Behavioral and neurophysiological correlates of auditory perception and memory : evidence from congenital amusia

Philippe Albouy

▶ **To cite this version:**

Philippe Albouy. Behavioral and neurophysiological correlates of auditory perception and memory : evidence from congenital amusia. Neurons and Cognition [q-bio.NC]. Université Claude Bernard - Lyon I, 2013. English. ⟨NNT : 2013LYO10316⟩. ⟨tel-01148531⟩

HAL Id: tel-01148531
https://theses.hal.science/tel-01148531

Submitted on 4 May 2015

N° d'ordre   316-2013   Année 2013

THESE DE L'UNIVERSITE DE LYON

Délivrée par

L'UNIVERSITE CLAUDE BERNARD LYON 1

ECOLE DOCTORALE NEUROSCIENCES ET COGNITION

DIPLOME DE DOCTORAT

(arrêté du 7 août 2006)

Mention : Neurosciences

soutenue publiquement le 19 décembre 2013

par

Philippe ALBOUY

Behavioral and neurophysiological correlates of auditory perception and memory: evidence from congenital amusia

Directeur de thèse : Dr Barbara Tillmann ; co-encadrement : Dr Anne Caclin

JURY : Mme Lauren Stewart (Rapporteur)

Mme Nathalie George (Rapporteur)

Mme Mari Tervaniemi (Examinateur)

Mme Barbara Tillmann (Directeur de thèse)

Mme Anne Caclin (Co-encadrant de thèse)

M. Fabien Perrin (Examinateur)

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université   **M. François-Noël GILLY**

Vice-président du Conseil d'Administration   M. le Professeur Hamda BEN HADID

Vice-président du Conseil des Etudes et de la Vie Universitaire   M. le Professeur Philippe LALLE

Vice-président du Conseil Scientifique   M. le Professeur Germain GILLET

Directeur Général des Services   M. Alain HELLEU

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard   Directeur : M. le Professeur J. ETIENNE

Faculté de Médecine et de Maïeutique Lyon Sud – Charles   Directeur : Mme la Professeure C. BURILLON
Mérieux

Faculté d'Odontologie   Directeur : M. le Professeur D. BOURGEOIS

Institut des Sciences Pharmaceutiques et Biologiques   Directeur : Mme la Professeure C. VINCIGUERRA

Institut des Sciences et Techniques de la Réadaptation   Directeur : M. le Professeur Y. MATILLON

Département de formation et Centre de Recherche en Biologie   Directeur : M. le Professeur P. FARGE
Humaine

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies   Directeur : M. le Professeur F. DE MARCHI

Département Biologie   Directeur : M. le Professeur F. FLEURY

Département Chimie Biochimie   Directeur : Mme le Professeur H. PARROT

Département GEP   Directeur : M. N. SIAUVE

Département Informatique   Directeur : M. le Professeur S. AKKOUCHE

Département Mathématiques   Directeur : M. le Professeur A. GOLDMAN

Département Mécanique   Directeur : M. le Professeur H. BEN HADID

Département Physique   Directeur : Mme S. FLECK

Département Sciences de la Terre   Directeur : Mme la Professeure I. DANIEL

UFR Sciences et Techniques des Activités Physiques et Sportives   Directeur : M. C. COLLIGNON

Observatoire des Sciences de l'Univers de Lyon   Directeur : M. B. GUIDERDONI

Polytech Lyon   Directeur : M. P. FOURNIER

Ecole Supérieure de Chimie Physique Electronique   Directeur : M. G. PIGNAULT

Institut Universitaire de Technologie de Lyon 1   Directeur : M. C. VITON

Institut Universitaire de Formation des Maîtres   Directeur : M. A. MOUGNIOTTE

Institut de Science Financière et d'Assurances   Administrateur provisoire : M. N. LEBOISNE

A Papa et Maman,

A Jeubine, Mimi, Ninou et Fifi,

A Ginette,

Remerciements

Mes premiers remerciements vont à mes directrices de thèse, Barbara Tillmann et Anne Caclin, qui m'accompagnent avec toujours autant d'enthousiasme depuis maintenant près de quatre ans. Je voudrais vous témoigner toute mon admiration et vous adresser ma profonde reconnaissance pour toutes les choses que vous m'avez apprises, pour votre patience, votre curiosité contagieuse, votre disponibilité, votre immense soutien et votre ouverture d'esprit. Merci aussi d'avoir facilité mes allers-retours montréalais avec autant de simplicité. Je pense que je ne le dirai jamais assez, merci Anne et Barbara!

Je tiens ensuite à remercier Isabelle Peretz, Robert Zatorre et Patrick Bermudez, qui m'ont encadré et chaleureusement accueilli au cours des différents séjours au BRAMS. Merci pour votre disponibilité et vos conseils avisés !

Je remercie sincèrement Mesdames Lauren Stewart et Nathalie George pour avoir accepté d'être rapporteurs de ce travail de thèse, ainsi que Mme Mari Tervaniemi et M. Fabien Perrin pour leur participation au jury.

Je remercie chaleureusement tous les membres de l'équipe CAP et de l'équipe DYCOG dans lesquelles j'ai passé tant de bons moments. Dans le désordre, merci à Olivier pour sa bonne humeur et sa gentillesse qui illuminent tout le laboratoire, à Jérémie pour ses conseils et son aide précieuse. Merci à mes sujets pilotes préférés (ils se reconnaîtront) qui ont du endurer de looooongues sessions de « mou-toux-loup » intercalées de « titititititititi ». Merci à Yohana et à Lesly, mes comparses de contrôlsie avec qui il est si agréable de travailler, merci à Floris pour sa joie de vivre, sa gentillesse et pour me laisser gagner de temps en temps au squash. Merci PierreM Superman pour ses pouvoirs transatlantiques. Merci à Mathilde, Jean-Philippe, Marcela et Juan pour leur aide dans la réalisation des protocoles intra. Merci à JB, àlaMargauxPerrin, Mani, Julien, Tatiana, Martine, Aline, Claude, Etienne, Framboise, Christina, Thibaut, Maïté, Alexandra, Patrick, Aymeric, Subliminal, Tomas, Dominique, Perrine, Suzanne, Laurie-Anne, Carolina, Niko, Fabien, Marie-Hélène et Marie-Anne pour leur gentillesse, leur humour et leur sympathie, qui font de ces laboratoires des lieux où il fait si bon vivre, et travailler. Merci à Aurélie et Sylvain pour leur générosité et leur aide, à Karim pour ses TGFs, son Dycog foot/frisbee et pour son aide précieuse dans mes recherches de post doc.

Un immense merci au cousin ManuDocMaby et son compère DocBouet pour leur aide constante et pour leur amitié. Merci à mes grands amis, à ce cher triouple de Gaëthein, Loulou et Matou ainsi qu'à Ludwig van Ludo, qui m'accompagnent si joyeusement depuis quelques années maintenant. Merci Alex, Anne-Lise, Flo, Jo, Caro, Fabfab, pour tous ces chaleureux moments passés ensemble. Je remercie mes plus vieux copains, Duny, Aude, Vivi, Titi, JC, Nobral, Nico, Gu, Doudi, Mine, GuiGui, Ayoun, pour leur amitié sans faille. Merci à AL, Oli, Fabi, Matoche, Pierre-Armand, Kathrin, Marion et Patrick qui ont donné à mes séjours répétés à Montréal ce goût de reviens-y.

Je remercie très affectueusement ma très chère famille. Merci Jeubine, Mimi, Ninou, Fifi, pour avoir toujours pris soin de votre frère et simplement parce que vous êtes trop sensass. Merci à GliGli, Augusto, Fouiny, Piapia, Poupée, Juju, Petit Paul et Croquette qui participent à rendre cette famille exceptionnelle.

Il n'y a enfin pas de mots pour remercier Papa et Maman (BonPa & Mamou). Merci d'avoir créé cet univers familial où l'on se sent si bien, merci pour votre éternel soutien, votre patience et votre douceur. Merci à Papa pour nos soirées orangina/clope(s) et merci à Maman pour avoir pris grand soin de ton fils pendant ces deux mois d'écriture.

Enfin je remercie ma jolie Ginette, probablement la personne la plus drôle et la plus sensass de la terre.

Abstract

The aim of this PhD thesis is to further our understanding about how humans perceive and memorize complex sound structures. We investigated the behavioral and cerebral correlates of auditory perception and memory for isolated tones, musical sequences, and verbal material both in typical listeners and in individuals presenting a lifelong disorder of music perception that has been referred to as congenital amusia. Using behavioral approaches, we demonstrated that amusic individuals' deficits in the pitch dimension are related to impairments both in the encoding of short tones and in the short term retention of pitch information. Using multimodal neuroimaging methods (MRI, MEG, fMRI) we observed anatomical and functional abnormalities in the amusic brain, mostly in the right frontal cortex and in the right auditory cortex. Functional abnormalities were observed at each level of processing in short-term memory tasks, that is for encoding, retention, and retrieval of the melodic information. In contrast, for verbal material, amusic participants recruited similar brain regions as those observed for controls, thus suggesting that separate neural resources support tonal and verbal memory. Based on the conclusions made on these first three studies, we explored two approaches aiming to boost pitch processing abilities in amusia; 1) by investigating whether implicit knowledge of the western tonal musical system could influence their short-term memory abilities, and 2) by exploring whether amusic individuals' altered encoding of short tones could be improved by audio-visual interactions These investigations were encouraging and provide the first steps toward designing tools of rehabilitation in this musical disorder. To conclude, it is worth underlining that these studies also improve our understanding of music processing in general, which is the subject of an increasing research domain that is often making the parallel to language processing.

Keywords: congenital amusia, auditory short-term memory, pitch, verbal memory, implicit knowledge, non-musician, multisensory integration, Magnetoencephalography, functional Magnetic Resonance Imaging

Résumé

Ce travail de thèse vise à améliorer notre compréhension des mécanismes sous-tendant la perception et la mémorisation des structures sonores complexes. Nous avons étudié les corrélats comportementaux et cérébraux de la perception et de la mémoire auditive pour des notes isolées, des séquences musicales et des séquences de mots, chez des auditeurs typiques et dans l'amusie congénitale, un trouble permanent de la perception de la musique. En utilisant des approches comportementales, nous avons démontré que les déficits des individus amusiques dans la dimension de la hauteur sont liés à des déficits à la fois durant l'encodage des sons courts et dans la rétention à court terme de l'information musicale. En utilisant des mesures neurophysiologiques (IRM, MEG, IRMf) nous avons démontré l'existence d'anomalies anatomiques et fonctionnelles dans le cerveau amusique, principalement dans les cortex frontal et auditif droits. De plus des anomalies fonctionnelles ont été observées à la fois durant l'encodage, la rétention et la reconnaissance de l'information mélodique lors d'une tâche de mémoire à court terme. En revanche, pour le matériel verbal, les participants amusiques recrutaient des régions cérébrales similaires à celles des contrôles, ceci suggérant que des ressources neuronales distinctes sous-tendent la mémoire tonale et verbale. A partir des conclusions de ces trois premières études, nous avons exploré deux approches visant à renforcer les capacités de traitement de la hauteur dans l'amusie: 1) en étudiant si des connaissances implicites du système musical occidental tonal pouvaient améliorer leurs capacités de mémoire à court terme pour les notes, et 2) en explorant si les capacités d'encodage altérées des participants amusiques pouvaient être améliorées par des interactions audio-visuelles. Ces études ont été encourageantes et pourraient guider la conception d'outils de réhabilitation dans ce déficit. Enfin, il convient de souligner que nos recherches contribuent également à améliorer la compréhension des mécanismes sous-tendant le traitement de la musique en général, un domaine de recherche émergeant qui fait le parallèle avec le traitement du langage.

Mots-clés: amusie congénitale, mémoire auditive à court terme, perception de la hauteur, mémoire verbale, connaissances implicites, non-musiciens, intégration multi-sensorielle, Magnétoencéphalographie, Imagerie par Résonance Magnétique fonctionnelle.

Contents

Preamble

Over the last decades, substantial efforts in cognitive psychology and neuroscience were made to study how humans perceive and memorize complex sound structures. Recently, these questions were assessed by using music as a tool to explore human cognition and brain functioning. Studies investigating music processing have explored multiple complex cognitive functions (such as perception, learning, memory, emotions, expertise, or brain plasticity) that have led to major advances in the understanding of auditory brain mechanisms. However, regarding the neural correlates of music processing, an open question remains: how much the functional network of music perception and memory is a distinct or shared system relative to networks processing other auditory materials (e.g., speech, timbre)?

While it is now well determined that musical sounds and other ecological sounds (such as speech) share most of the processing stages throughout the auditory neuraxis, several lines of evidence have led to the hypothesis of a specialized "musical network" within the human brain. This hypothesis came from seminal investigations in brain-damaged patients who exhibited specific musical disorders (while other cognitive functions were not altered). More recently, this hypothesis received support from the investigation of a congenital music-related disorder. This disorder has been referred to as tone-deafness, dysmusia, dysmelodia, and more recently, as congenital amusia. Congenital amusia is a lifelong disorder of music perception and production, with the most notable impairments along the pitch dimension. Given that this disorder seems to be music-specific, this condition offers an occasion to explore the cortical networks underlying music processing and the relationship between music and language in the brain. While the behavioral underpinnings of this deficit have started to get well defined, the cerebral correlates of the disorder, in particular related to pitch perception and memory, are not fully understood yet.

This thesis addresses two main points. First, we aimed to decipher the behavioral and neural correlates of auditory perception and memory for isolated tones, musical sequences, and verbal material by investigating these processes in congenital amusia as compared to the normal functioning brain. Second, we explored whether and how amusic individuals' pitch perception and memory abilities could be boosted by exploiting musical implicit knowledge and multisensory integration.

I. Theoretical Background

Chapter I
Psychological and Neurobiological Underpinnings of Auditory Perception

The aim of the present chapter is to introduce very general concepts about auditory perception, from methods of investigation used in cognitive psychology and neuroscience to a description of the different processing steps supporting auditory perception in the brain. Depending on the research domain, two main approaches are used to explore the functioning and architecture of the auditory system. First, psychologists consider auditory processing as a complex hierarchical system which includes different processing steps, from auditory sensory encoding to higher level auditory processes such as memory as well as the influence of attention. The second approach is supported by neurobiologists who represent the auditory system as a dynamic, hierarchical, connected system supported by interaction between multiple brain areas. The main aim of contemporary cognitive neuroscience is to decipher the relationship between the psychological processes and the functional organization of the auditory areas in the brain.

1.1. Perception and perceptual systems

Human beings experience the world through their senses. Information from the environment is captured by multiple sensory (or perceptual) systems (auditory, gustatory, somatosensory, visual and olfactory). The definition of perception was originally debated through two alternative theories. In 1785, Reid (Reid and Hamilton, 1850) defined perception as *an act of the mind*, that is systematically dealing with consciousness. This theory suggests that perceiving objects (internal or external) depends on two characteristics: (1) the existence of an internal conception of the perceived object (referred to as *mental object*) and (2) a belief that this object exists. This theory suggests that perception needs an obligatory intervention of intellectual processes, the *act of perceiving* and its corresponding '*sensation*' occurring simultaneously.

The alternative theory was introduced by Gibson (Gibson, 1966) who suggested that senses can obtain information about external information without construction of a *mental object*. Gibson proposed thus a more ecological theory considering that *objects* could be processed directly by the perceptual systems. He therefore distinguishes the input to the nervous system that evokes conscious sensation from the input that evokes perception (the stimulus information can determine perception without having to enter in consciousness in the form of sensation).

However, it is relevant to note that the frontier between these two theories seems not clearly delineated as both conscious and unconscious perceptions require sensory information to be processed by the perceptual systems. By adopting this simple view, seminal researches in cognitive psychology and neuroscience have tried to determine the functional brain mechanisms that underlie perception within the perceptual systems. Based on this general framework, a simple question remains: how can we define anatomically and functionally a perceptual or sensory system?

While seminal concepts in cognitive neuroscience have considered that perceptual systems in the brain are supported by isolated brain areas, recent theories go beyond this view notably by considering that perception is supported by highly hierarchically organized functional systems that involve dynamic interactions between brain areas.

This theory allows us to consider the brain as a constructive or predictive organ that actively generates inferences from its sensory inputs using an internal or generative model (Friston, 2003), these inferences being produced consciously or unconsciously (see Helmholtz's original writings on unconscious inference (Helmholtz, 1866/1962)).

The present PhD thesis is embedded in this contemporary framework that considers the brain as a computational system designed to make inferences about properties of a physical environment. We assume that these processes are supported by local and large scale interactions between several brain areas (Friston, 2003; Mumford, 1992; Rao and Ballard, 1999; Varela et al., 2001).

1.2. Methods of investigations

Two main approaches are currently used for the study of perception. First experimental psychology methods, such as psychophysics and cognitive psychology, are exploring human cognition as revealed by behavioral performance. The second approach consists in the exploration of the cerebral bases of perception by numerous methods, from electrophysiology in animals to cerebral imaging studies in Humans.

1.2.1. Experimental psychology

The purpose of experimental psychology is the investigation of human mental faculties including, among others, perception by using behavioral methods. This approach is based on three main components: 1) psychophysics, which aims to quantitatively analyze the relationship between a sensation and the physical characteristics of the stimulus, 2) cognitive psychology, which tries to decompose a mental faculty into small components that interact, and 3) neuropsychology, which uses behavioral studies in brain-damaged patients to make the parallel between psychology and brain functioning The following paragraphs will briefly present these branches of experimental psychology for the specific case of auditory perception.

Psychoacoustics

Psychoacoustics is the branch of psychophysics that aims to study the sensation raised by sounds, and its relationship with stimulus characteristics (intensity, frequency, etc.). Various questions are addressed in this domain by using different protocols: from the determination of perception thresholds for low intensity level stimuli, over discrimination thresholds of pitch, to the study of more complex phenomena, such as masking effects. In general, these studies allow making inferences about the early steps of the auditory processing (such as cochlear encoding) in the nervous system. A review of psychoacoustic approach could be found in Moore's book (Moore, 2003).

Auditory cognition

As described by McAdams and Bigand (McAdams and Bigand, 1993), auditory perception encompasses a broader set of mechanisms than those explored by psychoacoustic methods. To understand auditory perception as well as high-level cognitive processes (i.e. speech, music perception), the exploration of additional cognitive abilities such as memory or attention is needed. This approach proposes that perception is a way to acquire knowledge about our surrounding world, with this previously acquired knowledge then impacting on our perception of the environment. The distinction between psychoacoustic and auditory cognitive approaches is however artificial, and might be more considered as a continuum. In order to explore a given cognitive skill (e.g. perception), one branch of cognitive psychology and neuropsychology illustrates this skill by interactions between processing modules (in a determined order). The role of the psychologist is to determine the function of each of these modules and to explore their relationships with other modules through information pathways. Nowadays, cognitive psychology approaches have developed several other methodologies to explore cognitive processes, notably by modeling the behavior by computational systems, exploring and developing artificial intelligence or by proposing interactive models. In addition to these approaches, auditory perception can be investigated through the study of its cerebral bases.

Neuropsychology

Seminal advances in the exploration of human brain function were done with behavioral studies in brain-damaged patients. By studying the link between the location of the brain lesion and the observed patient's deficits at the behavioral level allows making inferences about the role of the damaged brain structure. However note that this approach cannot show directly the potential changes in functional organization (related to brain plasticity) that can be observed in these patients (Muller and Knight, 2006).

1.2.2. Access to auditory brain functions

Brain activity is the consequence of interactions between neurons: a given neuron receives information from other neurons via synaptic communication (at its dendrites), and then transfers a new message to its target neurons through its axon (Warren, 2008). The processing of information requires that populations of neurons coordinate their activity. A simple model is to consider that a homogeneous local brain region (such as cortical column) receives afferents from a number of other brain regions or sensory receptors. These afferents leads to functional communication between regions allowing the information to be integrated in a given local region (via interactions between local inter-neurons). The resulting information is then sent via efferent pathways to remote structures. According to Varela, this architecture composed of local and large neural networks in the brain allows the integration of a "*symphony of emotions, perceptions, thoughts and actions*" (Varela et al., 2001).

Within this global network, it is therefore interesting to investigate the underlying physiological processes of each step of the integration chain including single cell physiology, interactions between neurons in a local region, and the coordinated activity between distant regions (cortico-cortical and cortico-subcortical interactions). Several methods in cognitive neuroscience were elaborated to explore brain functioning. These approaches are briefly described in the following paragraphs.

Single unit electrophysiology

Neuronal activity can be studied with recordings of electrical phenomenon associated to cell activity (postsynaptic potentials on the dendrites and soma, and action potentials in axons). This is the goal of single unit and multiunit electrophysiology recordings in animals, which uses electrodes that are implanted in a given cerebral area of interest. This recording technique allows measuring the activity of a cell population, by measuring local field potential (LFP) that reflects dendrites and soma activities (post-synaptic activity) (Baker et al., 1999).

Human electrophysiology

Local recordings of neuronal activity are also possible in humans in the special case of patients affected by pharmaco-resistant epilepsy and who benefit of intra-cerebral electrodes implanted in the brain (used for functional clinical investigation before surgery). In this specific case, electrical signals could be collected when patients perform of perceptual or cognitive tasks (Stereotactic electroencephalography, SEEG; or intracranial electroencephalography iEEG)(Lachaux et al., 2012). In addition, indirect measures of neuronal electrical activity can be obtained with non-invasive approaches using different cerebral imaging methods. The most often used methods are electroencephalography (EEG) and magneto-encephalography (MEG) which measure electrical

potential and magnetic field changes in the brain respectively via sensors disposed on the scalp, or close to it.

The coordinated activity of a neuronal population with similar orientations (equivalent to a current dipole) generates a large signal that can be recorded on the scalp. These signals reflect a distant post-synaptic activity which can be recorded with fine temporal resolution (millisecond). However, the spatial resolution is quite coarse, and the source reconstruction of the generators that supports the scalp recorded activities is a complex problem (Gross et al., 2013). EEG and MEG are very similar but are not interchangeable. In particular, the magnetic signal is less distorted by the biological tissues, thus increasing MEG spatial resolution relative to EEG. However, the magnetic signal attenuates more rapidly as the distance from the source increases, in comparison to the electric signal, which makes the MEG quite insensitive to deep sources. In addition, note that MEG is mainly sensitive to tangential generators, corresponding to activities generated on cortical sulci (Pernier, 2007).

Metabolic cerebral imaging

These methods measure, at each brain point, a biological parameter describing a metabolic activity in the brain. Positron emission tomography (PET) measures the local concentration of a weak radioactive marker (H_2O or glucose) in the brain, in order to estimate the regional blood flow or local glucose consumption(Ollinger, 1987).

Another functional neuroimaging procedure using MRI technology allows measuring functional brain activity. Functional magnetic resonance imaging (fMRI) measures brain activity by detecting associated changes in blood flow. This technique, as does PET, relies on the fact that cerebral blood flow and neuronal activation are coupled. When a brain region is active, blood flow to that region increases. Functional MRI thus depends on the degree of blood oxygenation level in the brain (BOLD imaging: blood oxygen level dependent) (Huettel et al., 2009).

Note however that while PET and fMRI techniques show a high spatial resolution (mm) the temporal resolution is low for two reasons: 1) the time required to acquire an entire cranial volume is long (about 2s for fMRI), and 2) the measured metabolic changes are slow, and occurred relatively late after stimulation (it takes several seconds to the BOLD signal to reach its maximum after stimulation). These methods are thus not appropriate for measuring the time-course of neural events.

Recently, another method called functional near-infrared (fNIR) spectroscopy has started to be used in cognitive neuroscience (Obrig and Villringer, 1997). This method is based on the fact that the functional state of tissue can influence its optical properties (i.e. cyanosis in hypoxia and pallor in anemia). fNIR exploits the NIR spectrum light by taking advantage of the optical window in which skin, tissue, and bone are mostly transparent to NIR light, while hemoglobin and deoxygenated-hemoglobin are stronger absorbers of light. Differences in the absorption spectra of deoxy-hemoglobin and oxy- hemoglobin allow the measurement of relative changes in hemoglobin concentration through the use of light attenuation at multiple wavelengths. Note, however note that the spatial resolution of this technique is coarse and does not allow to measure reliably deeper brain structures.

1.3. Psychological aspects of auditory perception

In the following paragraphs, we will briefly present the psychological aspects of auditory perception by describing the different processing steps of the auditory information through the perceptual system. The first step in the processing of sounds is the transduction of the acoustic signals into sensory data that will be integrated by the perceptual systems. The sound is first filtered in the outer and middle ear to achieve the cochlea where a first frequency analysis is performed. The perceptual analysis is preliminary based on this first representation.

1.3.1. Auditory scene analysis

The auditory system has two main roles: 1) it has to decompose the sound information presented by different sources simultaneously and 2) it has to recognize (or identify) these sources. Acoustic waves emitted by different sources have the particularity to be additive, thus resulting in a mixture of sound that it derives from a complex natural environment. Auditory scene analysis is the processing whereby the auditory system captures this mixture and sorts it into packages of acoustic evidence in which each package has arisen from a single source of sound. The following paragraphs are largely inspired from Anne Caclin's thesis (Caclin, 2004) and from the book of McAdams and Bigand (1993).

Since Bregman's book (Bregman, 1990), it is widely accepted that sound analysis in the perceptual system begins by the segregation of auditory sensory information in several streams. Bregman argues that the process of auditory scene analysis is governed by both innate and learned constraints that allow the construction of an auditory representation. He distinguished two main mechanisms, one based on unlearned mechanisms called *primitive scene analysis* and those of the learned ones referred as *schema-driven analysis (also referred as schema-based construction of the description)*. Both processes are concerned with the decomposition of mixture of information so that the right combination of information can enter into the description of an environmental sound (Figure 1).

Primitive scene analysis

Primitive analysis is a simple process, probably innate, driven by the incoming acoustic data. This approach is based on sounds' general acoustic properties, which allow the mixture to decompose sources without a priori assumptions on the sources. The main rules of primitive scene analysis are:
- independent sources rarely start and stop emitting sounds simultaneously;
- sound variation tends to be gradual when the sound is emitted by a single source;
- vibrating objects (such as vocal chords) emit often harmonic sounds, and it is rather rare that two simultaneous sources present the same periodicity;
- changes in a source affect coherently the different components of the sound it is emitting.

```
                              ┌──────────────┐
                              │    Action    │
                              └──────────────┘
┌──────────────────┐                              ┌──────────────────────┐
│ Previously acquired│         ▲      ↖           │      Attention       │
│    knowledge       │                            │ Short-term Memory    │
│ (Long Term Memory) │                            │ Mental representation │
└──────────────────┘                              └──────────────────────┘
                    ┌──────────────────────────┐
                    │    Source Recognition    │
                    └──────────────────────────┘
                              ┌──────────┐
                              │  Echoic  │
                              │  Memory  │
                              └──────────┘
                    Extraction/Computation of sound
                         Properties/Attributes

                              Auditory
                              Grouping
                                          Sensory
                                        Transduction

                              ⬭ Sound ⬭
```
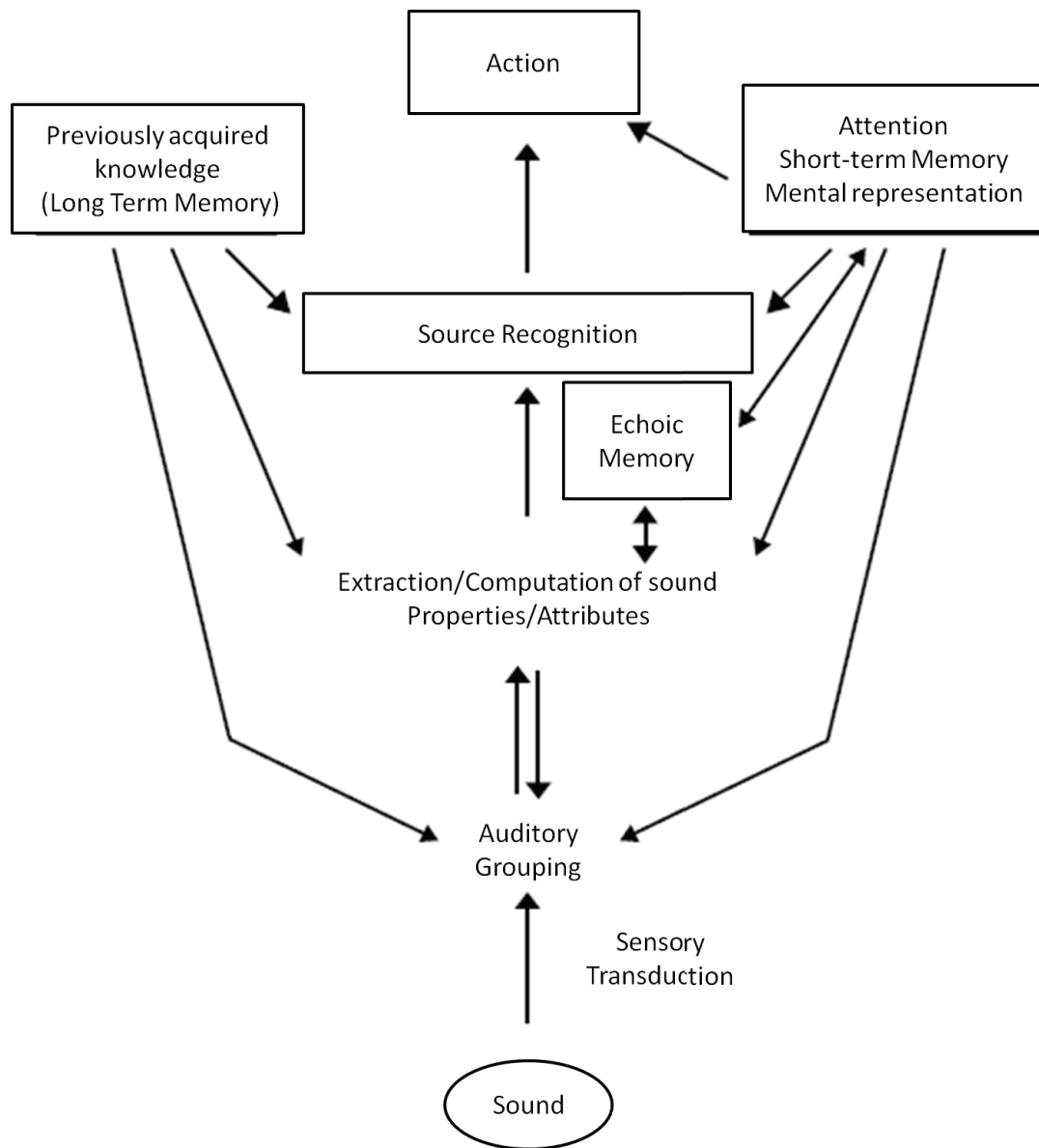
Figure 1: The modular psychological model of auditory perception. Each box represents a processing component and arrows represent pathways of information (adapted from McAdams & Bigand 1993).

Schema-driven analysis

In addition to the unlearned mechanisms supporting auditory analysis, stream segregation can also be constrained by listeners' previously acquired knowledge that leads to expectations. This hypothesis is based on the fact that information from the environment does not stimulate a completely naïve organism. Acquired knowledge interacts with the current sensory data to interpret the auditory stimulation. The schema-driven process is based on this assumption. This analysis is thus presumed to involve the activation of stored knowledge of familiar patterns or schemas of the acoustic environment. One simple example to illustrate this idea is that we are particularly sensitive to our first name. More generally, depending on the context, we could have expectation on the occurrence of a sound (or a type of sound), and then use pre-existing schemas (stored in long-term memory) to analyze

sensory inputs. In addition to these analyses, the auditory system has to decide which bits of sensory evidence have arisen from the same environmental events.

Sequential and simultaneous organizations

Bregman proposed that this grouping processing is done in at least two dimensions, across spectrum and across time. These two processes were defined as *the sequential organization* (across time), and the *simultaneous organization* (at each time point). Sequential and simultaneous organizations underlie many aspects of auditory processing. In music processing for example, sequential grouping is the foundation of rhythm and of many aspects of melodic form whereas simultaneous grouping is involved with such experiences as chord perception, timbre, consonance and dissonance. Finally, these organizations allow segregating different auditory streams exhibiting distinct temporal evolutions. These mechanisms are to some extent modulated by attention.

Along with the partitioning of the different streams, the auditory system has to extract the particular attributes of each stream. Traditionally the following attributes are distinguished: pitch, loudness, duration, timbre and spatial position. The extraction of auditory attributes especially related to pitch is presented with more details in page 34.

1.3.2. Recognition of sound sources

Recognition of sound sources is a complex problem that consists in focusing on one auditory stimulus while filtering out a range of other stimuli (Wood and Cowan, 1995). Interestingly, the auditory system succeeds in recognizing different individuals thanks to their voices and in inferring the internal properties of external auditory objects (i.e. type of stimulus, auditory material). Recognition of sound sources is often presented as a subsequent process of extraction of auditory attributes (McAdams, 1993). Studies investigating sound source recognition have tried to determine which acoustic cues are used by the subject to identify the source. The problem is the following: depending on the context, a same object could emit sounds quite differently. For example, the sound spectrum of a musical instrument will depend on the intensity with which it is played. It is therefore necessary to extract appropriate acoustic invariants, to overcome the contextual effects in the determination of the contribution of each source of an acoustic mixture (see McAdams, 1993). Finally, recognition of sound sources requires listeners to use their previously acquired knowledge and is thus related to the available information stored in memory.

1.3.3. Auditory Memories

Depending on nature of the duration of storage, the following systems are distinguished: 1) echoic memory, 2) auditory short-term memory, and 3) auditory long-term memory. Note that all these systems can be modulated by auditory attention.

Echoic memory refers to a sensory process that keeps a trace of recent auditory events in sensory memory. This concept is widely used in the field of brain evoked potentials (Winkler et al., 1992). Behavioral studies on echoic memory have shown that there are two different types of auditory sensory memory (Cowan, 1984), 1) a short auditory store, which extends the apparent duration of a stimulus up to about 300 ms and is used in stimulus recognition, and 2) a long auditory store, which retains auditory information of a sound or a sound sequence for at least several seconds.

Auditory short-term memory is the mechanism that allows the active retention in memory of recent auditory information. In some models, short-term memory is replaced by (or associated to) the notions of working memory. The differences between these two concepts will be discussed in Chapter II (see page 50, 58).

Auditory long-term memory allows the retention of auditory information over a long time period (minutes to years). This type of process allows recognizing familiar sounds such as the voice of a familiar individual. The exact forms of the information stored in long-term memory remains debated.

This brief presentation of auditory perception through the view of cognitive psychology would not be complete without mentioning auditory attentional processes. The study of attention is a broad domain, in such a way that it becomes difficult to define precisely attention processes (Jones and Yee, 1993). In general, attention aims to perform an active selection of given sensory information within multiple external stimulations. Several models could be found in the models of auditory attention. A first consideration is that orientation of attention could be voluntary or involuntary (also referred as exogenous or endogenous attention). In the first case (exogenous), an event that occurred suddenly, or occurring sufficiently saliently in comparison to a previous event (or context), can irresistibly attract attention. The second case (endogenous) refers to the internal state of the individual. Research in cognitive psychology tries to manipulate this endogenous process by asking the participant to perform a task that requires a deliberate attentional focus on a particular aspect of the stimulus. In voluntary attention, the distinction between divided and selective attentions can be made. In selective attention, the subject has to focus on a given stimulus characteristic while ignoring the others, whereas in divided attention, the subject musts focus his attention on several competing information.

Additionally to selective and divided attention, another aspect of auditory attention had been pointed out by Jones (Jones, 1976; Jones and Boltz, 1989; Large and Jones, 1999) in the *dynamic attending theory*. In this theory attention vary (*oscillate*) as a function of the temporal aspect of the stimulation. These attentional *oscillations* are supposed to have periods similar to the periods by which important stimuli appear in the environment, which could thus either have lower level attentional periods or high level attentional periods.

1.4.   Anatomical and functional architecture of the auditory system

After this brief presentation of auditory processing from the perspective of cognitive psychology, the following paragraphs will present the anatomical and functional architecture of the auditory system from the perspective of neurosciences. Auditory information is transmitted from the cochlea to the auditory cortex (and beyond), through different sub-cortical nuclei. Besides the upward path of auditory information, there are descending pathways which contain many levels, which are not described here (see Ehret and Romand 1997; Pickles 1988). For the description of the different sub-cortical processing levels, the data summarized below comes from single-unit electrophysiology studies, especially in the cat. The data relating to auditory cortex are derived from anatomical studies and cerebral imaging (metabolic and electrical) in Humans, as well as electrophysiological studies in primates. More details can be found in the books of Pickles (1988) and Ehret and Romand (1997).

1.4.1.   From sound wave to auditory nerve

Sounds are acoustic waves emitted by different objects in our environment. These waves are additive in the air to form an acoustic mixture that reach the ear. This acoustic signal enters in the inner ear through the outer and middle ear, where a first frequency filtering is performed.

Outer and middle ear

The outer ear includes the pinna and the ear canal. The pinna is a concave cartilaginous structure, which conveys sound waves towards the ear canal or external auditory meatus. The ear canal drives the sound waves towards the tympanic membrane. The middle ear is composed of the tympanic membrane and of a cavity, which contains the ossicular chain consisting of the three smallest bones in the body: the malleus, incus, and stapes. The tympanic membrane transmits the air vibrations from the outer to the middle ear via this ossicular chain.

Inner ear and cochlear encoding

The inner ear is composed of the cochlea and of the vestibular system. While they are both encased in the same bony capsule and share the same fluid systems, they do not play the same role. We describe here only the cochlear system as the vestibular apparatus does not intervene in auditory perception (its role is in maintaining equilibrium, balance). The hearing part of the inner ear is the cochlea that operates as a frequency analyzer were the signal is analyzed frequency band by frequency band and then transmitted to the auditory nerve. The cochlea corresponds to a helical structure composed of three fluid-filled chambers that extend along the length of the structure

The two outer chambers are filled with the perilymph that acts as a cushioning agent for the structures that occupy the center chamber. The third fluid filled chamber is the center chamber, referred as the cochlear duct, which secretes the endolymph. The cochlear duct contains the basilar

membrane upon which lies the Organ of Corti, a sensory organ which plays an important role in auditory perception. The Organ of Corti consists of approximately 30.000 finger-like projections of hair cells (with cilia emerging from the apical surface of hair cells). Each hair cell is connected to a nerve fiber that relays various impulses to the cochlear branch of the VIIIth cranial nerve (or auditory nerve). The frequency of the impulse relayed is dependent upon which areas of the basilar membrane, and hence, which portions of the Organ of Corti are stimulated. The apical portion of the basilar membrane transfers lower frequency impulses. The basal end relays higher frequency impulses. This frequency selectivity is thus a passive phenomenon as it is related to the mechanical properties of the cochlea. Note that additionally, active phenomenons are also observed, notably supported by the outer hair cells that operate as an active amplifier of the vibrations detected by the inner hair cells.

It should be note that each hair cell responds thus to a specific narrow frequency band leading to a tonotopic (spatial specificity as a function frequency) organization of the auditory system starting in the cochlea. This tonotopic organization is preserved in the next steps of processing, from the sub-cortical auditory nuclei (relay) to the auditory cortex. Behavioral measures such as frequency discrimination or intensity thresholds reflect directly or indirectly the responses of hair cells and the activity in the auditory nerve (Pickles, 1988).

1.4.2. Sub-cortical auditory relays

After the cochlea, the auditory nerve projects towards several different sub-cortical nuclei representing parallel processing paths (see Figure 2). Neurons of the different nuclei have strong frequency selectivity; however, this selectivity decreases gradually with increasing processing hierarchy.

From cochlear nucleus to colliculus

Fibers of the auditory nerve project on ipsilateral cochlear nucleus. The cochlear nucleus is divided into two parts: 1) the ventral part which presents similar response profiles than those observed in the auditory nerve and 2) the dorsal part where more complex responses emerge. Neurons of the ventral part project ipsilaterally and contralaterally to the superior olivary complex. This structure helps to encode the spatial position (direction) of the stimuli, the neurons of the superior olivary complex being sensitive to intensity and to temporal dynamic (timing) differences between the two ears. Neurons of the dorsal part of the cochlear nucleus and of the superior olivary complex (toward the lateral lemniscus) project toward the inferior colliculus.
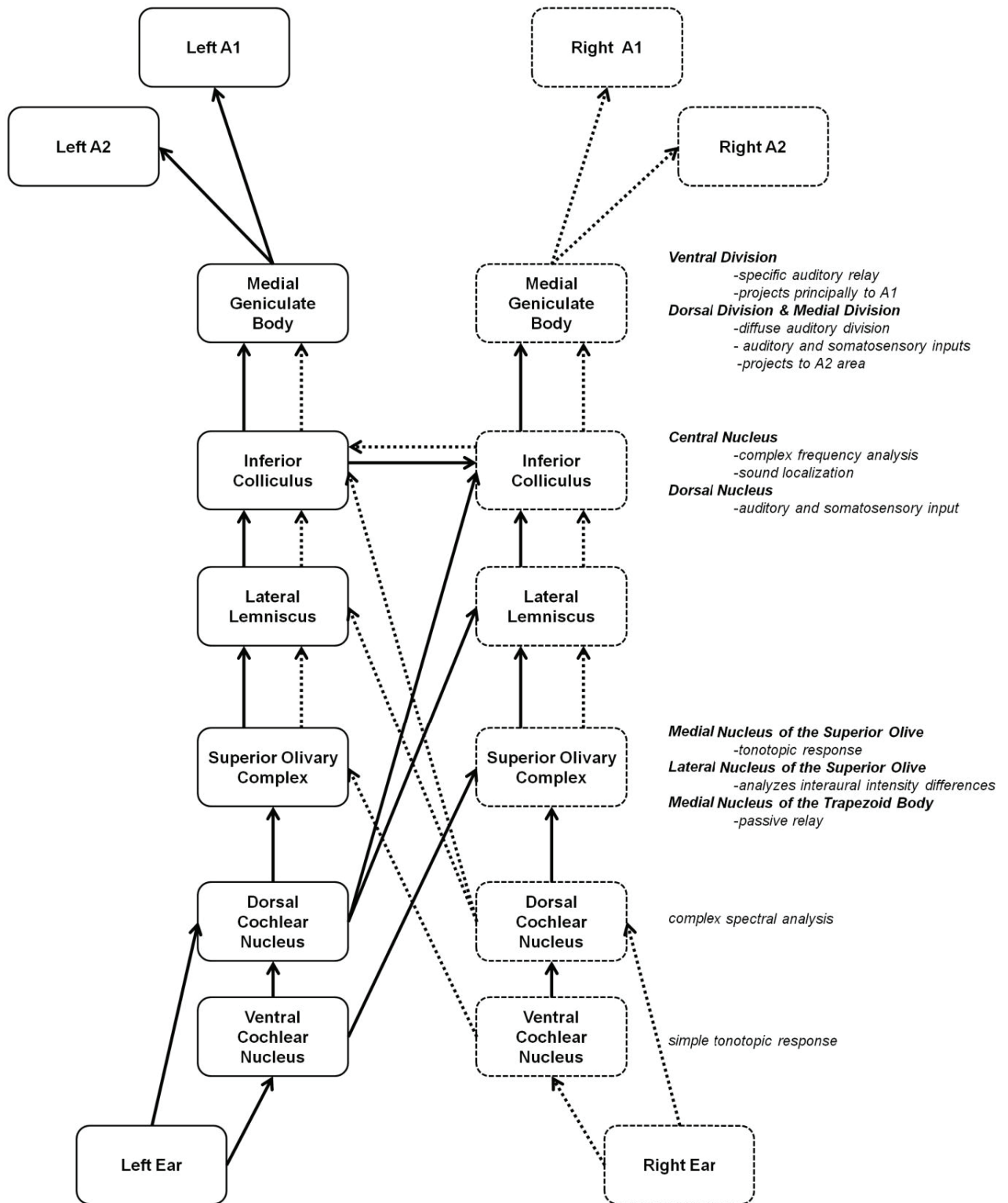
Figure 2 Overview of ipsilateral afferent pathways and controlateral interactions in the auditory brainstem. A1, Primary Auditory cortex; A2, Secondary Auditory Cortex; dotted lines, pathways from the right ear; full lines, pathways from the left ear. (based on Ehret and Romand, 1997).

From colliculus to thalamus

The major parts of the fibers projected by the auditory brainstem nuclei converge to the inferior colliculus. From the inferior colliculus (in cats), three pathways project to the cortex through the thalamus (including the Medial Geniculate Body): 1) a tonotopic pathway which rises to the primary auditory cortex and 2) two other pathways, one called diffuse pathways (non-tonotopic), the other called multisensory, projecting to the secondary auditory cortex. In the multisensory pathway, the auditory information is combined to somato-sensory information (Ehret and Romand, 1997). The responses of neurons in these various sub-cortical pathways are sophisticated and diversified as neurons are sensitive to different stimulus characteristics, such as modulation in intensity or frequency (Ehret and Romand, 1997).

1.4.3. Anatomical architecture of the auditory cortex

As detailed in the previous section, the auditory information reaches the cortex after a final relay in the thalamus Medial Geniculate Body. In humans, the auditory cortex is located in the supratemporal plane at the bottom of the Sylvian fissure, and extends to the superior temporal gyrus, to the parietal cortex and adjacent areas (Galaburda and Sanides, 1980; Rivier and Clarke, 1997). The auditory areas of the human superior temporal region have been delineated according to a variety of architectonic schemas. These schemas were based on the laminar organization of the cortex that allows making inferences on the functional organization of cortical areas.

Laminar organization of the auditory cortex

The neocortex presents a columnar arrangement of neurons across six cortical layers (layers I-VI) extending on 2-4 mm thick. The auditory cortex, including the primary field AI, has similar columnar organization, often described as functional macro columns (Scheich et al., 1993). The macro columns can be divided into three parts (see Figure 3): 1) the supragranular layers consisting in layers I to III. This part represents the origin and termination of intra-cortical connections, which are either associational (i.e., with other areas of the same hemisphere), or commissural (i.e., connections to the opposite hemisphere, primarily through the corpus callosum). This cortical portion is highly developed in humans and allows functional connections between a given cortical region and local or distant brain regions. 2) The internal granular layer, layer IV, which receives thalamo-cortical connections, especially from the Medial Geniculate Body of the thalamic nuclei. This is the most prominent layer in the primary auditory cortex. 3) The infragranular layers, layers V and VI, connecting primarily the cerebral cortex with sub-cortical regions.

By considering this cortical architecture, and based on the cellular population that constitutes each layer, inferences about the functional role of each part of the macro column could be made. In layer V, large numbers of pyramidal neurons are observed. These neurons show apical dendrites that span almost the whole thickness of cortex orthogonally to the layers.
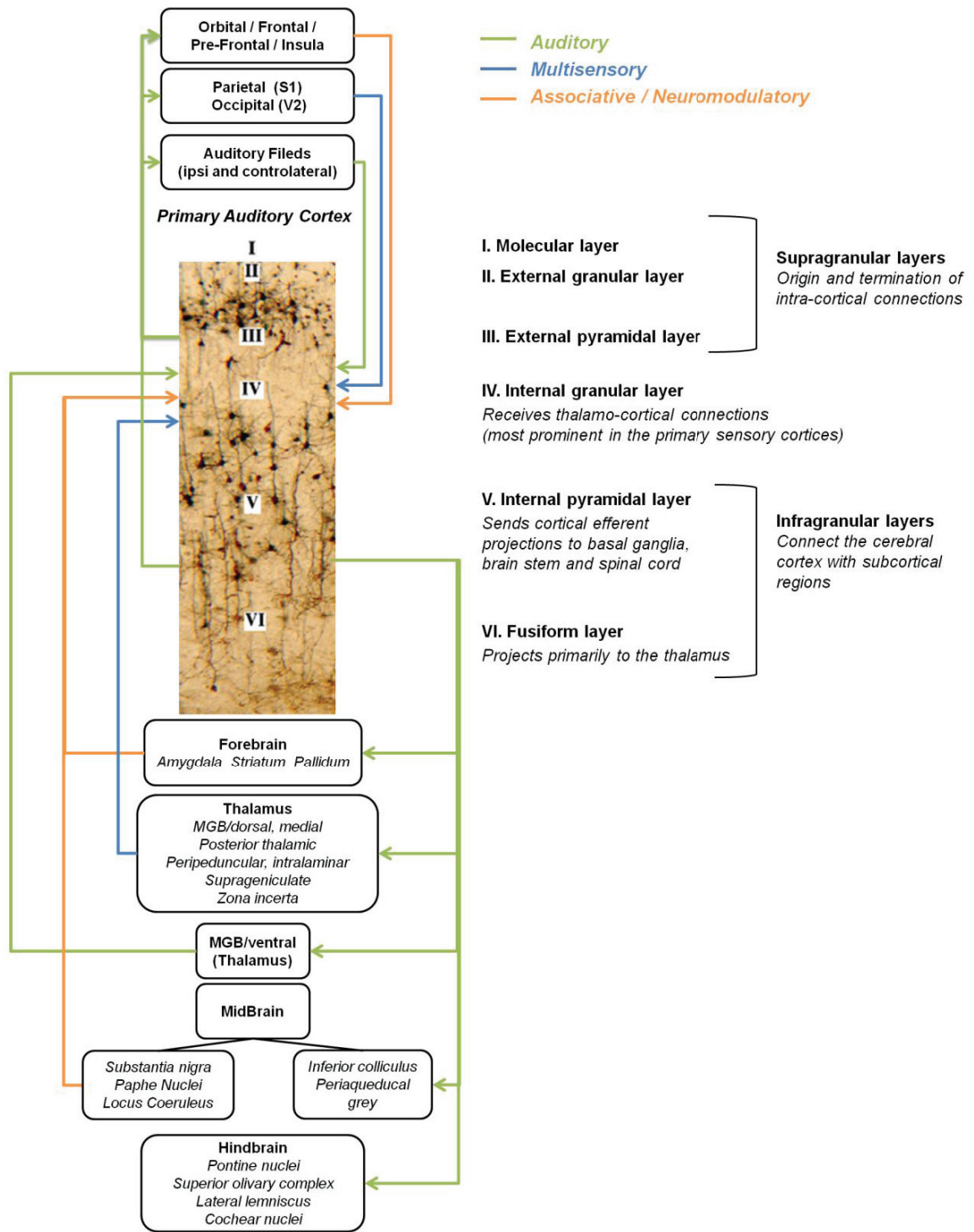
Figure 3. Schematic diagram of the connections of the primary auditory field (AI) adapted from (Scheich et al., 2007). Each box represents a brain structure which is connected with AI. Blue, green, and orange colors are used to identify the type of information which is most likely processed by these projections, i.e. auditory, other sensory/multisensory, or associative/cognitive/(neuro) modulatory information. Connections of AI with the medial prefrontal cortex include separate connections with anterior cingulate, prelimbic, and infralimbic areas. MGB, Medial Geniculate Body.

