory biases learned through a chronic disability in one or more sensory modalities, and task demands. Sequentially, pitch relations can be either explicitly or implicitly accessed that results in different behavioural performance and confidence level. With regards to congenital amusia, it is associated with limited access to pitch information within consciousness, thereby affecting amusic individuals' behaviours in pitch-related tasks. However, those people may have some implicit knowledge of pitch information that results in normal behaviours in tasks that not involve explicit processing of pitch. On the other hand, a long-term bias in weighing visual over auditory information due to the impairments in auditory modality, leads amusics to use visual information as additional cues to clarify or reinforce the pitch representation. That is, the audio-visual representation of pitch relations benefits amusic' auditory processing.

Concluding remarks

This thesis contributes to existing research on congenital amusia by providing behavioural and neurological evidence showing that individuals with congenital amusia have unstable or distorted spatial representations of pitch, which may lead to a failure in identifying pitch change direction at the perceptual stage of processing. These findings further support the "pitch awareness hypothesis", which emphasises congenital amusia is associated with a failure to access pitch information consciously. Drawing upon data from individuals with congenital amusia, it also sheds light on normal mechanisms underlying music perception and cognition, especially the importance of spatial representation during musical pitch processing.

Furthermore, this thesis contributes to the current understanding of audiovisual integration in support of the "optimal-integration hypothesis", which suggests that individuals tend to rely on one modality over the other depending on how reliable the information is. In this thesis, amusic individuals tend to make use of available visual cues in order to compensate for their difficulties in pitch perception on the basis of their daily experience, even when pitch differences are well above their threshold.

This finding, therefore, extends the "optimal-integration hypothesis" to circumstances where there is a long-term learned bias in weighing visual over auditory information, which generlises from a particular source of reliability in the sensory modality to the entire modality.

References

Albouy, P., Leveque, Y., Hyde, K. L., Bouchet, P., Tillmann, B., & Caclin, A. (2015a). Boosting pitch encoding with audiovisual interactions in congenital amusia. *Neuropsychologia, 67*, 111-120. doi: 10.1016/j.neuropsychologia.2014.12.006

Albouy, P., Mattout, J., Bouet, R., Maby, E., Sanchez, G., Aguera, P. E., . . . Tillmann, B. (2013). Impaired pitch perception and memory in congenital amusia: the deficit starts in the auditory cortex. *Brain, 136*(5), 1639-1661. doi: 10.1093/brain/awt082

Albouy, P., Mattout, J., Sanchez, G., Tillmann, B., & Caclin, A. (2015b). Altered retrieval of melodic information in congenital amusia: insights from dynamic causal modeling of MEG data. *Frontiers in Human Neuroscience, 9*, 20. doi: 10.3389/fnhum.2015.00020

Chen, J. L., Kumar, S., Williamson, V. J., Scholz, J., Griffiths, T. D., & Stewart, L. (2015). Detection of the arcuate fasciculus in congenital amusia depends on the tractography algorithm. *Frontiers in Psychology, 6*, 9. doi: 10.3389/fpsyg.2015.00009

Connell, L., Cai, Z. G., & Holler, J. (2013). Do you see what I'm singing? Visuospatial movement biases pitch perception. *Brain and Cognition, 81*(1), 124-130. doi: 10.1016/j.bandc.2012.09.005

Cousineau, M., Oxenham, A. J., & Peretz, I. (2013). The role of peripheral spectro-temporal coding in congenital amusia. *The Jouranl of the Acoustical Society of America, 133*(5), 3381. doi: 10.1121/1.4805825

Cousineau, M., Oxenham, A. J., & Peretz, I. (2015). Congenital amusia: A cognitive disorder limited to resolved harmonics and with no peripheral basis. *Neuropsychologia, 66*, 293-301. doi: 10.1016/j.neuropsychologia.2014.11.031i

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. doi: 10.1177/0956797612457374

Douglas, K. M., & Bilkey, D. K. (2007). Amusia is associated with deficits in spatial processing. *Nature Neuroscience, 10*(7), 915-921. doi: 10.1038/nn1925

Eitan, Z. (2013). Musical objects, cross-domain correspondences, and cultural choice: commentary on "Cross-cultural representations of musical shape" by

George Athanasopoulos and Nikki Moran. *Empirical Musicology Review, 8*(3-4), 204–207.