These neurons are surrounded by various types of non-pyramidal neurons in all layers and by smaller pyramidal neurons in infragranular (V–VI) and supragranular layers (I–III). These small pyramidal neurons are particularly observed in layer III, where they are all parallel to each other and show short apical dendrites that reach the surface layer I (Szentagothai, 1978).

Because of this geometry, apical dendrites of the different pyramidal neurons spatially overlap in the supragranular layers (more or less). Thus, the pyramidal cells from all layers could directly integrate all cortico-cortical inputs, as these inputs terminate also in the supragranular layers. In addition to this input integration observed across the layers by the apical dendrites, incoming inputs within the layers are also integrated by the basal dendrites of the pyramidal cells. This input constitutes a functional division between the cortico-cortical inputs to the superficial supragranular layers and the commissural inputs to the deeper infra-granular layers, particularly layer V. Finally, layer-specific outputs of the pyramidal cells are observed, via their extrinsic axonal projections. The laminar pattern of these axonal terminations within a given brain area as well as the laminar location of cells of origin within a given source region indicates the type of cortico-cortical connectivity between these two regions (i.e. the forward type, or backward type).[1]

Although the functional impact of this cortical architecture has not yet become fully determined, the study of the cortical and sub-cortical input sources to the different layers of the auditory cortex has allowed for the description of the different substructures that form anatomically the auditory cortex. As mentioned above, for example, the thalamo-cortical inputs to the auditory cortex from the ascending auditory pathway is predominantly observed in layer IV, thus allowing to determine within a complex cortical structure which cortical area corresponds to the primary sensory area.

Anatomical segmentation of the auditory cortex

Based on the assumption that primary sensory and associative auditory areas present different cytoarchitectonic organizations, researches in auditory neuroscience have tried to decipher the different substructures that constitute the auditory cortex. In Humans, primary sensory areas present several biological characteristics: 1) a koniocortical architecture (extensive accumulation of granular cells- mainly observed in cortical layer IV), 2) strong expression of cytochrome oxydase (metabolic enzyme), acetylcholinesterase (enzyme that deactivates neurotransmitter and neuromodulator) and parvalbumin (calcium binding protein), 3) and strong myelination. In contrast, secondary and associative areas are composed of large pyramidal neurons showing a diffuse myelination and exhibiting less metabolic activity.

In primates, the auditory cortex is located on the superior temporal cortex (Figure 3A). It is constituted of several auditory fields that follow a hierarchical organization consisting in three levels (Fullerton and Pandya, 2007): 1) two or three primary auditory areas that constitute the 'core' of the

---

[1] Note that we had this description of the laminar organization of the cortex to help the description (see Chapter page 75) of a connectivity model we have used in the present PhD work.

auditory cortex 2) seven to eight secondary auditory areas constituting a first 'belt' that surround the primary areas (also referred as 'root' areas) 3) associative areas that constitute a second belt ('parabelt'). The parabelt is adjacent to the secondary auditory areas along the Superior Temporal Gyrus. In the following paragraphs, we will present the anatomical organization of the auditory cortex in humans. Note, however, that the description of models derived from monkey studies have served as a starting point in understanding the structural and functional organizations of auditory cortical processing in Humans (Kelly et al., 2010; Rauschecker and Scott, 2009). It is therefore relevant to note that many questions concerning the homology of auditory-related cortical fields between human and nonhuman primates remain unsettled (Fullerton and Pandya, 2007; Hackett, 2003; Hackett, 2007; Hackett, 2008; Hackett et al., 2001; Hackett et al., 1999; Kaas and Hackett, 2000; Petrides and Pandya, 1999, 2002; Sweet et al., 2005).

Seminal studies by Brodmann (Brodmann, 1909) (Figure 3B) and Sarkisov et al. (Sarkisov et al., 1955) have defined the primary auditory cortex in the supratemporal plane as area 41, surrounded by areas 42 and 22 laterally, and area 52 medially, thus locating the primary auditory cortex in the postero-medial part of the transverse gyrus of Heschl. Later, Von Economo and Koskinas (von Economo and Koskinas, 1925) (Figure 3C) and von Economo and Horn (von Economo and Horn, 1930) defined the primary auditory area as TC, surrounded by areas TB, TD, TG, and TA. According to Galaburda and Sanides (1980 Figure 3D, see also (Wallace et al., 2002)) the primary auditory cortex is subdivided into medial koniocortical areas (KAm) and lateral (KAlt) components and by several surrounding regions as parakoniocortical fields.

More recently, using a densitometry method, Morosan et al. (Morosan et al., 2001) have outlined various fields in the transverse temporal gyrus of Heschl. They defined three highly granular areas with a developed layer IV, representing primary auditory cortex: Te1.1 most caudally, Te1.0 medially, and Te1.2 most rostrally (see Figure 5). These studies thus converge toward the existence of a compartmentalization in the primary auditory cortex along the Heschl's gyrus.

In addition to the primary sensory areas, several secondary and associative auditory regions were defined within the human auditory cortex 1) rostrally in the Planum Polare, 2) caudally in the Planum Temporale up to insular cortex and parietal cortices (supramarginal and angular gyri, parietal operculum), and 3) laterally in the superior Temporal Gyrus. According to cytoarchitectonic criteria and differential staining of brain tissue, different delimitations have thus been proposed in the secondary human auditory cortex, these delimitations having 3 to 8 regions (see Galaburda and Sanides 1980; Rivier and Clarke, 1997, Morosan et al. 2001; Fullerton & Dayna 2007). Wallace and colleagues (2002) concluded on the existence of six secondary auditory areas (A1 belt) in the supra-temporal plane:

- anterior auditory areas (AA) and medial areas (MA) located in the Planum Polare
- lateral (LA) and posterior (PA) auditory areas located in the Planum Temporale
- antero-lateral auditory area (ALA) located in the lateral part of the Heschl's gyrus

- superior temporal area (STA) located in the posterior part of the superior temporal gyrus. Note that the STA shows a different composition in comparison to other secondary auditory areas, as it shows a stronger concentration of pyramidal cells. This region might be related to higher level of integration as observed in macaque parabelt.
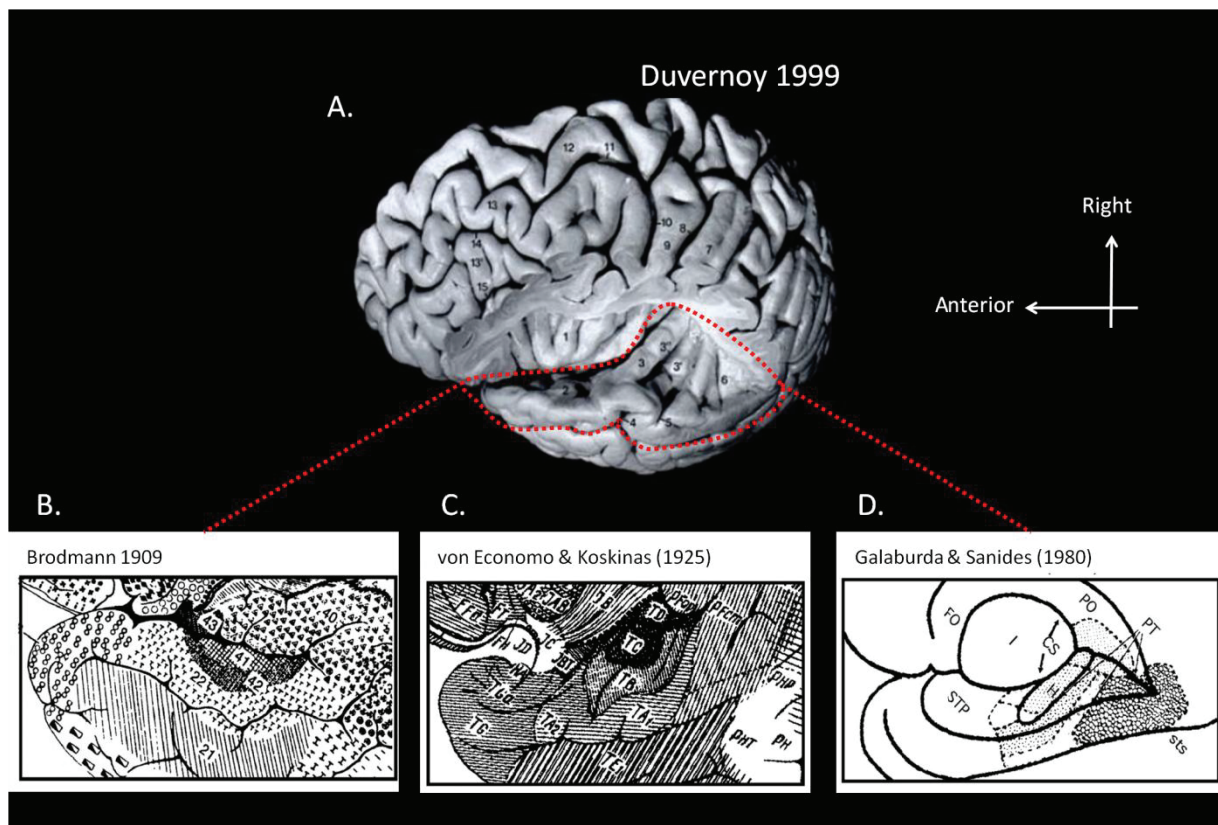
Figure 4. A. Superior aspect of left superior temporal gyrus after ablation of the lower part of the frontal and parietal lobes. 1 insula; 2 planum polare; 3 3' anterior and posterior transverse temporal gyri of Heschl; 3'' intermediate transverse temporal sulcus; 4 sulcus acousticus; 5 transverse temporal sulcus; 6 planum temporale; 7 postcentral gyrus; 8 central sulcus; 9 precentral gyrus; 10 superior precentral sulcus; 11 superior frontal sulcus; 12 superior frontal gyrus; 13 13' middle frontal gyrus, superior and inferior parts; 14 middle frontal sulcus; 15 inferior frontal sulcus. Adapted from Duvernoy 1999. B,C,D. Architectonic maps of lateral surfaces of cerebral hemisphere of human according to (B) Brodmann (1909), (C) von Economo and Koskinas (1925), and (D) Galaburda and Sanides (1980). The insets diagrams on the bottom are magnified views of the auditory-related areas in and around the lateral (Sylvian) fissure. Adapted from Fullerton & Pandya (2007) and Duvernoy (Duvernoy, 1999)

However, note that the number of secondary auditory regions vary as a function of the number of areas we consider as 'primary' areas. In Morosan et al. (2001) the most anterior (rostral) Te.1.2 area (see figure 5.B), could be associated to ALA region of Wallace et al. (2002, Figure 5.C).

Taken together, these studies lead to a tripartite organization of the auditory cortex: 1) areas on the supratemporal plane (core), 2) areas in the cortex of the circular sulcus (sulcus that runs all around

the insula, root), 3) and areas on the superior temporal gyrus (belt). More recently Fullerton and Pandya (2007) (see also Sweet et al., 2005) have further analyzed the core, lateral belt, and parabelt areas and have suggested that the cytoarchitecture of the auditory area changes in a stepwise manner from the koniocortical area, both toward the direction of the temporal polar proisocortex as well as toward the caudal temporal cortex. This architectonic dichotomy is consistent with differences in cortical and sub-cortical connections of superior temporal region and may be related to different functions of the rostral (lateral) and caudal (medial) temporal cortices.
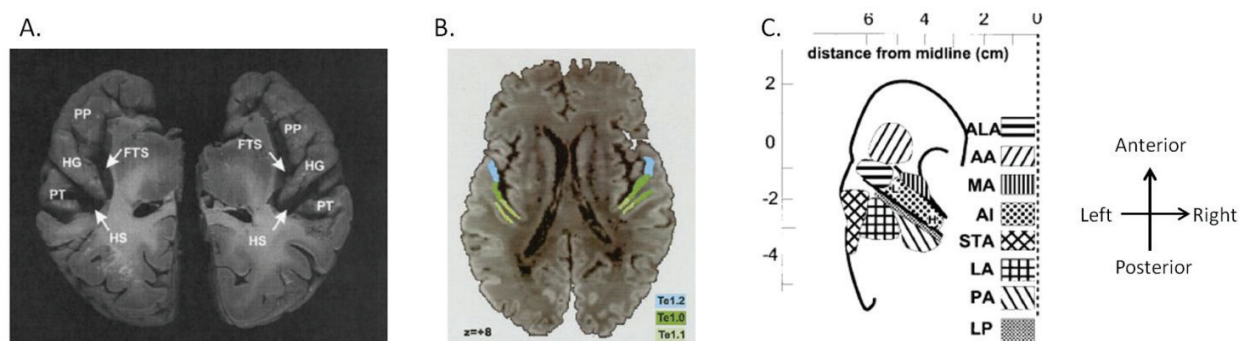
Figure 5. A. Dorsal view after removal of the frontal and parietal opercula of an individual brain showing macro-anatomical landmarks of the auditory regions. FTS, first transverse temporal sulcus; HS, Heschl's sulcus; HG, Heschl's gyrus; PP, planum polare; PT, planum temporale. B. Maps of areas Te1.2, Te1.0, and Te1.1 of the primary auditory cortex in an axial histological section. Adapted from Morosan et al (2001) C. Summary diagram of auditory areas on the superior temporal plane from Wallace et al. (2002). The core auditory areas have been shaded with dots and the presumed secondary areas by lines or hatching. ALA, anterior lateral auditory area; AA, anterior auditory areas; MA, Medial area; AI, primary auditory cortex; STA, superior temporal area; LA, lateral auditory area; PA, posterior auditory area; LP, Lateroposterior area

1.4.4. Electrophysiological architecture of the auditory cortex

Another way to investigate the different substructures constituting the auditory cortex is to explore functionally when and in which temporal order the different cortical auditory areas are activated after the presentation of a sound. This information about sound processing in the auditory cortex in humans can only be obtained by using high temporal resolution techniques. Electrophysiological studies (Picton et al., 1974; Wood and Wolpaw, 1982) have described that after stimulus onset, the early brainstem electrical evoked responses (occurring before 10 ms) are followed by different evoked components originating from the auditory cortex. The responses elicited between 10 and 70 ms after stimulus onset are termed middle latency components whereas later components are elicited in the range of 80-250 ms (referred as long latency components). Middle latency components are usually described (in EEG studies) as a succession of four waves labeled as Na (19ms), Pa (30ms), Nb (40ms) and Pb (also labeled P50 or P1, 50-70 ms). Middle latency components

are followed by the most prominent late latency component referred as N1 (or N100) having a latency of 100ms. These components have been identified using non-invasive EEG technique; it is thus relevant to note that these patterns of responses are quite different in terms of latency and profile when they are directly recorded using intra-cranial EEG.

Based on the identification of these brain responses, exploring the cortical generators of each of these components allows deciphering the functional compartmentalization of the auditory cortex in response to sound. Following this conceptual framework, several studies (Celesia, 1976; Gutschalk et al., 1999; Howard et al., 2000; Liegeois-Chauvel et al., 1994; Liegeois-Chauvel et al., 1991; Yvert et al., 2001) have used electrophysiological recordings to describe the functional anatomy of the auditory cortex. In line with Galaburda and Sanides (1980), Liégeois-Chauvel et al. (1994) have shown the existence of two primary auditory areas (the first area in the postero-medial part and the second in the central part) in the transverse temporal gyrus of Heschl using intracranial recordings. These regions elicited early potentials evoked by sound click (N13-P16-N30-P50). The first three potentials were generated in the most medial part of Heschl's gyrus while the last potential (P50) was generated in a more central part of the transverse temporal gyrus of Heschl. The following responses (about 60 and 75 ms) were generated in the lateral part of Heschl's Gyrus and the next response (100ms) was elicited by the Planum Temporale. These electrical dynamic patterns were also observed using simultaneous EEG/MEG recordings (Yvert et al., 2001). Yvert et al. (2001) performed source reconstruction of MEG data revealing serial processing of the auditory information along the medio-lateral axis of the supratemporal plane. In this study, the response *Pa* was observed in the medial part of Heschl's gyrus, the response Nb was elicited in the lateral part of the supra-temporal gyrus and the Pb potential was observed in the antero lateral portion of Heschl's gyrus.

These studies concluded that potentials evoked by sound clicks have shorter latency in the caudal part of the Heschl's gyrus while longer latencies were observed in the more rostral part. These studies thus suggest the existence of a serial and hierarchical organization of the auditory cortex for sound processing. This hypothesis was supported by other recent studies that have investigated the precise activation timings of each auditory cortical area using SEEG and MEG (Gueguin et al., 2007; Inui et al., 2006). In the study of Inui et al. (2006), activations were found in six cortical areas after the sound stimulus: 1) in the postero-medial part of Heschl's gyrus corresponding to the primary auditory cortex (AI), 2) in the antero-lateral (ALA) part of the Heschl's gyrus posterior to the transverse sulcus, the posterior parietal cortex (PPC), posterior (PA) and anterior parts (AA) of the superior temporal gyrus (STG), and 3) the planum temporale (PT).

In 2007, Guéguin et al. proposed the following conclusions 1) the processing of an auditory stream within the primary auditory cortex starts in the most medial (caudal) region of the Heschl's gyrus and extended later to the most lateral (rostral) region, 2) an unidirectional functional connection from the primary to secondary auditory cortex is present in human auditory cortex, 3) Heschl's sulcus plays a particular role in dispatching information to the different auditory areas. These findings suggest that cortical processing of auditory information is performed in hierarchical, serial, and parallel streams.

These data have shown that the auditory propagation is associated to a constant interaction between several primary and secondary areas that could reflect the large adaptive and plastic capacities of auditory cortex.

Taken together, cytoarchitectonic and histochemical studies have distinguished hierarchical and parallel organization of several substructures in the human auditory cortex. Functionally, the distributions of the latencies of auditory evoked potentials and magnetic fields are consistent with a serial processing of the auditory information along the medio-lateral axis of the supratemporal plane. This conclusion has received further support from other intracranial recording studies in humans more recently (Besle et al., 2008; Bidet-Caulet et al., 2007; Brugge et al., 2003; Brugge et al., 2008; Brugge et al., 2005; Crone et al., 2001; Garell et al., 2012; Reale et al., 2007; Steinschneider et al., 2011). Finally, note that the functional architecture of the auditory cortex has also been investigated through the exploration of the tonotopic organization of this structure. This part will be described in detail in the following section which investigates the particularities of the auditory system for pitch material (see section 1.5.3, pages 38 to 43).

1.4.5. Beyond the auditory cortex

In addition to these parallel processes observed in the different subdivisions of the auditory cortex, long-range connections constituting different cortical pathways allow the processing of auditory information. During the last decade, several studies have suggested that auditory cortical processing pathways are organized dually (Ahveninen et al., 2006; Alain et al., 2001; Altmann et al., 2007; Anourova et al., 2001; Barrett and Hall, 2006; Belin and Zatorre, 2000; Bidet-Caulet and Bertrand, 2005; Bushara et al., 1999; Clarke et al., 2005; Clarke and Thiran, 2004; Garell et al., 2012; Kaas and Hackett, 1999, 2000; Maeder et al., 2001; Rauschecker, 1998; Rauschecker and Scott, 2009; Rauschecker and Tian, 2000; Recanzone, 2011; Tardif et al., 2008; Zatorre et al., 2004), as it was earlier observed in the visual system (Mishkin et al., 1983). Using anatomical and functional approaches in primates (including humans), these processing pathways were defined as dorsal and ventral auditory streams. It has been suggested that the dorsal stream involves connections between posterior temporal auditory regions and the pre-frontal cortex via relays in parietal regions, while the ventral stream is supposed to involve projections from primary auditory sensory areas to anterior temporal and frontal cortices. As in the visual system, the dorsal parietal pathway was hypothesized to underlie auditory spatial processing ('where') while the ventral temporal pathway was supposed to be recruited by the identification of complex patterns or objects ('what') (Kaas and Hackett, 1999; Rauschecker and Scott, 2009). However, recent studies go beyond this view, and assume that more complex processes can be supported by these two pathways, as fronto-temporo-parietal interactions are known to be recruited by multiple cognitive functions such as language, memory and even music processing.

Anatomical evidence

Separate anterior (ventral) and posterior (dorsal) projection streams (Kaas and Hackett, 1999) can be observed anatomically, using anatomical tract tracing in monkeys (Rauschecker et al., 1997; Romanski et al., 1999), and more recently, using Diffusion Tensor Imaging in humans, a noninvasive method for investigating white matter fiber tracts in vivo using MRI (Catani and Thiebaut de Schotten, 2008; Frey et al., 2008).

In monkeys, these long-range connections were described as 1) a ventral route connecting rostral (lateral) auditory belt and parabelt areas, rostral (lateral) superior temporal gyrus and ventral prefrontal cortex, and 2) a dorsal route connecting caudal (medial) auditory belt and parabelt areas, and parietal and dorso-lateral prefrontal cortices (Romanski et al., 1999). In humans, diffusion studies have suggested that there may also be ventral and dorsal white matter tracts connecting auditory areas to frontal and parietal cortices (Catani and Thiebaut de Schotten, 2008; Croxson et al., 2005; Frey et al., 2008). Diffusion studies thus revealed that these pathways (Figure 6) are supported by: 1) for the ventral pathway:

- the uncinate fasciculus, that connects the anterior temporal lobe with the medial and lateral orbitofrontal cortex (Catani and Thiebaut de Schotten, 2008);
- the extreme capsule fiber tracts that connect posterior temporal cortex to inferior frontal regions (Croxson et al., 2005);

2) for the dorsal pathway:

- the arcuate fasciculus that connects the perisylvian cortex of the frontal, parietal, and temporal lobes (Catani and Mesulam, 2008)
- the superior longitudinal fasciculus II and III connecting parietal cortex to frontal cortex via two divisions, one in the inferior frontal gyrus (III), the other in the pre motor area (II) (Frey et al., 2008);

It is relevant to note, however, that although DTI provides information on the trajectories of associational white matter tracts, this method does little to illuminate the origin, terminal distribution, or transmission properties of these pathways. It is thus necessary to combine these anatomical explorations to functional measures in humans.

Functional evidence

Similarly to the visual system (Rauschecker, 1998), distinct auditory processing in ventral and dorsal pathways were observed with both electrophysiological research in animals and brain imaging studies in Humans. Electrophysiological studies in monkeys have suggested that these ventral/dorsal routes support two functionally distinct 'what' and 'where' auditory processes (Kaas and Hackett, 1999). In this model, anterior temporal and inferior frontal regions would process auditory ''object'' information whereas posterior temporal areas, parietal and frontal cortices would be involved in the

spatial processing of both stationary and moving sounds (Bendor and Wang, 2005; Poremba et al., 2003; Rauschecker and Tian, 2004; Rauschecker et al., 1995; Tian and Rauschecker, 2004).

In humans, numerous lesion (Clarke et al., 2005; Clarke et al., 2000; Clarke et al., 2002), functional neuroimaging (Alain et al., 2001; Baumgart et al., 1999; Bushara et al., 1999; Griffiths and Green, 1999; Griffiths et al., 2000a; Lewis et al., 2000; Maeder et al., 2001; Warren et al., 2002; Zatorre et al., 2002; Zatorre et al., 2004), and electrophysiological (Anourova et al., 2001; Ducommun et al., 2002; Kaiser and Bertrand, 2003; Lutzenberger et al., 2002; Xiang et al., 2002) studies have supported this hypothesis.
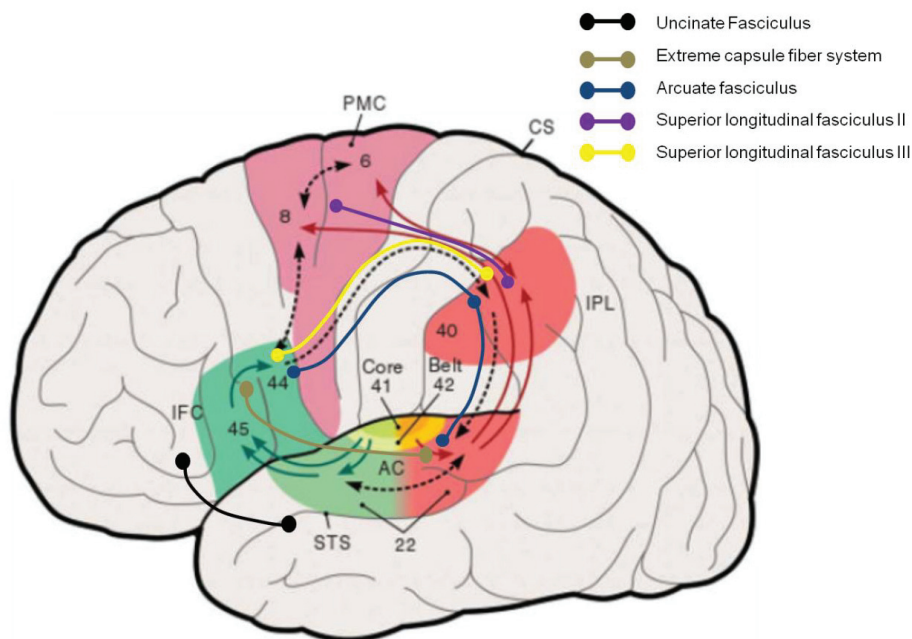
Figure 6. Dual auditory processing schema in the human cortex. This schema presents the functional and anatomical pathways that are supposed to support auditory perception. Antero-ventral stream is represented in green and postero-dorsal stream in red. Both streams originate from the auditory belt. The postero-dorsal stream interfaces with frontal areas and pivots around inferior parietal cortex. AC, auditory cortex; STS, superior temporal sulcus; IFC, inferior frontal cortex, PMC, premotor cortex; IPL, inferior parietal lobule; CS, central sulcus. Numbers correspond to Brodmann areas. Adapted from Rauschecker and Scott (2009).

These seminal studies have explored the predictions made by the dual-route model, notably showing that parietal lobe structures are involved in auditory spatial processing (Alain et al., 2001; Bushara et al., 1999; Maeder et al., 2001). However, the functional role of this dorsal pathway remains controversial (Rauschecker and Scott, 2009). Further studies have suggested that parietal areas do not seem to be specifically involved in auditory spatial processing and could also be activated during pitch tasks, speech or memory processing (Foster et al., 2013; Foster and Zatorre, 2010; Janata et al., 2002b; Maeder et al., 2001; Rauschecker and Scott, 2009; Zatorre et al., 2004; Zatorre et al., 1994). In a recent neuroimaging study, Foster and Zatorre (2010) observed that mental manipulation of pitch level

or temporal organization of melodies results in functional activation in the human intraparietal sulcus. They proposed that this parietal recruitment could be associated to an involvement of this structure in transforming auditory stimulus representations in a nonspatial context (see also Foster et al., 2013). This hypothesis is in line with the visual domain, where visuospatial mental transformations, including mental rotation, were shown to depend on a network including the posterior parietal cortex (Zacks, 2008). This area was described as a multimodal association region that receives visual, auditory and tactile information, and is connected with frontal working memory areas and motor planning centers (Catani and Thiebaut de Schotten, 2008; Frey et al., 2008).

In a related vein, the specific role of the ventral stream for object identification has been challenged by several studies. Alain et al. (2001) reported that several sites in the superior temporal gyrus (STG) are recruited, along with frontal and occipital activity, when subjects have to actively judge a pitch change as compared to a spatial change assuming that this pitch specificity may not reflect object-related processing per se. Maeder et al. (2001) reported a similar contrast in a task that involved recognizing a specific class of sounds (animal sounds) from among a complex background. Results implicated the anterior STG, but parietal, frontal, para-hippocampal, insular, and occipital cortices were also recruited. Although both these studies converge in reporting that anterior STG areas may be involved in auditory object processing, the areas reported are quite different, perhaps because these studies involved different and complex active tasks, which likely recruited a variety of cognitive mechanisms.

To conclude, the specific role of each stream is still unclear and controversial. It is now fairly admitted that the superior temporal gyrus, inferior frontal regions and parietal cortex are implicated in the processing of speech, music and language, and ignoring these reports and assigning an exclusively spatial function to the dorsal auditory stream and object identification function to the ventral stream would be not appropriate (Rauschecker and Scott, 2009). It is therefore essential to consider how the different substructures of the auditory cortex in the temporal lobe, the temporo-parietal junction and the inferior frontal cortex are involved in multiple cognitive functions such as speech, working memory, music, and whether we can assign a common computational function to the ventral and dorsal streams. Within the general framework of this PhD work, the following section will present the current knowledge about the specificities of auditory perceptual system for the special case of pitch perception.

1.5. Psychological and neurobiological underpinnings of pitch perception

As described above (page 18), many sound characteristics are related to auditory perceptual experience, resulting in a number of sound perceptual attributes or dimensions (pitch, loudness, duration, localization and numerous dimensions of timbre). Within these dimensions, pitch is the one that has received most interest in auditory perception research. Pitch is primarily related to frequency and is, according to de Cheveigné: "*the stuff of which music is made*" (de Cheveigné, 2010). In the following section, we briefly review the current knowledge in cognitive neurosciences about how pitch is encoded in the auditory system.

1.5.1. Definition of pitch

Among all auditory perceptual attributes, pitch seems to have a special status as it has been demonstrated that it plays a crucial role in music and speech perception, and in analyzing complex auditory scenes (Bregman, 1990). Concerning auditory scene analysis, pitch is an important component of both primitive and schema-driven analyses (see page 17) because it constitutes a cue to perceptually segregate sound sequences arriving simultaneously at the ears (see simultaneous organization of sound analysis). Pitch is also a central component of the Western musical system, which uses it as form-bearing dimension at the expense of other acoustic dimensions (Dowling and Harwood, 1986). The question of how pitch is encoded by the auditory perceptual system has a long and controversial history, with debates on the subject going back to the time of Ohm (Ohm, 1843) and Helmholtz (Helmholtz, 1863).

Extraction of pitch in the auditory system

There is still no consensus for the way by which pitch attribute is extracted, as different cues can provide frequency information to the perceptual system. According to Fourier's theorem, harmonic complex tones can be decomposed in a series of sinusoids (also referred to as partials) that occur at discrete frequencies corresponding to multiple integers of a fundamental frequency (F0). Because sounds at different frequencies are known to recruit different sets of auditory nerve fibers, cochlear encoding is often described as a set of bandpass filters. The frequency distance between the partials of a harmonic complex tone is equal to F0, nevertheless, it is relevant to note that the width of the auditory filters increases with frequency. While the lowest partials of complex sounds are sufficiently distant one from another to recruit different auditory filters (resolved harmonics), the higher order partials interact in auditory filters (unresolved harmonics). Resolved harmonics provide thus information about the F0 of the complex sound both in a place (or spectral code, associated to the tonotopic organization of the system) code and in a time code (associated to the firing pattern of the fibers (temporal dynamics). Unresolved harmonics provide temporal information about F0 (output of filters receive several unresolved harmonics beats at F0), but do not provide fine spectral information. Several studies have aimed to explain how pitch is encoded by the auditory system. These

Chapter I: Psychological and neurobiological underpinnings of pitch perception investigations were motivated by a simple question: how can the perceptual system associate the same pitch for sounds having very different spectra (which therefore recruit different groups of neurons in the auditory nerve and in the auditory nuclei of the brainstem). For example, a complex tone with an F0 of 200 Hz and harmonics 1-5 (200-1000 Hz) evokes the same pitch as a complex tone with an F0 of 200 Hz and harmonics 6-10 (1200-2000 Hz). In the latter case, while the component with a frequency equal to its F0 is absent, the perceived pitch is unaffected. In psychoacoustics, two alternative theories have aimed to explain pitch encoding: pattern matching models and auto-correlation models
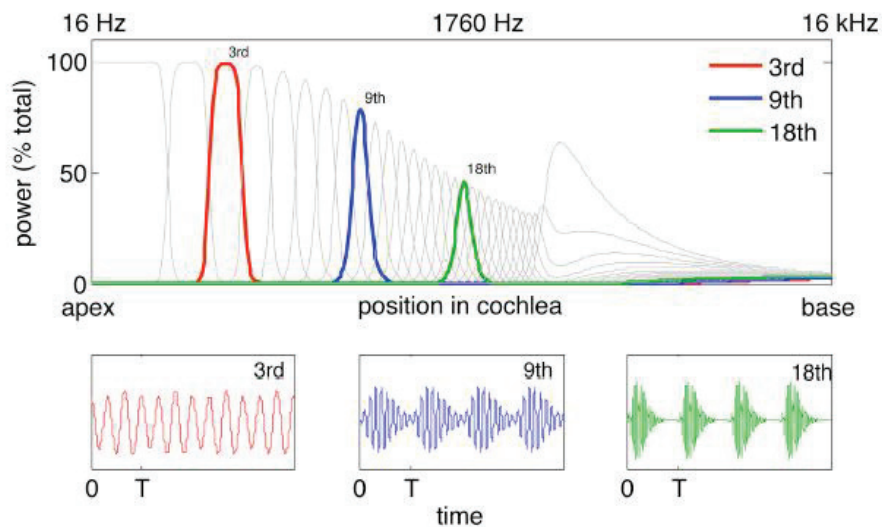
Figure 7: Upper panel: percentage of the signal power that can be attributed to a single harmonic at the output of a cochlear filter. This percentage decreases with increasing frequency (from the apex to the base of the cochlea). Lower panel: waveform at the output of cochlear filters tuned to the 3rd (left), 9th (middle) and 18th (right) harmonics. For low rank harmonics, the partial is resolved and the ouput of the filter is quasi-sinusoidal. For high rank harmonics, several partials interact within an auditory filter and the response is no longer sinusoidal. The response, however, contains information about the F0 of the sound in its temporal envelope (from de Cheveigné, 2010).

Pattern matching and autocorrelation models

Because the cochlea filters sounds into frequency bands with limited resolution, it has been proposed that the frequency and time domain can in principle provide distinct information allowing pitch encoding. Models of pitch perception have thus focused on the relative importance of temporal (timing) vs. spectral (frequency-based, place) pitch cues. Temporal theories assume that pitch is computed from the timing of action potentials that are phase-locked to the sound waveform, while spectral theories suggest that pitch is represented as the place of activation across the tonotopic map (i.e. the anatomical organization of frequency selectivity).

Along these lines, two competing explanations of pitch prevail: 1) pattern matching models (Goldstein, 1973; Shamma and Klein, 2000; Terhardt, 1974); and 2) autocorrelation models (Licklider,

1951; Meddis and Hewitt, 1991a, 1991b). As pitch is evoked by periodic stimuli (and its value depends on their period), these two approaches can be considered as different ways to extract the stimulus period.

Pattern matching models assume that the fundamental partial of the sound is the necessary correlate of pitch (note that it may nevertheless be absent if other parts of the pattern (harmonics often associated with it) are present). These models assume that the brain has stored 'templates' of the spectrum of harmonic tones, and that the perceived pitch corresponds to that of the template that best fits the stimulus. Pattern matching theories are appropriate considering the sharp peaks produced in the spectrum by resolved harmonics. However these models have difficulty to explain that complex tones comprising only unresolved harmonics evoke a weak pitch.

The autocorrelation models consider that pitch could be derived from a single temporal mechanism that acts on timing information from all frequency channels, regardless of resolvability. The basic idea is that the incoming neural signal is delayed in some pathways. Coincidence detectors are then recruited to compare the original and the delayed versions of the signal to compute the period at which the pattern is repeating. However, note that the estimation of the period of unresolved harmonics is too good (as unresolved harmonics are known to evoke a weak pitch).

Pattern-matching and auto-correlation models have many variants. Adaptations of both types of models have been claimed to be sufficient on their own to account for the perception of both resolved and unresolved harmonic complexes (Bernstein and Oxenham, 2005; Meddis and O'Mard, 1997).

Context effect

In the environment, sounds are rarely presented in isolation. It is thus relevant to consider the effect of context on auditory perception. These context effects were observed, for example, with discrimination and detection tasks. As an example, two context effects (among others) were described using psychoacoustic methods: 1) Forward masking effect (Plomp, 1964) 2) Backward masking effect (Massaro and Loftus, 1996). The forward masking effect corresponds to an increase in the detection threshold of a target sound when it is preceded by another masker sound. The rationale is that the amount of masking obtained is greater when the masker and the probe recruit the same auditory filter. Backward masking corresponds to an increase in discrimination threshold between two sounds when the second sound is temporally contiguous with the first sound. This effect was related to the observation that, for normal listeners, several hundreds of milliseconds (up to 300 ms) are needed for the construction of an 'internal representation' of the pitch of a sound (whatever the duration of that sound) (Demany and Semal, 2008).

Pitch processing has also been shown to be susceptible to context effects. Additionally to the forward and backward masking effects, the investigation of pitch memory has revealed an interference effect. When participants are required to discriminate two sounds that could differ in pitch, pitch discrimination abilities have been shown to be impaired when interfering tones were presented

between the two target sounds. More interestingly, this effect was more important when those distracting tones have a pitch that is similar to that of the target tones than when the distracting tones have very different frequency (Deutsch, 1974).

Physiological recordings also provide strong evidence in favor of context effects. The psychophysical forward masking described above was shown to result partly from a phenomenon of adaptation at different levels in the auditory pathway. Adaptation corresponds to the fact that electrical neural responses to a given stimulus are reduced by the repetition of the presentation this stimulus (Smith, 1977). Forward masking is thought to rely on the same mechanism because adaptation to a masker sound reduces the response to a subsequent probe.

Overall, these context effects show the importance of considering responses to both isolated sounds and sound sequences in order to get a full understanding of low level and high level aspects of auditory perception. Note that for the specific case of pitch in music, context effects will be presented in the following section.

1.5.2. Perception of pitch sequences

Local and global levels

Pitch in sequences play a major role in music perception that requires a complex perceptual analysis. In cognitive psychology, music perception was proposed to be supported by a modular system (Peretz and Coltheart, 2003). This model assumes that music perception is supported by two modules: 1) a pitch organization module, and 2) a time organization module. Evidence for distinct pitch and time modules comes from neuropsychological studies where double dissociations were observed; that is, time or pitch processing can selectively be impaired after brain damage, while leaving the processing of the other dimension unimpaired (Peretz, 1990). For both pitch and time organization, Peretz and Coltheart (2003) assumed that they are supported by separate local and global levels. For time organization, these levels correspond to rhythm (local) and meter (global) (Dowling and Harwood, 1986; Peretz, 1990).

For the pitch organization, two structural levels were defined: 1) The interval code that contains the exact pitch changes between notes, corresponding to the local level (Dowling, 1982; Peretz, 1990) and; 2) The contour, which corresponds to the pattern of ups and downs, corresponding to the global level. According to Dowling (1982) this two features in pitch organization are assumed to be processed independently. This dissociation between local and global levels was suggested after the observation that listeners could recognize melodies when all the intervals are changed to 1 semitone, so that the contour remained unchanged. Contour has been further investigated in studies using same/different paradigms in which pairs of novel pitch sequences could differ at a local level, where contour was preserved, or at a global level, where the overall contour was violated (Dowling and Fujitani, 1971). Moreover, transposition was found to impair interval perception but not contour perception. These results suggest that a hierarchical structure allows the processing of pitch sequences, where the global contour-processing system serves as a framework for decoding local stimuli.

Musical context

In addition to the global and local levels of pitch perception, another key aspect of pitch sequences in music is the tonal context. Pieces of western music are written in a given tonal context that corresponds to a key. A key is a set of scales in which the melody can evolve. A scale is a subset of pitches that is defined by the interval relation between them. In western system, scales generally consist of seven notes and repeat at the octave. Minor and major scales, for example, consist of the same notes except one. Contour is an especially important feature of melodies when a relationship between the melody and the scale schema has not been thoroughly established. This is the case in immediate recognition of short melodies, or in discrimination of atonal melodies (composed with the set of pitches which not respect the scales).

Interestingly, several studies have shown that previously acquired knowledge (in musician and non musicians) about tonal structure (tonal context) could influence pitch perception (Bigand and Tillmann, 2005; Francès, 1958; Marmel et al., 2008; Warrier and Zatorre, 2002). In non musician listeners, Marmel et al (2008) have shown the influence of tonal relatedness either on rating, priming of melodic comparison tasks. For example, for the discrimination task, they showed that participants' discrimination performance for small mistunings was better when the to-be-compared tones were tonally related to a melodic context. Similar effects were observed in Warrier and Zatorre (2002) study which suggests that the presence of the tonal structure improves pitch perception as it gives more cues that facilitate pitch extraction.

1.5.3. Pitch centers in the human brain

In addition to the particularities of pitch processing observed at the behavioral level with psychoacoustic methods and cognitive psychology approaches, several studies have investigated for pitch processing in the auditory system using neurophysiological and neuroimaging methods in humans. The determination of pitch-specific regions in the brain was first done in studies aiming to characterize the functional architecture of the auditory system. As presented in previous sections (page 21), encoding the pitch of a pure tone is, in theory, relatively easy. For a given sound level, the pitch can be derived as the place of maximal activation along the tonotopic map as early in the nervous system as the cochlea. Auditory information then reaches the primary auditory cortex via the tonotopic pathway which has its last relay in the Medial Geniculate Body of the thalamus.

Tonotopic organization

At the cortical level, several studies in monkeys and humans have shown a tonotopic organization in the different subparts of the auditory cortex (core, belt, parabelt). Studies in monkeys have revealed that in the core region, three fields could be defined thanks to their different systematic

Chapter I: Psychological and neurobiological underpinnings of pitch perception

representations of the cochlea (Kaas and Hackett, 2000) : 1) a most medial (caudal) AI, 2) a more lateral (rostral) field called R, and 3) an even more lateral temporal field, RT (Rostral Temporal).

In medial AI, neurons were shown to be mostly activated by high-frequency tones, whereas neurons in lateral AI were best activated by low-frequency tones. The tonotopic organization of R was shown to be the inverse from that of AI: low frequencies being represented medially, and higher frequencies being represented laterally. The tonotopic organization of RT is not well determined, but it appears that high frequencies are represented medially and low frequencies, laterally (Kaas and Hackett, 2000). More recently, a clear confirmation of this organization comes from high-field fMRI studies of tonotopic organization in the macaque (Petkov et al., 2006) revealing frequency gradients in the primary auditory region that matched the expected size and locations of AI, R, and RT, as known from previous invasive studies.

This tonotopic organization was also observed in monkeys' belt using microelectrode recordings. Although neurons in the belt showed less important responses to tones than those in the core, they elicited sufficient responses to show tonotopic gradients in belt areas, these responses being parallel to those of the adjacent core areas. Nevertheless, it has been shown that neurons in the lateral belt respond better to narrow-band noises than to pure tones (Kaas and Hackett, 2000). Thus, these divisions of the belt seem to be related to the connection patterns between the core and the belt. Similarly, tonotopic organization was difficult to demonstrate in the parabelt region, where neurons seem to depend more on belt inputs rather than thalamic inputs for auditory activation.
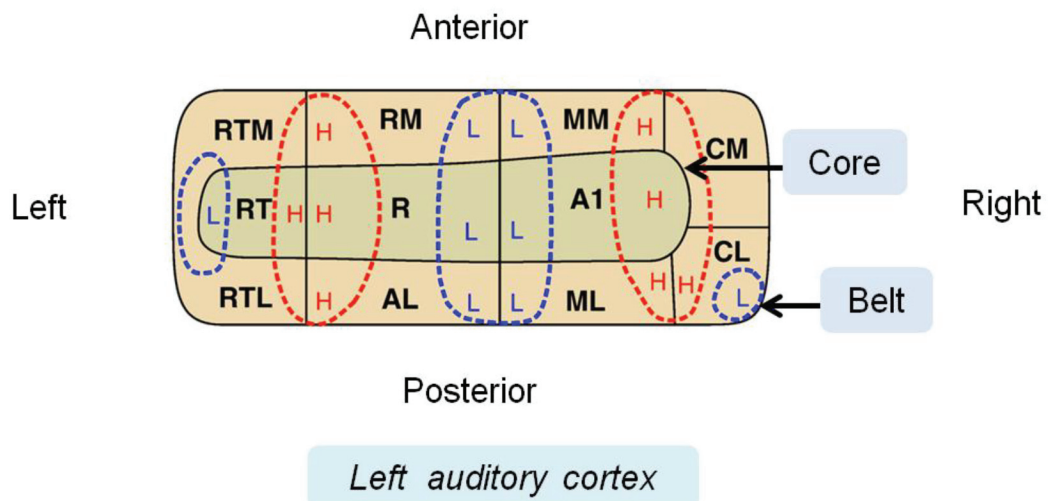
Figure 9: Functional areas identified in previous studies of non-human primate auditory cortex. Tonotopic gradients are represented by high (H) and low (L) endpoints. Contiguous high areas are marked in red and contiguous low areas in blue. C, Caudal, M, Medial; R, Rostral; L, Lateral; T, Temporal, A1, primary auditory region. From Humphries et al., 2010.

Chapter I: Psychological and neurobiological underpinnings of pitch perception

In humans, evidences for distinct tonotopic fields in cortical auditory regions were first demonstrated using MEG recordings (Romani et al., 1982). This study showed that the cortical activity in response to auditory stimuli was increased in depth with increasing tone frequency. The concept of tonotopic organization of the primary auditory region was then confirmed by numerous studies using electro-physiological approaches (Bertrand et al., 1991; Howard et al., 1996; Pantev et al., 1995; Pantev et al., 1996; Romani, 1986). Notably, Howard and al. (1996) performed single-unit recordings in humans to show tonotopic organization in the primary auditory region which was similar to that observed in monkeys (high-frequency sound recruiting the medial part of AI, and low frequency sound recruiting a more lateral part). Other evidence comes from the study of Pantev et al. (1995) who showed using MEG recordings that the sources of the Pa wave elicited a tonotopic organization in the transverse temporal gyrus of Heschl while the N1 component shows mirror organization in the planum temporale.

Later fMRI studies have improved our understanding of the tonotopic organization of human auditory cortex by exploring cortical activations not only for isolated extreme frequencies, but also for intermediate frequencies, allowing the identification of frequency gradients (spatial gradients with selective frequencies in auditory areas) (Formisano et al., 2003; Langers et al., 2007; Schonwiesner et al., 2002; Talavage et al., 2004; Upadhyay et al., 2007; Woods et al., 2009; Yang et al., 2000). Schonwiesner et al. (2002) found differences in activation between anterior and posterior Heschl's gyrus (as well as medial and lateral differences) as function of frequency. Formisano (2003) showed in line with seminal studies in monkeys (Kaas and Hackett, 2000), that two mirror-symmetric frequency gradients (high–low–high) extend along the axis of Heschl's gyrus. In Talavage et al. (2004) three consistent gradients were reported, but none of which clearly follow the long axis of Heschl's gyrus. Finally, Langers et al. (2007) reported a single high-to-low gradient extending from posterior medial (caudal) to anterior lateral (rostral) auditory areas.
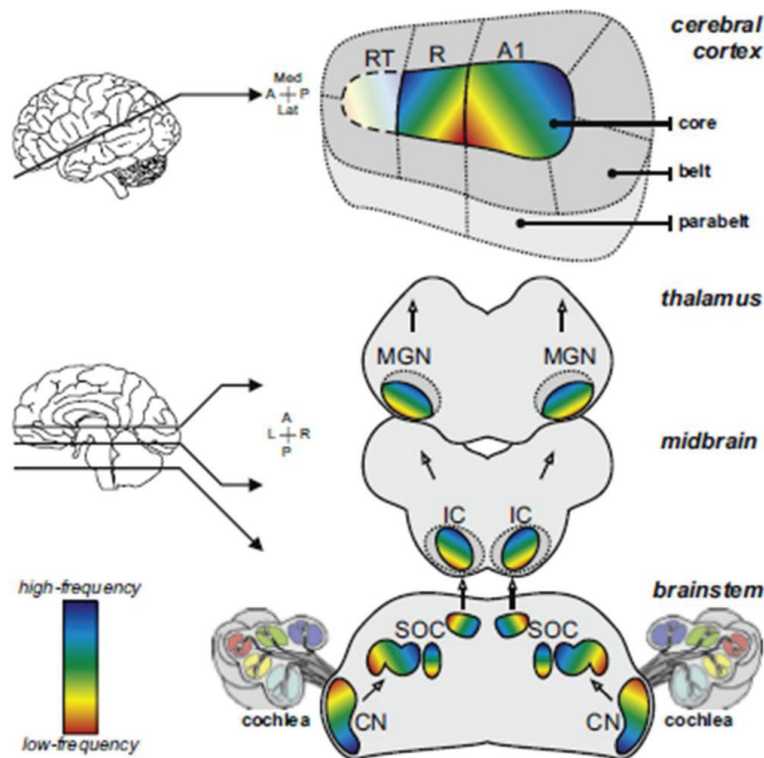
Figure 10. The central auditory pathway in humans, from Saenz and Langers (2013). All nuclei that form part of the classical lemniscal auditory pathway are tonotopically organised. These include various subdivisions of the cochlear nucleus (CN), superior olivary complex (SOC), inferior colliculus (IC), and medial geniculate nucleus (MGN). In the auditory cerebral cortex in the superior part of the temporal lobe, expected divisions of core, belt, and parabelt are based on the non-human primate model of auditory cortical organisation. Human neuroimaging consistently shows at least two primary tonotopic gradients ("high-low-high") in the auditory cortex, homologous to primary fields A1 and rostral field R in the monkey cortex. In some primate studies, a third rostrotemporal field RT is delineated, but neuroimaging evidence for a similar field in humans is sparse.

More recently, Humphries et al. (2010) proposed an alternative tonotopic organization. In this study, posterior and anterior to Heschl's gyrus were sensitive to higher frequencies, whereas an elongated zone on Heschl's gyrus responded more specifically to lower frequencies. They conclude that while the high-low-high frequency gradients remains present, the tonotopic progressions in the Heschl's gyrus is running perpendicularly across this areas, instead of running in parallel along it (see Figure 10). This data received subsequent supports from very recent studies (Da Costa et al., 2011; Herdener et al., 2013; Langers and van Dijk, 2012; Saenz and Langers, 2013) that similarly found gradients running across the Heschl's gyrus with fields AI and R on the medial and lateral parts of the Heschl's gyrus respectively. It has then been suggested that the direction of these tonotopic gradients is oriented diagonally, resulting in an angled V-shaped pair of frequency progressions (Saenz and Langers, 2013) that is consistent with the anterior-to-posterior V-shaped tonotopic axis in primates (Kaas and Hackett, 2000).

Regarding secondary auditory regions, belt areas have also been shown to elicit some degree of tonotopicity. On the planum temporale, an additional low-frequency focus has been observed (Humphries et al., 2010; Langers and van Dijk, 2012; Talavage et al., 2004). Moreover, it has been suggested (Saenz and Langers, 2013) that additional tonotopic gradients exist in the extreme lateral part of the superior temporal gyrus, close to the superior temporal sulcus and middle temporal gyrus. Further investigations are thus needed to understand clearly the tonotopic organization of secondary auditory areas.