Elkin, J., & Leuthold, H. (2011). The representation of pitch in auditory imagery: Evidence from S-R compatibility and distance effects. *Journal of Cognitive Psychology, 23*(1), 76-91. doi: 10.1080/20445911.2011.455251

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*(24), 429-433. doi:10.1038/415429a

Ernst, M. O., & Bulthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences, 8*(4), 162-169. doi: 10.1016/j.tics.2004.02.002

Gosselin, N., Jolicoeur, P., & Peretz, I. (2009). Impaired memory for pitch in congenital amusia. *Annals of the New York Academy of Sciences, 1169*, 270-272. doi: 10.1111/j.1749-6632.2009.04762.x

Hutchins, S., Gosselin, N., & Peretz, I. (2010). Identification of changes along a continuum of speech intonation is impaired in congenital amusia. *Frontiers in Psychology, 1*, 236. doi: 10.3389/fpsyg.2010.00236

Hyde, K. L., Zatorre, R. J., & Peretz, I. (2011). Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. *Cerebral Cortex, 21*(2), 292-299. doi: 10.1093/cercor/bhq094

Jiang, C., Lim, V. K., Wang, H., & Hamm, J. P. (2013). Difficulties with pitch discrimination influences pitch memory performance: Evidence from congenital amusia. *PLoS One, 8*(10), e79216. doi: 10.1371/journal.pone.0079216.t001

Küssner, M. B., Tidhar, D., Prior, H. M., & Leech-Wilkinson, D. (2014). Musicians are more consistent: Gestural cross-modal mappings of pitch, loudness and tempo in real-time. *Frontiers in Psychology, 5*, 789. doi: 10.3389/fpsyg.2014.00789

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. doi: 10.3389/fnhum.2014.00250

Lehmann, A., Skoe, E., Moreau, P., Peretz, I., & Kraus, N. (2015). Impairments in musical abilities reflected in the auditory brainstem: Evidence from congenital amusia. *The European Journal of Neuroscience, 42*(1), 1644-1650. doi: 10.1111/ejn.12931

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), 1189-1207. doi: 10.1037/0096-1523.33.5.1189

Liu, F., Maggu, A. R., Lau, J. C. Y., & Wong, P. C. M. (2015). Brainstem encoding of speech and musical stimuli in congenital amusia: evidence from Cantonese speakers. *Frontiers in Human Neuroscience, 8*, 1029. doi: 10.3389/fnhum.2014.01029

Loui, P., Alsop, D., & Schlaug, G. (2009). Tone deafness: a new disconnection syndrome? *Journal of Neuroscience, 29*(33), 10215-10220. doi: 10.1523/jneurosci.1701-09.2009

Loui, P., Guenther, F. H., Mathys, C., & Schlaug, G. (2008). Action-perception mismatch in tone-deafness. *Current Biology, 18*(8), R331-332. doi: 10.1016/j.cub.2008.02.045

Loui, P., Kroog, K., Zuk, J., Winner, E., & Schlaug, G. (2011). Relating pitch awareness to phonemic awareness in children: implications for tone-deafness and dyslexia. *Frontiers in Psychology, 2*, 111. doi: 10.3389/fpsyg.2011.00111

Loui, P., & Schlaug, G. (2009). Investigating musical disorders with diffusion tensor imaging. *Annals of the New York Academy of Sciences, 1169*, 121-125. doi: 10.1111/j.1749-6632.2009.04781.x

Lu, X., Ho, H. T., Liu, F., Wu, D., & Thompson, W. F. (2015). Intonation processing deficits of emotional words among Mandarin Chinese speakers with congenital amusia: an ERP study. *Frontiers in Psychology, 6*, 385. doi: 10.3389/fpsyg.2015.00385

Miller, J. (1986). Timecourse of coactivation in bimodal divided attention. *Perception & Psychophysics, 40*(5), 331-343.