Functional response to complex pitch

While the tonotopic organization of the auditory cortex starts to get well defined for pure tones, for complex sounds, there is no simple relation between pitch and frequency composition. It has thus been argued that for complex sounds, more sophisticated neural computations are necessary to determine sound periodicity (see Walker et al., 2011 for review). Considering that auditory nerve fibers are phase-locked to sounds properties, their firing patterns in response to a periodic sound are in consequence periodic. As described above, the frequency of a pure tone can be derived as the periodicity (or first-order autocorrelation see section 1.5.1 pages 35 to 36) of spikes observed across responsive auditory nerve fibers. For complex sounds, F0 is encoded within the all-order autocorrelation of spikes across fibers. It could thus be assumed that maps of stimulus periodicity exist in AI. Since pure tones have a well-defined periodicity and evoke a pitch that is comparable to that of their tone complex counterparts, these periodicity maps should be independent of sound spectrum and thus lie along the tonotopic map. However, while the idea of a periodotopic map in auditory cortex is emergent, the experimental evidence for such an arrangement, and its contribution to pitch perception, remains inconclusive (Walker et al., 2011).

Several studies have, nevertheless, tried to determine the cortical regions that could be specifically recruited by complex sounds. First of all, while it has been shown that the majority of sounds recruit the medial part of the Heschl's gyrus (primary auditory cortex), the activation of the adjacent auditory areas seems to vary as a function of sound spectro-temporal complexity. It is now well admitted that when the sound complexity increases, the extent of the recruitment of secondary auditory region increases, notably in the planum temporale. This assumption was confirmed by studies showing that pure tones activate primarily the core areas, whereas belt areas are more recruited by complex sounds, such as narrow-band noise bursts (Wessinger et al., 2001).

Later, several studies have tried to distinguish the relative contributions of auditory cortical fields in pitch processing. These studies have used low-pitch salience (unresolved harmonics), and a high-pass filter to minimize the spectral ripple that occurs in the stimulus so that the stimulus can be compared with a control noise with the same pass-band. This manipulation allows demonstrating responses that cannot be explained by changes in the time-averaged spectrum. Patterson et al. (2002) used fMRI to measure the blood oxygenation response in human listeners during the presentation of

Chapter I: Psychological and neurobiological underpinnings of pitch perception

iterated rippled noise and broadband noise bursts. They found that the lateral Heschl's gyrus was more strongly activated by pitch-evoking stimuli than by the aperiodic noises (Patterson et al., 2002).

Using intra cranial recordings in an epileptic patient, the lateral Heschl's Gyrus was shown to elicit a stronger response to iterated rippled noises than to aperiodic noises. Interestingly, note that the opposite result was found in medial portion of the Heschl's gyrus (Schonwiesner and Zatorre, 2008). More recently, two studies (Griffiths et al., 2010; Sedley et al., 2012) using intracranial and MEG recordings confirmed that maximal responses to iterated rippled noise are observed in lateral Heschl's gyrus, which have been interpreted as pitch mappings. Similarly, higher activation in lateral Heschl's gyrus has been observed when contrasting resolved harmonics (with high pitch salience) with unresolved harmonics (low pitch salience) that were presented in the same pass band (Penagos et al., 2004).

Overall these studies in Humans suggest a regional specialization for pitch in the lateral and posterior superior temporal plane. Finally, cortical specificity for pitch processing was investigated as a function of hemispheric recruitment. These researches were based on the observation of impaired pitch discriminations abilities in the case of right auditory cortical injuries, while these impairments were not observed after damage in the left hemisphere (Divenyi and Robinson, 1989; Johnsrude et al., 2000; Stewart et al., 2006).These studies have proposed a right lateralization of pitch processing in the human brain. This hypothesis has received further supports from fMRI studies in healthy listeners showing that activation in the right planum temporale is correlated with the size of frequency shifts presented between successive tones. Interestingly this correlation was not observed in the left auditory cortical activity (Hyde et al., 2008).

Functional response to pitch patterns and music processing

As described in the previous section natural stimuli from the environment are rarely composed of isolated pitch, but rather contain pitch in sequences such as speech or music. For the specific case of music processing, musical patterns are defined by pitch variation across sequences or in the form of glides. Neuropsychological studies in brain damaged patients have shown that deficits in the processing of pitch patterns are most often associated with lesions in the planum temporale, in the parieto-temporal junction, and in the supra-temporal plane anterior to Heschl's gyrus and in the anterior Superior Temporal Gyrus (STG) (Stewart et al., 2006). These studies thus suggested that lateral and posterior parts of auditory cortex (that are well known to play an important role in pitch representation during the early stages of pitch analyses) might also be recruited for more complex processing, such as the extraction of melodic features (i.e., melodic contour, interval etc.). By contrasting local and global pitch processing with fMRI, a recent study (Stewart et al., 2008) found that in comparison to global processing, additional activations were observed for local processing in planum temporale and posterior superior temporal sulcus (pSTS) thus supporting the hypothesis of hierarchical organization between global and local features in pitch sequences.

Additionally, the effect of varying pitch in sequences has been shown to activate a distributed set of cortical areas in the superior temporal lobe (Griffiths, 1999; Griffiths et al., 1999; Janata et al., 2002a; Janata et al., 2002b; Koelsch et al., 2009; Maess et al., 2001; Peretz and Zatorre, 2005; Schulze et al., 2011a; Schulze et al., 2011b; Tillmann et al., 2003; Tillmann et al., 2006; Zatorre et al., 2002; Zatorre et al., 1994). These studies revealed that in addition to the auditory cortex, pitch processing in the human brain involves the contribution of distant brain regions, notably frontal and parietal areas. In particular, inferior frontal areas have been shown to be strongly recruited during encoding melodic information (Griffiths et al., 1999; Overath et al., 2007; Zatorre et al., 1994). These studies will not be described in this chapter, but note that the specific network related to encoding of melodies in the specific case of short term memory processing will be presented with further details in Chapter II.

Conclusion

This chapter presents a brief overview about our current knowledge concerning of how auditory perceptual systems encode sound information. We have presented here above several steps of the processing from the sound wave to cortical processing by insisting on pitch processing and by presenting the conclusions of psychoacoustic, psychological and neuroimaging studies. In brief, these studies showed that several hierarchical organizations could be observed at each step of the auditory processing and that specific subparts of the human auditory system are related to pitch processing. However, while most of the previous research in this domain has largely focused on the role of specific cortical regions involved in pitch perception, memory, or other functions, it still remains an open question about how these regions work together in terms of network activity. As suggested by Sporn (2013), contemporary neuroscience emphasizes the idea that higher-order aspects of cognition emerge from the dynamic interactions between brain areas. To explore how connectivity patterns could underlie brain functioning and to understand complex cognitive processes requires considerable efforts at mapping the networks of the human brain with both functional and anatomical approaches (Sporns, 2013).

Chapter II Psychological and Neurobiological Underpinnings of Auditory Short-Term Memory

2.1. Memory

Once the external information has been processed by the perceptual systems, higher order mechanisms are setting up in the human brain to allow the elaboration of multiple complex cognitive functions. Among several complex processes such as emotion, attention, learning, motor programming etc., memory has received an extensive interest in the domain of cognitive psychology and neuroscience. According to Endel Tulving,

"*Memory is the ability of living organisms to acquire, retain and use information or knowledge*" (Tulving, 1987)

This ability has been described as playing an important role in governing our everyday life activities, by driving the construction of our identity, our knowledge and our perceptual abilities. While it is widely admitted that memory plays a ubiquitous role in human cognition, its underpinnings were historically debated along two main theories that have either considered it as an homogeneous entity, or as an arrangement of multiple different systems.

2.1.1. Models of memory

The first scientific investigation of memory was historically performed by Hermann Ebbinghaus (Ebbinghaus, 1885/1913), who has tested whether he could acquire and recall, at various periods of time (up to 31 days), new information in form of series of nonsense syllables. Ebbinghaus noted that he often had a ''*first fleeting grasp ... of the series in moments of special concentration*'' (p. 33). By this comment, he introduced the notion of 'immediate memory', but assumed that his *fleeting grasp* did not ensure that the series had been sufficiently memorized to allow its recall later on. A distinction between this 'immediate memory' supposed to be a limited-capacity system, and another type of large-capacity system was proposed by William James (James, 1890). On the basis of introspective analysis, he proposed that different types of memory exist: 1) a limited storage he called primary memory *"[that] comes to us as belonging to the rearward portion of the present space of time, and not to the genuine past"*;

and 2) an unlimited storage referred as secondary memory: *"[An object of secondary memory, by contrast,]. . . is one which has been absent from consciousness altogether, and now revives anew. It is brought back, recalled, fished up, so to speak, from a reservoir in which, with countless other objects, it lay buried and lost from view"* (James, 1890, pp. 646–647).

Later on, Drachman and Arbit (1966) have proposed an evolution of the concepts of primary and second memory by introducing the labelings of "short-term memory" and "long-term memory" (Drachman and Arbit, 1966):

"Short-term memory . . . deals only with subspan memoranda, evanescently, as long as the subject's attention is directed towards the memorandum . . . By contrast, "long-term" memory (storage) deals both with supraspan memoranda held for long or short intervals and with subspan memoranda recalled following the redirection of attention." (p. 59).

This dualistic view of memory has been later supported by different neuropsychological studies in brain-damaged patients. These studies have suggested that as a function of the location of the lesion, specific behavioral impairments were observed either in short or long-term memory processing. In 1966, Milner has reported that patients who had damages in temporal lobes and hippocampus showed a general impairment in the capacity of learning and remembering new material (verbal or visual). However, these patients seemed to present normal short-term memory abilities as measured by the digit span task (the capacity to hear and immediately repeat back a unfamiliar sequence of numbers) (Milner, 1966). In contrast, the opposite pattern of impairments were observed in another study (Shallice and Warrington, 1970): patients with damage in the perisylvian region of the left hemisphere showed near normal long-term memory abilities, but strong impairments in short-term memory capacities (digit spans limited to one or two items). Based on these observations, the dominant hypothesis in the literature proposed that memory might be supported by a short-term store and a long-term store. In a related vein, Atkinson & Shiffrin (1968) proposed a model of memory based on the duration for which the representation stays in the system (Figure 10). In this model, external information flows from the environment through a series of brief sensory memories, into a limited capacity short-term store (Atkinson and Shiffrin, 1968). They assumed that when the informatiopn is retained long enough in the short-term store, it could then be transferred to a long-term store.

Since these seminal studies, research in cognitive psychology has rather consistently considered memory as an arrangement of multiple systems. However, note that an alternative theory has been proposed by numerous researchers, who have considered memory more as a unitary (or monistic) system (Nairne, 1992; Neath, 1998; Ranganath and Blumenfeld, 2005). These studies have tried to define a single set of principles that could capture all different forms of memory, from the very immediate to the very long-term. However, although there are substantial differences between dualistic and monistic conception of memory, there have been few experimental efforts to directly contrast the two types of models (Ranganath and Blumenfeld, 2005). In the present PhD work instead of considering memory as a homogeneous entity, we will adopt the hypothesis of memory as an arrangement of multiple systems(Winkler and Cowan, 2005). This is in line with previous studies investigating auditory memory for music and language processing that have been based on this concept of multiple forms of memory.
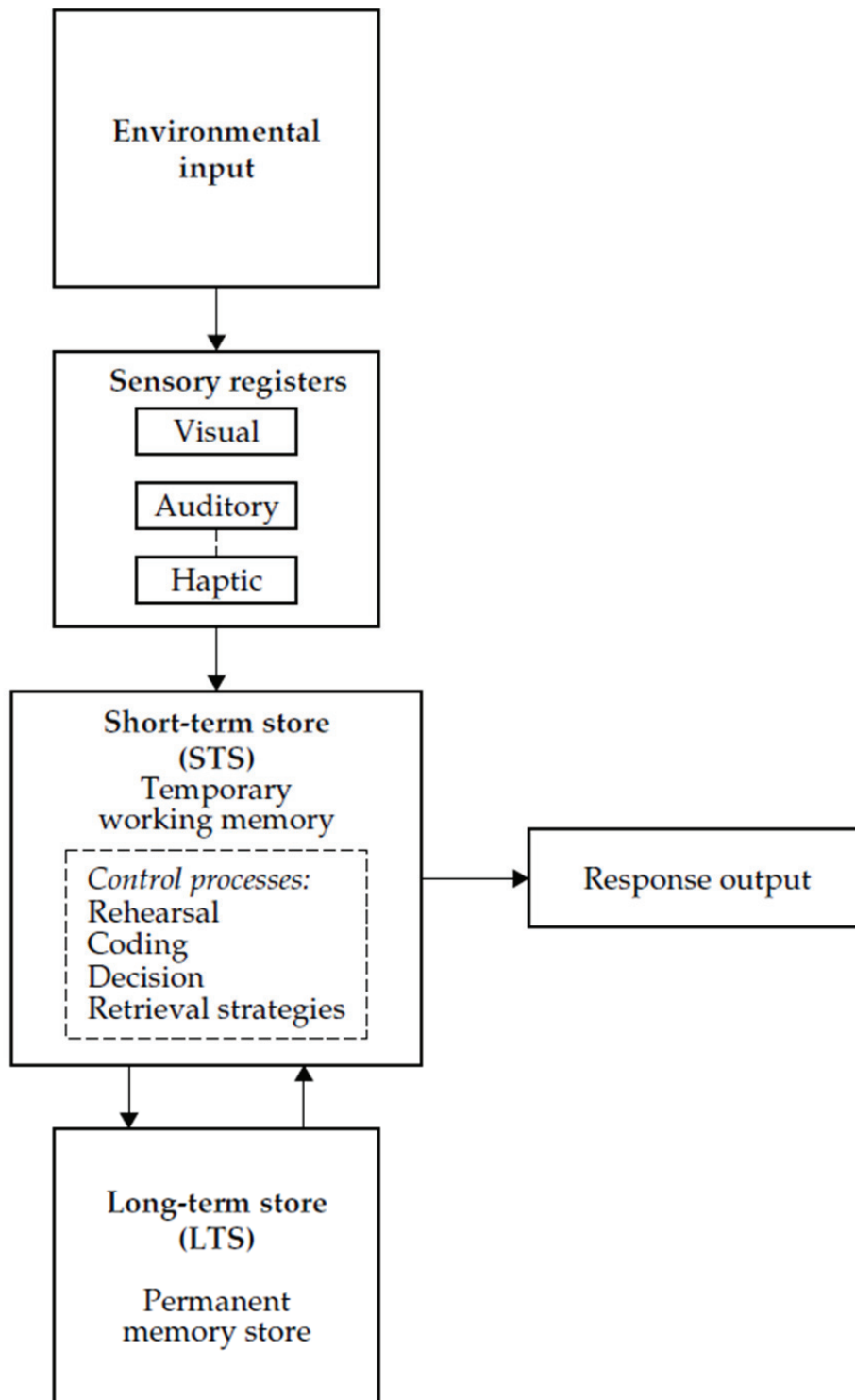
Figure 10: The model of Human memory from Atkinson & Shiffrin (1968). This model of memory comprises a succession of storage systems in which information flows from the environment, into a series of temporary sensory buffers, before being passed on to a limited capacity short-term memory store, which then feeds long-term memory. Atkinson and Shiffrin (1968) proposed that this short-term system acts as a working memory, controlling the flow of information into and out of long-term memory, and playing a crucial role in learning and in cognition more generally.

Since Atkinson & Shiffrin (1968), multiple forms of memory were defined: 1) sensory memory, 2) short-term and working memory 3) long-term memory.

First, sensory memory (see also echoic memory for auditory processing (see section 1.3.3, pages 18 to 19) allows the retention of information for a very short time period (<10 sec.) (Cowan, 1984). This type of very brief memory has often been related to the processing of information in the perceptual system. Second, short-term and working memories correspond to the 'immediate memory' or 'primary memory' described by James (1890). Short-term memory refers to the active retention of information for brief periods of time (maintenance), while working memory refers to the simultaneous storage and processing of information (maintenance and manipulation). The distinction between working and short-term memories (see below page 50, 58) is still under debate (Cowan, 2008; D'Esposito, 2007), however there is a general consensus to consider these memories as dedicated systems that aim to maintain, manipulate and store information over a short-term period (several seconds) (Baddeley, 2003b, 2010; Cowan, 2008). Finally, long-term memory has been described as a large capacity system allowing the storage of information for a long period of time (from hour to years). Long-term memory was classified into several sub-components such as declarative memory (memory for facts, events) and non-declarative memory (or procedural memory dedicated to perceptual or motor skills) (Cohen and Squire, 1980). The precise definitions of each of these memory systems are still under debate (Cowan, 2008; Schulze and Koelsch, 2012).

2.1.2. *Stages of memory and methods of investigation*

By exploring the different processing steps of memory, studies in cognitive psychology have often separated out three aspects seeming essential for any memory system: 1) Encoding, the processes whereby information is registered; 2) Storage or Retention, corresponding to the maintenance of information over time; and 3) Retrieval, which refers to the accessing of the information by recall or recognition.

Once the information has been processed by the perceptual systems, it has to be encoded into a form that the cognitive system can process. This phase corresponds to the encoding of the information (note that this step could share or contain several mechanisms as those involved in of sensory encoding (Demany and Semal, 2008)). Encoding is typically studied in psychology by varying the way that the material is processed during learning and by modifying the nature of the to-be-encoded material. For example, it is well known that the encoding of a specific characteristic of a stimulus (shallow encoding) leads to a much poorer subsequent recall (or recognition) than processing it in terms of meaning (deeper encoding) (Baddeley, 1976).

After the encoding step, the information has to be stored by the system (storage or retention). Retention is usually measured through the rate of loss of information (forgetting). Interestingly, retention capacities have been shown to vary as a function of the nature of memory stores, in other words, it varied as a function of how long the memory lasts for (duration), how much can be stored at any time (capacity) and what kind of information is held.

Finally, retrieval refers to getting information out of storage. Methods of investigation of memory retrieval include two principal aspects 1) recall, in which case the participant is required to reproduce the stimulus items, and 2) recognition, requiring the participant to determine during the recognition phase whether items have been presented or not during the encoding phase (yes/no recognition), or to choose the previously presented item from a set of two or more alternatives (forced-choice recognition).

Serial recall requiring participants to repeat a presented list in correct serial order is one retrieval methods. A simpler method is the free recall for which participants are required to reproduce as many items as possible in any order they wish. Note that, when the recall phase of a given series of items occurs immediately after the encoding phase, the probability of an item being recalled correctly is dependent on its serial position in the presentation. In the case of serial recall, recall starts very accurately, decreases throughout the list, and then improves towards the end of the list. One traditional interpretation of this phenomenon is that this pattern is indicative of both short-term and long-term memory systems. According to this view, the recency effect reflects the recruitment of a short-term store in which the most recently presented items remain at the time of recall. Similarly, the primacy effect corresponds to the preservation of items that had sufficient time to activate associated semantic knowledge in long-term memory. In addition to the primacy and recency effects, in case of long duration between the encoding and the recalling phases (from minutes to days), free recall has been shown to be dependent on variables such as the imageability, frequency and semantic associability of the items (thus showing an interaction between short-term and long-term memory).

The second method of investigation of retrieval, the recognition method, is more dependent on the participant's response strategy, in particular his degree of caution. For example, in a task for which participants are required to determine (categorize) if an item was presented during the encoding phase (yes/no), if a given participant responds 'yes' to every stimulus, his/her categorization scores would not reflect any memory processing. Note, however, that there are a number of analyses dealing with different degrees of caution among participants. A correction can be applied by deducting the proportion of false alarms from the proportion hit score. An alternative and more complex way is to use signal detection theory, which yields two measures, one representing the hypothetical strength of the memory trace (sensitivity index as measures by d'), and the other the response criterion (or degree of caution as measured by response bias c) employed by a given participant (Lockhart, 2000).

The investigation of the behavioral and cerebral mechanisms of encoding, storage and retrieval aspects of memory processing has allowed improving the definitions and classifications of the different memory sub-systems (sensory, short-term/working, long-term). The present chapter will not present all of them, but will describe in more details the psychological concepts and the cerebral correlates of short-term memory and working memory (also by presenting the potential differences between them) in the specific case of auditory processing.

It should be noted, however, that the different types of memory are not independent. Sensory and long-term memories are well known to interact with both short-term and working memory (Demany and Semal, 2008; Jeneson and Squire, 2012). Sensory memory can play a crucial role in the encoding of auditory information in short-term memory. This hypothesis is based on the fact that sounds produce physical entities that are extended in time and that once a sound end, it remains as memory traces (see Chapter I, page 18) that affect the encoding of future sounds. In this view, it is reasonable to state that, at least in the auditory domain, sensory memory and encoding of information in short-term memory are deeply interrelated (Demany and Semal, 2008).

In a related vein, and as shortly introduced above, long-term memory can also play a role in the processing of short-term memory and working memory. Numerous studies have shown that the use of familiar material (such as familiar spatial patterns or familiar verbal material) leads to better short-term memory performance than does unfamiliar material (i.e. non-sense words) (Bor et al., 2003; Savage et al., 2001). The underlying hypotheis here is that previously acquired knowledge can improve perception, encoding and retention of information in short-term or working memory. The interactions between different types of memory are important and will be presented in further details in the section 2.2 of the present chapter. Finally, note that the classification of different subsystems of memory we have adopted in the present manuscript is based on the duration of information storage, from very short (sensory memory) to very long time periods (long-term memory). It is, however, relevant to consider that other classifications of memory have been proposed, especially concerning the dichotomy between explicit and implicit information processing. The following sections will focus on explicit memory. The definition of implicit processing, and its differences with explicit processing in the special case of auditory perception and memory will be presented in further details in Chapter III.

2.2. Working & Short-Term Memory

The terms of short-term memory and working memory have not been used consistently in the literature (Baddeley, 2012; Cowan, 2008). *Short-term memory* has been used to refer to the simple temporary storage of information (maintenance), and working memory to refer to the maintenance and the manipulation of information (Baddeley, 1986; Engle et al., 1999). In the following sections, we will present the concepts of working memory and short-term memory through the view of cognitive psychology.

2.2.1. *Working memory*

Working memory is a term referring to memory as it is used to plan and carry out behavior (Miller et al., 1960). This term has become highly influential in cognitive psychology since Baddeley and Hitch (1974) proposed that the single short-term store of Atkinson & Shiffrin (1968) does not take

into account all kinds of 'immediate' memory. According to Baddeley and Hitch (1974), Atkinson & Shiffrin model encountered at least two problems (see also Baddeley 2003, 2010).

First, it does not consider that holding an item in short-term store does not guarantee efficient learning. For example the processing of a word in terms of its perceptual appearance is much less effective for subsequent learning than encoding the material on the basis of its meaning. According to this encoding effect, Craik & Lockhart (1972) have suggested that the probability of recall (or recognition) is directly related to the 'depth' to which an item has been processed. In this view, recall of a word list following semantic processing (considered as a 'deep encoding process as it requires participants to make a judgment about the meaning of the item) is far better than recall following superficial processing, such as coding of form or size (Craik and Lockhart, 1972).

The second problem of the Atkinson and Shiffrin's (1968) model is that it assumes that the short-term store is an obligatory step for long-term learning. This conclusion was challenged by the study of patients showing a very specific short-term memory deficit, but exhibiting normal long-term learning (Shallice and Warrington, 1970). According to Atkinson and Shiffrin's (1968) model, in the absence of an adequate short-term memory, information should be rapidly lost and hence such patients should not be able to learn. Furthermore, if Atkinson and Shiffrin's (1968) short-term store did indeed function as a working memory, patients with impaired short-term memory should be cognitively seriously impaired. However, while patients described in Shallice and Warrington (1970) study showed short-term memory impairments, they seemed to exhibit near normal cognitive abilities (comprehending, reasoning, etc.).

Baddeley and Hitch (1974) have thus attempted to explore this paradox by studying the effect of disrupting short-term memory on complex cognitive abilities, such as reasoning, comprehending and learning, in typical individuals. By combining such cognitive tasks with a concurrent activity, namely remembering and repeating back sequences of digits, the performance on the cognitive task should be progressively disrupted as the length of the digit sequence increases in the short-term memory task. Participants showed indeed a significant effect, with speed of performance in the complex tasks declining with increasing length of digit-sequences. However, even with very long digit sequences, error rates in the complex tasks were low and unchanged. Based on this later observation they proposed that different subsystems in the brain support these different mechanisms (while still sharing some mechanisms).

They thus proposed that the concept of a simple unitary short-term store could be replaced by a more complex system, labeled as working memory (Baddeley and Hitch, 1974). The construction of this model was based on the observation of distinct memory resources for the processing of visuospatial and verbal material. The claim of domain specificity was supported by findings of selective interference of concurrent visuospatial and verbal activities on the performance of same-modality tasks (Baddeley et al., 1975a; Baddeley and Lieberman, 1980; Brandimonte et al., 1992), and of selective impairment of either verbal or visuospatial short-term memory in neurological patients (Vallar and Baddeley, 1984). These observations have thus led to the highly influential model

(Baddeley, 1986) of working memory in which 3 components interact (Figure 11): 1) the phonological component, 2) the visual-spatial component and 3) the central executive corresponding to attention-related processes.
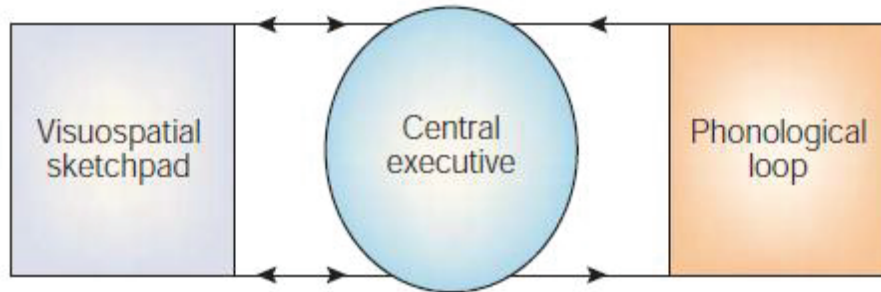
Figure 11: The three-component model of working memory form Baddeley and Hitch (1974). This model comprises a control system, the central executive, and two subsidiary storage systems, the visuospatial sketchpad and the phonological loop.

The phonological loop

The phonological loop is assumed to consist in a short-term store, which is subject of rapid temporal decay (Baddeley, 1992, 2003b, 2010, 2012; Baddeley, 1986; Vallar, 2006; Vallar and Baddeley, 1984; Vallar and Papagno, 2002). It assumes that linguistic inputs have obligatory access to a phonological store. This component was subdivided into two parts: 1) a passive storage component (phonological storage), which is assumed to maintain auditory or speech-based information for a couple of seconds (Baddeley, 1992, 2010); and 2) an active rehearsal mechanism, which is supposed to maintain information for longer time periods thanks to articulatory rehearsal processes (such as subvocal speech) (see Figure 12).

Multiple behavioral evidences have support this view, notably by showing that the storage of auditory and verbal information might be supported by the establishment of phonological codes in the brain.

1) The first evidence was labeled as the *phonological similarity effect*, which was related to the observation of impaired recall for verbal items that share similar phonological structure (Baddeley, 1966), whereas visual or semantic similarity showed very little effects.

2) The second evidence came from the *irrelevant speech effect*: it has been demonstrated that listening to irrelevant speech during serial recall tasks resulted in poorer retention (Colle and Welsh, 1976), than listening of pulsed noise (Salame and Baddeley, 1987). This observation suggests that irrelevant speech effect might be due to the obligatory access of phonological (but not other auditory) information during retention of the information.

3) The third evidence was named the *word-length effect*: it refers to the fact that an improved memory span and a superior recognition accuracy are observed for short words in comparison to long words (Baddeley et al., 1975b), thus implying that the phonological information is recruited for retention. This effect suggested a rapid temporal decay of the phonological representations in the short-term store.

4) The fourth evidence of the phonological component has led to the description of the *articulatory rehearsal process*, which has been evidenced by the effect of articulatory suppression. When typical individuals are prevented from rehearsing to-be-remembered items, by being required to repeat continuously an irrelevant sound (e.g., /the/) or naming the months of the year, performance is impaired. This suggests that articulatory rehearsal processes are a necessary mechanism of the phonological storage.
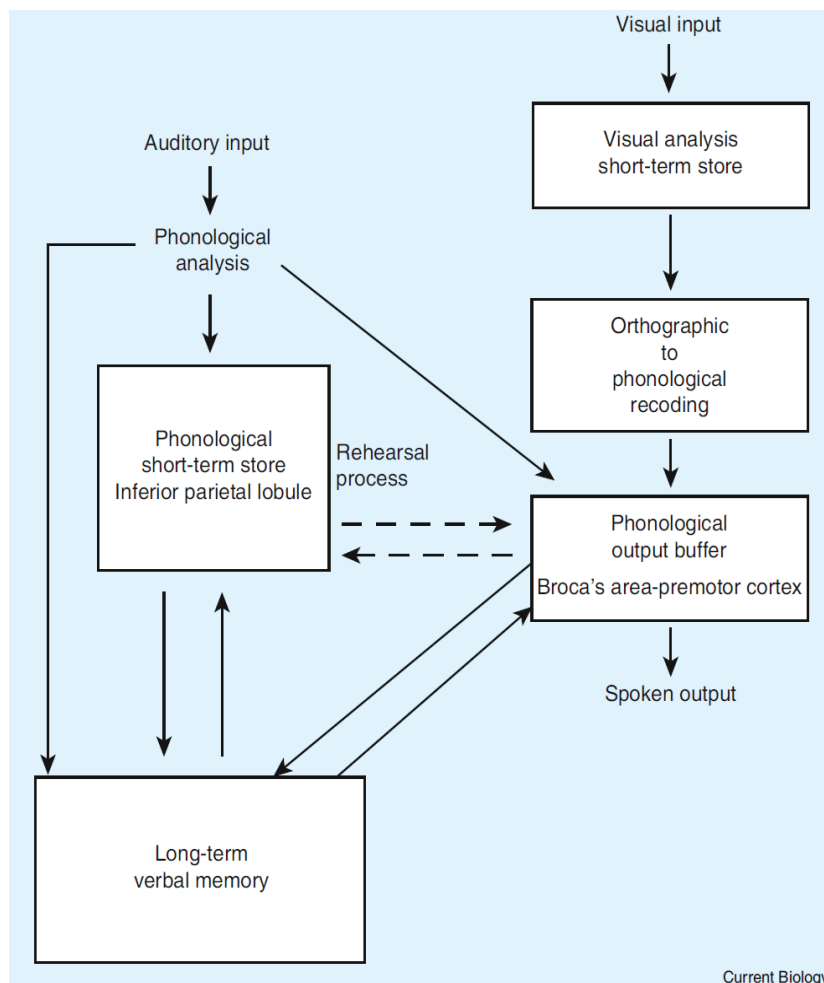
Figure 12. Detailed presentation of the phonological loop model based on both behavioral and neuropsychological evidence (Vallar, 2006).

Additionally to these four aspects, two supplementary characteristics of the phonological loop have been described (Baddeley, 2000).

5) The transfer of information between codes: it had been shown that typical individuals can subvocally rehearse visually presented items, thereby transferring the information from a visual code

to an auditory code (Figure 12). It should be noted that articulatory suppression prevents this code transfer as it removes the effect of phonological similarity for visually presented items, but not for auditorily presented items, as these are automatically registered in the phonological store.

6) The last evidence for the existence of a phonological loop was provided by the study of neuropsychological patients who showed specific deficits either of the phonological store or of rehearsal processes. Baddeley (2000) has pointed out that aphasic patients affected by dyspraxia are unable to set up the speech motor codes necessary for articulatory rehearsal processes. In contrast dysarthric patients, exhibiting peripheral speech difficulties, show a normal capacity for rehearsal, thus suggesting that articulatory rehearsal is the central rehearsal code of the phonological loop.

Taken together, all these effects have suggested that during the retention of verbal material, information is maintained by both a passive phonological store and an active articulatory rehearsal process, comparable to subvocal speech. Over the last decades, this model has proved capable of accommodating a great deal of experimental evidence from typical adult participants, children, and neuropsychological patients (Baddeley, 1997; Gathercole and Baddeley, 1993).

The visuospatial sketchpad

The visuospatial sketchpad is assumed to allow the active storage and manipulation of visual and spatial information. It has been suggested that memory of object shapes involves the retention of spatial arrangements and visual features and that spatial information can sometimes be encoded in the form of static visual patterns (Pickering et al., 2001; Smyth and Pendleton, 1989). These observations have led to embedding these two kinds of information (i.e., spatial and visual) into one component in the model of working memory. However, a breakdown between these two processes was then proposed based on an accumulation of evidence indicating dissociations between memory for spatial movements and for visual patterns (Klauer and Zhao, 2004). Note that this component will not be described in detail in the present chapter as the concerns of visuospatial short-term and working memory are far from the central themes of this PhD work.

Central executive

The third component of the model, the central executive, is assumed to constitute an attentional controller for the visuospatial and phonological subsidiary systems of working memory. It was described as playing an important role in dividing and switching attention, together with a role in connecting working memory with long-term memory (Baddeley, 1996). However, whether this system involves a single unitary controller, or an arrangement of multiple executive processes has not been clearly defined yet (Baddeley, 2010).

Baddeley (1996) suggested to define the central executive in parallel with the supervisory attentional system proposed by Norman & Shallice (1986). This supervisory attentional system proposed that attentional control of action is supported by well-learned habits and schemas, guided by environmental cues (Norman and Shallice, 1986). This concept includes a wide range of potentially

separable executive processes, that have shown to be altered in patients with frontal lobe damage (Baddeley, 1996). One way to decipher the different components of the central executive is to investigate its cerebral correlates with neuroimaging methods in humans. Note that neuroimaging studies exploring this component will be presented with more details in the following sections (see section 2.4.1.1 pages 64 to 70).

A fourth component of working memory: the episodic buffer

More recently, an additional component has been added to the Baddeley and Hitch (1974) model of working memory: the *episodic buffer* (Figure 13). According to Baddeley (2000), the episodic buffer is assumed to be capable of storing information in a multi-dimensional code. It thus provides a temporary interface between the subsidiary systems (the phonological loop and the visuospatial sketchpad) and long-term memory. It is assumed to be controlled by the central executive and it serves as a modeling space that is separate from long-term memory, but which forms an important stage in long-term episodic learning (Baddeley, 2000).

Note that the buffer had been added in the model following the observation of some inconsistencies of data with the original Baddeley and Hitch's (1974) model. For example, patients showing impaired short-term phonological memory exhibiting auditory memory span of only one digit, could even recall about four digits after visual presentation. Based on this evidence (among several other studies), the episodic buffer is assumed to explain short-term memory processing of features that did not match the other stores (particularly semantic information in memory).

Current Biology
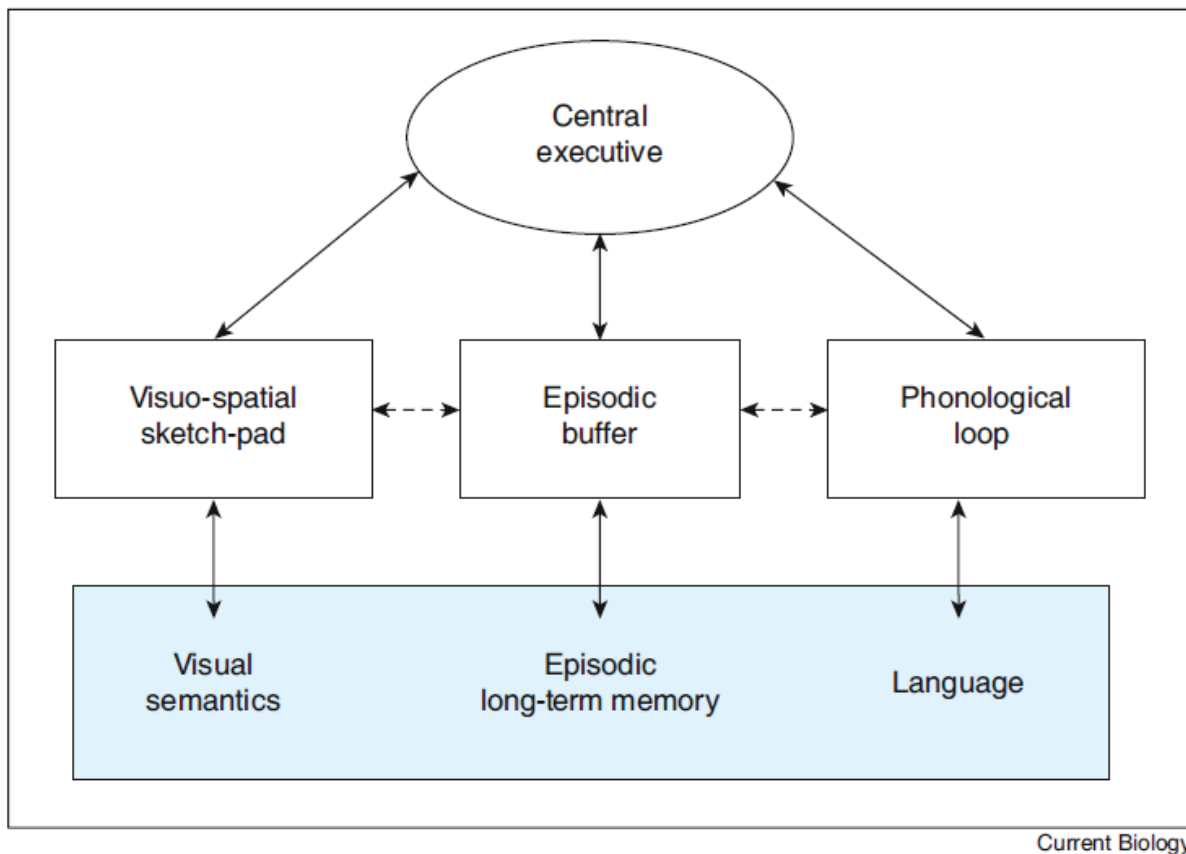
Figure 13. A further development of the working memory model of Baddeley and Hitch from Baddeley (2010)

Another model of working memory

An alternative model differing from Baddeley and Hitch's (1974) proposal was introduced by Cowan (1988). In his view, short-term memory refers to a subset of information in long-term memory that is temporarily activated above some threshold. Once the information has been 'activated', it rapidly returns to an inactive state unless it is refreshed by an attentional focus which is limited in capacity. In this model, working memory includes short-term memory as well as the attentional processes used to keep some short-term memory contents in an activated state (Figure 14). In contrast to Baddeley and Hitch, working memory does not rely on specialized dedicated storage buffers.
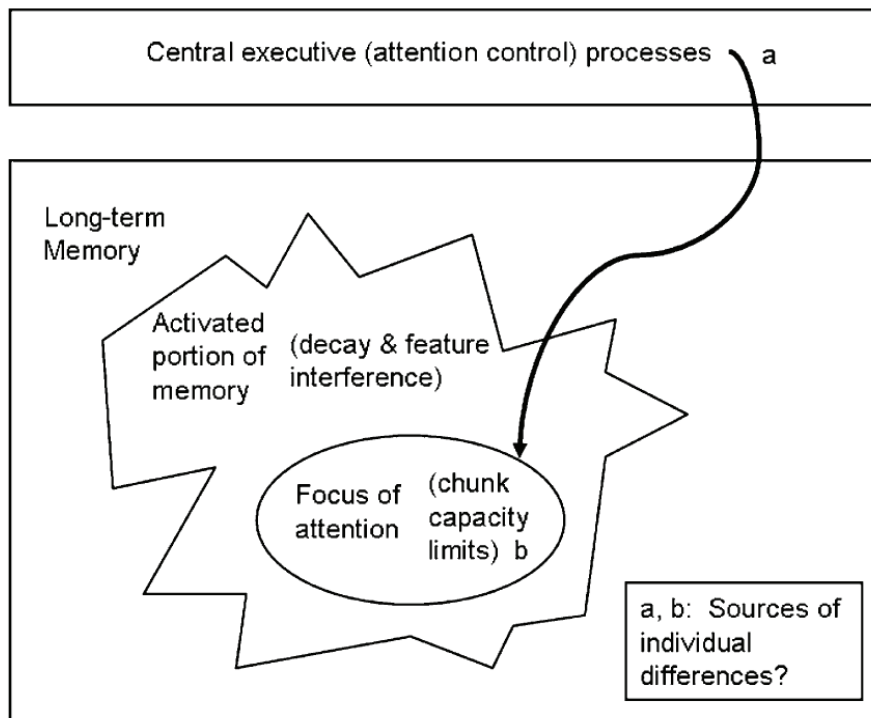
Figure 14. A depiction of the theoretical modeling framework of memory from Cowan (2008). Cowan proposes that the 'contents of working memory' are not maintained within dedicated storage buffers, but rather are simply the subset of information that is within the focus of attention at a given time. In this model working memory comes from hierarchically arranged faculties comprising long-term memory, the subset of working long-term memory that is currently activated and the subset of activated memory that is the focus of attention.

Later on, Engle et al. (1999) proposed an adaption of this model by stipulating that working memory could consists in " *(a) a store in the form of long-term memory traces active above threshold, (b) processes for achieving and maintaining that activation, and (c) controlled attention"* (p. 104). In this model, working memory thus refers only to the attention-related aspects of short-term memory (Engle, 2002).

As presented in the previous section, working memory has received an extensive interest in cognitive psychology and neurosciences over the last decades. Studies in patients and in typical individuals have allowed defining several sub-components of the system that were supposed to support the retention and the manipulation of information for a short-term period. However, and as mentioned above, the distinction between short-term memory and working memory seems to not be clearly defined in contemporary research in cognitive psychology and neuroscience. Notable several studies have used working memory (maintenance and manipulation) for tasks involving only maintenance processes (short-term memory)(Schulze and Koelsch, 2012). To present provide a more precise definition, the following section will thus briefly introduce the concept of short-term memory.

2.2.2. Short-term memory

The distinction between short-term memory and working memory is not well defined, this being due to several inconsistencies in the use of each term over the past decades. While seminal studies have considered that working memory is embedded in short-term stores (Atkinson and Shiffrin, 1968), other models have rather considered short-term memory as a sub-component of working memory (see Cowan, 2008).

Historically, short-term memory referred to the 'primary memory' of James (1890). However, this definition has been rapidly considered as too vague as primary memory could also be associated to sensory memory (Cowan, 2008). Later on, Atkinson and Shiffrin (1968) defined short-term memory as a short-term store that allows the temporary storage of a limited quantity of information. As described in the previous section, Baddeley and Hitch (1974) have opted for a multi-component system that could not be reduced to a unitary short-term store and used the term working memory to describe that entire system and its subcomponents. One option in the distinction of short-term and working memory is considering that the subsidiary systems (phonological loop, visuospatial sketchpad) proposed by Baddeley and Hitch (1974) could consist in short-term storage components (Cowan, 1988), these component being manipulated by the central executive processes. In this view, working memory thus corresponds to storage and processing of the information, and includes short-term stores and other processing mechanisms that help to make use of short-term memory.

However, whether short-term memory is a subcomponent of working memory or whether it encompasses working memory is still under debate. A unifying view (and simplest view) is to consider these two mechanisms as *distinct* systems. This hypothesis has received support in studies that have investigated participants' performance in tasks that were supposed to recruit short-term memory (storage-only) and working memory (storage and processing) (Cantor et al., 1991; Engle et al., 1999; Kail and Hall, 2001; Kane et al., 2004). These studies were based on two hypotheses. First, as short-term memory is supposed to be independent of working memory, tasks that recruit short-term memory processing should be distinguishable from tasks that measure working memory. Supporting evidence for this hypothesis could be found in Cantor et al (1991) and Engle et al. (1999) who have required participants to perform simple digit and word span tasks, while performing more complex span tasks. In the complex span tasks participants were required to read sets of sentences and concurrently remember the last word from each sentence in the set. Factor analyses revealed that the simple span tasks and complex reading span tasks were loaded on separate factors that were interpreted to reflect short-term and working memory, respectively.

The second hypothesis concerns the potential links between short-term and working memory and higher order cognitive mechanisms defined as cognitive aptitudes (such as reasoning, problem solving, and reading). As these 'cognitive aptitudes' are supposed to recruit more attentional and complex processes in comparison to a simple recall task, performance on these tasks should be predicted more accurately by working memory than by short-term memory.

Many investigators have compared the extent to which short-term memory and working memory could predict 'aptitudes' on cognitive tasks. In the studies by Engle and colleagues described above (Cantor et al., 1991; Engle et al., 1999), only working memory accounted for significant variance in verbal scores and in measures of fluid intelligence. According to Cowan (2008), as measures of working memory have been found to correlate with intellectual aptitudes (and especially fluid intelligence) better than measures of short-term memory, working and short-term memory are certainly separated. Taken together, these studies supported the hypothesis that working memory and short-term memory are distinct (but at some point interrelated) and that working memory plays a greater role than short-term memory in higher order cognitive processes.

The present PhD work is embedded in this general framework considering short-term memory and working memory are distinct. In this view, short-term memory refers to the storage of information and working memory reflects the use of measures that incorporate not only storage, but also processing. Along these lines, the following section will present the behavioral and cerebral correlates of short-term memory for auditory verbal and musical materials.

2.3.   Auditory Short-Term Memory for verbal and tonal material

The majority of research on auditory short-term memory has been carried out using verbal material, such as words, syllables or phonemes. More recently, several studies have investigated the memory mechanisms that support other kind of auditory material, such as music (tones, timbres, rhythms etc.) (Golubock and Janata, 2013; McKeown and Wellsted, 2009; Mercer and McKeown, 2010; Schulze and Tillmann, 2013). Auditory short-term memory for tones is a fundamental component of music perception and production, but also plays an important role in speech processing and comprehension. In the following section, we will describe in more details the current knowledge about the mechanisms that support auditory short-term memory for verbal and tonal material.

2.3.1.   Verbal short-term memory

It is now largely admitted that verbal short-term memory can be supported by the phonological loop proposed by Baddeley and Hitch (1974). As described above (see page 52), the phonological loop has been defined thanks to the observation of multiple behavioral effects (*the phonological similarity effect; the irrelevant speech effect; the word-length effect; the articulatory suppression effect*) leading to the conclusion that verbal material is stored in short-term memory through a transformation of the information into a phonological code. In addition to these mechanisms, the phonological loop has been demonstrated to be sensitive to timing effect. This effect was labeled the *temporal grouping effect*, which corresponds to the improvement of short-term memory performance when temporal pauses between to-be-remembered verbal items are presented during the encoding phase (Hitch et al., 1996).

Furthermore, it is also well admitted that verbal short-term memory does not operate in isolation, but within the context of a complex cognitive system. Several studies have investigated the potential impact of other processes on verbal short-term memory. The support available from knowledge stored in long-term memory is one such candidate process. Along these lines, several studies have shown the contribution of long-term knowledge on verbal short-term memory performance. Examples include the better recall of familiar words than nonsense syllables (Hulme et al., 1991), and the better retention of words in sentences than of unrelated words (Baddeley et al., 1987). This contribution of long-term knowledge was also demonstrated by numerous studies showing that better recall is observed for: 1) sound sequences presenting a higher probability of occurring in the lexicon, known as the phonotactic frequency effect (Pickering et al., 2001); 2) words that are more frequently used, corresponding to the word frequency effect (Hulme et al., 1997); 3) words for which it is easier to form a mental image, the imageability effect (Bourassa and Besner, 1994). In a related vein, it has been suggested that long-term knowledge could enhance the stability and quality of phonological representations themselves (Thorn et al., 2005).

Articulation rate is another mechanism found to be associated with verbal short-term memory (Cowan et al., 1998). It has been suggested that articulation rate may limit rehearsal in short-term memory, which is assumed to be a real-time process similar to subvocal speech. While these mechanisms start to get well understood for verbal information, it still remains an open question asking whether other inputs, such as music, are or not supported by the same component. One might suppose that non-verbal information is re-coded as verbal information to be processed by the phonological loop. Others have considered that additional specialized mechanisms are involved, thus introducing the concept of a 'tonal loop'. In the following section, we will consider this issue for the special case of musical (tonal) information.

2.3.2. *Tonal short-term memory*

The nature of memory for music has received increasing interest in cognitive psychology and neuroscience over the past decade (Tillmann et al., 2011b). Seminal studies have hypothesized that a specialized subsystem might be responsible for the temporary storage of tonal pitch (Deutsch, 1970b, 1975a; Deutsch, 1977). Deutsch explored the potential differential effect of interfering tones and spoken numbers on short-term memory performance for pitch (interfering sound between to-be compared pitches). While interfering tones were shown to importantly disrupt performance, the spoken numbers showed only little effect, thus leading to the hypothesis for the existence of a specific system supporting memory for pitch (Deutsch, 1970a, 1972, 1974, 1975b).

In order to challenge or support this view, Salame and Baddeley (1989) have explored whether the processing of linguistic items could be affected by the concurrent processing of musical stimuli and noise. While noise showed no detrimental effect, music significantly disrupted short-term memory performance in comparison with silence, although it was less disruptive than speech (Salame and Baddeley, 1989). More interestingly, this effect was shown to be sensitive to participants' task

training: For participants who had previously performed a separate memory experiment for an hour, only vocal music interfered with memory performance (but not instrumental). For less-trained participants, both vocal and instrumental music showed a detrimental effect. To discuss this point, Salame and Baddeley proposed that "*the short-term storage component of the articulatory loop system is not a general acoustic store protected by a filter, but is rather a detection and storage system that has evolved for processing speech-like material*" (p. 121). To explain why instrumental music interferes with speech, Salame and Baddeley proposed that "*the fact that we can hear and remember sounds that are very unlike speech means that there must be some-presumably additional-form of acoustic storage system capable of dealing with such material*" (p. 121). This hypothesis suggests thus that speech and music are process in a common working memory system, but might recruit different sub-systems (or loop).

Tonal loop

These studies have thus led to a model of working memory that contains an additional subsidiary system (or short-term store), referred as the tonal loop. This component was supposed to allow the processing and storage of musical stimuli separated from the verbal or visuospatial components (Berz, 1995; Pechmann and Mohr, 1992). However, this model has been weakened by Semal et al. (1996) who proposed that the absence of interference between the tones and the interpolated verbal material in Deutsch's 1970b study, could be explained by the fact that the frequency relations between the two materials were not controlled. By performing the same experiment while controlling the pitch of the interfering stimuli (words or tones) they showed that pitch similarity between the two materials exhibited a more important effect on short-term memory performance than the material itself (verbal or tonal) (Semal et al., 1996). They thus proposed that pitch for both verbal and tonal stimuli are processed in the same system.

In a related vein, Chan (Chan et al., 1998) reported that musical training increases verbal working memory performance, suggesting that overlapping mechanisms are underlying verbal and tonal short-term memory. Recently, Williamson and colleagues (2010) explored short-term recall for tones and letters and suggested that a comparable effect of the phonological similarity (observed in the verbal domain) could be transposed to the tonal material (Williamson et al., 2010a). Further support for similarities between auditory verbal and tonal short-term memory comes from an experiment using an articulatory suppression paradigm in musically trained participants (Schendel and Palmer, 2007). This study showed that musical (singing) and verbal suppression decreased recognition accuracy for both digit and tone sequences, thus indicating that the degree of separation between verbal or tonal short-term memory is quite small. Taken together, it might be considered that the nature of musical memory itself would be very similar or supported by the phonological loop of Baddeley's model with a short-term store and a control process based on inner speech, in this case, a musical inner speech (i.e., singing or humming).