Moreau, P., Jolicoeur, P., & Peretz, I. (2013). Pitch discrimination without awareness in congenital amusia: Evidence from event-related potentials. *Brain and Cognition, 81*(3), 337-344. doi: 10.1016/j.bandc.2013.01.004

Moreau, P., Jolicœur, P., & Peretz, I. (2009). Automatic brain responses to pitch changes in congenital amusia. *Annals of the New York Academy of Sciences, 1169*, 191-194. doi: 10.1111/j.1749-6632.2009.04775.x

Nieuwenhuis, S., Yeung, N., van den Wildenberg, W., Ridderinkhof, K. R., 2003. Electrophysiological correlates of anterior cingulate function in a go/no-go task: effects of response conflict and trial type frequency. *Cognitive, Affective, & Behavioral Neuroscience, 3*(1), 17-26. doi: 10.3758/CABN.3.1.17

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*(16), 6104-6108. doi: 10.1073/pnas.1322705111

Peretz, I., Brattico, E., Jarvenpaa, M., & Tervaniemi, M. (2009). The amusic brain: in tune, out of key, and unaware. *Brain, 132*(5), 1277-1286. doi: 10.1093/brain/awp055

Peretz, I., Brattico, E., & Tervaniemi, M. (2005). Abnormal electrical brain responses to pitch in congenital amusia. *Annals of Neurology, 58*(3), 478-482. doi: 10.1002/ana.20606

Pratt, C. C. (1930). The spatial character of high and low tones. *Journal of Experimental Psychology, 13*(3), 278-285. doi: 10.1037/h0072651

Proctor, R. W., & Vu, K.-P. L. (2002). Eliminating, magnifying, and reversing spatial compatibility effects with mixed location-relevant and irrelevant trials. In W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action: Attention and performance* (pp. 443-473). Oxford, UK: Oxford University Press.

Rusconi, E., Kwan, B., Giordano, B., Umilta, C., & Butterworth, B. (2006). Spatial representation of pitch height: the SMARC effect. *Cognition, 99*(2), 113-129. doi: 10.1016/j.cognition.2005.01.004

Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Attention, Perception & Psychophysics., 73*(4), 971-995. doi: 10.3758/s13414-010-0073-7

Stewart, L., Verdonschot, R. G., Nasralla, P., & Lanipekun, J. (2012). Action-perception coupling in pianists: Learned mappings or spatial musical association of response codes (SMARC) effect? *The Quarterly Journal of Experimental Psychology 66:37-50.* doi: 10.1080/17470218.2012.687385.

Taylor, J. E., & Witt, J. K. (2014). Listening to music primes space: pianists, but not novices, simulate heard actions. *Psychoogical Research 79:175-82.* doi: 10.1007/s00426-014-0544-x.

Thompson, W. F., Marin, M. M., & Stewart, L. (2012). Reduced sensitivity to emotional prosody in congenital amusia rekindles the musical protolanguage hypothesis. *Proceedings of the National Academy of Sciences of the United States of America, 109*(46), 19027-19032. doi: 10.1073/pnas.1210344109

Tillmann, B., Leveque, Y., Fornoni, L., Albouy, P., & Caclin, A. (2016). Impaired short-term memory for pitch in congenital amusia. *Brain Research, 1640*, 251-263.. doi: 10.1016/j.brainres.2015.10.035

Tillmann, B., Schulze, K., & Foxton, J. (2009). Congenital amusia: A short-term memory deficit for non-verbal, but not verbal sounds. *Brain and Cognition, 71*(3), 259-264. doi: 10.1016/j.bandc.2009.08.003

Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science, 21*(1), 21-25. doi: 10.1177/0956797609354734

Williamson, V. J., McDonald, C., Deutsch, D., Griffiths, T. D., & Stewart, L. (2010). Faster decline of pitch memory over time in congenital amusia. *Advances in Cognitive Psychology, 6*(1), 15-22. doi: 10.2478/v10053-008-0073-5

Williamson, V. J., & Stewart, L. (2010). Memory for pitch in congenital amusia: Beyond a fine-grained pitch discrimination problem. *Memory, 18*(6), 657-669. doi: 10.1080/09658211.2010.501339

Yeung, N., Botvinick, M. M., Cohen, J. D., 2004. The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychological Review*, *111*(4), 931-959. doi: 10.1037/0033-295X.111.4.931

Zendel, B. R., Lagrois, M. E., Robitaille, N., & Peretz, I. (2015). Attending to pitch information inhibits processing of pitch information: the curious case of amusia. *The Journal of Neuroscience, 35*(9), 3815-3824. doi: 10.1523/JNEUROSCI.3766-14.2015

Appendix A: Supplementary Materials

Chapter 5: Influences of visuospatial information on pitch processing

Supplementary Table 1. Summary of repeated-measures ANOVAs on task performance for AO task and AV implicit task evaluated by percent correct (%), and for AV explicit task evaluated by d-prime (d'). Effect size was estimated using partial eta-squared (η_p^2), and DF refers to the degrees of freedom. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

Supplementary Table 1.1 – The behavioural results (percent correct) for the AO task.