Effect of musical expertise

As described for the processing of verbal material, a contribution of long-term knowledge in short-term processing of music might also exist. In turn, short-term memory for music may thus be crucially related to individual's musical expertise.

Pechmann and Mohr (1992) asked musically trained and untrained participants to compare sequences of two tones separated by an interval that was filled with tonal, verbal, or visual material. Results revealed that while musicians' performance was only affected by the interpolation of tones the performance of non musicians was also affected by verbal and visual items. The differences in terms of performance between the two groups were interpreted in a differential efficiency of the recruitment of the articulatory rehearsal mechanisms (or articulatory loop). As the articulatory loop is assumed to work rather automatically, that is, without the requirement of central resources, the authors argued that musician participants' performance was not affected by verbal and visual stimuli, because of this automatic functioning. In contrast, for untrained participants they hypothesized that the tonal rehearsal may represent a much more controlled process that requires additional allocation of attention. These findings have led to the assumption that musical training significantly affects short-term memory performance.

Differences in terms of musical training were also assessed by exploring short-term memory performance for tone sequences in musician and non musicians (Schulze and Koelsch, 2012). Notably, Schulze et al. (2012) have reported better maintenance for tones in musician than in non-musician listeners. Furthermore, by comparing tonal and atonal melodies in the two groups, they observed that both musician and non-musicians exhibited improved performance for tonal sequences in comparison to atonal sequences, but the effect was larger in musicians (Schulze et al., 2012). For musicians, this effect was considered as an additional evidence of the impact of long-term memory on short-term memory performance, because of their developed expertise in the processing of tonally structured material. The results of non-musicians were rather considered as a marker of their implicit knowledge of the tonal structure, acquired via mere exposure to music in everyday life, which could improve their short-term memory performance for structured material (Bigand and Poulin-Charronnat, 2006; Francès, 1958; Tillmann, 2005; Tillmann et al., 2000). This hypothesis has also received support from numerous studies (see Tillmann, 2005) notably by observing that non-musicians exhibit, good singing capacities without training (Dalla Bella et al., 2007) and by the exploration of music perception capacities in infants (Hannon and Trehub, 2005).

According to this observation, explicit and implicit knowledge of tonality could improve memory processing for musical material. In this view, the storage could be improved in short-term memory thanks to the processing of structures stored in long-term memory (Berz, 1995). Several studies have supported this view that memory for pitch, even for individual pitches, is dependent of the tonal context (Cuddy, 1971; Dewar et al., 1977; Marmel et al., 2008; Sloboda and Parker, 1985).

Role of contour and interval information in melody recognition

Investigating tonal context effects has allowed defining the impact of the different melodic components (scale, contour, interval, see Chapter I, page 37) on short-term memory performance. A series of experiments (Bartlett and Dowling, 1980; Dowling, 1978; Dowling and Fujitani, 1971) have suggested that contour information plays a particular role in melody recognition when the melodies are either novel, transposed, or tonally weak. Some of these studies have used the transposition effect (changing the scale while keeping intact contour and interval), which is supposed to require listeners to locate the new area of the tonal space occupied by the melody. Modifying either contour or intervals of the transposed melody in a melody comparison task has allowed exploring whether these melodic characteristics are dependent on the tonal context. While it is difficult to recognize the exact interval relationships when transposition took place, it is equally easy to detect changes in contours of melodies regardless of whether transposition occurred. Edworthy (1985) explored the role of controu and interval as a function of the strength of the tonal context by varying melodies' lengths (long melodies containing stronger tonal framework than short melodies). She showed that, on transposition, contour information is immediately precise for short melodies, but lost as melody length increases. In contrast, she showed that interval information is initially less precise, but more resistant to forgetting in longer melodies. In this view, it has been proposed that contour could be encoded independently from tonal context and that interval becomes more precise as a tonal framework is established (Edworthy, 1985). Providing converging data, Dowling (1991) manipulated systematically the tonal strength of melodies; that is, how strongly the tonal framework was established by the used tones. In a continuous running memory paradigm, memory for exact intervals was observed only for strong tonal melodies. Dowling (1991) has thus suggested that interval, contour, and tonality are not encoded independently, but rather form an integrated whole leading to a stronger memory trace for tonal melodies over atonal melodies (as observed in Schulze et al., 2012), with atonal melodies missing the potential integration in a tonal framework (Dowling, 1991).

All together, these studies have supported the hypothesis of a role of articulatory rehearsal processes in maintenance of tonal information, that are recruited differentially as a function of participants' previously acquired knowledge about tonal structure. However, whether short-term memory for tones relies or not upon a dedicated short-term store is not clearly defined yet. According to d'Esposito (2007), it is counterintuitive that all temporarily stored information (such as verbal, tonal, tactile, visual, spatial) would require specialized dedicated buffers or systems. Progress in this theoretical question has been made by investigating these mechanisms with neuroscience methods. Investigating the neural mechanisms underlying short-term and working memory allow shaping and modifying cognitive models. These issues will be presented in the following section.

2.4. Cerebral correlates of Auditory Short-Term Memory

The investigations of cerebral correlates of short-term and working memory have tested the potential existence of specialized short-term stores (or buffers) and executive centers in the brain. These studies allow asking whether encoding, maintenance and retrieval processes are supported by activation of the brain regions that also support perception (primary areas) and if they involve multiple brain regions. In the following section, we will present a brief overview of the current knowledge of the cortical networks supporting working and short-term memory. We will then present a more precise description of the neuroanatomical and functional correlates of short-term and working memory for the specific case of auditory verbal and tonal materials.

2.4.1. Global cerebral network of working memory

By combining neuropsychological studies in brain lesion patients with neuroimaging and neurophysiological approaches, research in neuroscience has proposed the description of a global cortical network supporting working and short-term memory processing. These studies have supported the idea of an implication of various brain regions during maintenance of information in working memory (D'Esposito, 2007; Logie and D'Esposito, 2007; Owen, 2000). In the following paragraphs we briefly present this network, by focusing on the specific role of the prefrontal cortex (PFC).

2.4.1.1. Prefrontal Cortex

Theoretical and empirical research has hypothesized the role of the prefrontal cortex (PFC) in short-term, working and long-term memory (Champod and Petrides, 2007; Curtis and D'Esposito, 2003; D'Esposito, 2007; D'Esposito and Postle, 1999; Howard et al., 2003; Khan and Muly, 2011; Lachaux et al., 2012; Mainy et al., 2007; Marklund et al., 2007; Owen, 2000; Petrides et al., 2012). Jacobsen (Jacobsen, 1935, 1936) was the first to report a link between PFC and working memory by demonstrating deficits in a delayed response task in non-human primates after bilateral prefrontal lesions. He trained monkeys to choose one object (that had previously been rewarded) from a pair of objects. After large bilateral frontal lesions, monkeys were impaired in delayed response allowing Jacobsen to consider this effect as evidence for a memory deficit. However, subsequent research has challenged this hypothesis, postulating that the observed deficits on tests of delayed response (or even delayed matching-to-sample) could arise from deficits in multiple other processing steps, such as encoding, stimulus discrimination, or accessing recently acquired information (despite intact long-term memory) (D'Esposito and Postle, 1999).

Later on, thanks to several methodological advances, more circumscribed lesions of the cerebral cortex were possible and have allowed studying the role of different sub-regions of the PFC in different aspect of memory processing (Fuster, 1989). Before presenting the functional roles of the different sub-regions, it should be noted that the classification of these areas was first done by

anatomical and cyto-architectural studies. As presented in Chapter I (see section 1.4.3 pages 24, Figure 3), cyto-architectural studies are based on the investigation of the differences in cellular population, cell size and the arrangement of the neurons (differences in cell density, or the relative thickness of the layers) in various cortical layers. In addition to segmenting anatomical regions, the cyto-architectonic approach has allowed to hypothesize connection patterns between local and distant regions: a given region receives afferent connections from specific cortical and subcortical areas in the layer IV, while efferent connections of a particular frontal cortical area are sent to control information that is processed in other cortical and sub-cortical areas (well developed layers III and V). In the PFC, several regions have been isolated, but we will present here only the dorsolateral and ventrolateral prefrontal areas that are recruited by memory processing.

2.4.1.1.1. *Dorsolateral prefrontal cortex (DLPFC)*

The dorsolateral prefrontal (DLPFC) cortex includes Brodmann's areas 9 and 46 on the superior and middle frontal gyri. Area 46 occupies the middle frontal gyrus, whereas area 9 extends in both superior and middle frontal gyri. Brodmann's area 46 (Figure 15A.) shows a well-developed layer IV and a developed layer III, and is separated from area 8 on the middle frontal gyrus by area 9.

Brodmann has identified area 9 on the superior frontal gyrus above area 46 (see Figure 15A) and on the middle frontal gyrus, between area 46 and area 8. In 1994, Petrides and Pandya (Figure 15 C) proposed a more precise delimitation of area 9 by showing that its portion on the superior frontal gyrus contained a poor layer IV, whereas the area 9 on the middle frontal gyrus showed a well-developed layer IV. In this respect, they proposed that area 9 on the superior frontal gyrus is more similar to area 46 than to Brodmann's area 9 (see 9/46d on Figure 15C). In the middle frontal gyrus, the cortical Brodmann's area 9 (between area 46 and area 8) contains a strong concentration of granular cells (well-developed layer IV). Although this portion of the cortex on the middle frontal gyrus shares this characteristic with area 46, Petrides and Pandya (1999) have distinguished it from area 46 by observing in area 9 of the middle frontal gyrus the presence of large pyramidal neurons in the lower part of layer III (see area 9/46v on Figure 15C) (Petrides and Pandya, 1999, 2002; Petrides et al., 2012).

Brodmann (1909)   Petrides and Pandya (1994)
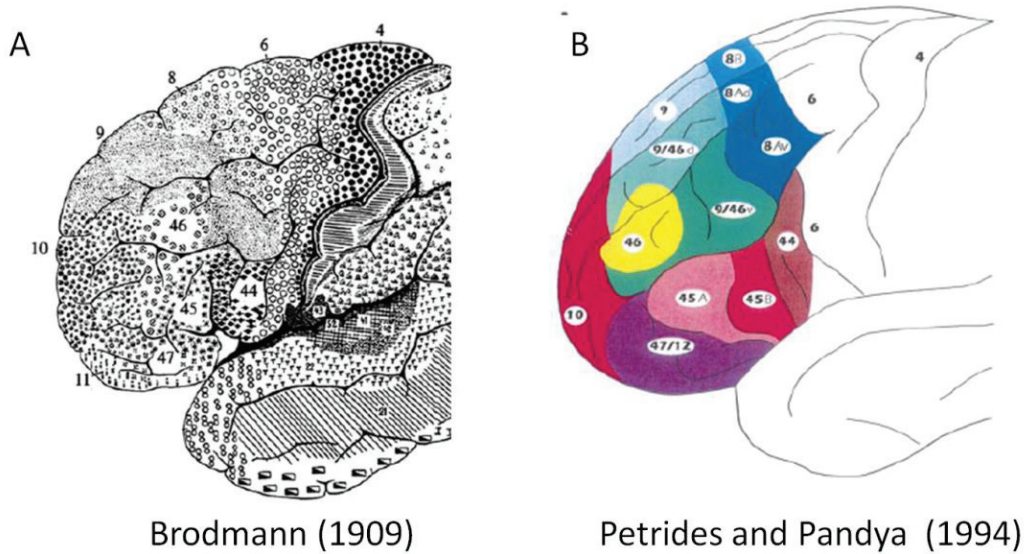
Figure 15: Architectonic maps of lateral surfaces of human cerebral hemispheres according to (A) Brodmann (1909), (B) Petrides and Pandya (1994).

2.4.1.1.2. Ventrolateral prefrontal cortex (VLPFC)

According to Brodmann (Figure 15A), the ventral part of the precentral gyrus is occupied by areas 4 and 6. In front of motor area 4 and premotor area 6, on the inferior frontal gyrus, three areas have been identified: 1) area 44 on the pars opercularis, 2) area 45 on the pars triangularis, 3) and part of area 47 on the pars orbitalis.

Area 44 (Figure 15) is observable just in front of premotor area 6 and was defined as a dysgranular region that exhibits a narrow and interrupted layer IV, and shows large pyramidal neurons in the lower part of layers III and V. Area 45 was shown to exhibit a very well-developed layer IV (in contrast to area 44) and large pyramidal neurons in the deepest part of layer III. According to Petrides and Pandya (2002), this external granular layer could be considered as a distinctive characteristic of area 45 as the pyramidal cells are unusually clustered. Finally, area 47 is a very large and architectonically heterogeneous zone. Brodmann (1909) has proposed that this area could be further subdivided as it includes the most ventral part of the ventrolateral frontal region, but also extends on the caudal orbital surface (see Figure 15A). In 2002, Petrides and Pandya have thus introduced the narrowest 47/12 area. They showed that the transition from area 45 to area 47/12 could be supported by strong differences in the deeper part of layer III. In area 45, this layer contains numerous large pyramidal neurons, whereas in area 47/12, layer III contains small and medium pyramidal cells in its upper part, and medium and larger pyramidal neurons in its lower part. Moreover, they observed additional differences between area 45 and area 47/12 in the layer IV. In area 47/12 this layer is developed and clearly separates layers III and V, but remains not as well developed as that of area 45. Finally, in layers V and VI of area 47/12, they observed densely clustered medium to small pyramidal cells leading to prominent infragranular layers that were not as developed as those of area 45.

Note that as in chapter I, the laminar organization of these regions will drive the Dynamic causal modeling approach that we will present at the end of this chapter.

2.4.1.1.3. *Sub-cortical and cortical connections of the PFC*

In addition to the study of the laminar organization of the PFC, autoradiographic techniques and DTI approaches have allowed describing the connection patterns of its different sub-regions (Frey et al., 2008; Petrides and Pandya, 1999, 2002; Petrides et al., 2012). These studies have revealed the existence of several long-distance pathways connecting the ventrolateral and dorsolateral regions to parietal and temporal cortices that we have previously described in Chapter I (see section 1.4.5, pages 30 to 34). These connections are supported by several white matter tracts in the human brain, such as the extreme capsule fasciculus, the arcuate fasciculus, the ucinate fasciculus and the different branches of the superior longitudinal fasciculus.

Additionally to these long distance cortico-cortical connections, several studies have defined the pathways that linked the PFC with sub-cortical regions (see Khan and Mully 2011, for a review). The thalamus provides the major inputs of the PFC (in layer IV) where most of the thalamocortical afferents originate from the mediodorsal nucleus. Another input to the PFC comes from the midbrain dopamine nuclei. Mesocortical dopaminergic projections primarily arise from different areas of? the tegmental ventral area, the retrorubral area and the substantia nigra. In addition to these inputs, the PFC sends efferent projections (outputs) to the basal ganglia, with projection to the striatum that have been suggested to participate in the modulation of behavior. Finally, the different sub regions of the PFC present local, intrinsic, connections. Intrinsic connections have been shown to play a role in the signal processing involved in several cognitive functions (such as memory). The PFC contains a diverse set of inhibitory interneurons influencing the response properties of prefrontal neurons. While the precise role of intrinsic prefrontal circuitry in working and short-term memory function is not elucidated clearly, it is widely accepted that these interactions have functional importance.[2]

2.4.1.1.4. *Functional evidences*

In addition to the anatomical explorations, neuroimaging studies in the human brain have provided evidence about the functional contribution of the dorsolateral and ventrolateral cortical regions of the PFC in short-term memory and working memory.

In monkeys, restricted lesions of the DLPFC (areas 9/46) (Blum, 1952) were shown to be sufficient to produce working memory deficits. These findings were complemented by neurophysiological studies in behaving monkeys using recordings from single neurons within the DLPFC. Using tasks that require a monkey to retain information over a brief period of time, these studies have consistently found persistent, sustained levels of neuronal firing during the retention interval (Funahashi et al., 1989; Fuster and Alexander, 1971). It has thus been proposed that the

---

[2] Note that we had this description of the laminar organization of the PPC to help the description in a future section (see Chapter II page 75) of a connectivity model we have used in the present PhD work.

DLPFC in the monkey's brain is necessary for the maintenance of information. In humans, this hypothesis has received support from functional neuroimaging studies that have found lateral PFC activity during short-term and working memory tasks (for a review, see Curtis & D'Esposito 2003).
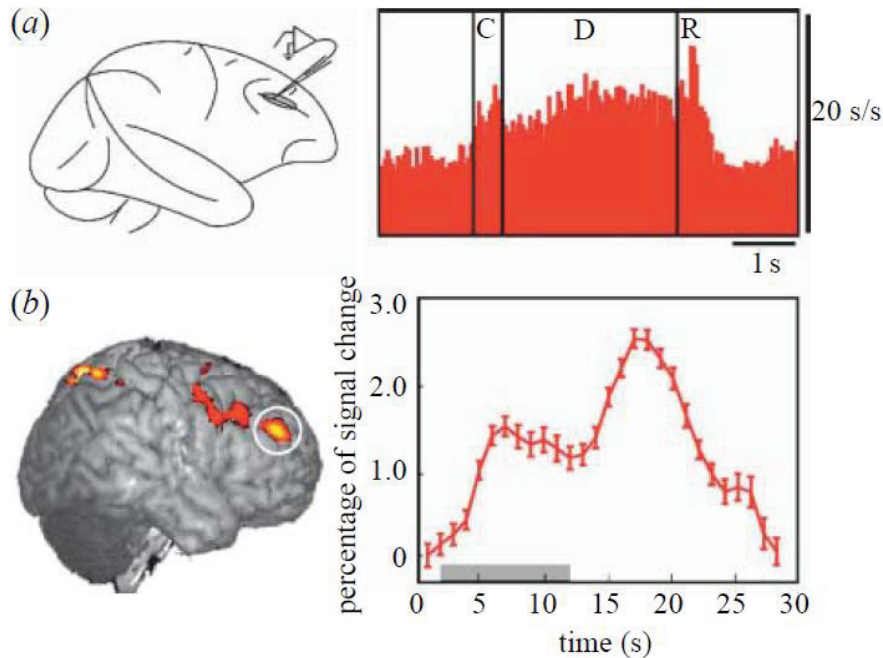
Figure 15. Neural activity in the monkey and the human lateral PFC during the retention interval of a spatial oculomotor delayed response (ODR) task from D'Esposito (2007), figure 1. (a) Macaque: average of single-unit recordings from 46 neurons with delay-period activity from the monkey lateral PFC (brain area (BA) 46. C, cue; D, delay; R, response. (b) Human: significant delay-period activity (left) and average (Gs.e.) fMRI signal (right) from right lateral PFC (BA area 46; circled) in a human participant performing an ODR task. The grey bar represents the length of the delay interval. Note how in both cases the level of PFC activity persists throughout the delay, seconds after the stimulus cue has disappeared.

Based on these first observations, studies in neuroscience have explored with numerous neuroimaging methods (such as fMRI, PET, MEG, EEG, SEEG) the potentially specific and/or general roles that the DLPFC and the VLPFC could play in working memory and short-term memory. These investigations have explored two major questions: 1) are these activations domain-specific? 2) do they support more global cognitive functions?

Initially, fMRI experiments in humans have clearly demonstrated activations in dorsal and/or ventral frontal regions during object and spatial working memory tasks relative to non-mnemonic control tasks (Curtis and D'Esposito, 2003; D'Esposito, 2007; D'Esposito et al., 1998a; D'Esposito and Postle, 1999; Logie and D'Esposito, 2007; Nystrom et al., 2000; Owen, 2000; Owen et al., 1996; Postle et al., 2000a; Postle et al., 2000b; Sala and Courtney, 2007). However, the characterization of

these mechanisms within the PFC has been strongly debated. Results from both monkeys (Levy and Goldman-Rakic, 2000) and humans (Courtney et al., 1998; Curtis and D'Esposito, 2003; Sala et al., 2003) have indicated that the DLPFC is mainly recruited during maintenance of spatial information, and the VLPFC cortex is preferentially involved during maintenance of non-spatial information (object-related, see also Chapter I, section 1.4.5 pages 30 to 34). These observations have supported the hypothesis of domain-specificity of these regions, in which information is segregated in short-term memory in the PFC according to the type of to-be-remembered material. This idea presented the advantage to fit nicely with the models described in cognitive psychology (Baddeley, 2003b, 2010; Baddeley, 1986), which have hypothesized the existence of specific neural systems (or buffers) for the storage of spaial and verbal information.

However, these conclusions were challenged by numerous studies that have failed to observe any dorsal-ventral specific organization within the PFC related to the maintenance of object versus spatial information (Curtis and D'Esposito, 2003; D'Esposito, 2007; D'Esposito et al., 1998b; Logie and D'Esposito, 2007; Owen, 2000; Petrides, 1995a, b; Postle and D'Esposito, 1999). Instead of considering that a given region of the PFC is specifically recruited for the storage of a given type of material, these studies have defined more general, but distinct, roles of the DLPFC and VLPFC for working memory and short-term memory respectively.

Notably, several studies have shown persistent delay-period activity in the VLPFC during active rehearsal (see Owen, 2000; D'Esposito 2007; Curtis & D'Esposito 2003; Khan and Muly 2011). Based on these observations, Owen (2000) has proposed that the VLPFC may play a general role in memory, in which it is supposed to trigger active low-level encoding strategies (such as rehearsal), and to initiate explicit retrieval of information from long-term memory. This hypothesis was also supported by the fact that the VLPFC (i.e. Broca's area) is often activated in working memory tasks where subvocal rehearsal is the main strategy for maintenance (Awh et al., 1996; Schulze and Koelsch, 2012). Rehearsal of non-verbal material like spatial locations has been more difficult to resolve, but has been shown to likely involve a similar process (Awh and Jonides, 2001). Taken together, the VLPFC has been considered as supporting low-level mnemonic processes, such as encoding, rehearsal and retrieval in short-term recognition (Owen, 2000).

In contrast, the DLPFC has been hypothesized to mediate more complex types of processing as it is mainly recruited when active manipulation and/or monitoring of the remembered information is required (Petrides, 1991, 1994). The DLPFC could be considered as a specialized region where stimuli or events, that have been first interpreted and maintained in other association cortical areas, can be monitored and manipulated (D'Esposito, 2007). This system could be related to the central executive of Baddeley and Hitch (1974) as the DLPFC is capable of monitoring multiple cognitive representations in working memory. The current opinion in cognitive neurosciences is thus to consider that the DLPFC and VLPFC could not simply be conceived as specialized working memory modules for specific information, but can rather be considered as complex systems for both low-level

Chapter II: Auditory short-term memory for verbal and tonal material
(encoding, rehearsal for the VLPFC) and high-level (monitoring and manipulation for the DLPFC) mechanisms supporting memory processing.

2.4.1.2. Association cortical areas

In addition to the important role of the different sub-regions of the PFC, several other areas have been shown to support working memory and short-term memory, suggesting that memory maintenance and manipulation involve a distributed network in the brain.

2.4.1.2.1. Anterior cingulate cortex

There is some evidence suggesting that the medial frontal lobe plays a role in memory and cognition, and in particular the anterior cingulate cortex. Several brain imaging studies in humans using tasks requiring complex or necessitating active control on behavior have reported lateral prefrontal and anterior cingulate cortex activations (Koski and Paus, 2000; Paus et al., 1993). In a PET study in monkeys, the anterior cingulate cortex has been shown to be activated in a working memory task (Inoue et al., 2004). However, this activity seems to be unspecific, as its recruitment has been shown to be strongly influenced by reward expectation and error detection (Khan and Muly, 2011).

2.4.1.2.2. Premotor cortex

Premotor areas have also been observed to participate in the working memory network (di Pellegrino and Wise, 1991; Wallis and Miller, 2003). However, as compared to the PFC, its activity was supposed to be more closely related to motor planning rather than intervening in the processing and memorization of the stimulus itself. In di Pellegrino and Wise (1991), using a double delay task, latencies of phasic and tonic activities was shorter in the PFC than in the premotor cortex during the retention period. These effects were interpreted as a transfer of information from PFC to premotor cortex. However, this may not be always the case, and premotor activity can have shorter latency than prefrontal activity (Khan and Muly 2011). Although the dominant hypothesis emphasizes a role of premotor areas in response planning (motor) rather than in active memory, its role in the retention of information during working memory tasks cannot be ruled out.

2.4.1.2.3. Posterior Parietal cortex

Many studies have tried to decipher the role of the posterior parietal cortex PPC (Brodmann's areas 40/7) in working memory and short-term memory. Initially, this region was identified as the short term store (corresponding to the buffers described in models of cognitive psychology) supporting maintenance (D'Esposito, 2007; Jonides et al., 1998; Paulesu et al., 1993; Postle et al., 1999; Smith and Jonides, 1998; Wendelken et al., 2008) and retrieval (Olson and Berryhill, 2009) of the information. However, later studies have pointed out that the location of the peaks in the PPC in functional studies were considerably widespread (Champod and Petrides, 2007, 2010) and involved: 1)

the angular and supramarginal gyri, 2) the superior parietal cortex and 3) the perisylvian parietal cortex.

Observing such a diffuse network has led to debates concerning the role of the PPC in working memory. Recently, another area of the PPC, the intraparietal sulcus (IPS) has been proposed to have a role in working memory (Champod and Petrides (2007, 2010). By investigating brain activity in the IPS and DLPFC during the realization of working memory tasks in which either monitoring or manipulating information was required, Champod and Petrides (2010) highlight the specific contribution of each region in working memory. The results show that the PPC was involved in manipulation processes, whereas activation of the DLPFC was more related to the monitoring of the information that was being manipulated (Figure 16).
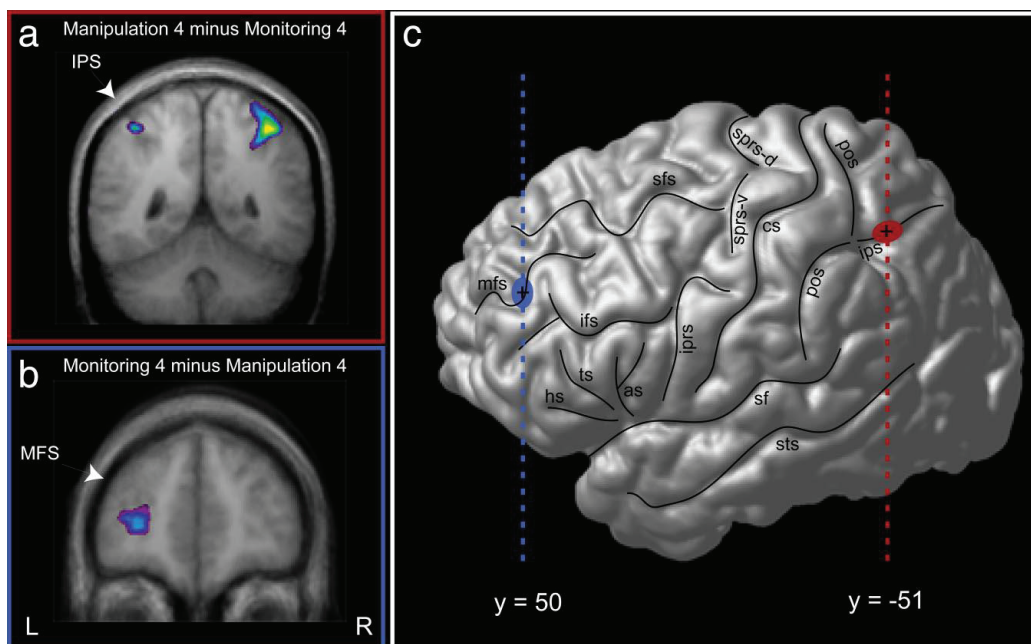
Figure 16: Figure and legend from Champod and Petrides (2010). a, Bilateral increased activity in the IPS region in the manipulation task for words minus monitoring task for word comparison. b, Increased activity in the left MDLFC in the monitoring task for words minus manipulation task of words comparison. MFS, Middle frontal sulcus. c, Cortical surface rendering in standard stereotaxic space of the left hemisphere of a subject's brain. The vertical blue line indicates the anteroposterior level of the coronal section illustrated in b, and the blue circle indicates the focus of activity in the MDLFC. The vertical red line indicates the anteroposterior level of the coronal section illustrated in a, and the red circle indicates the focus of activity in the depth of the IPS. L, Left hemisphere; R, right hemisphere; as, ascending sulcus; cs, central sulcus; hs, horizontal sulcus; ips, intraparietal sulcus; ifs, inferior frontal sulcus; iprs, inferior precentral sulcus; mfs, middle frontal sulcus; pos, postcentral sulcus; sf, Sylvian fissure; sfs, superior frontal sulcus; sprs-d, dorsal branches of the superior precentral sulcus; sprs-v, ventral branches of the superior precentral sulcus; sts, superior temporal sulcus; ts, triangular sulcus.

2.4.1.2.4. *Medial Temporal lobe*

In monkeys, entorhinal, perirhinal, and hippocampal neurons exhibit sustained discharges during the delay period of a working memory task like neurons in the PFC (Khan and Muly, 2011). However, it has been suggested that the medial temporal lobe does not seem to be necessary for working memory performance, but rather plays a role when novel stimuli are involved in the task. This hypothesis was supported by the observation that entorhinal cortical responses are resistant to interference when monkeys are required to remember the initial sample of a sequence of stimuli and disregard subsequent non-match stimuli. Moreover, lesions of parahippocampal cortices in monkeys (Eacott et al., 1994; Otto and Eichenbaum, 1992) lead to impairment in the performance on working memory tasks when novel stimuli are used, whereas no effect was observed with familiar stimuli.

2.4.1.2.5. *Striatum*

An implication of the striatum in working memory was supported by human fMRI studies that have found increase in activity in the caudate nucleus during a verbal working memory task (Lewis et al., 2004). Additionally, a study of stroke patients showed that basal ganglia lesions made patients more susceptible to irrelevant information in a working memory task though overall working memory capacity was unaffected (Baier et al., 2010). As in the PFC, sustained activity during the delay period in the caudate nucleus has been observed (Kawagoe et al., 1998). These results support a role for basal ganglia in working memory, but it appears to be less related to the maintenance of information, and more related to reward expectancy, which can be used as a filter to enhance the focus on relevant stimuli.

2.4.1.2.6. *Thalamus*

As described above, the PFC shows connections with the mediodorsal thalamic nucleus. Lesions of this nucleus produce impaired performance on delay tasks (Isseroff et al., 1982) and neurons in this nucleus show sustained activity during delay periods of working memory tasks (Tanibuchi and Goldman-Rakic, 2003). During working memory tasks, the delay period, activities in mediodorsal nuclei are similar to those seen in lateral prefrontal and in parietal cortices, however, more mediodorsal units seems to be more devoted to the motor aspects of the task (Watanabe and Funahashi, 2004a, b).

Taken together, studies in neuroscience have allowed the definition of a cortical network that supports global working memory and short-term memory. This network involves mainly regions of PFC and PPC that are supposed to interact to allow encoding, retention, manipulation, monitoring and retrieval of the information in memory. Numerous studies have investigated this network notably for the processing of verbal and visuo-spatial material. However few studies have explored its implication

in auditory short-term memory for tonal material. In the following section we will review such studies in particular to assess whether the network supporting memories for verbal and tonal material are shared or different.

2.4.2. Cerebral correlates of verbal short-term memory

As in studies of cognitive psychology, most studies investigating the cerebral correlates of auditory short-term and working memory have used verbal material. Initially, based on Baddeley's model (Baddeley, 1992), lesion studies exploring clinical syndromes (i.e. dysarthria) or individual case reports have tried to define the functional neuro-anatomy of the phonological loop. These studies were followed by neuroimaging investigations in typical participants that have explored if the two components of the loop, the phonological short-term store and the articulatory subvocal rehearsal process (see Chapter II, page 52), could be supported by specific regions in the human brain.

Articulatory subvocal rehearsal process

It has first been suggested that the left inferior frontal gyrus (area 44 VLPFC) is critical for verbal rehearsal. Active rehearsal is hypothesized to consist of the internal repetition of relevant representations to transiently reactivate and refresh this information. Subvocal articulatory mechanisms have been proposed to mediate the rehearsal of verbal memoranda (Baddeley 1986) as articulatory suppression (e.g. repeting 'the…the…the' during a retention interval), which interferes with rehearsal, degrades memory performance (see Chapter II, section 2.2.1 pages 52). Along these lines, neuroimaging studies have indicated that the VLPFC (i.e. Broca's area: Brodmann's area 44), in particular in the left hemisphere, is often activated in short-term memory tasks for which subvocal rehearsal is the main strategy for maintenance (Awh et al., 1996; Fiez et al., 1996; Gruber and von Cramon, 2003; Paulesu et al., 1993; Ravizza et al., 2004; Schulze and Koelsch, 2012). In addition to the VLPFC, neuroimaging studies have suggested that both the anterior insular area (Bamiou et al., 2003; Chein and Fiez, 2001; Paulesu et al., 1993) and the cerebellum (Chen and Desmond, 2005; Kirschen et al., 2005; Ravizza et al., 2004) are involved in the internal rehearsal of verbal information.

These results suggest the existence of similarities between the cortical networks supporting short term memory (or working memory) and speech perception and production. Therefore, it could be hypothesized that sensory-motor processes are involved in the representation and manipulation of verbal information in short-term memory. This hypothesis is in line the dual-stream model of speech perception (Hickok, 2009; Hickok and Poeppel, 2007; Rauschecker and Scott, 2009) assuming that the dorsal path of the auditory system (projecting to the parietal–temporal junction, the premotor cortex and VLPFC regions) is involved in sensory–motor integration.

Phonological store

In comparison to the functional correlates of internal rehearsal, the explorations of the cerebral correlates of the phonological store have been much less conclusive. As reported in the previous section (see page 70), the PPC, in particular the inferior parietal lobule (IPL) and the superior parietal lobule (SPL) have been suggested to support the maintenance of the verbal information (Awh et al., 1996; Chen and Desmond, 2005; Gruber and von Cramon, 2003; Jonides et al., 1998; Kirschen et al., 2005; Paulesu et al., 1993). This assumption was originally based on the observation that patients with conduction aphasia who also show storage-related deficits have lesions encompassing the inferior parietal cortex and the arcuate fasciculus (Fiez, 2001). In a related vein, Vallar et al. (1997) compared a patient (L.A) with a lesion involving the inferior parietal lobule and the superior and middle temporal gyri with a patient with lesions in sub-cortical premotor and rolandic regions (T.O). While patient L.A. was unable to retain auditory-verbal material, he performed normally on phonological judgments that involved the rehearsal process. In contrast, patient T.O. never made use of the rehearsal process but his memory capacity of the phonological short-term store was comparatively preserved (Vallar et al., 1997).

Although lesion studies have hypothesized a specific role of the PPC in the maintenance of verbal information, the localization of the phonological store in the parietal lobe remains controversial for multiple reasons. First, as described previously (see section 2.4.1.2.3 pages 70 to 71) the neural activity in the PPC might also reflect engagement of attentional resources and manipulation of the information (Cabeza and Nyberg, 2000; Champod and Petrides, 2007, 2010; Corbetta and Shulman, 2002). Second, the reported coordinates in the PPC differ between studies (Buchsbaum and D'Esposito, 2008). Considering these limitation, Schulze and Koelsch (2012) have proposed that the storage of verbal information may not be localized in only one region of the PPC, but rather be a function of a more complex network of anterior prefrontal and inferior parietal brain regions subserving non-articulatory maintenance of phonological information (Schulze and Koelsch, 2012).

Finally, the Spt (Sylvian–parietal– temporal, left posterior planum temporale) has also been proposed to support the temporary storage of verbal information during short-term memory tasks. This area is well known to be involved in speech processing and is activated during the delay period of a short-term memory task (Buchsbaum et al., 2005; Hickok et al., 2003). On the basis of these findings it has been proposed that area Spt acts as an auditory–motor interface for short-term and working memory (Buchsbaum and D'Esposito, 2008; Buchsbaum et al., 2005; Hickok et al., 2003). This proposition fits nicely with the hypothesis of a dual-stream model of speech processing (Hickok, 2009; Hickok et al., 2011; Hickok and Poeppel, 2007; Rauschecker and Scott, 2009; Schulze and Koelsch, 2012) in which a ventral stream supports speech comprehension while a left dominant dorsal stream, which comprises also area Spt, enables sensory–motor integration.

2.4.3. Cerebral correlates of tonal short-term memory

In comparison to verbal material, the cerebral correlates of short-term memory processing for tonal material have been less investigated by neuroimaging and neuropsychological studies. In a lesion study, Zatorre and Samson (1991) demonstrated that auditory short-term memory for pitch material is supported by the right auditory cortex. However, this result had been tempered by considering that a lesion in the auditory cortex leads to disruption in perceptual processing that would likely lead to difficulty in maintaining a perceptual trace of the auditory information over time (Peretz and Zatorre, 2005; Zatorre and Samson, 1991). In addition to the participation of auditory cortical regions, neuroimaging studies (Gaab et al., 2003; Griffiths et al., 1999; Holcomb et al., 1998; Zatorre et al., 1994) have uncovered a cortical network that was surprisingly similar to the network subserving the phonological loop described above. During the active retention of pitch, Zatorre et al. (1994) have reported activations in inferior frontal and insular cortex, the planum temporale, and the supramarginal gyrus (SMG) (Zatorre et al., 1994). A similar network was later observed by Gaab et al. (2003) who showed activation of the SMG, the SPL, the planum temporale, premotor regions encroaching on VLPFC, and cerebellar regions during a pitch memory task (Gaab et al., 2003). More recently, using a short-term memory task, Foster and colleagues have shown an important role of the IPS in transforming pitch information (Foster et al., 2013; Foster and Zatorre, 2010), thus suggesting that additionally to maintenance processes, this structure could allow the manipulation of information (see also Champod and Petrides 2007, 2010). These observations have largely converged on the idea that working memory and short-term memory for tones engages interactions between frontal cortical and posterior temporal areas, as it was described for other material in the previous section (see section 2.4.1 pages 64 to 73). According to Peretz and Zatorre (2005), the cortical regions recruited by working and short-term memory for pitch could be seen as a specialized subsystem within the framework of general working memory. Note that similar conclusion were proposed by a study using real music (Janata et al., 2002a).

Comparison between verbal and tonal short-term memory

One way to test if short-term memory for pitch recruit a specialized subsystem in the brain is to compare auditory short-term memory processing for verbal and tonal materials (Hickok et al., 2003; Koelsch et al., 2009; Schulze and Koelsch, 2012; Schulze et al., 2011a). Hickok et al. (2003) asked participants to internally rehearse verbal and tonal stimuli during fMRI acquisition. Results showed that internal rehearsal of both verbal and tonal material activated the area Spt, the area 44 of the VLPFC, and the left premotor regions (Hickok et al., 2003). In two others studies exploring the same question (Koelsch et al., 2009; Schulze et al., 2011) using recognition tasks, very similar activations were observed. In the study by Koelsch et al. (2009) during the verbal rehearsal, an activation pattern comprising the premotor cortex, the anterior insula, the SMG/intraparietal sulcus (IPS), the planum temporale, VLPFC, and the cerebellum was observed in both verbal and tonal processing. Schulze et

al. (2011) have found similarities of the pattern of brain activations underlying the internal rehearsal of verbal and tonal short-term memory. In line with previous studies, both verbal and tonal short-term memory activated areas typically reported in previous experiments on either verbal (Awh et al., 1996; Baddeley, 2003a; Gruber and von Cramon, 2003; Paulesu et al., 1993) or tonal memory (Gaab et al., 2003; Griffiths et al., 1999; Hickok et al., 2003; Zatorre et al., 1994). Based on this observation of a large overlap of neural resources underlying short-term and working memory for both verbal and tonal information they concluded that no specific regions in the brain support retention and rehearsal of tonal material.

Comparison between musicians and non-musicians

Another way to explore the cerebral network supporting memory processing for tonal information is to explore the potential neurophysiological differences between musician and non-musicians. Because the processing of verbal material is a fundamental human skill typically acquired during early life, non-musicians can be considered as trained participant for speech processing, and musicians can be considered as trained participants for both speech material and tonal material. Importantly, note that there is strong evidence for anatomical and functional differences between musically trained listeners and untrained listeners for auditory processing (Baumann et al., 2008; Herholz and Zatorre, 2012; Pantev et al., 1998; Pantev et al., 2001; Recanzone et al., 1993).

Schulze et al. (2011) investigated short-term memory processing for verbal and tonal material in highly trained musicians and non-musicians. The VLPFC, the left premotor cortex, the left insular cortex, the cingulate gyrus, and the left IPL, which were activated more strongly in non-musicians during verbal compared to tonal material, were activated more strongly in musicians compared to non-musicians for tonal material. The authors thus proposed that the functional network recruited by non-musicians for verbal information was also used by musicians for tonal information (but note that this analysis do not exclude that non-musician recruit the same brain regions but with less strength). Additionally, they assumed that musicians might have more elaborate sensory-motor codes underlying the internal rehearsal of tones compared to non-musicians due to functional plasticity induced by musical training. (Schulze and Koelsch, 2012; Schulze et al., 2011a).

2.4.4. *Investigating functional connectivity between cortical regions*

Taken together, the studies described in this chapter have provided evidence that active encoding, maintenance, rehearsal, monitoring, manipulation and rehearsal of task-relevant representations require the implication of numerous cortical regions in the brain. However, it should be noted that the implication of different brain regions within a functional network probably differ in their degree of participation depending on the context of the operation being actively performed. This assumption has received support in the observation that the same regions that are involved in temporarily maintaining a representation are often also engaged during the encoding and retrieval of that information (D'Esposito, 2007). It thus still remains an open question how different regions

Chapter II: Auditory short-term memory for verbal and tonal material implicated in complex cognitive processes, in particular for short-term and working memory, process work together in terms of network activity.

Along these lines, contemporary neuroscience is now trying to determine how neural communication occurs and how distinct neuronal populations exchange information and coordinate their activity within the brain networks (D'Esposito, 2007; Goldman-Rakic, 1988; Lachaux et al., 2012; Tallon-Baudry and Bertrand, 1999). It has been proposed that such interactions could support working memory maintenance processes via synaptic reverberations in recurrent circuits (Durstewitz et al., 2000; Wang, 2001) or synchronous oscillations between neuronal populations (Engel et al., 2001; Lachaux et al., 2012; Singer, 1999; Singer and Gray, 1995). To explore this issue, the majority of studies have used indirect methods to explore how working memory and short-term memory are implemented by the interaction of nodes within a neural network.

Due to the several methodological progresses made in functional neuroimaging approaches and electrophysiology in humans in the last decade, new approaches are now suited to explore network interactions. Over the last decade most studies have used functional connectivity measures, such as phase-synchronisation, temporal correlations or coherence between the activities of two areas at the scalp or source level Functional connectivity is used to establish statistical dependencies between time series and is useful because it does not require an underlying model or knowledge of the causal nature of the interactions. However, there are circumstances when the causal architecture of the interactions is the focus of interest. In the following paragraph, we will briefly present a method to explore causal architecture of cerebral networks, namely the Dynamic Causal Modeling (DCM) (Friston et al., 2003) method for electrophysiological data (used in the present PhD work).

Dynamic Causal Modeling

As opposed to functional connectivity, DCM uses the concept of effective connectivity, which refers explicitly to the influence one neuronal system exerts over another. This procedure models interactions among cortical regions and allows making inferences about system parameters and investigating the influence of experimental factors on these parameters.

This method uses neural mass models (David and Friston, 2003) to explain source activity in terms of the ensemble dynamics of interacting inhibitory and excitatory subpopulations of neurons (Jansen and Rit, 1995). It emulates the activity of a cortical source using three neural subpopulations, each assigned to one of three cortical layers (see Chapter I, page 23, Chapter II page 64): 1) an excitatory subpopulation in the granular layer; 2) an inhibitory subpopulation in the supra-granular layer and, 3) a population of deep pyramidal cells in the infra-granular layer.

In this model, the excitatory pyramidal cells receive excitatory and inhibitory input from local inter-neurons (via intrinsic connections, confined to the cortical sheet), and send excitatory outputs to remote cortical areas via extrinsic connections. Within this model, bottom-up or forward connections originate in the infra-granular layers and terminate in the granular layer. In contrast, top-down or backward connections link agranular layers and lateral connections originate in infra-granular layers

and target all layers. Additionally, the model considers that all extrinsic cortico-cortical connections are excitatory and are mediated through the axons of pyramidal cells. Exogenous inputs to the model have the same characteristics as forward connections.

By adopting this network architecture, DCM is able to assess how a given experimental manipulation activates a cortical pathway rather than a cortical area or source. This approach uses a biologically informed model that allows for inferences about the underlying neuronal networks generating evoked responses such as event-related potentials (ERPs) and fields (ERFs).

Conclusion: This chapter presents a brief overview of the behavioral and cerebral correlates of auditory working and short-term memory. We have presented the different forms of memory through the view of cognitive psychology and have asked whether verbal and tonal memory processes shared the same cerebral network. In the studies presented above, the distinction between the cerebral correlates of short-term memory for verbal and tonal material seemed weak, as the two types of information recruited similar cortical processes. This conclusion seems to contrast several lines of evidence that have led to the hypothesis of either a specialized "language network" or a specialized "musical network" within the human brain.

This hypothesis came from seminal investigations in brain-damaged patients who exhibited specific musical disorders (while other cognitive functions, including language, were not altered, e.g., Peretz et al., 1997). More recently, the hypothesis for music-specific networks has received support from the investigation of a congenital music-related disorder. This disorder has been referred to as tone-deafness, dysmusia, dysmelodia, and more recently, as congenital amusia. In the following chapter, we will present the behavioral and neurophysiological correlates of this lifelong deficit.

Chapter III Congenital amusia

As presented in Chapter II, the investigation of the behavioral and cerebral correlates of auditory short-term memory for tonal and verbal materials has led to disparate conclusions. While several studies have proposed that the maintenance of tonal information in memory is supported by the same networks as those recruited for verbal material, some evidence have started to challenge this view. It thus exists some evidences suggesting that short-term memory for music recruits a specialized subsystem in the brain. This hypothesis has emerged from seminal investigations in brain-damaged patients, who exhibited specific musical disorders while other cognitive functions (including language) seemed unimpaired. More recently, further support for this idea came from the investigation of a congenital music-related disorder that most research has referred to as congenital amusia (Peretz, 2001). In the following chapter, after a very brief presentation of cases of acquired amusia, we will present a state-of-the-art overview of the current understanding of congenital amusia[3].

3.1. Acquired Amusia

Acquired disorders refer to specialized cognitive impairments occurring after brain injury, mainly following stroke, tumor, ischemia or surgery in pharmaco-resistant epileptic patients. Such injuries are diverse and often affect multiple cognitive functions. However, in some cases, cognitive disorders can be highly selective. By exploring the link between the location of the lesion and its consequence estimated at the behavioral level, conclusions regarding brain organization and brain functioning can be made. For the specific case of music and verbal processing, studies in brain-damaged patients have allowed several advances in our understanding of their cerebral correlates, notably by proposing that distinct brain regions or networks underlie these cognitive abilities. Historically, one support for this view could be found in the famous case of Shebalin (Luria et al., 1965), a Russian composer who exhibited specific impairment in language processing (aphasia) after vascular accidents and who:

". . . nevertheless, . . . continued to compose, notably completing his Fifth Symphony, which Shostakovitch considered to be one of his most brilliant and innovative works". (Peretz, 2002) (Page 153). It should be considered that conclusions related to this specific case are limited, mainly because this composer was an over-trained expert musician. However, this limit was rapidly challenged by the observation of similar type of disorders in non-musician brain-damaged patients (Nicholson et al., 2003; Patel et al., 1998; Peretz et al., 1997; Peretz and Kolinsky, 1993; Peretz et al., 1994; Steinke et al., 2001).

---

[3] Note that several parts of this chapter are an adaptation and a re-organization of Tillmann, B., Albouy, P. and Caclin, A.'s chapter, *'Congenital amusias',* Tillmann, B., Albouy, P., and Caclin, A. (2013). Congenital Amusias. In Handbook of Clinical Neurology (Oxford, Elsevier), p. in press..

3.1.1. Acquired amusia in non-musician brain-damaged patients

Studies in non-musician patients have explored if music processing could be mediated by a specialized network in the brain. This was done by exploring patients' musical abilities (pitch perception and production, melody recognition, rhythm etc.) and by comparing them to their capacities to process other types of information (such as speech, prosody, environmental sounds). For the specific case of auditory short-term memory, dissociation between musical and verbal materials has been tested by recognition tasks of auditory patterns in patients (Peretz, 1993; Polster and Rose, 1998). Notably, two interesting cases reported by Peretz and collaborators (1993, 1996, 1997, 2002, Patel et al., 1998) were the patients I.R and C.N.

C.N. suffered from specific impairments in music processing after bilateral lesions into middle temporal gyri and temporal poles, the right insula, and to a limited extent, the right inferior frontal gyrus. Although this patient was unimpaired at discriminating pitch patterns, she was not able to recognize familiar music. This observation has led the authors to conclude that C.N.'s deficit was associated to difficulties in accessing long-term stored melodic representations, while she showed preserved short-term memory abilities for such material (Peretz, 1996). Additionally, C.N was able to discriminate musical versions of sentences presented in pairs (replacing syllables by tones, while maintaining pitch and timing information) and was able to extract prosodic information (thus based on pitch information) from a sentence (Patel et al., 1998).

The second patient, I.R., suffered from an asymmetric pattern of lesions in the two hemispheres. In the right hemisphere, she showed damages on the superior temporal gyrus and on the transverse gyrus of Heschl (see Patel et al., 1998 for details). Moreover, she showed an extensive lesion in the right inferior and middle frontal gyri, associated to damages in the precentral gyrus, insula, lateral orbitofrontal gyri and putamen. In the left hemisphere, she showed spared brain tissues on the posterior superior temporal gyrus and her transverse gyrus of Heschl was nearly all injured. Like C.N., I.R. was impaired at recognizing familiar music; however, she was also impaired in the short-term memory processing for pitch patterns (discrimination task). Additionally, I.R. showed deficits in tasks requiring the comparison of musical stimuli derived from sentences, but she was not impaired at extracting prosodic information from a sentence (Patel et al., 1998). By comparing these two patients Patel et al. (1998) have proposed that the right frontal cortex might be involved in the short-term retention and comparison of pitch and temporal patterns. This conclusion was based on the fact that I.R., who showed damage to the right frontal cortex, was impaired on pitch patterns and prosodic discrimination tasks (by assuming that prosody requires appreciation of pitch patterns, and thus shares the same mechanisms in the brain as those recruited for pitch processing), whereas C.N. did not. In another research group, by reproducing the study of Patel et al. (1998) in the patient K.B. (Nicholson et al., 2003), further conclusions concerning the cerebral correlates of short-term memory for pitch were proposed. KB suffered from specific music-related impairments and did not exhibit language deficits (no aphasia). K.B. was also able to perceive segmental speech and environmental

sounds (Nicholson et al., 2002; Nicholson et al., 2003; Steinke et al., 2001). He suffered from focal damage in fronto-parietal areas, cerebellum, and lenticular nucleus of the right hemisphere (Steinke et al., 2001) and was impaired on pitch discrimination tasks, as well as prosodic perception tasks. K.B. presented thus a different pattern of deficits than those presented by I.R., who was only impaired on pitch pattern discrimination tasks (Patel et al., 1998). While K.B. and I.R. exhibited damages in the right frontal cortex, only K.B. exhibited parietal injuries. The comparison between I.R. and K.B. thus led the authors to conclude that the right parietal cortex is involved in the extraction of pitch patterns and that right frontal regions may rather play a role in the short-term retention such patterns (Nicholson et al., 2003).