Effect	DFn	DFd	F	p	η_p^2
Group	1	30	3.82	0.06	0.11
Interval	1	30	90.64 ***	< 0.001	0.75
Interval × Group	1	30	14.78 **	0.001	0.33

Supplementary Table 1.2 – The behavioural results (percent correct) for the AV implicit task.

Effect	DFn	DFd	F	p	η_p^2
Group	1	30	10.47 **	0.003	0.26
Interval	1	30	132.17 ***	< 0.001	0.82
Congruence	1	30	57.86 ***	< 0.001	0.66
Interval × Group	1	30	16.73 **	0.001	0.36
Congruence × Group	1	30	10.24 **	< 0.01	0.25
Interval × Congruence	1	30	23.09 ***	< 0.001	0.44
Interval × Congruence × Group	1	30	3.49	0.07	0.10

Supplementary Table 1.3 – The behavioural results (d') for the AV explicit task.

Effect	DFn	DFd	F	p	η_p^2
Group	1	30	5.07 *	0.03	0.15
Interval	1	30	70.28 ***	< 0.001	0.70
Interval × Group	1	30	4.61 *	0.04	0.13

Supplementary Table 2. Summary of paired-sample t-tests on percent correct between AO and AV implicit task for each group. To adjust the multiple comparisons, the mean difference is significant at a significance level of 0.0125. DF refers to the degrees of freedom. *: $p < 0.0125$.

Group	Interval	Pairs	t	DF	p
Amusic	Small	AV congruent – AO	3.05 *	15	0.008
		AO – AV incongruent	3.25 *	15	0.005
	Large	AV congruent – AO	1.71	15	0.108
		AO – AV incongruent	3.01 *	15	0.009
Control	Small	AV congruent – AO	3.22 *	15	0.006
		AO – AV incongruent	1.56	15	0.140
	Large	AV congruent – AO	2.48	15	0.025
		AO – AV incongruent	-0.18	15	0.858

Supplementary Table 3. Summary of repeated-measures ANOVAs on the mean amplitudes (across trials with correct response for each condition) computed across the region of interest (CP3, CPZ, CP4, C3, CZ, and C4) within each selected time windows for the AO, AV implicit and AV explicit tasks. Effect size was estimated using partial eta-squared (η_p^2), and DF refers to the degrees of freedom. *: p < 0.05; **: p < 0.01; ***: p < 0.001.

Supplementary Table 3.1 – The ERP results (based on trials with correct response) for the AO task.

Time window	Effect	DF*n*	DF*d*	F	p	η_p^2
90-150 msec	Group	1	30	0.33	0.57	0.01
	Interval Size	1	30	8.14 **	< 0.01	0.21
	Interval Size × Group	1	30	< 0.001	0.99	< 0.001
200-380 msec	Group	1	30	0.52	0.48	0.02
	Interval Size	1	30	1.36	0.25	0.04
	Interval Size × Group	1	30	0.01	0.92	< 0.01

Supplementary Table 3.2 – The ERP results (based on trials with correct response) for the AV implicit task.

Time window	Effect	DF*n*	DF*d*	F	p	η_p^2
100-180 msec	Group	1	30	0.01	0.94	< 0.001
	Interval Size	1	30	13.91 **	0.001	0.32
	Congruence	1	30	0.01	0.94	< 0.001
	Interval Size × Group	1	30	0.34	0.56	0.01
	Congruence × Group	1	30	0.23	0.64	0.01
	Interval Size × Congruence	1	30	2.03	0.16	0.06
	Interval Size × Congruence × Group	1	30	0.23	0.64	0.01
260-380 msec	Group	1	30	1.70	0.20	0.05
	Interval Size	1	30	48.36 ***	< 0.001	0.62
	Congruence	1	30	7.43 *	0.01	0.20
	Interval Size ×Group	1	30	1.11	0.30	0.04
	Congruence × Group	1	30	4.75 *	0.04	0.14
	Interval Size × Congruence	1	30	2.67	0.11	0.08
	Interval Size × Congruence × Group	1	30	0.01	0.92	< 0.001