3.1.2. *Temporal lobectomies in pharmaco-resistant epileptic patients*

In addition to the description of cases of acquired amusia who showed deficits on multiple aspects of music processing, other studies have explored the underlying mechanisms of auditory perception and short-term memory in patients with temporal lobectomy. Such surgeries are realized for patients presenting pharmaco-resistant temporal epilepsy, for who the resection of several regions of the temporal lobe (such as the hippocampus, the amygdala, the temporal pole and the anterior part of the temporal gyri) was shown to be an efficient treatment (Bidet-Caulet et al., 2009). While these patients are not referred to as amusics (note however that in some cases, excisions have also been shown to disrupt music processing (Liegeois-Chauvel et al., 1998; Zatorre, 1985)), the investigation of their behavioral performance before and after surgery allows the exploration of brain functioning in relation to auditory perception and short-term memory.

For the specific case of music processing, temporal lobectomies have been shown to induce deficits in the perception of pitch direction (Johnsrude et al., 2000), pitch retention (Zatorre and Samson, 1991), as well as in the perception of melodies (Samson and Zatorre, 1992) and rhythm (Penhune et al., 1999). More recently, Bidet Caulet et al. (2009) investigated the role of the anterior temporal lobe regions in the encoding and short-term retention of non-verbal auditory material both before and after unilateral temporal excision. Interestingly, they showed that before and after surgery, in comparison to control participants, patients were not impaired in basic acoustic processing, but exhibited deficits in pitch retention, and in the identification and the short-term retention of environmental sounds (Bidet-Caulet et al., 2009). As the impairments were not related to the lobectomy, the authors argued that patients' deficits were related to pathological processes of their epilepsy. This study has thus pointed out a limitation of performing studies in patients, by proposing that the conclusions about brain functioning made on pathological brains could sometimes be difficult to transpose to typical individuals.

Taken together, while studies in patients have allowed several advances in cognitive psychology and neuroscience, these findings have to be considered with caution, notably because it is now widely admitted that the investigation of brain-damaged patients present several limitations (Muller and Knight, 2006).

3.1.3. *Limitations in the investigation of acquired disorders*

First, lesions following different diseases (tumors, ischemia etc.) are difficult to compare in terms of symptoms and anatomy. In a related vein, these lesions are rarely limited to a precise region, involving often several brain areas (see Bidet-Caulet et al., 2009; Patel et al., 1998) while only a small part of a disease-affected region can sustain the cognitive function investigated behaviorally. Second, the boundaries of lesions are rarely clear and in some cases, damage in a given brain area altered functions of brain tissue that macroscopically seems unimpaired.

Additionally Muller and Knight (2006) have pointed out three major problems in the investigation of brain-damaged patients, that encompass all types of lesions: 1) they do not consider the effect of potential re-organization in the brain after the damage (plasticity); 2) they do not allow the constitution of homogenous groups of patients; and 3) they have difficulties to consider and investigate that complex cognitive functions could be supported by interactions between brain regions instead of recruiting local and specialized brain areas.

A way to overcome these limitations is the investigation of congenital disorders. For music processing, such lifelong deficit has been referred to as congenital amusia (Kalmus and Fry, 1980; Peretz, 2003). In contrast to cases of acquired amusia presented above (Nicholson et al., 2002; Nicholson et al., 2003; Patel et al., 1998; Peretz, 1996, 2002; Peretz et al., 1997; Steinke et al., 2001; Tillmann et al., 2007) congenital amusia occurs without brain injuries or sensory deficits. This musical disorder occurs despite normal performance on tests of intelligence, auditory processing, cognitive functioning, language processing, and it is not due to a lack of environmental stimulation to music (Ayotte et al., 2002; Foxton et al., 2004b; Peretz et al., 2002). The research domain interested in congenital amusia is currently expanding. In addition to behavioral methods, neurophysiological methods are now increasingly used (Voxel-Based Morphometry, VBM; DTI; EEG; fMRI), aiming to define anatomic and functional brain anomalies of this deficit, thus going beyond the description of the phenomenon and contributing also to our understanding of normally functioning neural circuitry underlying music processing (perception, memory, production). In the following sections, we will describe this phenomenon by emphasizing congenital amusic individuals' (or amusic individuals here below) behavioral impairments for pitch perception and memory and the cerebral correlates of this condition.

3.2.   Origins of Congenital Amusia

Congenital amusia has been described as a deficit in music production *and* perception. Individuals affected by congenital amusia have difficulty detecting when someone sings out-of-tune (including themselves), recognizing familiar tunes without lyrics, detecting a wrong or out-of-tune note and memorizing even short melodies (Ayotte et al., 2002; Peretz et al., 2002). It is important to distinguish this disorder from the phenomenon that has been termed "poor singing" (Dalla Bella et al., 2007; Pfordresher and Brown, 2007), referring to individuals who only have production deficits, but who do not have perceptual difficulties. It needs also to be distinguished from the general non-musician population who might be lacking confidence in their musical capacities and refer to themselves as *tone-deaf*, but who are actually not (Cuddy et al., 2005; Sloboda et al., 2005).

Even though the disorder has been recognized for a long time (Allen, 1878), it has been systematically studied only relatively recently (Ayotte et al., 2002; Peretz, 2002; Peretz et al., 2002), mainly thanks to the development of the *Montreal Battery for the Evaluation of Amusia*, MBEA (Peretz et al., 2003, http://www.brams.umontreal.ca/plab/publications/article/57). In the MBEA, six sub-tests address various components of music perception and memory, notably the pitch dimension (or melodic dimension corresponding to the detection of an out-of-key note, a contour violation, or interval changes), the time dimension (rhythm and meter perception), and incidental memory (i.e., melodies used in preceding sub-tests) (Figure 17A). Another test used to define a population with a musical deficit is the *Distorted Tune Test* (DTT). This test investigates participants' capacity to discriminate intervals between tones in a melody (Drayna et al., 2001; Jones et al., 2009b; Kalmus and Fry, 1980). However, it has a limit as it uses well-known tunes (in North America) and is thus dependent on participants' long-term memory knowledge of a cultural musical repertoire (see Chapter II page 60). The MBEA is based on newly composed melodies, and it has been used successfully across various countries and cultures (North America, Europe (e.g., France, UK, Greece), China, and New Zealand). Because the MBEA is considered as the standard 'diagnostic' tool by the research community, this battery was used in the present PhD work.

A genetic origin

The musical deficit has been estimated to affect about 4% of the general population (Kalmus and Fry, 1980). This estimate has been criticized as it depends on the test used, how the cut-off scores are defined and the skew in the distributions (Henry and McAuley, 2010). This criticism is not specific to congenital amusia, but it also affects prevalence reports for other disorders, such as dyslexia, dyscalculia, or developmental prosopagnosia (Henry and McAuley, 2010).

Furthermore, there exists considerable uncertainty when assessments are only based on verbal reports without direct testing, notably because of individuals' underestimation of their capacities as well as response bias due to difficulty in admitting that they do not like or understand music. The hypothesis of a genetic origin of this disorder has received support by a family-aggregation study (Peretz et al., 2007) and a twin study (Drayna et al., 2001). Peretz et al. (2007) tested nine amusic families and ten control families. They demonstrated that 39% of first-degree relatives in the amusic families showed the same disorder (while only 3% did in the control families). The incidence of amusia as a heritable disorder is of similar magnitude compared to that reported for other cognitive functions (e.g., specific language impairment, absolute pitch). Note that the calculation of the risk factor is based on the information from the siblings and the hypothesis of a population prevalence of 4% (thus also subjected to the criticism of Henry and McAuley, 2010). Nevertheless, these findings provide the basis for future research aiming to map the genetic loci for hereditary amusia.

Recently, congenital amusia has been also reported in childhood, thus providing some further evidence for the definition as a lifelong disorder (Lebrun et al., 2012; Mignault Goulet et al., 2012). Furthermore, investigating congenital amusia has been promoted as a possibility to study the interaction between influences due to genes, environment, and behavior, in particular for music cognition and its relation to other cognitive capacities, such as, in the framework of the present PhD work, verbal processing (Peretz, 2008).

3.3. A deficit on the pitch dimension: Perception and Memory

Currently, the most extensively investigated hypothesis about the main deficit in congenital amusia concerns the pitch dimension, notably as pitch is a major component of musical structures. As we shall see below, this predominance of pitch processing studies in the congenital amusia literature originates from seminal studies showing that pitch was more affected than rhythm in this disorder (e.g., Allen, 1878; Peretz et al., 2003). While the first studies have focused on pitch perception deficits (i.e., pitch discrimination), more recent studies have pointed to deficits regarding pitch memory – even in cases with unimpaired pitch discrimination capacities (see Figure 17). Congenital amusia thus represents a unique condition for the investigation of both the behavioral and cerebral correlates of auditory perception and short-term memory for musical material.

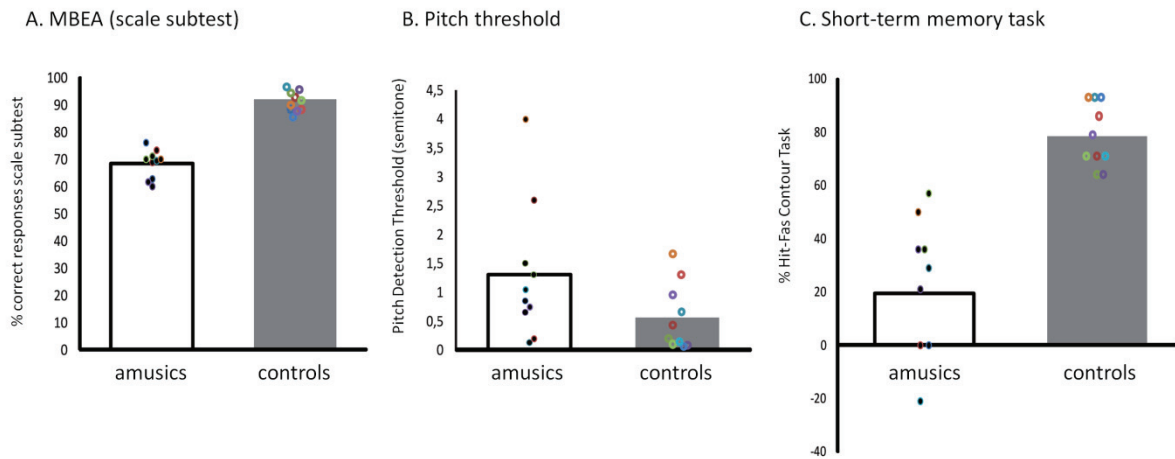A. MBEA (scale subtest)   B. Pitch threshold   C. Short-term memory task

Figure 17 Group data for amusic and control participants together with individual data points (10 participants per group) for three tasks requiring the processing of the pitch dimension: A) Performance (percentage of correct responses) for the scale subtest of the MBEA (Peretz et al., 2003); B) Pitch detection threshold (task measuring the smallest pitch difference, in semitones, that participants can detect, see procedure in Tillmann et al., 2009); C) Performance (Hits minus False Alarms, FAs) in a short-term memory task: indicating whether two five-tone melodies (S1, S2) were the same or different; with different trials being changed on the melodic contour (Tillmann et al., 2009). Note that results of the amusic group (average performance) are significantly worse than the results of the control group for all three tasks, while individual data show considerable overlap between the two groups in the pitch threshold task, but not in the MBEA and the short-term memory task. Note that the same participants from Tillmann et al. (2009) have performed the 3 tasks (A, B, C)

3.3.1. Pitch Perception

Seminal studies of congenital amusia (Ayotte et al., 2002; Foxton et al., 2004b; Hyde and Peretz, 2004; Peretz et al., 2002) led to the assumption that this condition is based on a disorder of fine-grained pitch processing, notably pointing to fundamental deficits in pitch discrimination. Documenting the first reported case of congenital amusia, *Monica*, Peretz et al. (2002) showed that she had difficulties in detecting pitch changes as large as two semitones in tone sequences and tone pairs.

The first group study of individuals afflicted with this developmental disorder (Ayotte et al., 2002) further defined this deficit by studying the discrimination, perception, and memory of musical material as compared to other materials (i.e., language, environmental sounds). While amusic participants performed as well as controls for non-pitch related material (i.e. language, environmental sounds), all amusic participants were impaired in detecting pitch changes in both familiar and unfamiliar melodies. For example, amusic individuals' pitch processing abilities were impaired in an anomalous pitch detection task: participants were asked to detect a pitch anomaly within a melody (a note shifted by a semi-tone; thus creating an out-of-key violation).

To investigate the pitch-processing deficit outside a musical context, Hyde and Peretz (2004) required amusic individuals to detect a pitch change (of 0.25 to 3 semitones) within the context of a

five-tone sequence (monotonic and isochronous). Amusic individuals' performance was impaired for pitch changes smaller than two semitones, while controls showed ceiling performance for changes as small as 0.25 semitones (see Figure 18). These seminal studies led to the hypothesis that abnormalities in pitch perception underlie the phenomenon of congenital amusia (Hyde and Peretz, 2004; Peretz and Hyde, 2003). They proposed that this developmental disorder would arise from a failure to encode pitch with sufficient precision to allow the acquisition of knowledge about the pitch structure of the musical system (Bigand and Poulin-Charronnat, 2006; Tillmann et al., 2000; Tillmann et al., 2007), thus impairing tonal coding and also leading to general difficulties in pitch-related processing (i.e., amusic individuals' inability to perceive dissonance (Ayotte et al., 2002; Cousineau et al., 2012).

To further investigate this hypothesis, subsequent studies have used auditory threshold-based testing with other cohorts of amusic individuals (Foxton et al., 2004b; Jones et al., 2009b; Liu et al., 2010; Tillmann et al., 2009; Williamson and Stewart, 2010). Foxton et al. (2004) have used a range of pitch tasks with forced-choice procedures to evaluate amusic individuals' deficit in pitch perception. Deficits were identified in "low-level" pitch perception tasks (detection of segmented and gliding frequency changes) and in pitch pattern perception tasks, which ranged from determining the direction of pitch changes to detecting differences between pitch sequences. These later findings have also led to the hypothesis of a major deficit concerning the processing of pitch direction (Stewart, 2011). In contrast, the perceptual organization of sounds into different streams on the basis of pitch was found to be unimpaired in amusic individuals (Foxton et al., 2004b).
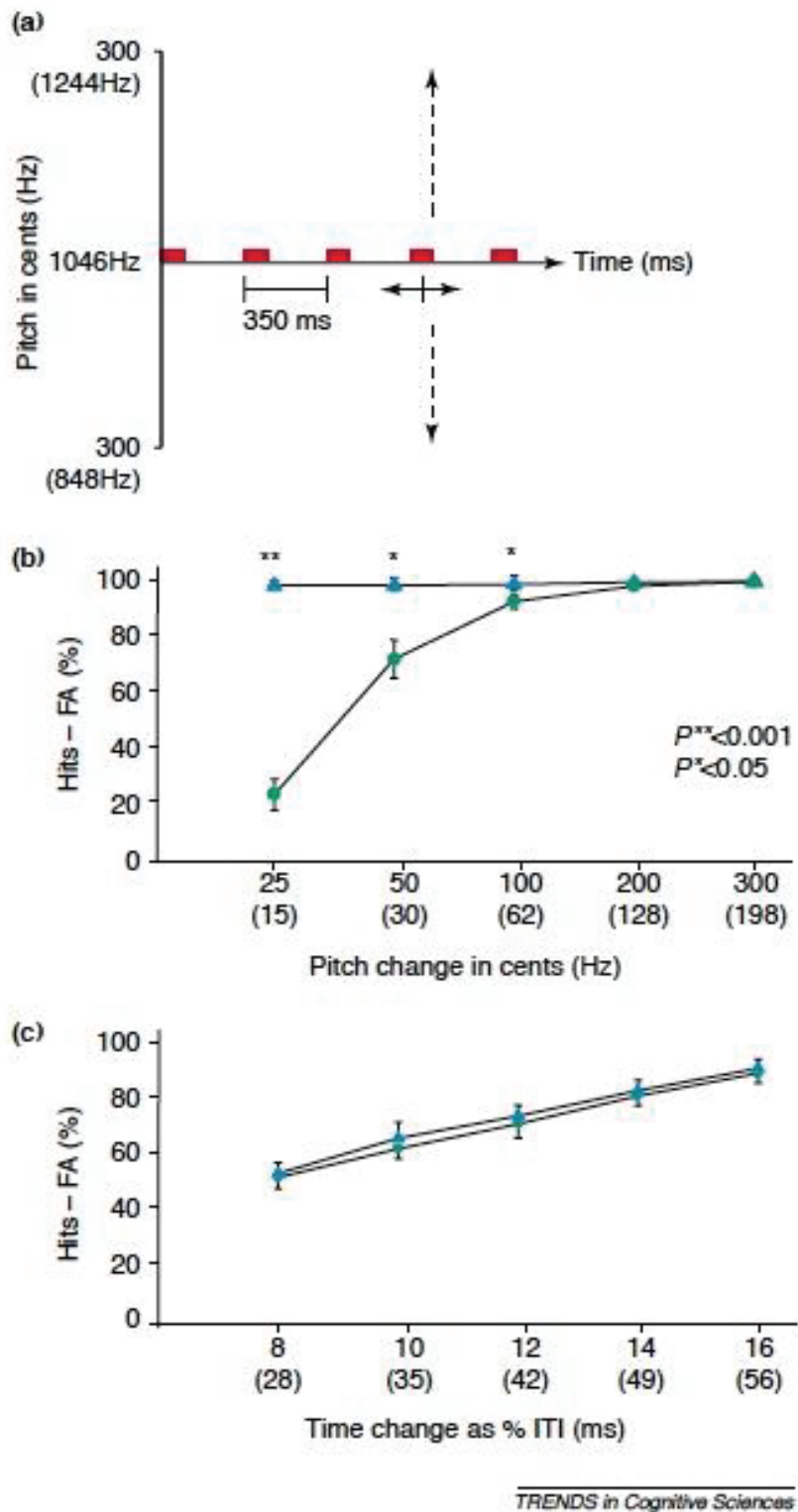
Figure 18 A) Presentation of the experimental task for pitch and time dimensions used in Hyde and Peretz (2003). For the pitch task, the fourth tone was altered by one of five pitch distances, ranging from 25 to 300 cents (100 cents corresponding to 1 semitone). For the time task, the fourth tone was changed by temporal increments from 8 to 16%. B) Hits minus False alarm (FA) rates for amusic and control participants in the pitch condition. C) Hits minus False alarm (FA) rates for amusic and control participants in the time condition. Figure reprinted from Peretz and Hyde (2003), Figure 1.

These data have provided evidence for pitch perception deficits in congenital amusia, notably by showing that amusic individuals exhibit elevated pitch discrimination thresholds (in comparison to controls). However, several studies reported overlap in pitch thresholds between amusic and control groups (for pitch discrimination: Foxton et al. 2004; Tillmann et al., 2009; Jones et al., 2009b; Liu et al., 2010; and for pitch direction: Foxton et al., 2004; Liu et al., 2010; Williamson and Stewart 2010), showing that some amusic individuals have thresholds within the range of control participants (see Figure 17B), and thus suggesting the need to consider additional impaired processes in this condition.

3.3.2. *Memory for Pitch*

A first hypothesis has suggested that amusic individuals' elevated pitch discrimination thresholds would lead to impaired detection of small pitch changes in tone sequences, such as in the melodies of the MBEA (Peretz et al., 2003), the battery used as diagnostic tool for amusia. However, it is relevant to point out that the pitch tests of the MBEA (scale, contour, and interval tests, respectively) also require short-term memory processing as two melodies separated by a silent retention period need to be compared. Several studies have provided evidence for amusic individuals' impairment of short-term memory for pitch, even for amusic individuals without elevated pitch discrimination thresholds, suggesting that the core deficit in congenital amusia might be related to impaired pitch memory (see Figures 17 B and 17C).

Memory deficits for pitch patterns in congenital amusia were explored by Foxton et al. (2004) who asked amusic and control participants to compare two pitch sequences (of four or five pure tones) that were separated by a 1100 ms gap (same/different paradigm). They used three different tasks for which the comparison sequences were either contour violated, contour preserved, or contour violated within a transposed sequence (shifted up or down by half an octave, thus changing the tonal context, participants being requested to ignore the transposition). They demonstrated that amusic individuals were strongly impaired in short-term memory performance in comparison to controls, without overlap in performance between the two groups (thus contrasting with the observed overlap in the pitch discrimination thresholds, see Figure 17B).

Subsequent studies used comparison tasks for single tones (Gosselin et al., 2009; Williamson et al., 2010b) and for tone sequences (Gosselin et al., 2009; Tillmann et al., 2009) to further investigate the extent of the deficit in short-term memory for pitch.

For single tones, these studies asked amusic participants to compare two tones separated by a silent retention period, the delay between the to-be-compared tones being either unfilled or filled with irrelevant tones (Gosselin et al., 2009; Williamson et al., 2010b). While control participants showed good performance for all pitch intervals (1 to 3 tones in Gosselin et al., 2009; and 1 tone in Williamson et al, 2010) and in both conditions (filled, unfilled delays), congenital amusic individuals performed worse than controls when irrelevant tones were presented during the delay. A short-term memory deficit for single tones had been also shown when testing the effect of the duration of the silent retention period between the two to-be-compared tones (0 to 15s): amusic participants showed

faster decline in performance with increasing retention delay than did controls (Williamson et al., 2010). Further evidence for amusic individuals' short-term memory deficit for pitch, suggesting a weaker and more unstable memory trace, was revealed with amusic individuals' sensitivity to memory load: when participants had to compare two tone sequences (separated by a retention interval of 2 s), amusic individuals' performance was lower for five tone-sequences than for three-tone sequences, while controls' performance was only slightly affected by load (Gosselin et al., 2009).

The short-term memory deficit was observed for musical material (tones, timbres), but did not affect spoken words: when requested to compare sequences of tones, timbres, or words (separated by a retention delay of 3 sec), amusic participants performed worse than controls for tone and timbre sequences, but not for words (Tillmann et al., 2009) (see Figure 19).
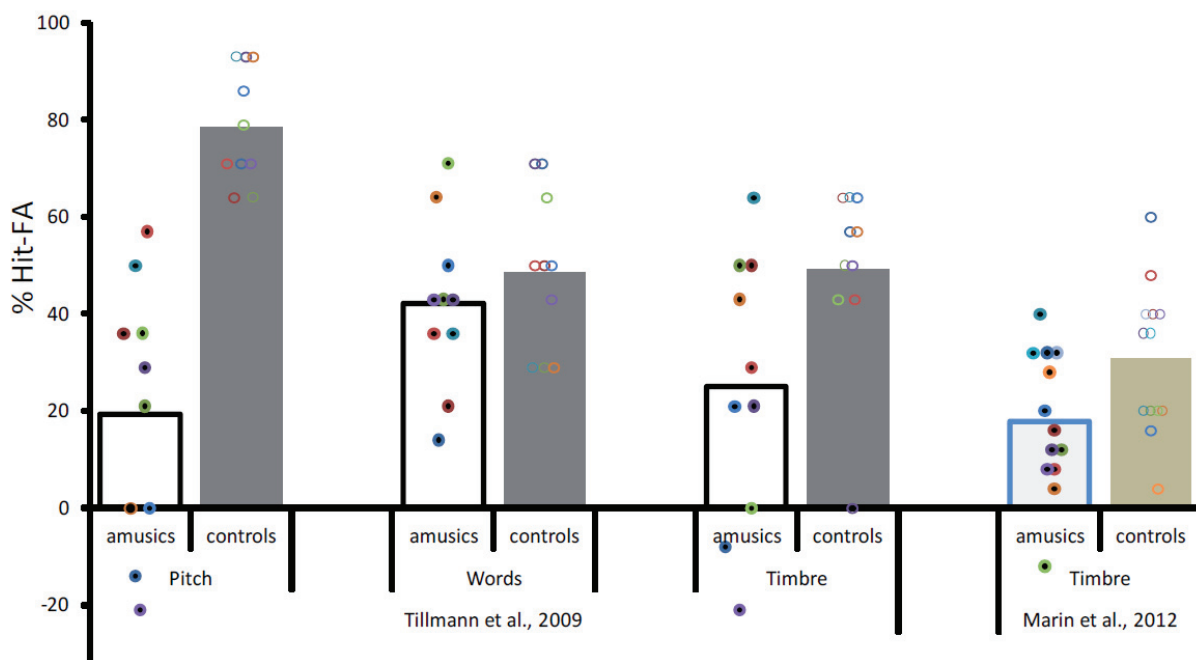
Figure 19 Performance of amusic and control individuals in terms of Hits minus False Alarms (FA), presented as a function of the material used in a short-term memory task (pitch, timbre, words). Left, Data from Tillmann et al., 2009 (10 participants per group); right, data from Marin et al., 2012 (13 participants per group) (collapsed across sequence lengths from 4 to 8 items). Note that for the same sequence length (5 items) as used in Tillmann et al. (2009), average performance of amusic individuals were similar in Marin et al. (2012) (Hits minus FA rates were 25% in Tillmann et al., 2009, and 17% in Marin et al., 2012). Bars indicate group average, and circles individual data points. Note that the amusic group was significantly impaired in comparison to the control group for pitch and timbre tasks, but not for verbal material. While individual data show overlap between the two groups in the timbre short-term memory tasks, no overlap was observed for the pitch task (as in Figure 17C with the same participants of Tillmann et al., 2009).

The observed deficit was stronger for the tone material than for the timbre material (see Figure 19). Impaired memory for timbre was confirmed by Marin et al. (2012) who replicated results of Tillmann et al. (2009) using timbre sequences (composed of four to eight events), and further showed that this impairment is not related to a deficit in timbre discrimination.

The specificity of the deficit for musical material has been further supported by Williamson and Stewart (2010) comparing a digit span test and a tone span test: congenital amusic individuals' had significantly smaller tone spans in comparison to controls, whereas their digit spans were of similar size to those of controls. Importantly, in all these studies, using either isolated tones (Gosselin et al., 2009; Williamson et al., 2010) or tone sequences (Gosselin et al., 2009; Tillmann et al., 2009), the pitch changes exceeded amusic participants' psychophysically measured pitch difference detection thresholds pointing to a memory-based deficit for pitch that extends beyond a perceptual discrimination impairment. In addition, it is worth noting that whereas some amusic individuals exhibited normal pitch discrimination thresholds, all amusic individuals displayed impairments in pitch memory (Figure 17C) (Foxton et al., 2004; Tillmann et al. 2009).

By considering the deficits in fine-grained pitch processing and in short-term memory for pitch, one question now is to determine whether deficits in pitch discrimination and pitch memory constitute separated or combined elements in congenital amusia. One hypothesis would suggest that the pitch discrimination deficit might be related to a pitch memory deficit, but with amusic individuals differing in the extent of this memory impairment. As pitch discrimination measures require the comparison of two or more events (or testing intervals), a strong pitch memory deficit could also affect more perceptual measures, such as discrimination thresholds.

3.4. Neural correlates of congenital amusia

In parallel to the investigation of behavioral expressions of the deficit, brain-imaging studies have started to decipher the cerebral correlates of this neuro-developmental disorder using anatomical and functional investigations. This research contributes not only to our understanding of the deficits underlying congenital amusia, but also furthers our understanding of neural correlates underlying normal functioning of music perception and memory.

3.4.1. Anatomical correlates

The seminal studies have primarily investigated anatomical correlates of congenital amusia, and have explored their link with the behavioral impairments. The first study (Hyde et al., 2006) used VBM in two independent cohorts of amusic individuals to explore whether structural abnormalities for white and grey matter concentration could underlie the disorder. In comparison to controls, the amusic brain showed decreased white matter concentration and increased grey matter concentration in the right Inferior Frontal Gyrus (rIFG, Brodmann Area (BA) 47, see purple circles on Figure 20 upper

panel). To extend this observation with another method, Hyde et al. (2007) used cortical thickness measures and revealed abnormalities in the amusic brain in terms of grey matter integrity in the rIFG (BA47) and in the right Superior Temporal Gyrus (rSTG): amusic individuals showing thicker cortex than controls in these regions (Hyde et al., 2007; see green circles on Figure 20 upper panel).

These group differences were also supported by correlation analyses between anatomical measures and behavioral data: the performance in musical tests (scale and memory test of the MBEA) correlated positively with white matter concentration in the rIFG (Hyde et al., 2006); and the global score of the MBEA correlated negatively with cortical thickness in the rIFG (and rSTG) (Hyde et al., 2007). In another sample of amusic individuals (Mandell et al., 2007), positive correlation between MBEA scores and grey matter concentration were observed in the left IFG (BA47, for the melodic subtests) and left STG (BA22, for the rhythmic subtest) (see orange circles on Figure 20, upper panel).

The findings suggesting abnormalities in white and/or grey matter concentrations are in line with structural anomalies in adult brains having experienced long-lasting deficiencies such as language and speech disorders (Watkins et al., 2002). It has thus been argued that the observed abnormalities in the amusic brain have a genetic basis (Peretz et al., 2007), leading to cortical malformations that may compromise the normal development of a right fronto-temporal pathway. In agreement with this hypothesis, abnormalities of white matter have been observed with DTI, showing reduced right Arcuate Fasciculus (see Chapter I, page 31) connectivity in amusic individuals (Loui et al., 2009). In addition, the volume of the superior Arcuate Fasciculus branches correlated positively with pitch discrimination abilities in controls and amusic individuals (see the region of interest used by Loui et al. (2009) to perform the tractography on Figure 20 upper panel).

The observed structural differences suggest cortical malformations of a right fronto-temporal pathway in the amusic brain resulting in pitch-processing impairment. This hypothesis is consistent with neuroimaging data with typical individuals showing that the right superior temporal cortex is involved in perceptual analyses of melodies and that pitch processing recruits neural networks including the right frontal cortex (Griffiths et al., 1999; Maess et al., 2001; Peretz and Zatorre, 2005; Tillmann et al., 2003; Tillmann et al., 2006; Zatorre et al., 1994) (see Figure 20, lower panel).
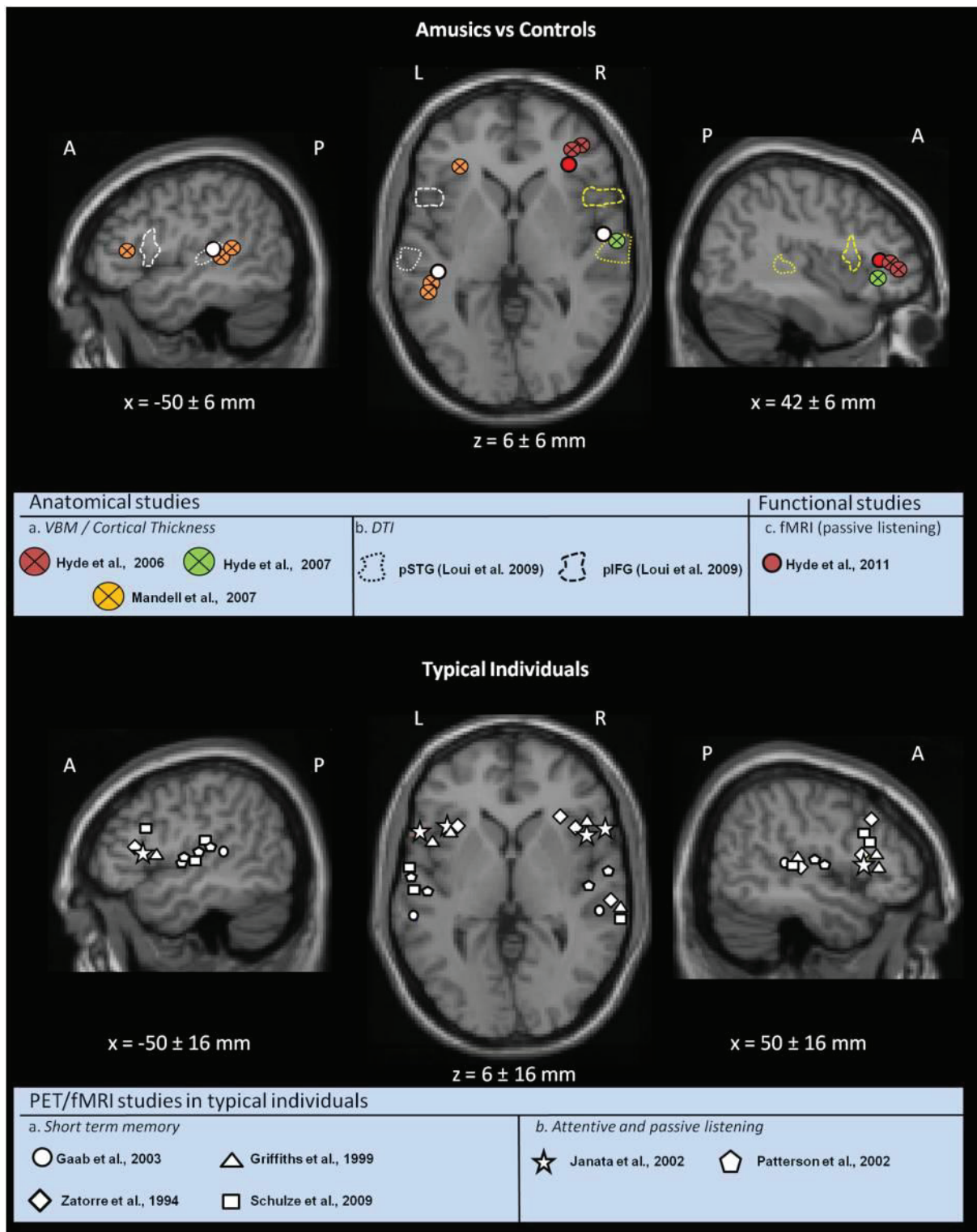
Figure 20 Upper panel: Brain regions observed in anatomical and functional studies comparing control participants and amusic individuals (white: no group difference; color: group difference). a) Coordinates of the anatomical group differences observed with VBM (Hyde et al., 2006 (purple); Mandell et al., 2007 (orange); and Cortical thickness (Hyde et al., 2007 (green)) approaches. b) Approximate location of the posterior inferior frontal gyrus (pIFG) and posterior superior temporal gyrus (pSTG) seed regions used in the tractography study of the arcuate fasciculus by Loui et al. (2009) (bold dotted lines and plain dotted line, respectively). Note that these regions were manually

selected by a neurologist who was blind concerning the protocol and previously reported activations. Group differences were observed only in the right hemisphere (yellow color). c) Coordinates of the functional group differences observed with fMRI in the right inferior frontal gyrus (Hyde et al., 2011, red circle), and of areas in the auditory cortices where both amusic individuals and control exhibit similar BOLD activity to pitch variations (Hyde et al., 2011, white circles).

Lower panel: activation peaks observed in studies using PET (Positron Emission Tomography) or functional MRI to investigate music processing in typical individuals. Coordinates of activations obtained with pitch material for a) auditory short-term memory tasks (Zatorre et al., 1994, diamonds; Griffths et al., 1999, triangles; Gaab et al., 2003, circles; Schulze et al., 2009, squares); b) attentive listening (Janata et al., 2002, stars) and passive listening (Patterson et al., 2002, pentagons).

All activations are displayed on the single subject T1 template provided by SPM8. The figure displays activations with coordinates that were 6 mm up or down the central slice and 6 mm left or right of the lateral slices for the upper panel, and ± 16 mm for the lower panel. R = right; L = left; A = anterior; P = posterior).

3.4.2. *Functional correlates*

To further link the anatomical abnormalities to the behavioral deficit, recent studies investigated the cerebral correlates of pitch perception and memory in congenital amusia on a functional level (using EEG and fMRI). In the first EEG study (Peretz et al., 2005), amusic and control participants performed a pitch change detection task during EEG recording, and auditory evoked responses were analyzed for the deviant and standard tones. Amusic individuals' N1 component was normal and its topography was consistent with a localization of the generators in the secondary auditory cortex (Näätänen and Picton, 1987). Investigating the neural processing of musical pitch incongruities in congenital amusia, Peretz et al. (2009) showed that the amusic brain can track small pitch changes (i.e., a quarter tone), exhibiting an early right-lateralized negative brain response, even though the amusic participants mostly failed to report explicitly these pitch incongruities. Moreover, a recent Mismatch Negativity (MMN) study (Moreau et al., 2013) provided congruent evidence of normal pre-attentive pitch change detection in amusic individuals.

These findings suggest near-normal cerebral processing of musical material in the amusic brain, notably showing a near-normal functioning of the auditory cortex, despite explicit deficits on pitch tasks at a behavioral level. This was further supported by fMRI data showing that amusic individuals' auditory cortices respond normally to pitch during passive listening to pure-tone melody-like patterns: activity increased in the auditory cortex as a function of increasing pitch distance (from 0.25 to 2 semitones) (Hyde et al., 2011). In addition, abnormal BOLD deactivation in the rIFG was associated with decreased right fronto-temporal functional connectivity between the rIFG and the rSTG in amusic individuals in comparison to controls. These findings thus led to the hypothesis that the functional correlates in congenital amusia are related to impaired anatomical *and* functional fronto-temporal connectivity, with intact functioning of the auditory cortex.

These findings are consistent with the involvement of right temporo-frontal regions in pitch and melody processing, as observed in normal participants with functional cerebral imaging (Gaab et al., 2003; Griffiths, 1999; Griffiths et al., 1999; Janata et al., 2002a; Koelsch et al., 2005; Koelsch et al., 2009; Peretz and Zatorre, 2005; Schulze et al., 2009; Schulze and Koelsch, 2012; Schulze et al., 2011a; Schulze et al., 2011b; Tillmann et al., 2003; Tillmann et al., 2006; Zatorre et al., 2002; Zatorre et al., 1994) (see Figure 20 lower panel) and are congruent with anatomical anomalies observed along the auditory-frontal pathway in the amusic brain (Hyde et al., 2007; Hyde et al., 2006; Hyde et al., 2011; Loui et al., 2009; Mandell et al., 2007) (see Figure 20 upper panel).

However, it should be noticed that all these previous studies in amusia have explored only passive listening or simple detection tasks to investigate the functional correlates of the disorder, and, to our knowledge, none has explored the cerebral correlates of the short-term memory deficits in congenital amusia. Additionally, observing intact functioning of the auditory cortex in this disorder is surprising, notably by considering the particularly important role of this region in pitch perception and memory (see Chapter I, section 1.5.3 pages 38 to 44; Chapter II, section 2.4.2 pages 75 to76). Within this framework, one purpose of the present PhD work is to explore these mechanisms, first to improve our understanding of the deficits underlying congenital amusia, and second to propose further insights on the brain mechanisms supporting perception and short-term memory for pitch in typical individuals.

3.5. Are the processing deficits in congenital amusia pitch specific?

The dominant hypothesis of regarding the core deficit in congenital amusia refers to the pitch dimension, notably with impaired pitch perception and memory. This focus can be explained by the fact that pitch information is the main form-bearing dimension in Western tonal music. Some research has addressed amusic individuals' processing capacities for other types of information, either related to music (temporal structures) or pitch processing related to other materials (speech, space). Of interest for the present PhD work, these studies thus allow characterizing whether amusic individuals' deficits are widespread or in contrast specialized to music processing. The latter case would confirm that congenital amusia represents a unique condition allowing for the study of the behavioral and cerebral correlates of perception and memory of tonal (impaired) and verbal (preserved) in the human brain.

3.5.1. Temporal processing

In addition to the pitch dimension, the second form-bearing dimension of the Western tonal musical system is the time dimension, that is, the rhythmic and metrical structures (McAdams, 1989). Temporal processing is also tested in the MBEA, revealing that in contrast to the subtests of the pitch dimension (scale, contour, interval), performance in the temporal subtests (i.e., rhythm and meter

processing) is more variable, with about half of the amusic individuals showing a deficit (Ayotte et al., 2002; Peretz et al., 2003).

Some amusic individuals are impaired in discriminating melodies by their rhythms, and have difficulties in keeping time to a musical beat (Ayotte et al., 2002). Ayotte et al. (2002) proposed the hypothesis that this deficit in temporal processing might be a cascade effect of impaired pitch processing and not a temporal processing deficit per se. This hypothesis has received partial support: when tested for temporal processing in a non-musical context (i.e., detecting a time deviation from isochrony in a monotonous sequence), amusic individuals were as accurate as control participants – even though these same amusic individuals were severely impaired when testing pitch processing in a nonmusical context (detecting a pitch deviation from monotony in an isochronous sequence, see Figure 18) (Hyde and Peretz, 2004). Converging evidence has been reported by another study (Foxton et al., 2006): When amusic individuals had to compare two sound sequences, which were either monotonic or randomly varying in pitch, for the presence of a changed temporal interval, the pitch changes interfered with the processing of the rhythmic patterns. Amusic individuals' performance was decreased for the sequences that varied in pitch compared to the sequences that did not, while controls' performance did not differ between the two conditions. Pitch and time processing were then further investigated in a non-musical context, suggesting that amusic individuals also have a deficit in the integration of pitch and time information in a combined mental representation (Pfeuty and Peretz, 2010).

While cases of acquired amusia have shown selective deficits of either the pitch or the time dimension (Peretz and Kolinsky, 1993), selective deficits on the time dimension have been reported only recently for congenital amusia (Phillips-Silver et al., 2011; Thompson, 2007). Phillips-Silver et al. (2011) reported a case study with a deficit in beat finding in the context of music, while showing near-normal synchronization with a metronome (i.e., no pitch changes). The participant (who did not show any other cognitive or perceptual deficits, without any neurological or psychiatric anomalies) scored in the normal range of the MBEA, except for the meter subtest. His data pattern led the authors to propose a new form of congenital amusia, based on time deficits, which calls for future investigations and extensions to group studies.

3.5.2. *Speech processing*

More interestingly for the proposal of the present PhD work, both acquired and congenital amusias represent an opportunity to investigate the hypothesis of the modularity of music and language processing (Peretz, 2002; Peretz et al., 1997; Piccirilli et al., 2000): are neural correlates independent or shared for the processing of these two structured systems?

Individuals with congenital amusia have been reported to be unimpaired in language tasks, such as learning and recognizing lyrics (Ayotte et al., 2002; Peretz et al., 2002) as well as short-term memory tests for words or digits (Tillmann et al., 2009; Williamson and Stewart, 2010). In earlier studies, amusic individuals have also been reported as unimpaired for prosody tasks, such as

Chapter III: Are the processing deficits in congenital amusia pitch specific?

classifying a spoken sentence as statement or question based on final (falling or rising) pitch information as well as identifying or discriminating stressed words in sentences (Ayotte et al., 2002; Patel et al., 2005; Peretz et al., 2002). More recent studies, however, revealed deficits also for prosody processing in amusia (see below), as it was previously observed in cases of acquired amusia (Patel et al., 1998; see Chapter III page 80).

Because of the initially reported spared intonation processing, Peretz and Hyde (2003) suggested that the difference between speech and music perception might be linked to the size of relevant pitch variations (Figure 21). In speech intonation of non-tonal languages, variations in fundamental frequency (F0), in particular those indicating statements and questions, is typically coarse (up to more than 12 semitones for the pitch rise of the final word in a question) (Patel et al., 2008). In music, however, the pitch variations are typically more fine-grained (1 or 2 semitones) (Vos and Troost, 1989). Accordingly, amusic individuals' pitch deficit would affect music more than speech because music is more demanding in pitch resolution, and congenital amusia would represent a music-relevant deficit, not necessarily a music-specific deficit.
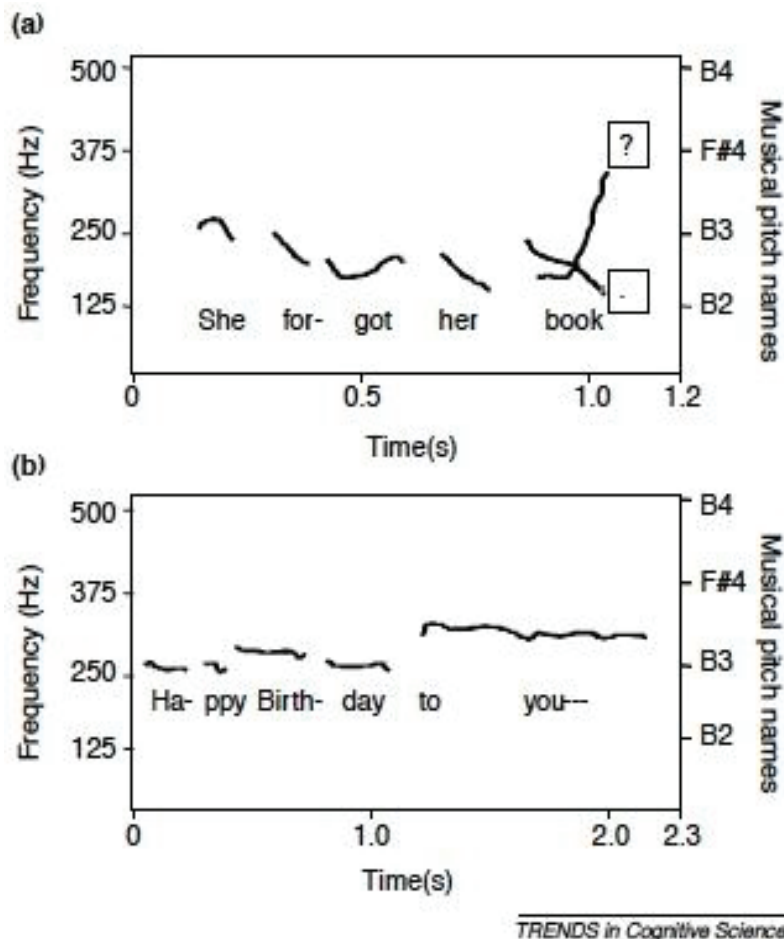

Figure 21 A) Spectrograms illustrating pitch variations in two sentences (a question, a statement) spoken by a female English speaker. B) Spectrogram illustrating the pitch variations in the song "Happy Birthday" sung by the same voice as spoken material in (A). Figure reprinted from Peretz & Hyde (2003), Figure 2.

However, amusic individuals' deficit persisted for large pitch differences when these were derived from speech intonation patterns and presented as musical analogues (imitating gliding pitch changes or transformed into discrete steps). Amusic individuals were impaired when asked to discriminate imitations of spoken intonation (without words). The deficit was limited to the non-speech context because amusic individuals were normal at discriminating the same changes in the context of speech (Ayotte et al., 2002; Patel et al., 2005). Ayotte et al. (2002) suggested a beneficial effect of the linguistic cues that could serve as anchoring points for poor pitch processing abilities. Another possible account for this speech advantage is that labeling the word with the salient pitch change might provide a strategy that decreases the memory load of the task – a strategy that is not possible for the tone analogues. In the latter case, the pitch memory problem experienced by amusic individuals (Gosselin et al., 2009; Tillmann et al., 2009; Williamson et al., 2010) may explain why they fail to discriminate intonation patterns in musical analogues even though they contain relatively large pitch changes.

The hypothesis that amusic individuals cannot fully compensate for their pitch deficit by using speech-based strategies is supported by other data sets, revealing subtle deficits in speech processing. Amusic individuals showed mild deficits in processing speech intonation (questions vs. statement) in their native language (English or French) (Liu et al., 2010; Patel et al., 2008), or processing pitch contrasts in tone language words (Mandarin or Thai) - whether for native Mandarin speakers (Jiang et al., 2010, 2012; Nan et al., 2010) or native French speakers (Nguyen et al., 2009; Tillmann et al., 2011a). When tested with natural speech, which involves multiple acoustic cues, Mandarin amusic individuals were not impaired (Liu et al., 2012a; Liu et al., 2012b). Note also that it has been recently suggested that amusic individuals who are native speakers of Hong Kong Cantonese show better pitch processing abilities than amusic individuals who are native speakers of non-tonal languages (English, French) (Wong et al., 2012). For non-tonal language speakers, it has also been shown that amusic individuals performed better with speech than with musical analogues, especially among those amusic individuals with high pitch discrimination thresholds (over one semitone; Tillmann et al., 2011a) (Figure 22). Thus, speech may enhance pitch processing in amusia, even though it does not necessarily restore normal performance levels.
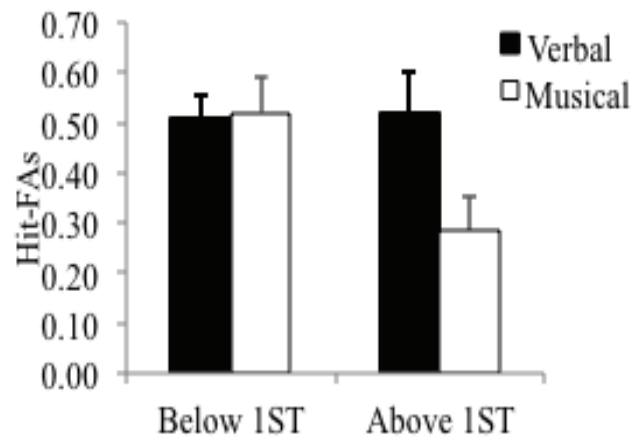
Figure 22 Performance [expressed as Hits − False Alarms (FA)] for amusic participants in Experiment 2 of Tillmann et al. (2011a) tested with Thai speech material and its musical analogs. Amusic participants were separated into two groups: eight amusic individuals with thresholds below one semitone (mean of 0.68, ranging from 0.13 to 0.97) and seven amusic individuals with thresholds above one semitone (mean of 2.41, ranging from 1.3 to 4). For amusic individuals with thresholds above one semitone, performance was significantly better for the verbal material than the musical material; six of the seven amusic showed higher mean performance for the verbal material. This advantage was not observed for amusic individuals with thresholds below one semitone.

As the main deficit of congenital amusic individuals seems to concern pitch processing, amusic individuals have been tested not only for the processing of speech intonation, but also for the processing of emotional processing, which was shown to be impaired too (Thompson et al., 2012). Pitch processing capacities and their deficits have been recently linked to phonological and phonemic awareness abilities (Jones et al., 2009a; Loui et al., 2011), leading to the hypothesis that there might be shared neural basis for pitch-related impairments in amusia/tone-deafness and phonemic awareness in dyslexia (Loui et al., 2011) (phonemic awareness refer to the ability to manipulate phonemes and syllables in spoken words).