Supplementary Table 3.3 – The results (based on trials with correct response) for the AV explicit task**.**

Time window	Effect	DFn	DFd	F	p	η_p^2
100-180 msec	Group	1	30	0.06	0.81	< 0.01
	Interval Size	1	30	11.19 **	< 0.01	0.27
	Congruence	1	30	3.23	0.08	0.10
	Interval Size × Group	1	30	< 0.001	1.00	< 0.001
	Congruence × Group	1	30	5.84 *	0.02	0.16
	Interval Size × Congruence	1	30	< 0.01	0.98	< 0.001
	Interval Size × Congruence × Group	1	30	0.19	0.67	< 0.01
260-380 msec	Group	1	30	1.03	0.32	0.03
	Interval Size	1	30	30.40 ***	< 0.001	0.50
	Congruence	1	30	29.08 ***	< 0.001	0.49
	Interval Size × Group	1	30	0.15	0.70	< 0.01
	Congruence × Group	1	30	0.01	0.93	< 0.001
	Interval Size × Congruence	1	30	0.43	0.52	0.01
	Interval Size × Congruence × Group	1	30	3.32	0.08	0.10

Supplementary Table 4. Summary of independent-sample t-tests on behavioural (percent correct) and ERP data (the mean amplitudes computed across the region of interest: CP3, CPZ, CP4, P3, PZ, and P4) for the VO task between amusic and control groups. DF refers to the degrees of freedom.

	DFn	**DFd**	**t**	**p**
Behavioural	1	30	1.75	0.10
ERP (120-180 msec)	1	30	0.84	0.41
ERP (200-380 msec)	1	30	1.84	0.08

Supplementary Table 5. Summary of repeated-measures ANOVAs on the mean amplitudes (across all experimental trials for each condition) computed across the region of interest (CP3, CPZ, CP4, C3, CZ, and C4) within each selected time windows for the AO, AV implicit and AV explicit tasks. Effect size was estimated using partial eta-squared (η_p^2), and DF refers to the degrees of freedom. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

Supplementary Table 5.1 – The ERP results (based on all experimental trials) for the AO task.

Time window	Effect	DFn	DFd	F	p	η_p^2
90-150 msec	Group	1	30	0.81	0.37	0.03
	Interval Size	1	30	11.30 **	< 0.01	0.27
	Interval Size × Group	1	30	< 0.01	0.97	< 0.001
200-380 msec	Group	1	30	0.20	0.66	0.01
	Interval Size	1	30	3.79	0.06	0.11
	Interval Size × Group	1	30	0.12	0.73	< 0.01

Supplementary Table 5.2 – The ERP results (based on all experimental trials) for the AV implicit task.

Time window	Effect	DFn	DFd	F	p	η_p^2
100-180 msec	Group	1	30	< 0.001	0.99	< 0.001
	Interval Size	1	30	25.40 ***	< 0.001	0.44
	Congruence	1	30	0.46	0.50	0.02
	Interval Size × Group	1	30	0.26	0.61	0.01
	Congruence × Group	1	30	1.18	0.29	0.04
	Interval Size × Congruence	1	30	3.13	0.09	0.10
	Interval Size × Congruence × Group	1	30	0.73	0.40	0.02
260-380 msec	Group	1	30	2.39	0.13	0.07
	Interval Size	1	30	51.54 ***	< 0.001	0.63
	Congruence	1	30	12.83 **	0.001	0.30
	Interval Size ×Group	1	30	0.01	0.94	< 0.001
	Congruence × Group	1	30	5.75 *	0.02	0.16
	Interval Size × Congruence	1	30	2.90	0.10	0.09
	Interval Size × Congruence × Group	1	30	0.01	0.93	< 0.001

Supplementary Table 5.3 – The results (based on all experimental trials) for the AV explicit task.