3.5.3. Space processing

As exposed above, congenital amusia is thought to result from a pitch-processing deficit. This prevailing view has been challenged by Douglas and Bilkey (2007) who reported that amusic individuals were impaired in a classic mental rotation task and were less influenced by the spatial layout of response keys in a pitch-judgment task, compared to non-amusic participants (Douglas and Bilkey, 2007). The hypothesis of a link between pitch and space processing has been previously supported by interactions between pitch processing and visual space representations (e.g., stimulus-response compatibility effects) (Lidji et al., 2007; Rusconi et al., 2006) as well as enhanced visuo-spatial processing in musicians (Brochard et al., 2004). According to Douglas and Bilkey (2007), amusic individuals' deficit of spatial processing suggests that pitch processing

". . . depends on the cognitive mechanisms that are used to process spatial representations in other modalities" (page 915).

According to this view, amusic individuals' disorder might arise as a consequence of a general spatial-processing deficit. However, this study has been criticized, leading to a call for:

". . . a clear need for replication and extension before such a view is accepted" (page R893) (Stewart and Walsh, 2007).

Two further studies investigating spatial processing in congenital amusia failed to show any spatial processing deficit in congenital amusic individuals with a variety of spatial tasks (e.g., Shepard-Metzler mental rotation task, bisection task, Corsi block task, visual patterns test) (Tillmann et al., 2010; Williamson et al., 2011). These findings thus suggest that amusic individuals' deficit in pitch processing does not co-occur with a deficit in spatial processing, and further suggest largely independent pitch representations and visuo-spatial representations (see Chapter II, section 2.2.1 pages 52).

3.6. Pitch perception and production

Congenital amusia is characterized by pitch perception and memory deficits, more or less strongly accompanied by deficits along the time dimension (i.e., deficits in the processing of rhythm or meter). These perception and memory deficits impact most strongly on musical material, with some consequences also in the speech realm.

However, the colloquial understanding of amusia is that people suffering from this condition sing notoriously out-of-tune. Some studies have thus attempted to characterize musical production deficits in congenital amusia: Initial studies of musical production have shown that amusic participants are impaired when asked to synchronize to music in tapping tasks (Ayotte et al., 2002; Dalla Bella and Peretz, 2003). The vast majority of studies investigating musical production in amusia have focused on singing abilities. As expected, most of the amusic individuals exhibit poor singing, as assessed both by peer ratings (Ayotte et al., 2002) and acoustic analyses of vocal performance (Dalla Bella et al., 2009; Hutchins et al., 2010a; Loui et al., 2008; Tremblay-Champoux et al., 2010; Williamson et al., 2012). Singing impairments are more pronounced along the pitch dimension than the time dimension (Dalla Bella et al., 2009; Tremblay-Champoux et al., 2010), reminiscent of the pattern observed for perceptual deficits. Along the pitch dimension, most amusic individuals make interval and contour errors (Dalla Bella et al., 2009) and perform poorly on pitch matching tasks (Hutchins et al., 2010), whereas about half of them sing on time (Dalla Bella et al., 2009). Amusic individuals are extremely impaired when requested to sing a song without lyrics (e.g. on /la/), with for example half of the

amusic sample of Dalla Bella et al. (2009) being unable to produce complete performances on /la/. These data thus suggest a benefit of the language dimension for the musical production.

It should be noted that not all poor singers suffer from perceptual pitch deficits: poor singing is estimated around 10-15% of the general population (Dalla Bella et al., 2012), and congenital amusia per se is estimated at about 4% of the general population (e.g., Kalmus & Fry, 1980; see above), thus excluding perceptual deficits as the sole possible origin of poor singing. Poor singing without pitch perception deficits is sometimes referred to as "vocal amusia" or "oral-expressive amusia". Such a condition might arise from failures at any component of a "sensori-motor vocal loop": auditory-motor integration, motor production, short- and long-term memory (Berkowska and Dalla Bella, 2009). Experiments with a "slider" allowing to perform pitch-matching with a non-vocal motor output supported the assumption of a variety of causes for poor pitch singing (Hutchins and Peretz, 2012b), notably poor vocal motor control and difficulties to match a target pitch when the model has a different timbre. It is generally assumed that in congenital amusic individuals, poor pitch singing arises from deficits in pitch perception and memory, which render difficult to adequately monitor the acoustic feedback from the individual's own voice. In support of this claim, Dalla Bella et al. (2009) found that poor pitch singing was correlated with the ability to detect pitch changes (as tested in Hyde and Peretz, 2004; Figure 18). Memory components seem also important in singing performance, as the limited ability of amusic individuals to sing a song without lyrics is correlated with the score in the memory subtest of the MBEA (Dalla Bella et al., 2007; Tremblay-Champoux et al., 2010).

Most interestingly, studies of vocal performance have revealed that some amusic individuals exhibit normal singing despite their pitch perception deficit (Ayotte et al., 2002; Dalla Bella et al., 2009; Tremblay-Champoux et al., 2010). Taken together, these observations in poor singers (pure vocal amusia) and congenital amusic individuals suggest a dissociation between pitch perception and production. The most striking result supporting this claim was the report that congenital amusic individuals can accurately reproduce the direction of pitch intervals whereas they cannot classify the same intervals as being rising ("up") or falling ("down") (Loui et al., 2008). However, a replication study with a larger sample of amusic participants (Williamson et al., 2012), led to a more nuanced result pattern: only two out of 14 amusic participants had better production than perception thresholds, and furthermore, amusic participants performed differently when the perceptual task used (or not) the "up/down" labels to designate the pitch intervals. Using a different approach, based on the "pitch-shift reflex", which is an automatic pitch adjustment following an altered feedback of one's own voice, Hutchins and Peretz (2013) found that four out of nine amusic individuals exhibited pitch shifts similar to that of controls after both small and large pitch changes. They labeled this sub-group "Production-Preserved (Congenital) Amusic individuals". Pitch shifts were predicted by pitch matching accuracy (Production), but not pitch discrimination abilities (Perception), which was interpreted as further evidence for a dual-route neural architecture for the processing of pitch (see below).

Further support for a dissociation between pitch perception and production has been reported for speech materials. While amusic individuals are impaired at imitating pitch contours in statements and questions, their production abilities exceed their perceptual abilities with the same sentences (Liu et al., 2010). Two other studies point to an even clearer dissociation between amusic individuals' pitch perception and production abilities for speech: in a first study amusic speakers of a tone language (Mandarin) were impaired in distinguishing lexical tones, but not in their production Nan et al. (2010), in a second study amusic individuals were impaired only when discriminating but not imitating pitch variations embedded in sentences (Hutchins and Peretz, 2012a).

Overall, pitch perception and production seem to be partly decoupled in congenital amusia, although it does not seem possible to advocate a complete dissociation. The cerebral underpinnings of this partial dissociation are not yet understood, but it has been proposed that several pathways exist for pitch encoding, resulting in an abstract code (more consciously accessible, see Chapter II, section 2.2.1 pages 58) and a vocal-motor code (Hutchins and Peretz, 2012a, 2013). Compensation through somatosensory cues has also been proposed to account for intact production abilities (Nan et al., 2010). The findings of anatomical abnormalities in auditory and frontal areas (Hyde et al., 2006, 2007; Mandell et al., 2007), as well as a reduced fiber connectivity between these areas (Loui et al., 2009), has first been taken as evidence that congenital amusia is an auditory-motor feedback disorder (Mandell et al., 2007). However, this hypothesis cannot easily account for all types of congenital amusia, notably when there are preserved pitch production abilities (Hutchins et al., 2010b).

It remains a challenge for future studies to clearly establish the links between pitch perception/production abilities and abnormalities in auditory and frontal areas and their connections.

Finally, recent research has started to investigate whether poor pitch-singing in congenital amusia might be improved by training. Amusic individuals seem to benefit from singing in unison when singing without lyrics (Tremblay-Champoux et al., 2010), and from a seven-session training program, leading to improved vocal performance of familiar songs (Anderson et al., 2012).

3.7.   Implicit pitch processing in congenital amusia

The partial dissociation between pitch perception and production has sometimes been taken as evidence of some implicit processing of pitch and its representations in congenital amusia (e.g., Dalla Bella et al., 2009). Similarly, preserved emotional judgments (Gosselin et al., 2011; Paquette et al., 2011) (posters presented at The Neurosciences and Music-IV Learning and Memory Conference; unpublished data) have also been interpreted in line with this hypothesis, although musical emotions could also be based on other musically-relevant information than pitch (rhythm, timbre, tempo, etc.).

More generally, pitch processing has been further characterized in congenital amusia using implicit investigation methods. Such approaches are of interest because they are thought to have the potential to reveal unconscious or degraded representations (e.g., knowledge, memory) in the absence of awareness, thus contrasting with explicit knowledge, which can be consciously recollected and verbalized. Indirect investigation methods have provided evidence for the influence of spared implicit knowledge on perception and action in the presence of severe impairments in tasks requiring explicit processing. The dissociation between implicit and explicit task performance has been reported in various neurological disorders (e.g., alexia, agraphia, aphasia, prosopagnosia), including acquired and congenital amusia. The priming paradigm is one of the implicit investigation methods and can be used to probe participants' musical knowledge (Tillmann, 2005).

In the musical priming paradigm, a musical context (e.g, a chord sequence) is used to generate expectations about upcoming events. These musical expectations arise from the regularities of the musical system to which listeners typically become sensitive via mere exposure to music of their culture (Tillmann et al., 2000). Performance (notably speed of processing) in an unrelated task (e.g., judgments on the tuning, created sensory dissonance, or used timbre) concerning the last chord of a chord sequence (i.e, the target chord) depends on the musical relatedness of this chord with its preceding context, with related targets leading to faster processing. This paradigm proved efficient in revealing implicit musical knowledge in a case of acquired amusia following a stroke (Tillmann et al., 2007) and also in congenital amusia (Tillmann et al., 2012): The musical priming paradigm revealed amusic individuals' implicit sensitivity to harmonic relationships between chords belonging to a given key.

Further evidence for implicit pitch representations in congenital amusia was found using a melodic priming paradigm (Omigie et al., 2012b). The probability of notes embedded in melodies (that is, their expectedness based on various dimensions, including melodic contour, tonal structures etc.) modulated performance in an incident timbre task in amusic individuals and controls. However, amusic individuals were impaired relative to controls for explicit judgments of the expectedness of target notes in the same melodies.

Note however, that not all attempts to find evidence for implicit pitch representations in congenital amusia were successful: Pfeuty and Peretz (2010) found no evidence of implicit pitch processing in a pitch-time interference task.

Additional evidence for unconscious pitch-tracking abilities in the amusic brain comes from ERP studies. Peretz et al. (2009) showed that amusic participants exhibit reliable ERPs following tones mistuned by a quarter-tone in tonal melodies, although they were not able to explicitly report the mistuning. More specifically, when controls exhibited both an early (N200) and a late (P600) ERP component following quarter-tone mistunings or out-of-key notes, amusic individuals only exhibited the early deviance-selective response, mainly for mistuned (but not out-of-key) notes (Peretz et al., 2009). Converging evidence for unimpaired early stages of auditory processing in passive listening paradigms with pitch deviances was obtained in an MMN paradigm with deviances as small as an eighth of a tone (Moreau et al., 2013), with amusic participants exhibiting normal MMN despite very poor explicit detection of deviants. This study also revealed impaired long-latency deviance ERPs (P300) in congenital amusia, even when the participant could explicitly detect the deviant tones (but see Braun et al. for a different result pattern (Braun et al., 2008)). Recently, Omigie et al. (2013) have explored amusic individuals' brain abilities in tracking pitch probabilities in melodies (degree of expectedness of a tone within a melody, thus exploring their long-knowledge on musical structures) with electrophysiological approaches. They investigated whether, within a context of real melodies having a given structure (context), the amusic brain's responses can differ from that of controls when processing tones with low probability of expectedness in the melody in comparison to tones with high probability of expectedness. Their data revealed that both amusic and control groups presented brain responses to tones that did not respect the melodic expectations given by the melodic context, and that these responses were modulated by the degree of expectedness of these tones (delayed P2 component, see Chapter II) (Omigie et al., 2013).

Evidence for intact implicit pitch representations in congenital amusia, as revealed by priming paradigms and passive ERP protocols, together with the behavioral expressions of the deficit, raise questions about the extent to which the learning mechanisms allowing the brain to extract the regularities in the auditory environment, most notably in music and language, are preserved in congenital amusia. This internalization of rules is thought to occur involuntarily through implicit learning, also referred to as statistical learning (Saffran, 2003; Tillmann et al., 2000). In congenital amusia, statistical learning studies with new artificial musical systems have yielded mixed results. Omigie and Stewart (2011) reported intact statistical learning capabilities in amusic individuals for both linguistic (Peretz et al., 2012), and tone materials (Omigie and Stewart, 2011), albeit amusic individuals exhibited less confidence in their responses, again suggesting a dissociation between their implicit capacities and their explicit knowledge or confidence. However, other studies (Loui and Schlaug, 2012; Peretz et al., 2012) failed to report statistical learning for various types of tone materials in amusic participants

.

Discrepancies in data sets might be due to the different definitions of the test phases: Omigie and Stewart (2011) used more salient violations to study the acquired artificial knowledge than did Peretz et al. (2012).

The data sets reported above point to some spared implicit processing of the pitch dimension in congenital amusia, as revealed by the musical priming paradigm, ERPs, and –in some cases– statistical learning paradigms. The studies on implicit processing of pitch, tonal structures as well as on implicit learning suggest that congenital amusic individuals can also have acquired partial musical knowledge about statistical regularities of the musical system. Their implicit discrimination performance might be sufficient to capture some of the regularities of associations between pitches (see Chapter I, section 1.5.2 pages 37), leading to some tonal knowledge (including the harmonic relations tested in the musical priming paradigm), even if this knowledge remains sparse and functional only at an implicit level.

Conclusion

In the present chapter, we reviewed research on congenital amusia with its deficits in music perception and production as well as the potential extensions of the deficit to the processing of other materials (in particular, speech). Research on congenital amusia investigating the underlying neural correlates provides further insights not only into the impaired neural networks of music processing for this particular disorder, but integrates also in research investigating the normal functioning of the non-amusic brain. The main deficit that has been studied for this phenomenon relates to the pitch dimension, in particular to pitch perception and pitch memory. Depending on the studies, the current emphasis has been on either pitch discrimination, pitch direction processing, or pitch memory. A unifying hypothesis could suggest the pitch memory deficit as the basis of the impairment as memory capacities are required in both pitch discrimination as well as pitch direction tasks (i.e, comparing tones presented in a given trial). However, future research is necessary to confirm this hypothesis and/or disentangle the contribution of the different capacities to eventually define different subtypes of congenital amusia.

While most amusic cases are impaired on the processing of the pitch dimension, some amusic cases are also impaired on the time dimension – and recent research further suggests amusic cases solely impaired on the time dimension. These observations thus contribute to the suggestion that 'congenital amusia' is not a unified disorder, but might be a more general terminology to cover various forms of music perception and production deficits.

The subtypes of congenital amusia might be further revealed by self-reports as acquired by questionnaire-based data (McDonald and Stewart, 2008; Omigie et al., 2012a; Peretz et al., 2008). For example, recent research on musical experience in congenital amusia suggests that amusic individuals might differ in their levels of music perception abilities as well as their music appreciation (and thus their music engagement in everyday life, leading to different attitudes towards music) (McDonald and Stewart, 2008; Omigie et al., 2012b). Subjective reports will help us to develop the investigation of

further refinements of different subtypes of amusia, and also to indicate future directions to investigate whether congenital amusia might be related to other deficits (e.g., dyslexia) and/or have consequences for the processing of other materials (as partially addressed in research on spatial processing). Finally, the investigation approach based on music experience in everyday life further shows the need to bring this research area back to more real music (not just highly controlled and thus reduced material focusing for example on the pitch dimension) and emotional experience with music.

In sum, research on congenital amusia is a currently growing domain with the aim to understand a deficit that has been known for a long time, but was missing systematic investigation. In the context of the present PhD work, studying the phenomenon of congenital amusia contributes not only to the understanding of the disorder itself, but also of normally functioning cognitive and neural correlates related to music processing, in particular pitch perception and memory.

Objectives

Studies presented in Chapters I, II and III have described our current understanding of the behavioral and cerebral correlates of auditory perception and memory for pitch (or tonal) material, and have compared it to the processing of verbal material. However, these studies have pointed out several unresolved questions.

Studies presented in Chapter I have proposed the existence of hierarchical organizations in the brain for the processing of auditory information, and have shown that specific subparts of the human auditory system are related to pitch processing. However, while most of these studies have focused on the role of isolated cortical regions, it remains an open question about how these regions communicate together to support cognition.

This latter point has also emerged in studies described in Chapter II, which proposed that different forms of memory exist and that these diverse forms could be supported by different brain networks. More precisely, we described the cortical network (involving fronto-temporo-parietal regions) implicated in auditory short-term memory processes. This chapter has also pointed out that the studies exploring the cerebral correlates of auditory short-term memory for tonal and verbal materials had difficulties in defining whether these materials recruit or not the same brain networks. While several studies have proposed that these materials recruit a common cortical network, several studies have challenged this view. This later point was further discussed in Chapter III, by the description of both acquired and congenital specific music-related impairments. This last introductive chapter presented our current knowledge concerning congenital amusia, a lifelong disorder of music perception and production.

Studies in Chapter III showed that the most notable impairments in congenital amusia seem to concern the pitch dimension in particular, the short-term memory processing for pitch. By considering this fact, and by noting that short-term memory for verbal material seems preserved in this disorder, the study of congenital amusia offers a unique occasion to explore the cortical networks underlying music processing and the relationship between music and language processing in the brain. Furthermore, this chapter showed that the description of the cerebral correlates of this deficit is in a preliminary phase, and that further investigations are necessary to decipher the functional underpinnings of this disorder.

Based on these first conclusions, the first objective of the present PhD was to decipher the behavioral and neural correlates of auditory perception and memory for musical sequences, and verbal material by investigating these processes in congenital amusia as compared to the normally functioning brain. This was done by three studies using multimodal neuroimaging methods (VBM, MEG, fMRI) together with behavioral approaches. Study 1 combined MEG and structural MRI recordings to test the hypothesis of functional and anatomical abnormalities in amusic individuals' brain related to pitch perception and memory deficits. To further explore the cerebral correlates of the

deficits in the pitch dimension, Study 2 investigated short-term memory for tones and verbal material in congenital amusic individuals and their matched controls with fMRI. This study thus additionally assessed whether amusic individuals' cerebral network underlying normal-level performance with verbal material is similar to the network engaged by matched control participants. Additionally to the conclusions that these studies allow for the understanding of congenital amusia itself, they also allow testing the potential existence of a specialized memory system in the brain related to tonal material (by comparing amusic individuals and matched-controls data for both verbal and tonal memories).

Study 3 tries to pursue the characterization of amusic individuals' deficits by specifically exploring their pitch encoding abilities. We manipulated sound characteristics that have been previously shown to facilitate the encoding of the information (Demany and Semal, 2008) to assess whether amusic individuals could benefit (or not) of such manipulations in pitch discrimination and short-term memory tasks.

Based on conclusions made on these three studies, the second part of this PhD aimed to explore how and whether congenital amusic individuals' pitch perception and memory abilities could be boosted.

Study 4 explored previously reported implicit capacities of amusic individuals related to the processing of pitch and tonal structures, to investigate whether such implicit knowledge could boost their short-term memory abilities for tones. Finally, Study 5 is related to the conclusions made on Study 3 and aimed to boost pitch-encoding abilities in congenital amusic individuals, by proposing a new approach in the domain, which consisted in exploring the potential effectiveness of audio-visual interactions.

All together, these studies aim to improve our understanding about auditory and music processing in general, by exploring the behavioral and cerebral correlates of pitch perception and short-term memory congenital amusic individuals as compared to typical listeners.

Chapter IV Impaired pitch perception and memory in congenital amusia: The deficit starts in the auditory cortex

Objectives

The investigation of the cerebral correlates of congenital amusia has started only recently (Chapter III, section 3.4 pages 90 to 93). While anatomical abnormalities were reported in the right fronto-temporal pathway using various methodologies (Hyde et al., 2006, 2007, Mandell et al., 2007, Loui et al., 2009), functional investigations exploring the cerebral basis of pitch perception in congenital amusia (Peretz et al., 2005, 2009, Moreau et al., 2013; Hyde et al., 2011) have suggested near-normal cerebral functioning for pitch processing in this musical disorder (including the auditory cortex). These findings were surprising in light of the behavioral deficits in amusia as the right auditory cortex has been shown to play a critical role in various aspects of pitch processing, such as melody perception, memory or discrimination (see Chapter I, section 1.5.3 pages 38 to 44 and Chapter II, section 2.3.4 page 75). To pursue the characterization of the cerebral correlates of the deficits, the first study of the present thesis investigated short-term memory for tone sequences in congenital amusia by combining behavioral, Magnetoencephalography, and Voxel-Based Morphometry approaches.

Article 1
Impaired pitch perception and memory in congenital amusia: The deficit starts in the auditory cortex

Brain 2013: 136; 1639–1661

BRAIN
A JOURNAL OF NEUROLOGY

Impaired pitch perception and memory in congenital amusia: the deficit starts in the auditory cortex

Philippe Albouy,[1,2,3] Jérémie Mattout,[2,3] Romain Bouet,[2,3] Emmanuel Maby,[2,3] Gaëtan Sanchez,[2,3] Pierre-Emmanuel Aguera,[2,3] Sébastien Daligault,[4] Claude Delpuech,[2,3,4] Olivier Bertrand,[2,3] Anne Caclin[2,3,*] and Barbara Tillmann[1,3,*]

1 CNRS, UMR5292; INSERM, U1028; Lyon Neuroscience Research Centre, Auditory Cognition and Psychoacoustics team, Lyon, F-69000, France
2 INSERM, U1028; CNRS, UMR5292; Lyon Neuroscience Research Centre, Brain Dynamics and Cognition Team, Lyon, F-69000, France
3 University Lyon 1, Lyon, F-69000, France
4 CERMEP, MEG Department, Bron, F-69000, France

*These authors contributed equally to this work.

Correspondence to: Philippe Albouy,
INSERM U1028 - CNRS UMR5292, Centre de Recherche en Neurosciences de Lyon,
Equipe Dynamique Cérébrale et Cognition,
Centre Hospitalier le Vinatier,
Batiment 452, 95 Bd Pinel,
Bron, F-69500,
France
E-mail: philippe.albouy@inserm.fr

Congenital amusia is a lifelong disorder of music perception and production. The present study investigated the cerebral bases of impaired pitch perception and memory in congenital amusia using behavioural measures, magnetoencephalography and voxel-based morphometry. Congenital amusics and matched control subjects performed two melodic tasks (a melodic contour task and an easier transposition task); they had to indicate whether sequences of six tones (presented in pairs) were the same or different. Behavioural data indicated that in comparison with control participants, amusics' short-term memory was impaired for the melodic contour task, but not for the transposition task. The major finding was that pitch processing and short-term memory deficits can be traced down to amusics' early brain responses during encoding of the melodic information. Temporal and frontal generators of the N100m evoked by each note of the melody were abnormally recruited in the amusic brain. Dynamic causal modelling of the N100m further revealed decreased intrinsic connectivity in both auditory cortices, increased lateral connectivity between auditory cortices as well as a decreased right fronto-temporal backward connectivity in amusics relative to control subjects. Abnormal functioning of this fronto-temporal network was also shown during the retention interval and the retrieval of melodic information. In particular, induced gamma oscillations in right frontal areas were decreased in amusics during the retention interval. Using voxel-based morphometry, we confirmed morphological brain anomalies in terms of white and grey matter concentration in the right inferior frontal gyrus and the right superior temporal gyrus in the amusic brain. The convergence between functional and structural brain differences strengthens the hypothesis of abnormalities in the fronto-temporal pathway of the amusic brain. Our data provide first evidence of altered functioning of the auditory cortices during pitch perception and memory in congenital amusia. They further support the hypothesis that in neurodevelopmental disorders impacting high-level functions (here musical abilities), abnormalities in cerebral processing can be observed in early brain responses.

Introduction

About 4% of the population is estimated to experience a lifelong deficit in music perception and production that cannot be explained by hearing loss, brain damage, or cognitive deficits (Ayotte *et al.*, 2002; Peretz *et al.*, 2002; Stewart 2006, 2008). Individuals afflicted with congenital amusia are unable to detect out-of-key tones, and are unaware when others (or themselves) sing out of tune. Thanks to the Montreal Battery of Evaluation of Amusia (Peretz *et al.*, 2003), this disorder has been systematically studied. Behavioural investigations have revealed that the impairment is linked to a deficit in pitch perception (Foxton *et al.*, 2004; Hyde and Peretz, 2004) and memory (Gosselin *et al.*, 2009; Tillmann *et al.*, 2009; Williamson *et al.*, 2010). However, only few studies have investigated the neural bases of congenital amusia from structural or functional cerebral perspectives.

Using voxel-based morphometry (VBM), cortical thickness and diffusion tensor imaging methods, anatomical abnormalities were reported in the right fronto-temporal pathway. In comparison with matched control subjects, amusics' brains showed decreased white matter concentration in the right inferior frontal gyrus associated with increased grey matter concentration in the same region (Hyde *et al.*, 2006), and in the right superior temporal gyrus (Hyde *et al.*, 2007). The hypothesis of an abnormal right fronto-temporal pathway in the amusic brain has received support by the observation of reduced fibre connectivity in the right arcuate fasciculus (Loui *et al.*, 2009).

To relate anatomical anomalies to behavioural expressions, functional investigations have explored the cerebral basis of pitch perception in congenital amusia. Electrophysiological (Peretz *et al.*, 2005, 2009; Moreau *et al.*, 2009) and functional MRI investigations (Hyde *et al.*, 2011) have shown that the auditory cortex of amusic individuals responds normally to pitch, and that the amusic brain can track small pitch changes (i.e. a quarter tone), suggesting near-normal cerebral processing of musical material (Moreau *et al.*, 2009; Peretz *et al.*, 2009). These findings were surprising in light of the observed behavioural deficits in amusia as the right auditory cortex has been shown to play a critical role in various aspects of pitch processing, such as melody perception and discrimination (Peretz, 1990; Johnsrude *et al.*, 2000).

Our study investigated for the first time the cerebral correlates of pitch perception and short-term memory in congenital amusia by combining behavioural, magnetoencephalography (MEG), and VBM approaches. During MEG recording, amusics and matched control subjects performed two melodic tasks, in which two six-tone sequences were compared (same/different paradigm); a contour task and an easier transposition task. We expected impaired short-term memory performance in amusic participants for the contour task, but not for the transposition task, which did not require melodic contour memorization. The easier transposition task was designed as a control condition, for which amusic participants perform as well as control subjects, to rule out the interpretation that brain activity differences between the two participant groups for the contour task might be mere correlates of motivation or attention-level differences between the groups.

We analysed brain responses during encoding, retention and retrieval of melodic information using source modelling of MEG data to investigate whether altered functioning of the auditory and/or frontal cortices might underlie the disorder. We focused on (i) event-related responses for each note of the melodies for the encoding and retrieval parts of the tasks; (ii) sustained evoked responses for the encoding part of the tasks; and (iii) oscillatory gamma activities during the maintenance of musical information. Oscillations in the gamma frequency band have been reported to contribute to the neuronal underpinnings of working memory (Tallon-Baudry and Bertrand, 1999; Tallon-Baudry *et al.*, 1999; Kaiser *et al.*, 2003; Kaiser and Lutzenberger, 2005; Jensen *et al.*, 2007; Mainy *et al.*, 2007). Finally, using VBM in the same participants, we expected to observe anatomical abnormalities in amusic participants (compared with control subjects) in right frontal and temporal areas, as previously shown by Hyde *et al.* (2006, 2007), and we further aimed to compare the locations of these anatomical abnormalities with that of possible functional abnormalities revealed by MEG source analyses.

Material and methods

Participants and behavioural pretests

Nine amusic adults (five females; mean age, 31.55 ± 8.50 years, ranging from 20 to 44; mean education, 14.77 ± 1.71 years; mean musical education 1.16 years, ranging from 0 to 2.5) and nine matched non-musician control subjects (five females; mean age, 31.33 ± 7.31 years, ranging from 24 to 47; mean education, 16.11 ± 2.57 years; mean musical education 0.77 years ranging from 0 to 3) participated in the study. Each group was composed of seven right-handed and two left-handed participants. Severe peripheral hearing loss was excluded using standard audiometry and all participants reported normal hearing and no history of neurological or psychiatric disease. Participants gave their written informed consent, and were paid for their participation. Ethical approval was obtained from the French ethics committee on Human Research (CPP Sud-Est II, #2006-018/A-1).

To be considered as amusic, participants had to obtain an average score two standard deviations (SD) below the average of the normal population on the Montreal Battery of Evaluation of Amusia (i.e. a cut-off score of 23, maximum score = 30; Peretz *et al.*, 2003). In the Montreal Battery of Evaluation of Amusia, six subtests assess various components of music perception and memory by considering that musical material must be processed along a melodic dimension (i.e. sequential variations in pitch) and a temporal dimension (i.e. sequential variations in duration). All participants were tested with the Montreal Battery of Evaluation

of Amusia; the average scores of the amusic group (mean = 20.90, SD = 1.70, ranging from 18 to 22.83) differed significantly from the scores of the control group [mean = 27.61, SD = 0.83, ranging from 26.3 to 28.6, $t(16) = 10.59$, $P < 0.0001$].

To determine pitch discrimination thresholds, a two-alternative forced-choice task was used with an adaptive tracking, two-down/one-up staircase procedure (see Tillmann et al., 2009 for details). Observed pitch discrimination thresholds of the amusic group were higher (mean = 1.07, SD = 1.20, ranging from 0.14 to 4 semitones) than that of the control group [mean = 0.31, SD = 0.30, ranging from 0.07 to 0.95, $t(16) = -1.83$, $P = 0.042$, one-tailed]. The observed overlap in pitch thresholds between amusic and control groups was in agreement with previous findings (Foxton et al., 2004; Tillmann et al., 2009).

Voxel-based morphometry

All participants, except one amusic participant (because of claustrophobia), underwent a 3D anatomical MPRAGE T_1-weighted MRI scan on a 1.5 T Siemens Magnetom scanner (Siemens AG) equipped with an 8-channel head coil (repetition time = 1970 ms; echo time = 3.93 ms; inversion time = 1100 ms; flip angle: 15°, field of view 256 × 256 mm, voxel size 1.0 × 1.0 × 1.0 mm). The anatomical volume consisted of 176 sagittal slices with 1 mm³ voxels, covering the whole brain.

Preprocessing and segmentation

All image preprocessing and voxel-by-voxel statistical analyses were performed using the VBM functions of SPM8 (Wellcome Trust Centre for Neuroimaging, http://www.fil.ion.ucl.ac.uk/spm/, London, UK). Before preprocessing, all images were checked for artefacts and automatically aligned so that the origin of the coordinate system was located at the anterior commissure. Using the unified segmentation procedure implemented in SPM8 (Ashburner and Friston, 2005), the images were segmented into grey matter, white matter and CSF. For each participant, this resulted in a set of three images in the same space as the original T_1-weighted image, in which each voxel was assigned a probability of being grey matter, white matter and CSF, respectively, as well as a normalized version of these images (using the T_1-template from the Montreal Neurological Institute, provided by SPM8). Finally, normalized-segmented images were spatially smoothed with an 8-mm full-width at half-maximum isotropic Gaussian kernel.

Voxel-based morphometry statistical analyses

Group comparisons with two-sample t-tests at each voxel were performed separately for grey matter and white matter. Grey and white matter concentrations were compared for each brain voxel between amusics ($n = 8$) and control subjects ($n = 8$) (the control participant matched to the amusic participant without T_1-weighted MRI was excluded from the analysis). Given the limited sample size, an exploratory whole-brain analysis was not possible, and we focused our analyses on bilateral inferior frontal gyrus and superior temporal gyrus, where abnormalities have been reported previously in congenital amusia (Hyde et al., 2006, 2007; Mandell et al., 2007; Loui et al., 2009), and which are key regions in

pitch processing and memorizing. We thus adopted a lenient statistical threshold of $P < 0.05$ (uncorrected), and only report group differences that were found <1 cm away (in any direction) from the coordinates in the right inferior frontal gyrus and right superior temporal gyrus observed in previous studies, which directly compared amusics and control subjects with VBM (Hyde et al., 2006) or cortical thickness measures (Hyde et al., 2007). We also assessed symmetric locations in the left hemisphere.

Short-term memory tasks for tone sequences

Material

Participants performed two melodic tasks, both requesting to compare two six-tone sequences (S1, S2) separated by a silent retention period of 2000 ms. All sequences were composed of six 250 ms piano tones presented successively without interstimulus interval. The two tone sequences could be either the same or different. For 'different' trials, the second sequence differed by a single tone in the contour task and by all tones [i.e. transposed an interval of seven semi-tones (a fifth) up or down] in the transposition task (Fig. 1). One hundred and ninety-two different melodies (sequences) were created using eight piano tones differing in pitch height (Cubase software); all used tones belonged to the key of C Major (C3, D3, E3, F3, G3, A3, B3, C4). These 192 sequences were used as S1 and were the same for the contour and transposition tasks. In order to strengthen the tonality, the tone C occurred twice in each sequence, and half of the melodies ended on the tone C and the other half on the tone G (i.e. two functionally important tones in the key of C Major). Identical tones were not repeated consecutively in a sequence. For S2 in different trials of the contour and transposition tasks, variants of these sequences were created. For the contour task, one tone (in positions 2 to 5) was replaced by a different tone of the set to create a contour-violation in the melody. For the transposition task (i.e. the transposed sequences), two supplementary tone sets were created for the pitch-up condition (i.e. D4, E4, F4#, G4), and the pitch-down condition (i.e. F2, G2, A2, B2b), respectively.

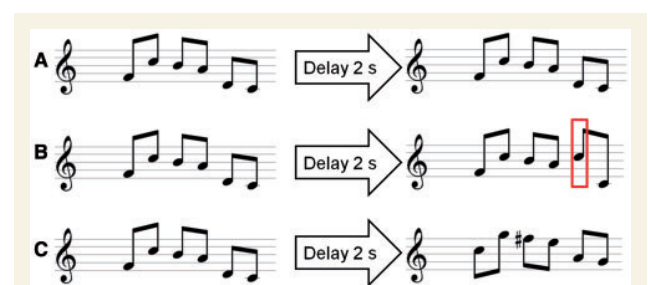**Figure 1** Examples of the musical stimuli used. (**A**) 'Same' trial in the contour task or the transposition task. S1 was repeated as the second melody of the pair (S2) after a 2 s delay. (**B** and **C**) For 'different' trials, the second melody of the pair changed only for one tone in contour task (**B**, red square) and was a transposition of S1 in transposition task (example of the pitch-up condition in **C**).

Procedure

Participants performed the contour and transposition tasks during the MEG recording. Presentation software (Neurobehavioral systems) was used to run the experiment and to record button presses. For each trial, participants listened binaurally to the first six-tone sequence with a total duration of 1.5 s (S1), followed by a silent retention period of 2 s, and then the second sequence (S2, 1.5 s). Participants had to decide whether S2 was identical to S1 or different from S1. There were six blocks of each of the two tasks. The blocks were separated by a break of 2–3 min. The two tasks were presented in alternation (counterbalanced across participants). Participants were informed of task order and asked to indicate their answers by pressing one of two keys with their right hand after the end of S2. They had 2 s to respond before the next trial, which occurred 2.5–3 s after the end of S2. No feedback was given during the experiment. In each block, 32 trials were presented (16 same pairs, 16 different pairs), resulting in 192 trials for each task. Within each block, the trials were presented in a pseudo-randomized order, with several constraints: the same trial type (i.e. same, different), or melodies ending by the same tone (C or G), could not be repeated more than three times in a row. Before the MEG recording, participants performed for each task a set of 15 practice trials without feedback.

Magnetoencephalography

Recordings

The recordings were carried out using a 275-channel whole-head MEG system (CTF-275 by VSM Medtech Inc.) with continuous sampling at a rate of 600 Hz, a 0–150 Hz filter bandwidth, and first-order spatial gradient noise cancellation. Horizontal and vertical electrooculograms and electrocardiogram were acquired with bipolar montages. Head position was determined with coils fixated at the nasion and the preauricular points (fiducial points). Head position was acquired continuously (continuous sampling at a rate of 150 Hz) and checked at the beginning of each block to ensure that head movements did not exceed 0.5 cm (this was confirmed by additional offline checking before the data analyses). Participants were seated upright in a sound-attenuated, magnetically-shielded recording room, and listened to the sounds presented binaurally through air-conducting tubes with foam ear tips. Before the MEG recording, participants' sound detection thresholds (using G3, the tone in the centre of the tone set for S1) were determined for each ear, and the level was adjusted so that the sounds were presented at ~50–55 dB sensation level with a central position (stereo) with respect to the participant's head. After the MEG session, participants' subjective reports regarding their strategies were collected.

Outline of the data analyses

The analyses reported here focused on event-related fields evoked by S1 (i.e. during the encoding and memorization of the melodic pattern); induced gamma oscillations during the delay period, corresponding to the retention in memory of the melodic information; and event-related fields evoked by the changed tone in S2 for the contour task (i.e. during the retrieval of melodic information).

MEG data were first analysed in sensor space using CTF tools (VSM Medtech Inc.) and the ELAN software package developed in the Brain Dynamics and Cognition team (Lyon Neuroscience Research Center, http://elan.lyon.inserm.fr/; Aguera *et al.*, 2011). Source reconstruction was performed with SPM8 (Wellcome Trust Centre for Neuroimaging, London, UK; Friston *et al.*, 2008; Litvak *et al.*, 2011) for event-related fields, and the FieldTrip toolbox (Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands, http://www.ru.nl/fcdonders/fieldtrip) for oscillatory activities, both using MATLAB 7.6 (Mathworks Inc.).

Data preprocessing

For two amusic participants, the original MEG recordings were contaminated by ferromagnetic artefacts caused by metallic dental prostheses, which created a temporally stationary artefact at the participant's respiratory frequency (artefact frequency; mean = 0.24 Hz, ranging from 0.22 to 0.29 Hz, SD = 0.02 Hz for one amusic participant; and mean = 0.31 Hz, ranging from 0.27 to 0.39, SD = 0.05 Hz for the second amusic participant). In addition, the MEG raw data of another amusic participant were contaminated by heartbeat signal (mean artefact frequency = 1.41 Hz, ranging from 1.38 to 1.44; SD = 0.01 Hz). As congenital amusics represent a limited participant pool, we applied a signal correction model with a signal space projection method provided by CTF (Tesche *et al.*, 1995; Ramirez *et al.*, 2011). This method allows uncovering a signal-space vector that describes the global topography of the artefact in sensor space (magnitude varied between 2000 fT and 2500 fT for the dental prosthesis artefacts, and between 20 000 fT and 22 000 fT for the heartbeat artefact). A projection of the MEG data onto the space orthogonal to this vector was then applied to remove the artefact. The efficiency of this correction to faithfully recover the cerebral signals was further assessed with additional recordings of auditory event-related fields in a control participant with and without an artefact-generating piece of metal attached to his head (not shown).

Individual MEG trials were automatically inspected from −100 ms to 5500 ms with respect to the onset of the first S1 tone (i.e. a time window covering S1, the delay, S2 and an additional 500 ms after S2). Trials with ranges of values exceeding ± 3000 fT within a 1000 ms sliding time-window at any sensor site (± 100 μV at electrooculograms channels) were excluded from the analysis; as a result, between 90 and 165 trials were kept for each participant and condition.

Event-related fields

Averaging was done separately for the two experimental conditions (contour task and transposition task) and a −100 to 0 ms interval before the first tone in S1 was used for baseline correction. Grand-average event-related fields were plotted at the sensor level using ELAN and were examined to determine topographical patterns reflecting sources in the auditory cortices (Fig. 4D). Accordingly, S1 analyses focused on (i) transient evoked responses: P50m and N100m (the magnetic counterparts of P50 and N100 waves observed with EEG) for the first tone, and N100m for subsequent tones (an average of the responses to tones 2 to 6 was created for that purpose); and (ii) the sustained

evoked response during the entire melody. The two types of responses were dissociated using two different second-order Butterworth filters (12 dB/octave slope): for transient evoked responses, a band-pass filter (cut-off frequencies at 2 and 30 Hz) was used to eliminate the sustained evoked responses and high-frequency noise (Fig. 4B); for the sustained evoked response a low-pass filter with a 2 Hz cut-off frequency was applied (Fig. 4C).

To analyse event-related responses following the 'changed' tone in S2 of the contour task, two additional averages were performed for each participant (note that we always kept the baseline in the −100 to 0 ms interval before S1); firstly, an average of all correctly detected changed tones, in a −100 to 700 ms time-window around the onset of the change (this event-related field thus combined data for differences in all possible positions in S2), and secondly, an average of tones from correctly classified same trials in the contour task with, for each participant, the same number of tones in position 2, 3, 4 or 5 was used for the event-related fields of the different trials. The change-specific response was then assessed as the difference between these two event-related fields. Note that because control participants had more correct responses than amusic participants, we used for each control participant the same number of trials as his/her matched amusic participant (selected randomly from the entire set of correct response trials). For the analysis of the change-specific responses, filtering was done with a band-pass filter between 0.5 and 30 Hz. As for S1, the frequency band of interest for S2 was chosen after observation of grand-average event-related responses.

Analyses were performed at the source level for the transient evoked responses to determine and characterize the generators involved in the encoding of S1 and the retrieval of melodic information in S2. For the sustained evoked responses, analyses were performed at the sensor-level (see below).

Source reconstruction of transient responses (S1 and S2)

For these analyses, we used the original MEG signals for all participants but one, for whom artefact-corrected data were used because of the heartbeat artefact. Distributed cortical source reconstruction of event-related fields were performed for the brain responses to the first tone and the average responses to the five subsequent tones of S1 separately, in a post-stimulus time window from 25 to 175 ms. The source reconstruction for responses to the changed tone in S2 (contour task) was performed for the difference waveform for correct responses (different trial – same trial, see above) in a −100 to 700 ms time window relative to the onset of the changed tone. For each participant (except the amusic participant without individual MRI), we created individual head meshes (20 484 vertices or dipolar sources) describing the boundaries of different head compartments (scalp, inner skull and cortical sheet) based on the participant's own structural brain scans (note that the same set of vertices were used for all participants, but the mesh was deformed to account for each participant anatomy). For the amusic participant who did not undergo the 3D MRI scan, we used the MNI template provided with SPM8. We then performed landmark-based co-registration of MEG data and MRIs using the locations of nasion and pre-auricular points (Mattout *et al*., 2007). Individual inverse solutions were obtained using the empirical Bayesian approach

implemented in SPM8 (Mattout *et al*., 2006). For each participant, and using the multiple sparse prior model, the two conditions for S1 (contour task, transposition task) were inverted together to increase the sensitivity of the ensuing statistical comparison between them (Friston *et al*., 2008). For S2, only the contour task was inverted as we were interested in the changed tone of the melodic contour task.

Definition of the regions of interest for transient responses in S1 and S2

Having completed the inversion for S1, we divided the source-reconstructed data in two separate time windows of interest, corresponding to the P50m peak (40–70 ms after the onset of the first tone); and to the N100m components (70–160 ms after the onset of each tone), respectively. For S2, the time window of interest was the 0–600 ms period after the onset of the changed tone. For each participant and time sample in each time window of interest, we extracted the cortical points (region of interest) that were significantly different from baseline according to the posterior probability maps provided by the Bayesian inversion. A posterior probability map provides the probability or confidence that the activation in a particular cortical vertex exceeds some specified threshold, given the data (Friston and Penny, 2003). We considered a zero-value threshold. As the data had been baseline-corrected prior to the inversion, this analysis resulted in identifying vertices with activity that significantly differ from baseline. This approach has two main advantages: (i) it provides inference at the 3D cortical surface of each individual, as a consequence of the Bayesian inversion; and (ii) it is not submitted to the multiple-comparison problem because the probability that an activation has occurred, at any particular cortical vertex, is the same, irrespective of whether only that vertex or the entire brain is analysed (Friston and Penny, 2003). Moreover, as each individual mesh is mapping in a one-to-one fashion to the MNI–template mesh (Mattout *et al*., 2007), individual statistics can be combined to perform group-level inference. For each participant and time sample in each time window of interest, we used posterior probability maps with a 99.98% threshold (i.e. 5% of false positive at most, Bonferroni-corrected across participants). This step provided for each participant and each time sample a list of cortical dipoles (vertices) where activity emerged significantly from baseline. We only analysed vertices with activity emerging for at least one-third of the duration of the time window of interest, in at least three participants (without considering the group factor). This provided separate sets of activated cortical vertices for the P50m for the first tone, the N100m for the first tone, the N100m for the tones 2 to 6, and the responses elicited by the changed tone in S2.

Amplitude and latency analyses of source data for transient responses in S1 and S2

Source activities were analysed for each region of the cortical mesh where activity was significantly different from baseline (Tables 1 and 2), an average of the activity of all vertices was created for each region in each hemisphere. For the amplitude analysis of the source data for S1, we conducted an ANOVA on the source time course at each time sample and for each

Table 1 Frontal and Temporal Generators of the N100m for tone 1 and tones 2–6

Tone(s)	Lobe	Region	Hemisphere	x	y	z	mm^2	Number of vertices
Tone 1	Frontal	Inferior frontal gyrus, pars opercularis	Right	54	8	6	38	4
			Left	−51	4	5	14	2
	Temporal	Heschl's gyrus/superior temporal gyrus/planum temporale	Right	45	−19	6	44	8
			Left	−44	−18	5	85	11
Tones 2 to 6	Frontal	Inferior frontal gyrus, pars opercularis	Right	54	8	6	38	4
			Left	−51	4	5	14	4
	Temporal	Heschl's gyrus/superior temporal gyrus/planum temporale/planum polare	Right	49	−8	5	20	5
			Left	−49	−12	5	81	11

We report regions where activity was significantly different from baseline ($P < 0.05$ Bonferroni corrected across participants), for at least one-third of the time window of interest, for at least three participants, as assessed from the obtained posterior probability maps (see 'Materials and methods' section for details). Coordinates correspond to the vertex with maximal amplitude within each region (coordinates are in MNI space). For each vertex, amplitude data of the window of interest were averaged across all participants and tasks to determine the coordinates of the cortical vertex showing the highest peak amplitude.

region with 'Group' as between-participants factor and 'Task' as within-participant factor. To correct for multiple comparisons, only effects lasting for >15 ms (i.e. 9 consecutive samples) were considered significant (Guthrie and Buchwald, 1991; Caclin *et al.*, 2008). For the S2 analyses, two sample *t*-tests were performed to compare amusic and control groups with the same duration criterion as for S1 (i.e. effects lasting >15 ms).

To estimate the peak latency of the P50m (for the first tone) and N100m (for the first tone and for the tones 2 to 6), we extracted the latency of the maximum of absolute amplitude of the source activities in the 25–70 ms and 90–160 ms time windows, respectively. For the P50m of the first tone, the latency values of the maximum amplitude were analysed with a $2 \times 2 \times 2$ ANOVA, with Group as between-participants factor (amusics, control subjects), and Task (contour task, transposition task) and Hemisphere (right, left) as within-participant factors. For the N100m, latency values were analysed with a $2 \times 2 \times 2 \times 2 \times 2$ ANOVA with Tone Rank (tone 1, tones 2 to 6), region of interest (Heschl's gyrus or inferior frontal gyrus), as additional within-participant factors. The latency analysis focused on the vertices that were significantly different from baseline for both the first tone and the tones 2 to 6 (resulting in two sets for each hemisphere: a set of three vertices with a peak at coordinates in Heschl's gyrus; and a set of two vertices at coordinates in the pars opercularis of the inferior frontal gyrus). We did not perform latency analyses of the brain responses associated with the changed tone of S2 (contour task) as they tended to present several local maxima within the time window of interest.

Dynamic causal modelling of transient responses in S1

To test whether group differences in event-related responses associated with the encoding of tones in S1 could be explained by changes in effective connectivity between sources, we used a dynamic causal modelling approach as implemented in SPM8 (David and Friston, 2003; David *et al.*, 2006; Garrido *et al.*, 2007; Kiebel *et al.*, 2009). Dynamic causal modelling attempts to explain event-related potentials using a network of interacting cortical sources and waveform differences in terms of coupling changes among sources. We focused our analyses on the N100m component in response to tones 2 to 6. We excluded tone 1 for which no major

difference could be observed between control subjects and amusics in the first steps of the analyses.

To characterize the two participant groups with a high signal-to-noise ratio and to investigate the putative differences in effective connectivity between them, we compared and modelled the difference between the grand average data of the control subjects (average over the nine participants and the two tasks) and the grand average data of the amusics (average over the nine participants and the two tasks).

All the models we compared were based on the same network architecture, which was motivated by (i) the results of our classical source reconstruction analysis of the N100m component revealing sources in a bilateral fronto-temporal network (see Table 1 for regions of interest that were significantly different from baseline); and (ii) the hypothesis of impaired fronto-temporal connectivity and interhemispheric connectivity observed in amusia with both functional (Hyde *et al.*, 2011) and anatomical approaches (Loui *et al.*, 2009). We thus assumed four sources, modelled as equivalent current dipoles, over left and right primary auditory cortices (A1), left and right pars opercularis of the inferior frontal gyrus (Table 1). Using these sources, we constructed the following dynamic causal model: an extrinsic input entered bilaterally to the primary auditory cortices (A1), which were connected to their ipsilateral inferior frontal gyrus. Interhemispheric (lateral) connections were placed between left and right A1. All connections were reciprocal (i.e. connected with forward and backward connections or with bilateral connections; Fig. 8B).

Given this network architecture, we used a factorial design and a family inference approach to assess various types of connection modulations to explain the group difference in auditory evoked responses (control subjects defined the baseline). Factor 1 refers to the modulation of intrinsic connectivity in bilateral auditory cortices [and includes two families (or levels), corresponding to models where these intrinsic connections were modulated or not between the two groups]. Factor 2 refers to the modulation of lateral connections between the two auditory cortices (two families: models which included modulation or not). Factor 3 refers to the type of connections between auditory and frontal areas that are modulated (namely forward, backward, or both forward and backward connections) or not (four families). Factor

4 refers to the hemispheric location of the above modulated connections, either in the right hemisphere, the left hemisphere, or both (three families). We thus fitted and compared 48 models. Assuming equal prior probabilities over models, we used Bayesian model selection to compare model families (Penny et al., 2010). This rests upon the free energy (or approximate marginal likelihood or evidence) for each model, and yields a posterior probability associated with each model family.

Sustained evoked response analyses (S1)

For the analyses of sustained evoked responses, we used artefact-corrected data for the three amusic participants with respiratory or heartbeat artefacts. These analyses were performed within a 1000 ms time window during S1 (500–1500 ms after the onset of the first tone) as this window covered the maximum amplitude of the responses in all participants and avoided the rising slope and the descending slope of the sustained response. Sensor topography indicated a distribution that was consistent with bilateral sources in the auditory cortices during S1. Analyses were performed at the sensor level because of complications for source reconstruction of the artefact-corrected data of the two amusic participants with artefacts at the respiratory frequency (because of residual artefacts in the most frontal sensors). For each hemisphere, the three sensors with the highest positive amplitude and the three sensors with the highest negative amplitude within the 500 to 1500 ms time window were selected. Thereby, negative responses in the right hemisphere corresponded with the right temporal site, and positive responses to the right frontal site, and conversely for the left hemisphere (negative/frontal site; positive/temporal site) (Fig. 9). To determine whether there were group differences in the amplitude of the sustained responses during S1, we computed the absolute values of the mean of these event-related fields over the 500 to 1500 ms time window. These amplitude values were analysed with a $2 \times 2 \times 2 \times 2$ ANOVA, with Group as between-participants factor (amusics, control subjects), and Task (contour task, transposition task), Hemisphere (right, left) and Sensor site (frontal, temporal) as within-participant factors.

Sensor-level analysis and source reconstruction of induced gamma activity (retention period)

We were interested in investigating the potential group and task differences in oscillatory activities in the gamma-band induced by the musical information during memory maintenance. Because neither latency nor frequency of these oscillatory bursts was known a priori, we first assessed sensor-level data using a method that preserved both types of information: the time-frequency representation based on a wavelet transform of the signals (Tallon-Baudry et al., 1996). The MEG signal was convoluted by complex Morlet's wavelets having a Gaussian shape both in the time domain (SD σt) and in the frequency domain (SD σf) around its central frequency f0. The wavelet family was defined by (f0/σf) = 10, with f0 ranging from 24 to 80 Hz in 2 Hz steps. The time–frequency wavelet transform was applied to each trial after subtraction of the evoked response at each MEG sensor and then averaged across trials, resulting in an estimate of oscillatory power at each time sample and at each frequency between 24

and 80 Hz. For all participants, except the one amusic participant with the heartbeat artefact, original MEG signals were used for this analysis (corrected data were used for the amusic with the heartbeat artefact). We investigated oscillatory gamma activities during the retention period, with respect to a prestimulus baseline, by subtracting for each frequency the mean power computed over the period defined as -1000 to 0 ms preceding the first tone of S1.

To identify the maximal increase of induced power within the gamma band relative to baseline, Wilcoxon tests were computed for each frequency and each time sample across all participants to compare the mean power (average of both tasks) with the mean power in the prestimulus baseline (defined between -1000 and 0 ms before the onset of the first tone of S1). This analysis allowed us to select the 30–40 Hz frequency band over a 1000 ms time window covering the central part of the retention period (2000 to 3000 ms).

To localize the neural sources in the frequency band of interest and explore the potential differences between tasks and groups within the retention period, we applied an adaptive spatial filtering or beamforming technique [Dynamic Imaging of Coherent Sources (DICS) as implemented in Fieldtrip] (Gross et al., 2001; Bouet et al., 2012). Based on individual MRIs (except for the amusic participant who did no undergo the T_1-MRI scan), a grid of $5 \times 5 \times 5$ mm^3 volume elements covering the whole brain was constructed, and the source power was computed for each of those elements. For each source location, a linear spatial filter was computed that passes activity from that location with unit gain while maximally suppressing activity from other sources. This filter depended on the signal spectral properties extracted for each participant through a fast Fourier transformation.

For each trial, we estimated the power in the 30–40 Hz frequency band at each volumetric element for both active (retention) and baseline periods, and we computed the log ratio between these two power values to determine the extent of the emerging gamma oscillations. After normalization of the brain in the MNI space, statistical comparisons were performed to compare groups for each task, and to compare tasks for each group with two sided t-tests corrected for multiple comparison using cluster-level statistics, as implemented in Fieldtrip.

Correlations between behavioural, functional and anatomical data

Five types of analyses were performed: (i) correlation between the behavioural data in the short-term memory task (Hit-False-Alarms, reaction times) and data from the pretests (Montreal Battery of Evaluation of Amusia, pitch discrimination threshold); (ii) correlation between the different measures of the behavioural data in the short-term memory task; (iii) correlation between the different measures (white matter, grey matter concentrations) of the VBM data; (iv) correlation between the different measures of the MEG data (event-related fields amplitude and latency, gamma oscillations; connectivity values); and (v) correlations between behavioural, functional and anatomical data. To explore differences explained not only by the group differences defined by the Montreal Battery of Evaluation of Amusia (Peretz et al., 2003), we present here only correlations that were significant over all

participants and for at least one of the two groups separately (amusic group or control group). Note that, except for the VBM data, none of the correlations met these criteria.

Results

Behavioural data

Performance was significantly above chance (i.e. 0% of Hits–False Alarms), for each condition in each group (*t*-tests, all *P*'s < 0.0001). Hits-False Alarms (Fig. 2A) were analysed with a 2 × 2 ANOVA with Group (amusics, control subjects) as between-participants factor and Task (contour task, transposition task) as within-participant factor. The main effects of Group [$F(1,16) = 74.82$, $P < 0.0001$], and Task [$F(1,16) = 300.08$, $P < 0.0001$] were significant, as well as their interaction [$F(1,16) = 119.15$, $P < 0.0001$]. Fischer's LSD *post hoc* tests indicated that the performance of the amusic group was lower than that of the control group for the contour task ($P < 0.0001$), but not for the transposition task ($P = 0.28$). All amusic participants exhibited a deficit in the contour task in comparison with the control participants (see individual data in Fig. 2A). This is in line with previous results (Gosselin *et al.* 2009; Tillmann *et al.*, 2009), suggesting a short-term memory deficit for tone sequences in congenital amusia. It is noteworthy that there was no overlap in performance between the two groups in the contour task, despite a substantial overlap between groups in pitch discrimination thresholds (see 'Materials and methods' section).

For the contour task, to assess the potential effect of pitch interval size in the changed tone of S2 (relative to the corresponding tone in S1) on performance, percentages of correct responses for different trials were extracted as a function of interval sizes. Within 96 different trials of the contour task, there were 36.46% of trials with a change of small interval size (with equal proportions of 1.5 tone, 2 tones and 2.5 tones); 35.41% of trials with a

change of medium interval size (3 tones, 3.5 tones and 4 tones) and 28.13% of trials with a change of large interval size (4.5 tones, 5 tones, 5.5 tones, 6 tones). These data were analysed with a 2 × 3 ANOVA with Group (amusics, control subjects) as between-participants factor and Interval (small, medium, large) as within-participant factor. The main effects of Group [$F(1,16) = 75.37$, $P < 0.0001$], and Interval [$F(2,32) = 32.38$, $P < 0.0001$] were significant, as well as their interaction [$F(2,32) = 4,36$, $P = 0.02$]. *Post hoc* tests indicated that the performance of the amusic group was lower than that of the control group for all interval sizes (all *P*'s < 0.01). Moreover, in amusics, increasing performance was observed with increasing interval changes (all *P*'s < 0.01 for small/medium, medium/large, and large/small comparisons). This increased performance was also observed in control subjects (all *P*'s < 0.05), except for the medium/large comparison ($P = 0.47$), which was probably due to ceiling performance.

Reaction times for correct responses (relative to the end of S2) were analysed by a 2 × 2 × 2 ANOVA with Group (amusics, control subjects) as between-participants factor, and Task (contour task, transposition task) and type of trial (same, different) as within-participant factors. The ANOVA revealed a significant main effect of Task [$F(1,16) = 19.63$, $P < 0.001$], with shorter reaction time for the transposition task (mean = 479.18 ms, SD = 49.19 ms) than for the contour task (mean = 585.37 ms, SD = 61.21 ms), and a significant interaction between task and group [$F(1,16) = 7.60$; $P = 0.01$]. Fischer's LSD *post hoc* tests revealed that only for the amusic group ($P < 0.001$), reaction times were longer for the contour task (mean = 629.84 ms, SD = 39.94 ms) than for the transposition task (mean = 479.71 ms, SD = 25.55 ms). This was not the case for control subjects (mean = 529.85 ms, SD = 77.84 ms for the contour task and mean = 489.74 ms, SD = 66.80 for the transposition task, $P = 0.25$). Moreover, the analysis revealed a significant interaction between task (contour task, transposition task) and type of trial [$F(1,16) = 5.36$; $P = 0.01$]; only for the transposition task
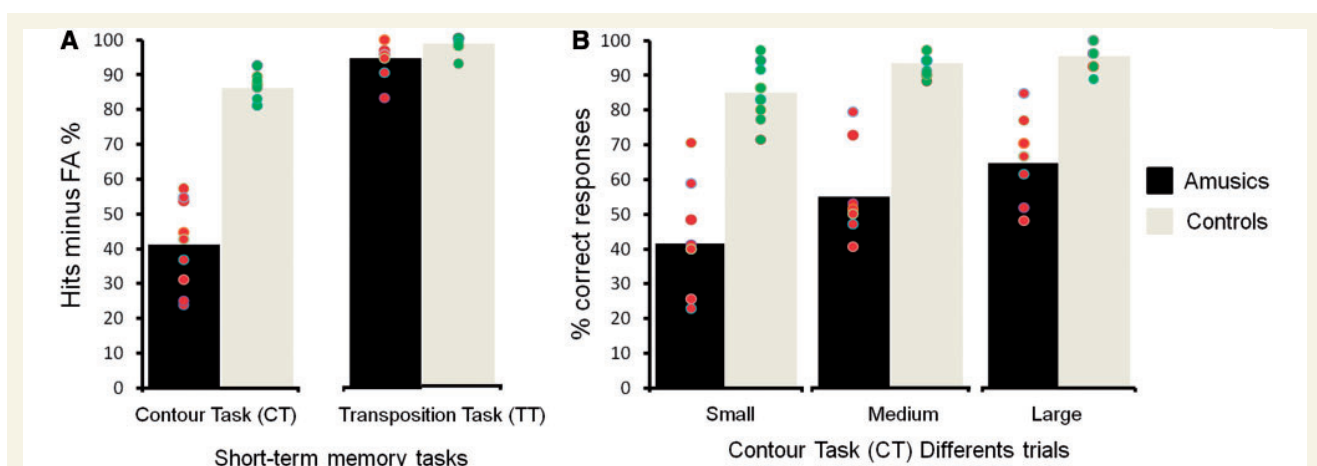
Figure 2 (**A**) Performance of amusic and control groups (grey = control subjects; black = amusics) in terms of Hits minus False Alarms (FA), presented as a function of the task (contour task, transposition task) and group (amusics *n* = 9; control subjects *n* = 9). (**B**) Performance of amusic and control groups for the contour task in terms of per cent of correct responses for different trials as a function of interval size (small, medium, large). Green circles: control subjects' individual performance; red circles: amusics' individual performance.

($P = 0.04$), reaction times were shorter (mean = 451.51 ms, SD = 46.76 ms) for different trials than for same trials (mean = 506.59 ms, SD = 51.93 ms). Analyses of reaction times for correct responses as a function of interval size for the contour task [with a 2×3 ANOVA with Group as between-participants factor and interval size (small, medium, large) as within-participant factor] revealed no significant effects of group ($P = 0.91$) or interval size ($P = 0.14$) and no interaction between them ($P = 0.35$).

Subjective reports

Differences were observed in participants' reports after the session: while all control subjects consistently reported to repeat the melody back in their mind during the delay (eight control participants by internal singing and one by visual mental imagery), amusics did not report to have elaborated any particular strategy (three amusic participants reported to have tried at the beginning to repeat S1 in their mind during the delay, but stopped doing so because of its inefficiency).

Voxel-based morphometry

In the right inferior frontal gyrus, in line with previous findings (Hyde *et al.*, 2006, 2007), between-group differences were found in both grey and white matter (Fig. 3). The analyses of grey matter showed larger concentrations in the right inferior frontal gyrus (bordering the right middle frontal gyrus) in amusics than in control subjects [maximum peak at $x = 36$, $y = 46$, $z = 6$; $t(14) = 2.68$, $P = 0.009$; K_E (cluster extent) = 192 mm^3]. The maximal difference was <1 cm away from that observed with VBM in the two independent samples of amusic and control participants in Hyde *et al.* (2006) ($x = 37$, $y = 42$, $z = -2$ and $x = 40$, $y = 43$, $z = -2$, respectively). Conversely, less white matter concentration in amusics than in control subjects was observed in the right inferior frontal gyrus [in particular in a cluster with a maximum peak at $x = 40$, $y = 30$, $z = 28$; $t(14) = 4.37$, $P < 0.0001$; $K_E = 1328$ mm^3, which encompassed an additional maximum at $x = 36$, $y = 44$,

$z = -6$, close to the coordinates previously reported in Hyde *et al.* (2006): $x = 37$, $y = 42$, $z = -2$ and $x = 42$, $y = 44$, $z = -2$]. This inverse relationship between white and grey matter concentrations in amusics' right inferior frontal gyrus is in line with Hyde *et al.* (2006, 2007). In symmetrical regions of the inferior frontal gyrus in the left hemisphere, between-group differences were observed for white matter only [maximum peak at $x = -34$, $y = 40$, $z = 2$; $t(14) = 3.71$, $P = 0.001$; $K_E = 5344$ mm^3], with less white matter concentration for amusic participants.

The right superior temporal gyrus showed less grey matter concentration in amusic participants than in control subjects (maximum peak of the cluster at $x = 44$, $y = -16$, $z = -18$; $t(14) = 3.86$, $P = 0.001$; $K_E = 17\,936$ mm^3; additional peak in the right superior temporal gyrus at $x = 60$, $y = -14$, $z = 6$). This result contrasts with cortical thickness data (Hyde *et al.*, 2007), showing more grey matter concentration in the right auditory cortex of amusics than of control subjects at approximately the same location (they reported a maximal difference at $x = 60$, $y = -12$, $z = 5$). In the vicinity of the grey matter difference [maximum peak at $x = 44$, $y = -8$, $z = 16$; $t(14) = 3.19$, $P = 0.003$; $K_E = 5520$ mm^3], we also observed less white matter concentration in amusic participants than in control subjects. In symmetrical regions of the left hemisphere, we observed only small between-group differences in the white matter (in two clusters of 464 and 760 mm^3, respectively). Note that there were no previous reports of white matter anomaly in the superior temporal gyrus for congenital amusia.

The correlation between grey matter concentration and white matter concentration was not significant for the right inferior frontal gyrus, [$r(14) = -0.01$, $P > 0.69$], but was significant for the right superior temporal gyrus for all participants [$r(14) = 0.94$, $P < 0.01$], as well as for amusic participants [$r(6) = 0.87$, $P < 0.01$] and control participants [$r(6) = 0.96$, $P < 0.01$], separately. In line with the abnormal lack of fibre connectivity along the right arcuate fasciculus (connecting right inferior frontal gyrus and
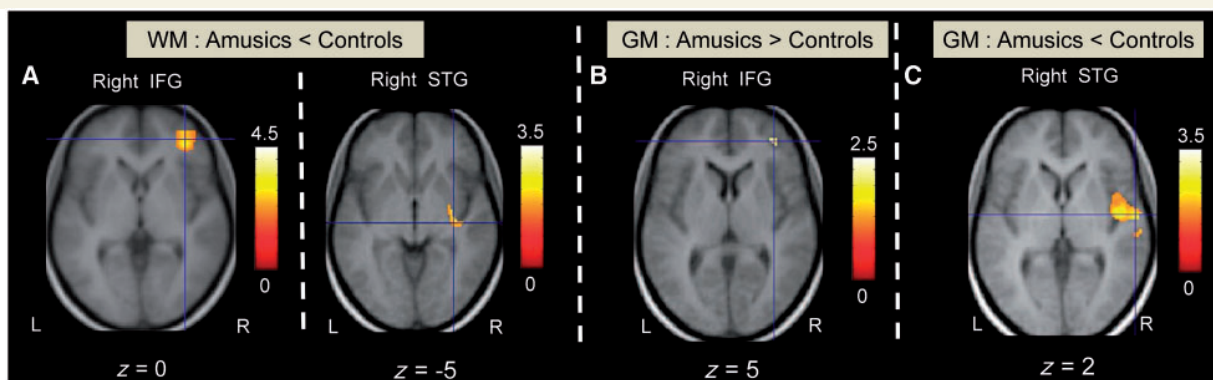
Figure 3 VBM group comparison of white matter (WM) and grey matter (GM) concentration differences: Each brain image corresponds to a statistical map ($P < 0.05$, uncorrected, T-threshold = 1.76) superimposed on the average anatomical MRI of all participants (R = right; L = left). (**A**) Control participants showed more white matter concentration relative to amusics participants in the right inferior frontal gyrus (IFG) and right superior temporal gyrus (STG). (**B**) Amusic participants showed more grey matter concentration relative to control subjects in the right inferior frontal gyrus. (**C**) Control participants showed more grey matter concentration relative to amusics participants in the right superior temporal gyrus.

right superior temporal gyrus) described by Loui *et al*. (2009) in the amusic brain, we observed a positive correlation between white matter concentration of right inferior frontal gyrus and white matter concentration of right superior temporal gyrus. This correlation was observed over all participants [$r(14) = 0.79$, $P < 0.01$] but was significant only for the amusic participants [$r(6) = 0.81$, $P < 0.01$] and marginally significant for the control participants [$r(6) = 0.62$, $P = 0.09$].

Magnetoencephalography

Auditory evoked responses were observed after each tone in all participants. The sensor plots indicated clear P50m, N100m and sustained evoked responses during S1 and S2. The topography of these responses was consistent with bilateral sources in the auditory cortices (Fig. 4).

Encoding of S1 melodies: transient evoked responses (P50m and N100m)

P50m and N100m represent the activity of generators located in the auditory and frontal cortices (Näätänen and Picton, 1987; Alcaini *et al*., 1994; Giard *et al*., 1994; Pantev *et al*., 1995;

Yvert *et al*., 2001). Typically, the evoked response for the first tone of a sequence is different and larger than that of subsequent tones (Fig. 4A and B), not only because of neural refractoriness, but also because the first tone can be considered as an 'infrequent/relatively unexpected' orienting stimulus because it is presented after a period of silence. In contrast, subsequent tones in a sequence (here, tones 2 to 6, presented without interstimulus interval) can be considered as 'frequent/expected' stimuli. The first tone should thus recruit more 'orienting' (non-specific) components (Alcaini *et al*., 1994; Demarquay *et al*., 2011) than tones 2 to 6. In comparison with the first tone, these subsequent tones should recruit more strongly the cortical areas implicating high-level processing or memory representation (relative to the recruitment of non-specific components). Source analyses thus explored the generators of P50m and N100m as a function of group and task, separately for the first tone and the subsequent tones of S1.

Source analyses of transient responses in S1 (P50m and N100m): regions of interest

For the P50m of tone 1, activity was significantly different from baseline in bilateral auditory regions (Heschl's gyrus, left hemisphere, $x = -48$, $y = -18$, $z = 4$; cluster surface = 53 mm²; right
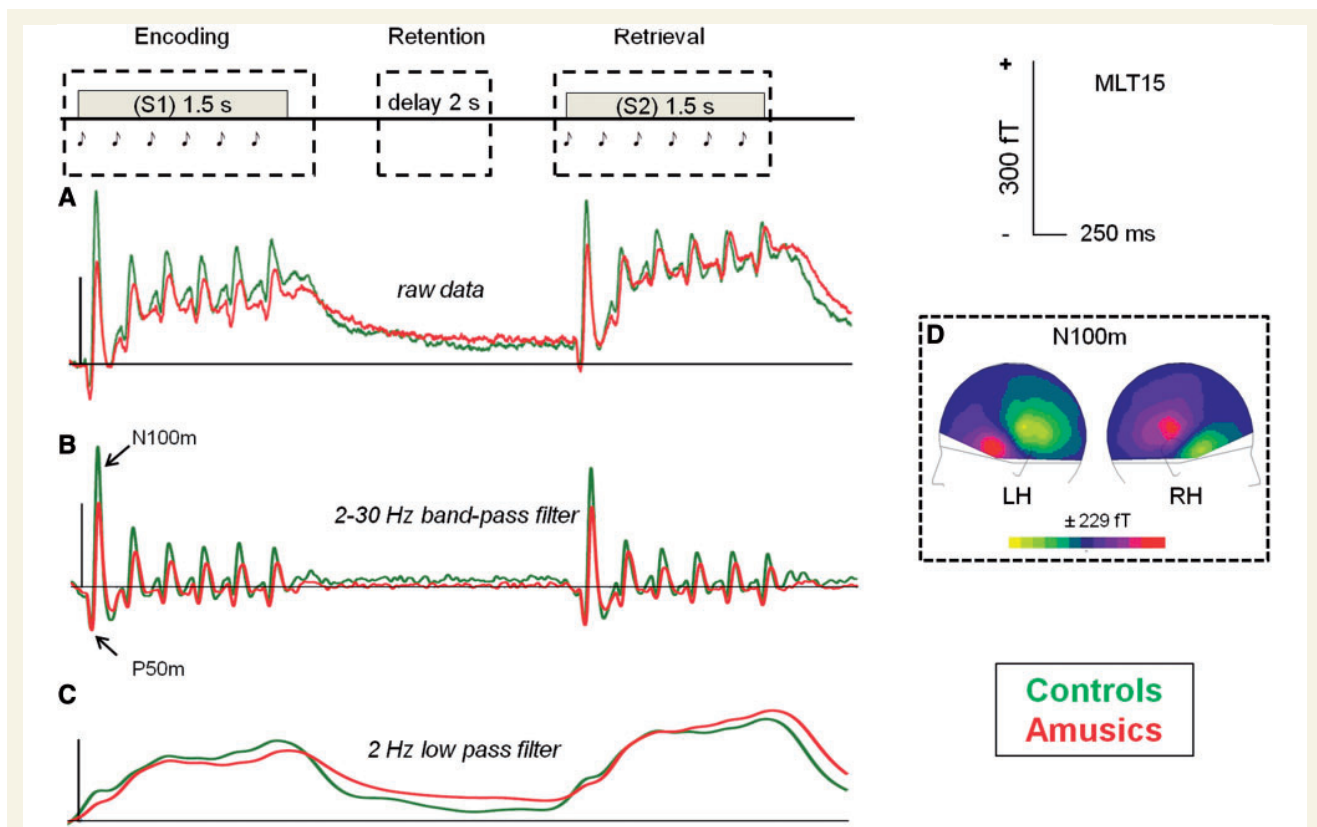
Figure 4 Grand average of a left temporal MEG sensor (MLT15, 'MEG left temporal') for a trial time window (-100 to 5500 ms), for the control group (green) and amusic group (red), collapsed across conditions (contour task, transposition task). (**A**) Event-related fields based on original MEG signals. (**B**) Event-related fields filtered with a 2–30 Hz band-pass filter to uncover P50m and N100m responses. (**C**) Event-related fields filtered with 2 Hz low-pass filter to uncover sustained evoked responses. (**D**) Sensor plot of the mean event-related fields in the 90–140 ms time-window for the first tone, averaged across groups and tasks. LH = left hemisphere; RH = right hemisphere.
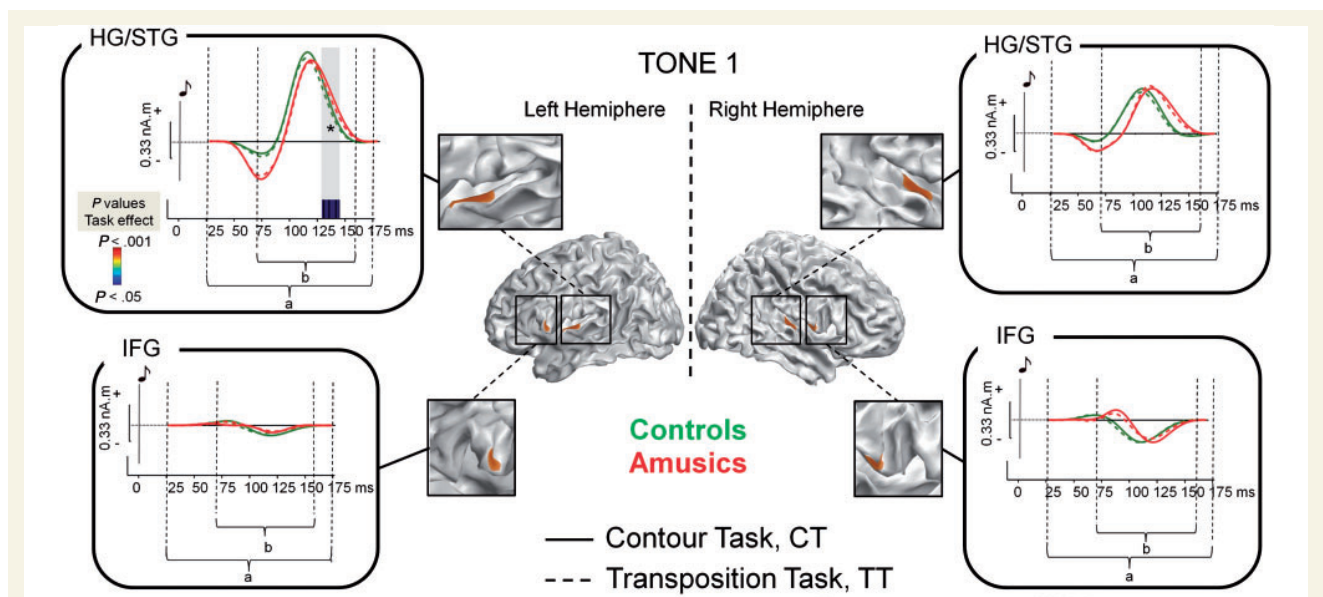
Figure 5 Amplitude data for the N100m evoked by the first tone of S1. Cortical meshes show bilateral regions that were significantly different from baseline (as indicated by the brown areas) for the time window of tone 1. These regions were bilateral Heschl's gyrus/superior temporal gyrus (HG/STG) activation in the Heschl's gyrus extending to the planum temporale and the planum polare) as well as the opercular part of the inferior frontal gyrus (IFG) (at the frontier with the rolandic operculum) (Table 1). The surrounding panels correspond to the grand average of source data for each region and for the time window where the inversion was performed [25–175 ms after the tone onset, as indicated by 'a' for the control group (green) and amusic group (red), for the contour task (full lines) and the transposition task (dotted lines)]. For N100m analysis, ANOVAs were performed at each time sample and for each region on source amplitude in the 70 to 160 ms time window (indicated by 'b'), in the two groups of participants. P-values for the main effects are reported across time below source amplitudes. Note that only effects lasting > 15 ms were reported. Colour bar represents the P-values for the task effect with blue for $P < 0.05$; green for $P < 0.01$; and red for $P < 0.001$.

hemisphere, $x = 51$; $y = -11$; $z = 4$; cluster surface = 25 mm^2). For the N100m, bilateral fronto-temporal regions with activity significantly different from baseline for tone 1 and for the average of tones 2 to 6 are presented in Table 1, together with the coordinates of emergent vertices for the highest peak amplitudes. Note that for the auditory regions, the vertex of maximum amplitude was located in Heschl's gyrus, but that the activation further extended to the planum temporale, planum polare and superior temporal gyrus. Note that, when we performed the same PPM analyses for each group separately, the same regions with the same peaks were emergent in control subjects and amusics, but with a smaller extent in the amusic group.

Source analyses of transient responses in S1 (P50m, N100m): amplitude

For P50m (tone 1), no significant effects or interaction were observed. Note that there were no differences between contour task and transposition task; and that we did not attempt to analyse the P50m of tones 2 to 6 because their amplitudes were rather small as a consequence of refractoriness (P50-gating), and because of the overlap with the end of the components evoked by the preceding tones.

For N100m (tone 1), whereas there was no significant effect of group, the main effect of task was significant: both amusics and control subjects showed larger activation in the contour task in comparison to the transposition task in the left auditory cortex (129 to 145 ms).

The ANOVAs for N100m (average of tones 2 to 6) revealed group differences in source amplitude, with higher amplitudes for control participants than for amusic participants in the four regions of interest and the following time windows (Fig. 6): (i) right Heschl's gyrus/superior temporal gyrus (together with planum temporale and planum polare), 80–120 ms; (ii) right inferior frontal gyrus, 105–122 ms; (iii) left inferior frontal gyrus, 107–127 ms; and (iv) left Heschl's gyrus/superior temporal gyrus (together with planum temporale and planum polare), 95–126 ms. Moreover, the inverse difference (amusics > control subjects) was observed in the right Heschl's gyrus/superior temporal gyrus for the 135–153 ms time window. However, this inverse difference for the right Heschl's gyrus/superior temporal gyrus seems to be related to an increased latency of the N100m in the amusic brain (see below). No significant effect of task and no interaction between group and task were observed.

Source analyses of transient responses (P50m, N100m): latency

For P50m (tone 1), there were no significant effects or interaction. In analyses for N100m (tone 1 and average of tones 2 to 6) there
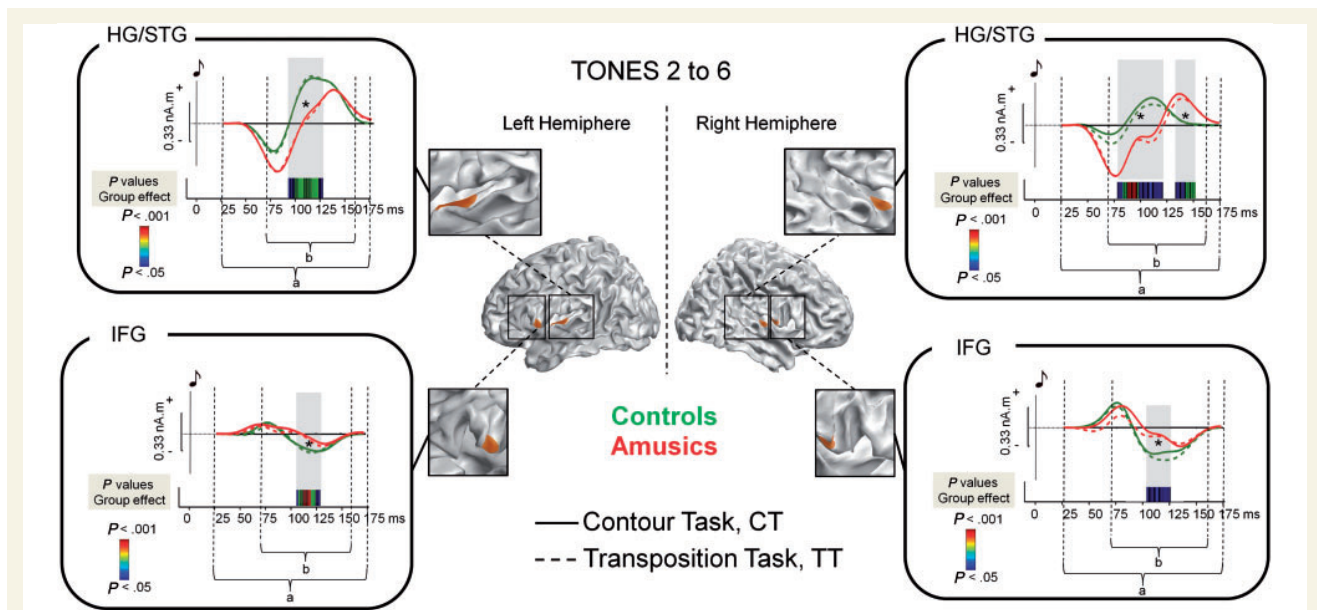
Figure 6 Amplitude data for the N100m evoked by the tones 2 to 6 of S1. Cortical meshes show bilateral regions that were significantly different from baseline (as indicated by the brown areas) for the time window of tones 2 to 6. These regions were bilateral Heschl's gyrus/superior temporal gyrus (HG/STG) (activation in Heschl's gyrus extending to the planum temporale and the superior temporal gyrus) and the opercular part of the inferior frontal gyrus (at the frontier with the rolandic operculum) (Table 1). The surrounding panels correspond to the grand average of source data for each region and for the time window where the inversion was performed [25 to 175 ms after the tone onset, as indicated by 'a' for the control group (green) and amusic group (red), for the contour task (full lines) and the transposition task (dotted lines)]. For the N100m analysis, ANOVAs were performed at each time sample and for each region on source amplitude in the 70 to 160 ms time window (as indicated by 'b') in the two groups of participants. P-values for the main effects are reported across time below source amplitudes. Note that only effects lasting longer than 15 ms were reported. Colour bar represents the P-values for the group effect with blue for $P < 0.05$; green for $P < 0.01$; and red for $P < 0.001$.

were significant main effects of: (i) group [$F(1,16) = 8.54$; $P < 0.01$], reflecting increased N100m latency of source activity for amusics in comparison to control subjects; (ii) tone rank [$F(1,16) = 19.73$; $P < 0.0001$], with increased latency for tones 2 to 6 in comparison with tone 1; and (iii) a marginally significant effect of region [$F(1,16) = 4.3$; $P = 0.054$], with increased latency in the inferior frontal gyrus in comparison with the Heschl's gyrus. The interaction between tone rank and task [$F(1,16) = 4.85$; $P = 0.042$] was significant. For tone 1, the latency did not differ between the two tasks ($P = 0.21$), but for tones 2 to 6, the latency was slightly longer for the contour task than for the transposition task, even though only marginally significantly ($P = 0.087$). In addition, the interaction between tone rank and region was significant [$F(1,16) = 9.33$; $P = 0.007$], and, most importantly, it was modulated by group, as indicated by the three-way interaction between tone rank, region and group [$F(1,16) = 5.17$; $P = 0.03$].

For tone 1 (Fig. 7), the increased latency in bilateral inferior frontal gyrus relative to bilateral Heschl's gyrus was more pronounced for amusics (reflecting a delayed frontal responses, $P < 0.003$) than for control subjects (where the difference was only marginally significant, $P = 0.06$). The latencies of the responses for tones 2 to 6 were similar to the latency of tone 1 (Fig. 7), except that, critically, for amusics, the Heschl's gyrus response for tones 2 to 6 was delayed by ~20 ms relative to control subjects ($P < 0.001$).

Dynamic causal modelling of the transient responses in S1 for tones 2 to 6

Figure 8A shows the results of the comparisons between model families (exceedance probabilities obtained for each family inference). Figure 8B shows the network architecture of the winning model as well as the conditional estimates of the connection strengths associated with the connections that proved to be significantly modulated to explain the amusic response compared to the control response.

Posterior estimates obtained with the winning model enabled us to conclude that, compared with control subjects, amusic participants showed an abnormally increased lateral connectivity between the two A1, decreased intrinsic modulations in both auditory cortices, and decreased backward connectivity between the right inferior frontal gyrus and the right auditory cortex.

Sustained evoked responses during S1

Few studies have explored the role of the sustained evoked responses in the auditory domain, but some of these data sets have demonstrated that: (i) sustained evoked responses are involved in the cortical representation of behaviourally relevant sounds (Picton *et al.*, 1978; Bidet-Caulet *et al.*, 2007); and (ii) they are modulated by attentional processes (for a review, see Picton *et al.*, 1978). In line with Bidet-Caulet *et al.* (2007) who reported sustained evoked
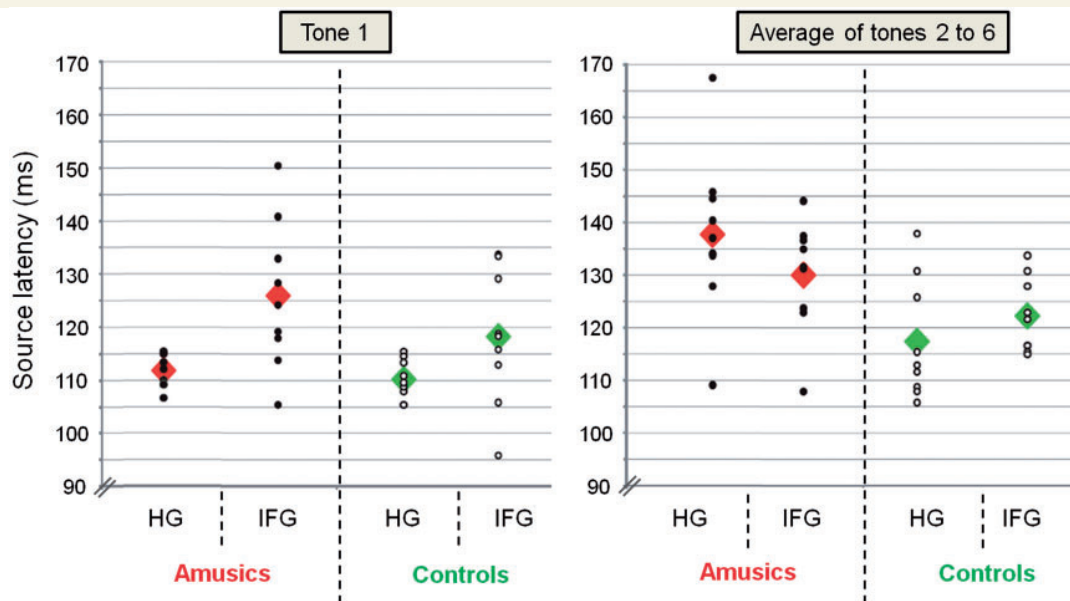
Figure 7 N100m latency (in ms) of amusic and control participants, calculated from the reconstructed source signal [amusics *n* = 9; control subjects *n* = 9 for bilateral Heschl's gyrus (HG) and bilateral inferior frontal gyrus (IFG); Green = control subjects; red = amusics] presented as function of tone rank (tone 1 and average of tones 2 to 6) and region (Heschl's gyrus, inferior frontal gyrus). Diamonds indicate average latency for each group and task (in ms); circles indicate participants' individual latency (in ms). Note that latencies were averaged across tasks (contour task, transposition task).
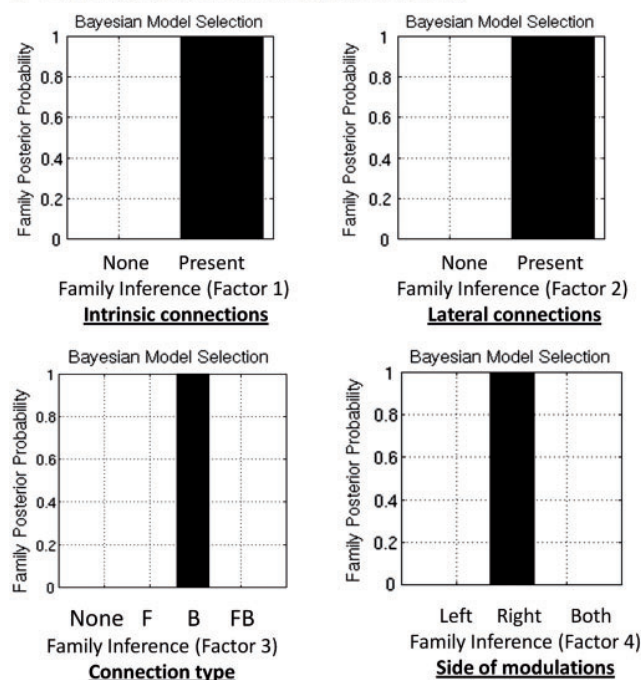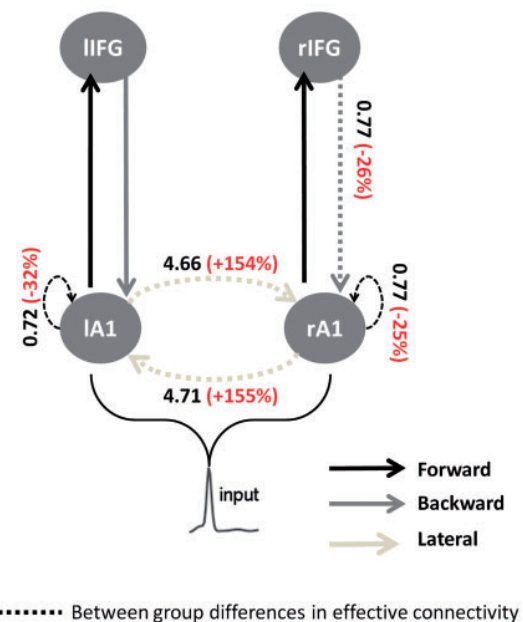
Figure 8 Dynamic causal modelling results. (**A**) Results of the four family-wise inferences. For each inference, the posterior probability of each model family is depicted. For each comparison, the family associated with a high posterior probability (*P* > 0.99) could be retained as the winning family. As each comparison yields a clear winning family, we could identify a winning model. (**B**) Winning model. Dashed arrows indicate modulated connections (i.e. connections that differ between groups) and solid arrows indicate fixed connections. Significant changes in effective coupling between control subjects and amusics are specified (in black: amount of coupling change between groups; in red: corresponding relative coupling with amusics coupling expressed in per cent of control coupling). B = Backward; F = Forward; FB = Forward Backward; lIFG = left inferior frontal gyrus; rIFG = right inferior frontal gyrus; lA1 = left primary auditory cortex; rA1 = right primary auditory cortex.
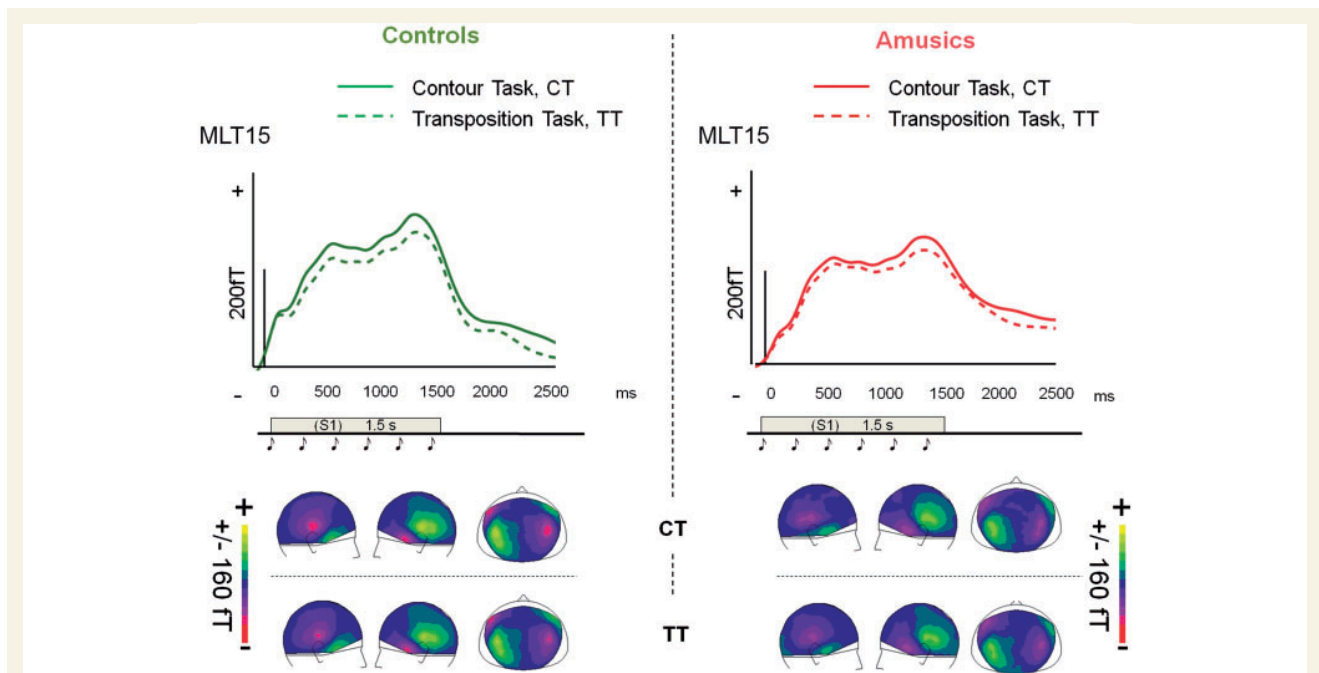
Figure 9 *Top:* Event-related fields of the 2 Hz low-pass filtered data showing the sustained evoked responses, at a left temporal MEG sensor (MLT15, MEG left temporal 15) for a time window (−100 to 2500 ms) including S1 (0 to 1500 ms), for the control group (green, on the left) and amusic group (red, on the right), and for each task (plain line: contour task, dotted line: transposition task). *Bottom:* Sensor plots of mean event-related fields for a 500 to 1500 ms time window (used for statistical analyses), for each group and each task.

responses in secondary auditory areas during long-lasting stimuli, the topographies of sustained evoked responses in the present data indicated bilateral sources in the auditory cortices for all participants. Sensor mean amplitude analyses revealed (i) a significant main effect of sensor site [$F(1,16) = 10.06$; $P = 0.005$], reflecting higher amplitude over temporal sensors in comparison to the frontal sensors; (ii) a significant main effect of hemisphere [$F(1,16) = 8.61$; $P = 0.009$], with higher amplitude in the left hemisphere in comparison with the right hemisphere; and (iii) a marginally significant main effect of task [$F(1,16) = 4.16$; $P = 0.058$], with increased mean amplitudes for the contour task in comparison with the transposition task. There was no significant main effect of group and no interaction implicating the group factor.

Retention of the melodic information: time–frequency analyses of the delay period

Gamma-power was analysed in the 30–40 Hz frequency band over the time window corresponding to the central part of the delay (2000–3000 ms). After source reconstruction, statistical comparisons were performed to compare groups for each task, and tasks for each group using two-sided *t*-tests (corrected for multiple comparisons using cluster-level statistics implemented in Fieldtrip). The results are depicted in Fig. 10.

Group effect

For the contour task, control subjects showed increased gamma synchronization (relative to baseline) in the right dorso-lateral prefrontal cortex [Brodmann area (BA) 9/46; $x = 45$, $y = 31$, $z = 25$;

$K_E = 6795 \text{ mm}^3$] in comparison with the amusic participants. For the transposition task, amusics showed increased gamma synchronization in the left dorso-lateral prefrontal cortex (BA9/46; $x = −37$, $y = 26$, $z = 29$; $K_E = 7870 \text{ mm}^3$) and the left temporo-parietal junction (BA39; $x = 46$, $y = −69$, $z = 28$; $K_E = 3430 \text{ mm}^3$) in comparison with the control participants.

Task effect

For control participants, increased gamma synchronization was observed for the contour task in comparison with the transposition task in the right inferior frontal gyrus (opercular part: $x = 56$, $y = 14$, $z = 20$; $K_E = 4220 \text{ mm}^3$) and in the left inferior frontal gyrus (opercular part: $x = −59$, $y = 6$, $z = 9$; $K_E = 5590 \text{ mm}^3$). For the amusic participants, no significant modulations of gamma oscillations with the tasks were observed.