Time window	Effect	DFn	DFd	F	p	η_p^2
100-180 msec	Group	1	30	< 0.001	0.99	< 0.001
	Interval Size	1	30	16.16 ***	< 0.001	0.35
	Congruence	1	30	3.41	0.08	0.10
	Interval Size × Group	1	30	0.04	0.85	0.001
	Congruence × Group	1	30	2.14	0.15	0.07
	Interval Size × Congruence	1	30	0.06	0.81	0.002
	Interval Size × Congruence × Group	1	30	0.04	0.85	0.001
260-380 msec	Group	1	30	0.65	0.43	0.02
	Interval Size	1	30	26.05 ***	< 0.001	0.47
	Congruence	1	30	41.57 ***	< 0.001	0.58
	Interval Size × Group	1	30	0.02	0.88	0.001
	Congruence × Group	1	30	0.66	0.42	0.02
	Interval Size × Congruence	1	30	0.71	0.41	0.02
	Interval Size × Congruence × Group	1	30	1.55	0.22	0.05

Supplementary Table 6. Summary of repeated-measures ANOVAs on task performance for AO task and AV implicit task evaluated by d' (considering trials with "went up" tones as signals, and those with "went down" tones as noises). Effect size was estimated using partial eta-squared (η_p^2), and DF refers to the degrees of freedom. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

Supplementary Table 6.1 – The behavioural results (d') for the AO task.

Effect	DFn	DFd	F	p	η_p^2
Group	1	30	2.79	0.11	0.09
Interval	1	30	117.70 ***	< 0.001	0.80
Interval × Group	1	30	13.06 **	0.001	0.30

Supplementary Table 6.2 – The behavioural results (d') for the AV implicit task.

Effect	DFn	DFd	F	p	η_p^2
Group	1	30	9.28 **	0.005	0.24
Interval	1	30	159.00 ***	< 0.001	0.84
Congruence	1	30	74.93 ***	< 0.001	0.71
Interval × Group	1	30	12.25 ***	0.001	0.29
Congruence × Group	1	30	10.29 **	0.003	0.26
Interval × Congruence	1	30	8.82 **	0.006	0.23
Interval × Congruence × Group	1	30	0.01	0.95	< 0.001

Appendix B: Ethics Approval

MACQUARIE UNIVERSITY

Office of the Deputy Vice-Chancellor (Research)

Research Office
C5C Research HUB East, Level 3, Room 324
MACQUARIE UNIVERSITY NSW 2109 AUSTRALIA

Phone   **+61 (0)2 9850 7850**
Fax   +61 (0)2 9850 4465
Email   ethics.secretariat@mq.edu.au

19 November 2014

Professor Bill Thompson
Department of Psychology
Faculty of Human Sciences
Macquarie University
NSW 2109

Dear Professor Thompson

Reference No: 5201200890

Title: Examination of disorders of music and language

Thank you for your correspondence dated 31 October 2014 submitting an amendment request to the above study. Your proposed amendment was reviewed and approved by the HREC (Medical Sciences) Executive at its meeting held on 17 November 2014.

I am pleased to advise that ethical approval of the following amendments to the above study has been granted:

1. The addition of a new participant information statement and consent form that provides necessary information to participants participating in EEG experiments.

2. Modifications to the current information in the participant information statement and consent form for the behavioural study, as some tasks are no longer conducted.

3. The removal of Dr Kirstin Robertson-Gillam and Dr Lena Quinto from the personnel involved in the project.

Details of this approval are as follows:

The following documentation submitted with your email correspondence has been reviewed and approved by the HREC (Medical Sciences):

Documents reviewed	Version no.	Date
Macquarie University HREC Request for Amendment Form	2.0	Received 31/10/2014
Participant Information Statement & Consent Form: EEG study		
Participant Information Statement & Consent Form: Behavioural study		

The HREC (Medical Sciences) Terms of Reference and Standard Operating Procedures are available from the Research Office website at:

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics

Please do not hesitate to contact the Ethics Secretariat should you have any questions regarding your ethics application.

The HREC (Medical Sciences) wishes you every success in your research.

Yours sincerely

Professor Tony Eyers
Chair, Macquarie University Human Research Ethics Committee (Medical Sciences)

This HREC is constituted and operates in accordance with the National Health and Medical Research Council's (NHMRC) *National Statement on Ethical Conduct in Human Research* (2007) and the *CPMP/ICH Note for Guidance on Good Clinical Practice.*