Source analyses of the transient responses of the changed tone in S2

To investigate if amusics' altered encoding and retention of melodies in memory could be further associated to altered retrieval of melodic information, we analysed the brain responses evoked by the changed tone in S2 of different trials (for correct responses). The difference wave (different trial − same trial) observed at the sensor level revealed that for control subjects, the processing of the changed tone was associated to two evoked responses. The first one was elicited ~150 ms after the onset of the changed tone; and the second one peaked at 500 ms after the tone onset (Fig. 11A). This biphasic response was not observed in amusics (Fig. 11B). To compare the
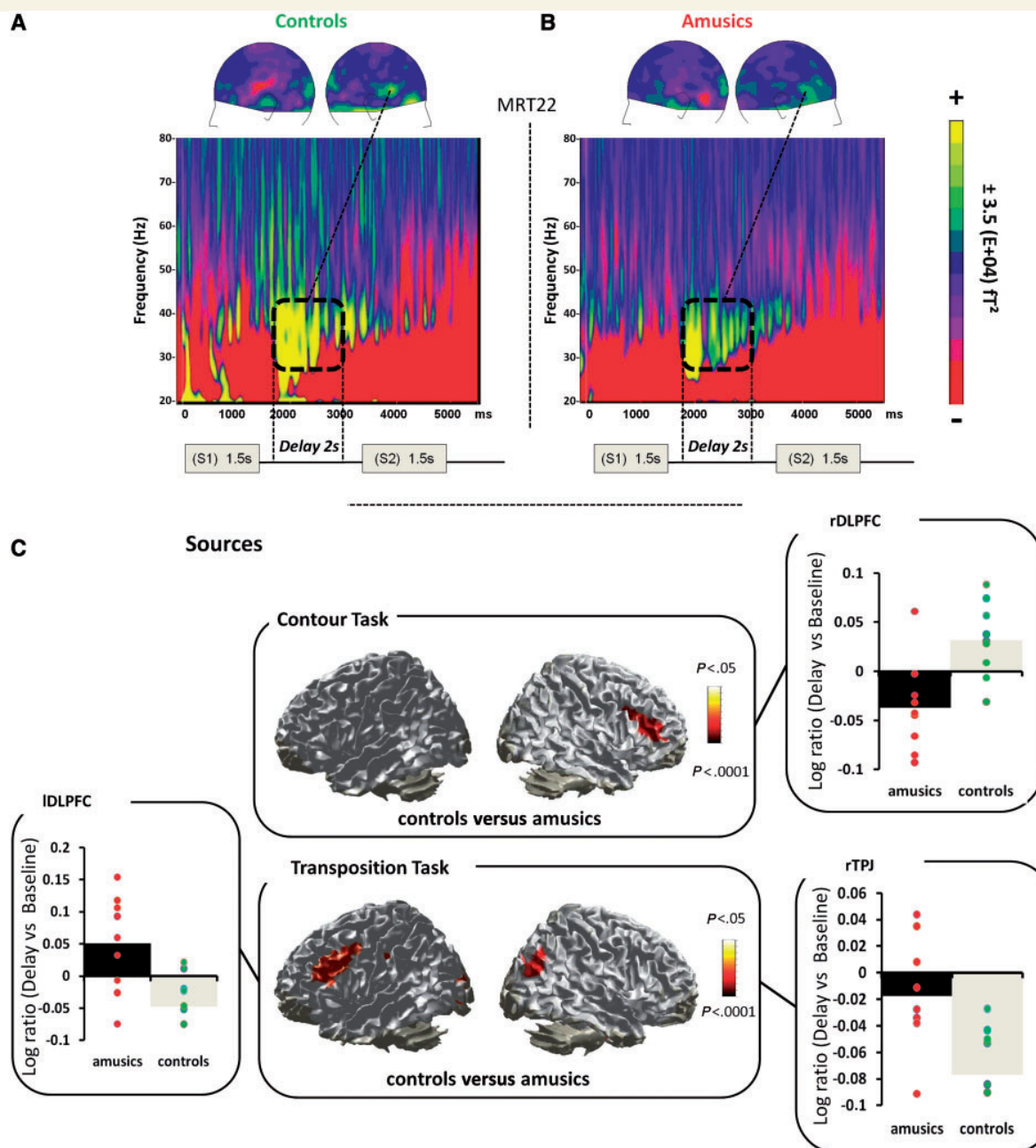
Figure 10 (**A** and **B**) Time–frequency plot of a right temporal MEG sensor (MRT22, 'MEG right temporal') for a trial time window (−100 to 5500 ms) including S1 (0 to 1500 ms), delay (1500 to 3500 ms) and S2 (3500 to 5000 ms) collapsed across conditions (contour task, transposition task), for the control group (**A**) and the amusic group (**B**). The time–frequency power values are plotted after subtraction of the mean power values of the baseline for each frequency. (**C**) Cortical meshes showing the statistics of two sided *t*-tests (corrected for multiple comparisons using cluster-level statistics) for the group comparison for each task (*upper* panel: contour task; *lower* panel: transposition task). *P*-values for the group effect are colour-coded with yellow for $P < 0.05$, red for $P < 0.01$ and black for $P < 0.001$. The surrounding panels correspond to the grand average of source data for each group and for each region. Black bars = amusics; grey bars = control subjects; circles = individual data; rDLPFC = right dorsolateral prefrontal cortex; lDLPFC = left dorsolateral prefrontal cortex; rTPJ = right temporo-parietal junction.

two participant groups for the processing of the changed tone in S2 and to investigate whether the same network was recruited by the two groups, we performed source modelling of the difference waveform (i.e. between different trials and same trials for correct responses). The bilateral fronto-temporal regions where activity was significantly different from baseline are presented in Table 2. Note that for the auditory regions, the vertex of maximum amplitude was located more laterally in the superior temporal plane than

Figure 11 (**A** and **B**) Grand average of a left temporal MEG sensor (MLT42) for a 0 to 700 ms time window after the onset of the changed tone in S2 for the contour task for each group and each type of trial. (**A**) For control subjects, green dotted line = different trials, correct responses; blue dotted line = same trials, correct responses; green plain line = difference wave (different trials − same trials for correct responses). (**B**) For amusics, red dotted line = different trials, correct responses; purple dotted line = same trials, correct responses; red plain line = difference wave (different trials − same trials for correct responses). Two-sample *t*-tests were performed at each time sample on sensor amplitudes in the 0 to 700 time window in the two groups of participants. *P*-values are reported across time in the lower panel with blue for $P < 0.05$, green for $P < 0.01$ and red for $P < 0.001$. Note that only effects lasting >15 ms were reported. (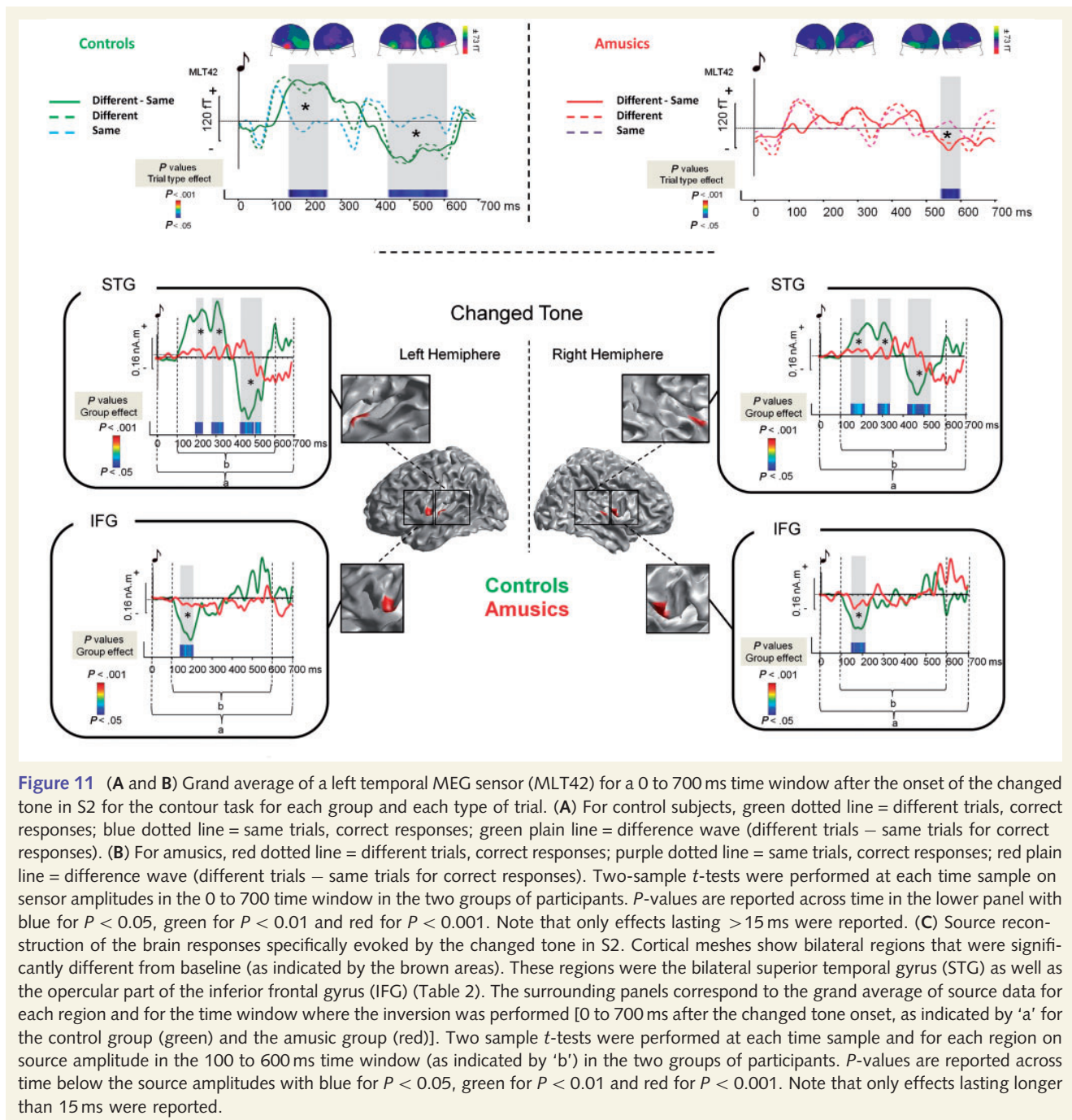**C**) Source reconstruction of the brain responses specifically evoked by the changed tone in S2. Cortical meshes show bilateral regions that were significantly different from baseline (as indicated by the brown areas). These regions were the bilateral superior temporal gyrus (STG) as well as the opercular part of the inferior frontal gyrus (IFG) (Table 2). The surrounding panels correspond to the grand average of source data for each region and for the time window where the inversion was performed [0 to 700 ms after the changed tone onset, as indicated by 'a' for the control group (green) and the amusic group (red)]. Two sample *t*-tests were performed at each time sample and for each region on source amplitude in the 100 to 600 ms time window (as indicated by 'b') in the two groups of participants. *P*-values are reported across time below the source amplitudes with blue for $P < 0.05$, green for $P < 0.01$ and red for $P < 0.001$. Note that only effects lasting longer than 15 ms were reported.

Table 2 Frontal and temporal generators of the change-specific response within S2

Lobe	Region	Hemisphere	x	y	z	mm²	Number of vertices
Frontal	Inferior frontal gyrus, pars opercularis	Right	55	4	6	49	10
		Left	−54	3	6	65	12
Temporal	Superior temporal gyrus/planum temporale	Right	55	−12	5	182	25
		Left	−52	−12	5	117	16

We report regions where activity was significantly different from baseline ($P < 0.05$, Bonferroni corrected across participants), for at least one-third of the time window of interest, for at least three participants, as assessed from the obtained posterior probability maps (see 'Materials and methods' section for details). Coordinates correspond to the vertex with maximal amplitude within each region (coordinates are in MNI space).

observed for the transient responses in S1. For the frontal generators, the coordinates were close to the coordinates for S1, but the activations were more extended. In addition, note that when we performed the same posterior probability map analyses for each group separately, the four clusters were emergent in control subjects (but only right and left frontal generators were emergent in amusics).

Two sample *t*-tests revealed group differences in source amplitude, with higher amplitudes for control participants than for amusic participants in the four regions of interest and in the following time windows: (i) right superior temporal gyrus, 150–210 ms, 290–340 ms, 425–520 ms; (ii) right inferior frontal gyrus, 160–210 ms; (iii) left inferior frontal gyrus, 150–205 ms; and (iv) left superior temporal gyrus, 190–230 ms, 280–340 ms, 425–530 ms (Fig. 11C).

Discussion

The present study investigated the cerebral correlates of pitch perception and memory in congenital amusics and matched control subjects. Using MEG, we investigated the amusic brain during encoding, retention, and retrieval of melodic information. The major finding is that pitch deficits in congenital amusia can be traced down to early brain responses. During the encoding of a melody (S1), auditory N100m components were observed after each tone for both groups of participants, but were strongly reduced and delayed in amusics. Source reconstruction analyses provided evidence of an altered recruitment of frontal and temporal N100m generators in the amusic brain during the encoding of the S1 melodies. Dynamic causal modelling of the N100m revealed an abnormally increased effective connectivity between the right and left auditory cortices in amusics, reduced intrinsic connections within the bilateral auditory cortices, as well as reduced backward connectivity between the right inferior frontal gyrus and the right auditory cortex in comparison with control subjects. During the retention of the melodic information, gamma oscillations revealed an altered recruitment of the right dorsolateral prefrontal cortex in amusics for the more difficult memory task (contour task). Finally, these altered responses observed in both encoding and retention of the melodic information of S1 were associated to an altered retrieval of the melodic information: amusics showed reduced brain responses elicited by the changed tone of the second melody (S2) in a bilateral fronto-temporal network. As predicted on the basis of previous findings (Hyde *et al*., 2006, 2007; Loui *et al*., 2009), we observed brain morphological anomalies in terms of white matter concentration and grey matter concentration in the right inferior frontal gyrus as well as in the right superior temporal gyrus of amusic participants in comparison to control participants. The convergence between functional and structural brain differences provided evidence for abnormalities in a fronto-temporal pathway (including functional anomaly in the auditory cortex) associated with the pitch encoding and short-term memory dysfunctions in congenital amusia.

Short-term memory deficit in congenital amusia

Control participants' performance was high for both the contour task and transposition task. Amusic participants' performance was unimpaired for the transposition task, but was strongly impaired for the contour task (all amusic participants exhibited a deficit in comparison with control participants). Pitch discrimination deficits can be excluded as the sole origin of the impaired performance as several amusics' pitch thresholds were comparable to control subjects' pitch thresholds (see also Tillmann *et al*., 2009). Additionally, we observed impaired short-term memory in amusics in comparison to control subjects even for changed tones of large interval sizes in the contour task (i.e. the large changes were superior to all amusics' pitch discrimination threshold), thus allowing us to argue that the deficit in short-term memory performance for the contour task was not merely a consequence of increased pitch discrimination threshold in amusics. The impaired performance of amusics in the contour task is in agreement with previously reported deficits in the short-term memory of tones (Gosselin *et al*., 2009; Tillmann *et al*., 2009; Williamson and Stewart, 2010; Williamson *et al*., 2010) and suggests that amusic individuals experience difficulties in maintaining the memory trace of melodic contour information.

Structural abnormalities in the amusic brain

VBM analyses revealed decreased white matter concentration and increased grey matter concentration in the right inferior frontal gyrus (BA 47) of amusics (relative to control subjects). These findings are in agreement with Hyde *et al*. (2006, 2007), who suggested that cortical abnormalities in the amusic brain might occur consequently to an anomaly in cortical development. Amusics' white matter abnormalities are hypothesized to reflect anomalous connectivity between auditory and frontal cortical areas (Hyde *et al*., 2006, 2007, 2011; Loui *et al*., 2009). This hypothesis was also supported by the positive correlation between white matter concentration in right inferior frontal gyrus and right superior temporal gyrus in amusic participants of our present study. Moreover, our present data revealed less grey matter concentration in the right superior temporal gyrus for amusics in comparison with control subjects; this result contrasts with cortical thickness data of Hyde *et al*. (2007). This opposite result pattern can be related to the methodological differences between cortical thickness and VBM. In other domains, several studies have demonstrated opposite result patterns in a given brain region when comparing cortical thickness or VBM measures. For example, Park *et al*. (2009) found a significant reduction of grey matter concentration in the primary and associative visual cortices in a group of blind participants using VBM. However, using cortical thickness and surface area measures in the same group of participants, they found a thicker cortex in the same areas. The authors attributed these volumetric atrophies to the decreased cortical surface area despite increased cortical thickness. They proposed that the two measures are complementary and might differently reflect morphological alteration during the developmental period (see Jiang *et al*., 2009, for converging

evidence of opposite results when comparing data obtained with the two methods). Combining these different approaches of cortical anatomy in amusia remains to be done within the same participants, but nevertheless, the present data agreed with Hyde *et al.* (2007) in revealing structural abnormalities in the right superior temporal gyrus of the amusic brain for grey matter concentrations.

Encoding of melodies in congenital amusia

Source reconstruction of the N100m allowed us to disentangle activity in bilateral auditory regions (in Heschl's gyrus/superior temporal gyrus/planum temporale/planum polare) and in the pars opercularis of the inferior frontal gyrus (BA 44). This observation is in line with numerous studies showing the major role of these areas in music perception and memory (Zatorre *et al.*, 1994; Griffiths *et al.*, 1999; Schulze *et al.*, 2009; see also Griffiths, 2001; Peretz and Zatorre, 2005 for reviews). For the N100m component, the observed frontal generators are in line with previous EEG data (Näätänen and Picton, 1987; Alcaini *et al.*, 1994; Giard *et al.*, 1994; Pantev *et al.*, 1995; see Trainor and Unrau, 2012 for review), suggesting that separate frontal and temporal neural systems mediating different processes could be activated during the N1-time range. Further evidence comes from intracranial recordings (Edwards *et al.*, 2005), which indicated that during the N1-time range, the most strongly activated brain regions were not only the superior surface of the temporal lobe, but also some areas of the frontal lobe.

Note that in comparison with tone 1, tones 2 to 6 recruited a more lateral part of Heschl's gyrus (and planum temporale). This finding is in agreement with Patterson *et al.* (2002) showing that the lateral part of the Heschl's gyrus is recruited by melodies in comparison to fixed-pitch sequences, suggesting that this part of the auditory cortex plays a role in melodic processing. It is well known that bilateral auditory regions (including superior temporal gyrus, Heschl's gyrus, planum polare and planum temporale) play a critical role in the encoding of acoustic features of individual tones (Griffiths, 1999, 2001) and in melody perception (Zatorre *et al.*, 1994; Liégeois-Chauvel *et al.*, 1998) and are more strongly recruited during high-load conditions, which require active rehearsal (Zatorre *et al.*, 1994; Griffiths *et al.*, 1999), such as retaining a pitch in memory while subsequent tones are presented. In addition to the participation of auditory cortices, Griffiths (1999, 2001; Griffiths *et al.*, 2000) has suggested that higher-order auditory patterns and information are analysed by distributed networks including temporal and frontal lobes, which are necessary for online maintenance and encoding of tonal patterns. For frontal regions, the inferior frontal gyrus has been further suggested to play a role in integrating sequential auditory events and in the encoding of tonality (Zatorre *et al.*, 1994; Griffiths *et al.*, 1999; Gaab *et al.*, 2003; Peretz and Zatorre, 2005; Schulze *et al.*, 2009, 2011).

For the first tone, source analyses revealed near-normal cortical distribution and activity, both for the P50m (e.g. Yvert *et al.*, 2001) and the N100m components in the amusic brain. As also reported by Peretz *et al.* (2005, 2009) and Hyde *et al.* (2011), the functioning of amusics' auditory cortices was unimpaired for a relatively simple acoustic task (the first tone of the sequence was presented after a period of silence, and should thus largely recruit non-specific components). However, latency analyses of the N100m revealed slightly delayed inferior frontal gyrus responses in amusics in comparison to control subjects, thus providing new functional evidence of an altered fronto-temporal network, even for the encoding of the first tone of a sequence.

For tones 2 to 6, the N100m differed between amusics and control subjects in amplitude and latency in Heschl's gyrus/superior temporal gyrus and inferior frontal gyrus, thus placing the deficit of the amusic brain early in the auditory processing stream. Amplitude data revealed that amusic participants recruited less strongly a bilateral fronto-temporal network than did control subjects. This network has been previously shown to allow auditory information to be processed, encoded, maintained on-line and related to previous elements of a sequence (Zatorre *et al.*, 1994, 2002; Griffiths, 2001; Janata *et al.*, 2002*a, b*; Gaab *et al.*, 2003; Peretz and Zatorre, 2005). Observing an abnormal recruitment of frontal (inferior frontal gyrus) and temporal (Heschl's gyrus/superior temporal gyrus) regions during the encoding of tones 2 to 6 in amusics can be taken as functional correlates of the deficit in perception and memory of tone sequences in congenital amusia. It might be argued that decreased attention could explain the decreased N100m amplitude of amusics in the contour task. However, two aspects of the data allow us to reject this argument: (i) if amusics were paying less attention because, for example, they found the contour task too difficult, the amplitude of their N100m should have been attenuated in comparison with the transposition task. However, no interaction between Group and Task was observed for the N100m amplitude, thus suggesting that participants' attentional or task-related strategies did not have a major influence on the amplitude differences in the present N100m data; (ii) the sustained evoked responses data revealed that task-related (the difficult contour task versus the easier transposition task) attentional modulation is preserved in congenital amusia. Indeed, this task effect on sustained evoked responses is in line with the hypothesis that the contour task should involve more attentional processes than the transposition task, as these responses are known to be modulated by attention (Picton *et al.*, 1978; Bidet-Caulet *et al.*, 2007). As there were no significant main effect of group and no significant interactions involving the group factor, the data pattern suggests that amusic participants maintained a sustained attention level, even for the difficult contour task.

For tones 2 to 6, increased latency of the N100m was observed in frontal and temporal regions for the amusic group (relative to the control group). It can be interpreted as reflecting delayed encoding of S1 in the amusic brain as well as impaired high-level processing and stimulus representations (e.g. delayed responses associated with difficulties to maintain tone sequences in memory). The most interesting point of these data is the increased Heschl's gyrus latency for tones 2 to 6 in the amusic group. Together with the amplitude data, this finding provides the first evidence of functional abnormalities in the auditory cortex in congenital amusia.

To further assess the differences between the two participant groups during the encoding of the melodies, we used dynamic causal modelling to investigate the effective connectivity between the frontal and temporal sources of the N100m. Most approaches of connectivity in the MEG/EEG literature use functional measures, such as phase-synchronization, temporal correlations or coherence, to establish statistical dependencies between activities in two different cortical areas. Although functional connectivity can be used to establish statistical dependency, it does not provide information about the causal architecture of the interactions. In dynamic causal modelling, this influence is parameterized in a causal model, which can then be estimated using Bayesian inference. Using the same underlying cortical architecture, but differing between participant groups in the modulations of specific connection types, we observed decreased intrinsic connectivity for amusics in both auditory cortices relative to control subjects, as well as increased connectivity in amusics between the two auditory cortices, as previously observed with functional MRI (Hyde et al., 2011). These abnormalities observed in both intrinsic and lateral connections in the auditory cortex provide additional evidence of functional anomaly of the auditory cortex in amusia (together with the observed delayed N100m responses). More precisely, the observed hyper-connectivity between the two auditory cortices in the amusic brain might be a marker of the primary deficit, as also observed in other developmental disorders (see Wolf et al., 2010 for dyslexia; Hyde et al., 2011 for converging evidence in amusia) or rather reveal compensatory mechanisms of the amusic brain. The latter would suggest that amusics might compensate for an impoverished processing in the right auditory cortex by recruiting the contralateral auditory cortex. These functional abnormalities in both auditory cortices were associated to decreased right frontal-to-temporal connectivity in amusics. This decreased right fronto-temporal connectivity is in agreement with the functional MRI data from Hyde et al. (2011), showing decreased connectivity between right inferior frontal gyrus and right superior temporal gyrus in amusics in comparison with control subjects during passive listening. Note that our present data, which are based on the effective connectivity approach (as opposed to functional connectivity used in Hyde et al., 2011), allow us to more precisely ascribe abnormal brain responses in amusics to reduced backward connections, but not forward connections.

Maintenance of melodic information in memory

To investigate neural correlates of maintenance of tone information in memory, we have analysed oscillatory synchronization in the gamma frequency during the retention period. We observed enhanced gamma-power in the right dorsolateral prefrontal cortex in control subjects in comparison with amusics for the difficult contour task. This finding is in agreement with previous data showing that high-level representations of task-relevant information are reflected in the gamma-frequency (Jensen et al., 2007), and that increased gamma-frequency power in fronto-temporal areas is related with short-term memory processing (Kaiser et al., 2003). This hypothesis is also supported by the task

comparison for control subjects, showing that the contour task recruited more strongly a bilateral frontal network involving the inferior frontal gyrus. Moreover, this right-lateralized group effect is in agreement with neuroimaging data showing that the dorsolateral prefrontal cortex participates in short-term memory and working memory (Glahn et al., 2002; Jerde et al., 2011). Note, however, that these modulations with tasks were not observed in amusics, thus providing another evidence of altered processing of melodic information in the amusic brain. In addition, for the easier transposition task, increased gamma synchronization was observed in the left inferior frontal gyrus and in the right temporo-parietal junction in amusics in comparison with control subjects, suggesting that the relative involvement of each hemisphere might differ between the two groups. The interaction between frontal and temporo-parietal regions is known to be implicated in working memory for musical material (Zatorre et al., 1994; Jerde et al., 2011), and the present finding thus suggests that, in order to perform as well as control subjects for the easier transposition task, amusics had to recruit more strongly this network. Following this hypothesis, we suggest that for the difficult contour task, this compensatory strategy was not sufficient to overcome the deficit in pitch memory.

Retrieval of melodic information

Source reconstruction of the event-related fields elicited by the changed tone (difference wave between different trials and same trials) allowed us to observe activity in bilateral superior temporal gyrus and in the pars opercularis of the inferior frontal gyrus (BA 44) in the control group. Observing that the same network is recruited during encoding and retrieval of melodic information is in line with research showing the role of these areas in music perception and memory (Zatorre et al., 1994, 2002; Griffiths, 2001; Janata et al., 2002a, b; Gaab et al., 2003; Peretz and Zatorre, 2005). The present data revealed an abnormal recruitment of frontal (inferior frontal gyrus) and temporal (superior temporal gyrus) regions during the retrieval of melodic information in the amusic brain; data that further reflect the functional correlates of the deficit in short-term memory of melodic information in congenital amusia.

Convergence between functional and anatomical data

MEG source locations revealing group differences were congruent with bilateral fronto-temporal networks well-known to be involved in short-term memory for pitch sequences (Zatorre et al., 1994; Griffiths et al., 1999; Schulze et al., 2009), and attentive listening to musical sequences (Janata et al., 2002b) (Table 3 and Fig. 12). In the right superior temporal gyrus, anatomical group differences were spatially congruent with the present MEG data, and activations peaks previously reported in positron emission tomography (PET)/functional MRI studies investigating the active processing of pitch sequences (i.e. short-term memory, attentive listening). However, for frontal regions, anatomical and functional group differences were observed in different parts of the inferior frontal gyrus. VBM analyses (present data and data of Hyde et al., 2006) revealed abnormal grey and white matter concentrations

Table 3 MNI coordinates (mm) of fronto-temporal activations reported in previous functional MRI/PET and DTI studies

Reference	Lobe	Region	Contrasts	Hemisphere	x	y	z
Zatorre *et al.*, 1994	Frontal	Insula/inferior frontal gyrus pars opercularis	Tone judgement– passive melodies	Right	38	20	5
				Left	−31	22	8
	Temporal	Superior temporal gyrus	Passive melodies–scanner noise	Right	62	−25	3
Griffths *et al.*, 1999	Frontal	Inferior frontal gyrus, pars opercularis	Pitch memory task– resting	Right	30	26	2
				Right	40	18	8
				Left	−48	10	4
	Temporal	Heschl's gyrus		Right	66	−30	6
Janata *et al.*, 2002*a*	Frontal	Inferior frontal gyrus, pars opercularis	Listen–resting state	Right	45	15	5
				Right	−38	19	5
	Temporal	Superior temporal gyrus	Listen–resting state	Right	60	19	5
			Attend–resting state	Left	−60	−15	5
				Left	−52	−19	5
Hyde *et al.*, 2011	Frontal	Inferior frontal gyrus, pars opercularis	Amusics-control subjects, pitch changed – fixed pitch	Right	34	32	2
	Temporal	Heschl's gyrus/planum temporale	Pitch co-variation in control subjects	Right	56	−12	4
		Planum temporale	Pitch co-variation in amusics	Right	60	−18	8
		Heschl's gyrus/planum temporale		Left	−44	−30	8
Loui *et al.*, 2009	Frontal	Inferior frontal gyrus, pars opercularis	Seed regions for tractography	Right and left			
	Temporal	Posterior superior temporal gyrus		Right and left			

The first three studies were run with typical individuals, whereas the last two studies compared amusics and control groups.

in the amusic brain in the anterior part of the right inferior frontal gyrus (BA 45/47; Fig. 12), whereas the MEG data revealed functional abnormalities in more posterior parts of bilateral inferior frontal gyrus (BA 44) during encoding and retrieval of melodic information (S1 and S2), and in a more dorsal part (BA9/46) during the retention. The anterior part of the right inferior frontal gyrus has also been shown to exhibit abnormal blood oxygen level-dependant deactivation in the amusic brain during passive listening to tone sequences (Hyde *et al.*, 2011). The posterior part of the inferior frontal gyrus was used by Loui *et al.* (2009) as a seed region for DTI tractography revealing an anomalous connectivity in the right arcuate fasciculus in congenital amusia (Fig. 12 and Table 3).

For music processing, these different subregions of right inferior frontal gyrus have been previously reported in typical individuals' brains (e.g. Maess *et al.*, 2001; Koelsch *et al.*, 2002 for BA 44; Levitin and Menon, 2003 for BA 45/47; Tillmann *et al.*, 2003, 2006). Hyde *et al.* (2006, 2011) discussed their VBM and functional MRI data of amusic participants in link with two subregions involved in music perception and structure processing. They suggested that BA 44 and BA 45/47 are part of the same network of right inferior frontal cortex and are involved in pitch sequence perception and integration. Similarly, Hagoort (2005) discusses the role of the left inferior frontal gyrus in language processing by regrouping BA 44, 45 and 47, notably their role in enabling the integration of structural information. While for language processing, subregions have also been specified for their respective roles in syntactic processing and phonological memory (BA 44) and semantic processing (BA 45/47) (Friederici, 2002, 2012), research in music processing still needs to investigate more specifically the respective roles of these subregions in right inferior frontal gyrus for pitch perception and memory in both normal-functioning and amusic brains.

Conclusion

In congenital amusics, N100m components were abnormal and strongly delayed in bilateral inferior frontal gyrus and Heschl's gyrus/superior temporal gyrus during the encoding of melodies, frontal gamma synchronization was decreased during the retention of melodic information, and fronto-temporal responses were altered during the retrieval of this information. These functional anomalies were related to abnormal grey matter and white matter concentrations in the same brain regions and to a deficit in memory processing observed at the behavioural level. This data set is in agreement with current hypotheses about the role of frontal and temporal structures (including auditory cortices) and of the fronto-temporal pathway in music processing, as well as its impairment in this 'musical handicap'. The present study
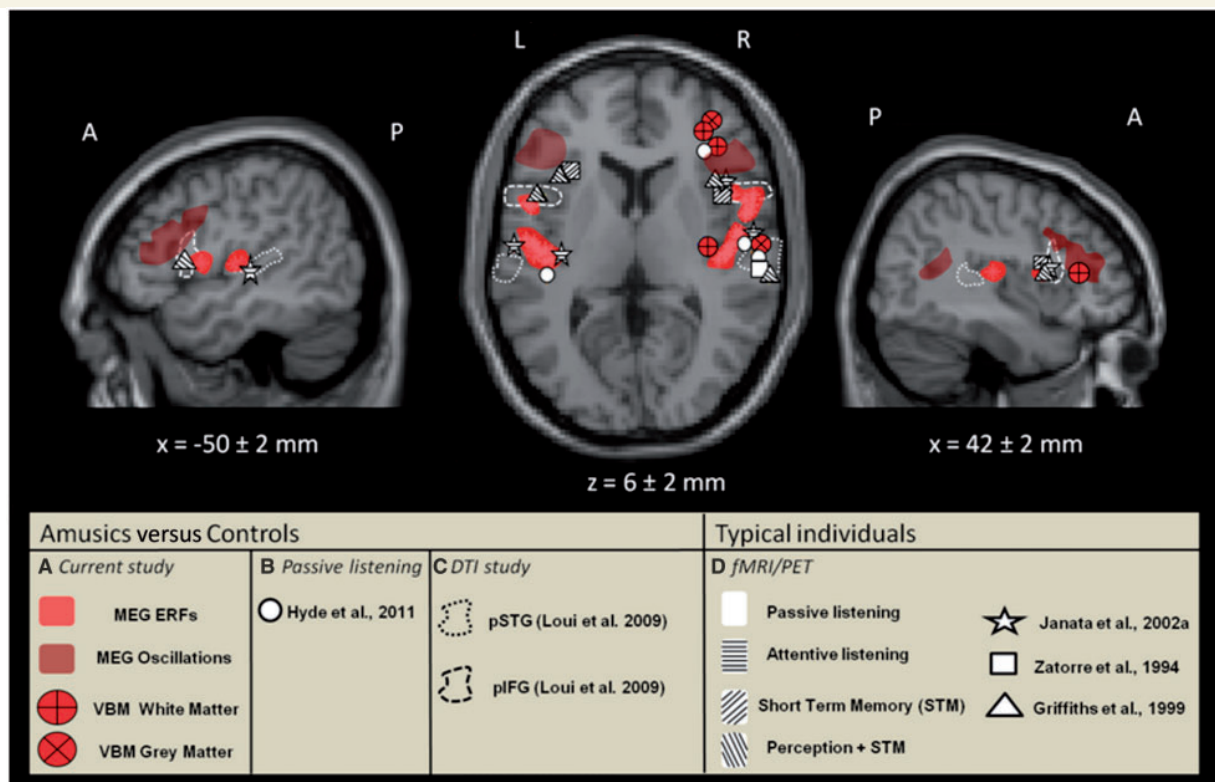
Figure 12 Comparison of MEG and VBM group differences observed in the present study (**A** in red and purple) with previous functional MRI (**B**) and DTI (**C**) studies comparing control subjects and amusics and with activation peaks observed in previous studies using PET or functional MRI in typical individuals (**D**) (Table 3). (**B**) Coordinates of the functional group difference in the right inferior frontal gyrus and of pitch co-variation data for both amusics and control subjects in the auditory cortex (Hyde *et al.*, 2011, circles; Table 3). (**C**) Approximate location of the posterior inferior frontal gyrus (pIFG) and posterior superior temporal gyrus (pSTG) seed regions used in the tractography study by Loui *et al.* (2009) (bold dotted lines and plain dotted line, respectively). Note that these regions were manually selected by a neurologist who was blind concerning the protocol and previously reported activations. (**D**) Coordinates of activations obtained with pitch material during auditory short-term memory tasks (Zatorre *et al.*, 1994, squares; Griffths *et al.*, 1999, triangles), attentive listening (Janata *et al.*, 2002, stars), and passive listening (Zatorre *et al.*, 1994, squares). Right-oriented bars = short-term memory; left-oriented bars = pitch perception and memory; horizontal bars = attentive listening; white background = passive listening. Activations are displayed on the single subject T_1 image provided by SPM8. The figure depicts activation with coordinates that were 2 mm up or down the central slice. R = right; L = left; A = anterior; P = posterior); PET = positron emission tomography.

improves our understanding not only of congenital amusia itself, but also, more generally, of typical brain functioning related to auditory perception and memory, as well as music processing. In particular, control participants' data reveal that the same bilateral fronto-temporal areas were recruited during the different stages of encoding, maintenance and retrieval of auditory information, suggesting the same networks are involved in these different aspects of melody processing and memory representation. Research investigating the neural correlates of music processing is an expanding research domain that also addresses the question of shared neural correlates with language processing (Patel, 2003, 2008). This comparison can also be extended to the understanding of developmental language impairments in parallel to the present musical impairments (as previously suggested by Hyde *et al.*, 2006). Beyond providing new data for the understanding of the deficits in the amusic brain, our data raises issues about the cerebral correlates of developmental disorders in general, notably by

suggesting that cerebral correlates of related deficits can be observed already in early brain responses. This later point represents a major contribution to the comprehension of neural correlates underlying congenital amusia as previous data have suggested that the functional neural anomaly mainly lies outside the auditory cortex (Peretz *et al.*, 2005, 2009; Moreau *et al.*, 2009; Hyde *et al.*, 2011).

Acknowledgements

References

Aguera PE, Jerbi K, Caclin A, Bertrand O. ELAN: a software package for analysis and visualization of MEG, EEG, and LFP signals. Comput Intell Neurosci 2011; 2011: 158970.

Alcaini M, Giard MH, Thevenet M, Pernier J. Two separate frontal components in the N1 wave of the human auditory evoked response. Psychophysiology 1994; 31: 611–15.

Ashburner J, Friston KJ. Unified segmentation. Neuroimage 2005; 26: 839–51.

Ayotte J, Peretz I, Hyde KL. Congenital amusia: a group study of adults afflicted with a music-specific disorder. Brain 2002; 125: 238–51.

Bidet-Caulet A, Fischer C, Besle J, Aguera P, Giard MH, Bertrand O. Effects of selective attention on the electrophysiological representation of concurrent sounds in the human auditory cortex. J Neurosci 2007; 27: 9252–61.

Bouet R, Jung J, Delpuech C, Ryvlin P, Isnard J, Guenot M, et al. Towards source volume estimation of interictal spikes in focal epilepsy using magnetoencephalography. Neuroimage 2012; 59: 3955–66.

Caclin A, McAdams S, Smith BK, Giard MH. Interactive processing of timbre dimensions: an exploration with event-related potentials. J Cogn Neurosci 2008; 20: 49–64.

David O, Friston KJ. A neural mass model for MEG/EEG: coupling and neuronal dynamics. Neuroimage 2003; 20: 1743–55.

David O, Kiebel SJ, Harrison LM, Mattout J, Kilner JM, Friston KJ. Dynamic causal modeling of evoked responses in EEG and MEG. Neuroimage 2006; 30: 1255–72.

Demarquay G, Caclin A, Brudon F, Fischer C, Morlet D. Exacerbated attention orienting to auditory stimulation in migraine patients. Clin Neurophysiol 2011; 122: 1755–63.

Edwards E, Soltani M, Deouell LY, Berger MS, Knight RT. High gamma activity in response to deviant auditory stimuli recorded directly from human cortex. J Neurophysiol 2005; 94: 4269–80.

Foxton J, Dean J, Gee R, Peretz I, Griffiths T. Characterization of deficits in pitch perception underlying 'tone deafness'. Brain 2004; 127: 801–10.

Friederici AD. Towards a neural basis of auditory sentence processing. Trends Cogn Sci 2002; 6: 78–84.

Friederici AD. The cortical language circuit: from auditory perception to sentence comprehension. Trends Cogn Sci 2012; 16: 262–8.

Friston K, Harrison L, Daunizeau J, Kiebel S, Phillips C, Trujillo-Barreto N, et al. Multiple sparse priors for the M/EEG inverse problem. Neuroimage 2008; 39: 1104–20.

Friston KJ, Penny W. Posterior probability maps and SPMs. Neuroimage 2003; 19: 1240–9.

Gaab N, Gaser C, Zaehle T, Jancke L, Schlaug G. Functional anatomy of pitch memory—an fMRI study with sparse temporal sampling. Neuroimage 2003; 19: 1417–26.

Garrido MI, Kilner JM, Kiebel SJ, Stephan KE, Friston KJ. Dynamic causal modelling of evoked potentials: a reproducibility study. Neuroimage 2007; 36: 571–80.

Giard MH, Perrin F, Echallier JF, Thévenet M, Froment JC, Pernier J. Dissociation of temporal and frontal components in the human auditory N1 wave: a scalp current density and dipole model analysis. Electroencephalogr Clin Neurophysiol 1994; 92: 238–52.

Glahn DC, Kim J, Cohen MS, Poutanen VP, Therman S, Bava S, et al. Maintenance and manipulation in spatial working memory: dissociations in the prefrontal cortex. Neuroimage 2002; 17: 201–13.

Gosselin N, Jolicoeur P, Peretz I. Impaired memory for pitch in congenital amusia. Ann NY Acad Sci 2009; 1169: 270–2.

Griffiths TD. Human complex sound analysis. Clin Sci (Lond) 1999; 96: 231–4.

Griffiths TD. The neural processing of complex sounds. Ann N Y Acad Sci 2001; 930: 133–42.

Griffiths TD, Johnsrude I, Dean JL, Green GG. A common neural substrate for the analysis of pitch and duration pattern in segmented sound? Neuroreport 1999; 10: 3825–30.

Griffiths TD, Penhune V, Peretz I, Dean JL, Patterson RD, Green GG. Frontal processing and auditory perception. Neuroreport 2000; 11: 919–22.

Gross J, Kujala J, Hamalainen M, Timmermann L, Schnitzler A, Salmelin R. Dynamic imaging of coherent sources: studying neural interactions in the human brain. Proc Natl Acad Sci USA 2001; 98: 694–9.

Guthrie D, Buchwald JS. Significance testing of difference potentials. Psychophysiology 1991; 28: 240–4.

Hagoort P. On Broca, brain, and binding: a new framework. Trends Cogn Sci 2005; 9: 416–23.

Hyde KL, Lerch JP, Zatorre RJ, Griffiths TD, Evans AC, Peretz I. Cortical thickness in congenital amusia: when less is better than more. J Neurosci 2007; 27: 13028–32.

Hyde KL, Peretz I. Brains that are out of tune but in time. Psychol Sci 2004; 15: 356–60.

Hyde KL, Zatorre R, Griffiths T, Lerch J, Peretz I. Morphometry of the amusic brain: a two-site study. Brain 2006; 129: 2562–70.

Hyde KL, Zatorre RJ, Peretz I. Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. Cereb Cortex 2011; 21: 292–9.

Janata P, Birk JL, Van Horn JD, Leman M, Tillmann B, Bharucha JJ. The cortical topography of tonal structures underlying Western music. Science 2002a; 298: 2167–70.

Janata P, Tillmann B, Bharucha JJ. Listening to polyphonic music recruits domain-general attention and working memory circuits. Cogn Affect Behav Neurosci 2002b; 2: 121–40.

Jensen O, Kaiser J, Lachaux JP. Human gamma-frequency oscillations associated with attention and memory. Trends Neurosci 2007; 30: 317–24.

Jerde TA, Childs SK, Handy ST, Nagode JC, Pardo JV. Dissociable systems of working memory for rhythm and melody. Neuroimage 2011; 57: 1572–9.

Jiang J, Zhu W, Shi F, Liu Y, Li J, Qin W, et al. Thick visual cortex in the early blind. J Neurosci 2009; 29: 2205–11.

Johnsrude I, Penhune VB, Zatorre RJ. Functional specificity in the right human auditory cortex for perceiving pitch direction. Brain 2000; 123: 155–63.

Kaiser J, Lutzenberger W. Human gamma-band activity: a window to cognitive processing. Neuroreport 2005; 16: 207–11.

Kaiser J, Ripper B, Birbaumer N, Lutzenberger W. Dynamics of gamma-band activity in human magnetoencephalogram during auditory pattern working memory. Neuroimage 2003; 20: 816–27.

Kiebel SJ, Garrido MI, Moran R, Chen CC, Friston KJ. Dynamic causal modeling for EEG and MEG. Hum Brain Mapp 2009; 30: 1866–76.

Koelsch S, Gunter TC, v Cramon DY, Zysset S, Lohmann G, Friederici AD. Bach speaks: a cortical "language-network" serves the processing of music. Neuroimage 2002; 17: 956–66.

Levitin DJ, Menon V. Musical structure is processed in "language" areas of the brain: a possible role for Brodmann Area 47 in temporal coherence. Neuroimage 2003; 20: 2142–52.

Liegeois-Chauvel C, Peretz I, Babai M, Laguitton V, Chauvel P. Contribution of different cortical areas in the temporal lobes to music processing. Brain 1998; 121: 1853–67.

Litvak V, Mattout J, Kiebel S, Phillips C, Henson R, Kilner J, et al. EEG and MEG data analysis in SPM8. Comput Intell Neurosci 2011; 2011: 852961.

Loui P, Alsop D, Schlaug G. Tone deafness: a new disconnection syndrome? J Neurosci 2009; 29: 10215–20.

Maess B, Koelsch S, Gunter TC, Friederici AD. Musical syntax is processed in Broca's area: an MEG study. Nat Neurosci 2001; 4: 540–5.

Mainy N, Kahane P, Minotti L, Hoffmann D, Bertrand O, Lachaux JP. Neural correlates of consolidation in working memory. Hum Brain Mapp 2007; 28: 183–93.

Mandell J, Schulze K, Schlaug G. Congenital amusia: an auditory-motor feedback disorder? Restor Neurol Neurosci 2007; 25: 323–34.

Mattout J, Henson RN, Friston KJ. Canonical source reconstruction for MEG. Comput Intell Neurosci 2007; 2007: 67613.

Mattout J, Phillips C, Penny WD, Rugg MD, Friston KJ. MEG source localization under multiple constraints: an extended Bayesian framework. Neuroimage 2006; 30: 753–67.

Moreau P, Jolicoeur P, Peretz I. Automatic brain responses to pitch changes in congenital amusia. Ann NY Acad Sci 2009; 1169: 191–4.

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

Pantev C, Bertrand O, Eulitz C, Verkindt C, Hampson S, Schuierer G, et al. Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. Electroencephalogr Clin Neurophysiol 1995; 94: 26–40.

Park HJ, Lee JD, Kim EY, Park B, Oh MK, Lee S, et al. Morphological alterations in the congenital blind based on the analysis of cortical thickness and surface area. Neuroimage 2009; 47: 98–106.

Patel AD. Language, music, syntax and the brain. Nat Neurosci 2003; 6: 674–81.

Patel AD. Music, language and the brain. Oxford: Oxford University Press; 2008.

Patterson RD, Uppenkamp S, Johnsrude IS, Griffiths TD. The processing of temporal pitch and melody information in auditory cortex. Neuron 2002; 36: 767–76.

Penny WD, Stephan KE, Daunizeau J, Rosa MJ, Friston KJ, Schofield TM, et al. Comparing families of dynamic causal models. PLoS Comput Biol 2010; 6: e1000709.

Peretz I. Processing of local and global musical information by unilateral brain-damaged patients. Brain 1990; 113: 1185–205.

Peretz I, Ayotte J, Zatorre R, Mehler J, Ahad P, Penhune V, et al. Congenital amusia: a disorder of fine-grained pitch discrimination. Neuron 2002; 33: 185–91.

Peretz I, Brattico E, Jarvenpaa M, Tervaniemi M. The amusic brain: in tune, out of key, and unaware. Brain 2009; 132: 1277–86.

Peretz I, Brattico E, Tervaniemi M. Abnormal electrical brain responses to pitch in congenital amusia. Ann Neurol 2005; 58: 478–82.

Peretz I, Champod S, Hyde K. Varieties of musical disorders: the Montreal Battery of Evaluation of Amusia. Ann NY Acad Sci 2003; 999: 58–75.

Peretz I, Zatorre RJ. Brain organization for music processing. Annu Rev Psychol 2005; 56: 89–114.

Picton TW, Woods DL, Proulx GB. Human auditory sustained potentials: part I. The nature of the response. Electroencephalogr Clin Neurophysiol 1978; 45: 186–97.

Ramirez RR, Kopell BH, Butson CR, Hiner BC, Baillet S. Spectral signal space projection algorithm for frequency domain MEG and EEG denoising, whitening, and source imaging. Neuroimage 2011; 56: 78–92.

Schulze K, Gaab N, Schlaug G. Perceiving pitch absolutely: comparing absolute and relative pitch possessors in a pitch memory task. BMC Neurosci 2009; 10: 106.

Schulze K, Mueller K, Koelsch S. Neural correlates of strategy use during auditory working memory in musicians and non-musicians. Eur J Neurosci 2011; 33: 189–96.

Stewart L. Congenital amusia. Curr Biol 2006; 16: R904–6.

Stewart L. Fractionating the musical mind: insights from congenital amusia. Curr Opin Neurobiol 2008; 18: 127–30.

Tallon-Baudry C, Bertrand O. Oscillatory gamma activity in humans and its role in object representation. Trends Cogn Sci 1999; 3: 151–62.

Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J. Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. J Neurosci 1996; 16: 4240–9.

Tallon-Baudry C, Kreiter A, Bertrand O. Sustained and transient oscillatory responses in the gamma and beta bands in a visual short-term memory task in humans. Vis Neurosci 1999; 16: 449–59.

Tesche CD, Uusitalo MA, Ilmoniemi RJ, Huotilainen M, Kajola M, Salonen O. Signal-space projections of MEG data characterize both distributed and well-localized neuronal sources. Electroencephalogr Clin Neurophysiol 1995; 95: 189–200.

Tillmann B, Schulze K, Foxton JM. Congenital amusia: a short-term memory deficit for non-verbal, but not verbal sounds. Brain Cogn 2009; 71: 259–64.

Tillmann B, Janata P, Bharucha JJ. Activation of the inferior frontal cortex in musical priming. Ann N Y Acad Sci 2003; 999: 209–11.

Tillmann B, Koelsch S, Escoffier N, Bigand E, Lalitte P, Friederici AD, von Cramon DY. Cognitive priming in sung and instrumental music: activation of inferior frontal cortex. Neuroimage 2006; 31: 1771–82.

Trainor LJ, Unrau AJ. Development of pitch and music perception. In: Werner L, Fay RR, Popper AN, editors. Springer handbook of auditory research: human auditory development. New York: Springer; 2012. p. 223–54.

Williamson VJ, McDonald C, Deutsch D, Griffiths TD, Stewart L. Faster decline of pitch memory over time in congenital amusia. Adv Cogn Psychol 2010; 6: 15–22.

Williamson VJ, Stewart L. Memory for pitch in congenital amusia: beyond a fine-grained pitch discrimination problem. Memory 2010; 18: 657–69.

Wolf RC, Sambataro F, Lohr C, Steinbrink C, Martin C, Vasic N. Functional brain network abnormalities during verbal working memory performance in adolescents and young adults with dyslexia. Neuropsychologia 2010; 48: 309–18.

Yvert B, Crouzeix A, Bertrand O, Seither-Preisler A, Pantev C. Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. Cereb Cortex 2001; 11: 411–23.

Zatorre RJ, Belin P, Penhune VB. Structure and function of auditory cortex: music and speech. Trends Cogn Sci 2002; 6: 37–46.

Zatorre RJ, Evans AC, Meyer E. Neural mechanisms underlying melodic perception and memory for pitch. J Neurosci 1994; 14: 1908–19.

Chapter V
When memory for tones differs from memory for words: Evidence of separated neural resources for tonal and verbal materials in congenital amusia

Objectives

In Study 1, we have shown that the amusic brain exhibited abnormal and strongly delayed responses in bilateral Ventro-Lateral Pre-Frontal Cortex and Superior Temporal Gyri during the encoding of melodies in a short-term memory task, along with right-lateralized functional anomalies in the same areas and in the Dorso-lateral Pre-Frontal Cortex and Posterior Parietal Cortex during retention and retrieval of melodic information. These abnormalities were in agreement with previously reported functional and anatomical anomalies observed along the auditory-frontal pathway in the amusic brain (Hyde et al., 2007; Hyde et al., 2006; Hyde et al., 2011; Loui et al., 2009; Mandell et al., 2007) and with the network that was described as supporting pitch perception and memory in typical individuals (see Chapter II). These findings have improved our understanding about the underpinnings of the disorder and the functioning of the normal functioning brain for pitch processing.

Interestingly, while amusic individuals exhibit short-term memory deficit for pitch, their performance in similar tasks for verbal material are preserved (Tillmann et al., 2009). Along these lines if verbal and tonal short term memory are supported by a similar cortical network in the brain (see Chapter II),

a question remains opened: how can we explain amusic individuals' preserved abilities for verbal memory?

Within this framework, Study 2 of this thesis used fMRI to investigate the cerebral correlates of short-term memory for tonal and verbal material in congenital amusic individuals and their matched controls in order to 1) further our understanding of the relationship between the cortical networks supporting each material and 2) pursue the characterization of the cerebral correlates of memory deficits for pitch in congenital amusia.

Article 2
When memory for tones differs from memory for words: Evidence of separated neural resources for tonal and verbal materials in congenital amusia

In preparation

When memory for tones differs from memory for words: Evidence of separated neural resources for tonal and verbal materials in congenital amusia

Authors: Philippe Albouy [1, 2, 3], Barbara Tillmann [1, 3], Patrick Bermudez [4, 6], Robert Zatorre [4, 5], Isabelle Peretz [4], and Anne Caclin [2, 3]

Affiliations: [1] CNRS, UMR5292; INSERM, U1028; Lyon Neuroscience Research Center, Auditory Cognition and Psychoacoustics team, Lyon, F-69000, France

[2] INSERM, U1028; CNRS, UMR5292; Lyon Neuroscience Research Center, Brain Dynamics and Cognition Team, Lyon, F-69000, France

[3] University Lyon 1, Lyon, F-69000, France

[4] International Laboratory for Brain, Music and Sound Research (BRAMS) and Center for Research on Brain, Language and Music (CRBLM), Department of Psychology, University of Montreal, Montreal, QC, Canada H3C 3J7

[5] The Auditory Cognitive Neuroscience Laboratory, Montreal Neurological Institute, Mcgill University, 3801 University Montreal, QC Canada H3A2B4

[6] Centre de Recherche en Neuropsychologie et Cognition, CERNEC, Université de Montréal, Montreal, QC, Canada H3C 3J7

✱Authors contributed equally to this work

Corresponding author: Philippe Albouy, INSERM U1028 - CNRS UMR5292, Brain Dynamics and Cognition Team, Lyon Neuroscience Research Center, Le Vinatier Hospital (Bât. 452); 95, Bd Pinel, 69675 Bron Cedex, France. Phone: +33 4 72 13 89 36. Fax: +33 4 72 13 89 01. E-mail: philippe.albouy@inserm.fr

Abstract

Congenital amusia is a lifelong disorder of pitch perception and memory. The pitch-specificity of the short-term memory deficit has been demonstrated recently, notably by showing that while amusic individuals exhibit strongly altered memory performance for tonal material, they show preserved memory performance for verbal material. Studying congenital amusia is thus a unique occasion to investigate whether verbal and tonal memories share the same neural resources. The present study investigated the behavioral and functional cerebral correlates of tonal and verbal short-term memory in congenital amusic individuals and their matched controls. During fMRI recordings, participants performed perception and memory tasks with tones and words. In both tasks, two sequences separated by a silent delay were presented. In the memory task, participants were required to indicate whether the two sequences were the same or different. In the lower-level perception task they were required to determine if the last two elements of the second sequence were the same or different. As expected, behavioral results showed that amusic individuals exhibited strongly impaired short-term memory performance in comparison to controls for tone sequences, but not for word sequences. Functional data analyses contrasted, for each type of material (verbal, tonal), the brain activations related to the memorization of the information with the brain activations related to the perception of the information. For tonal memory, amusic individuals exhibited functional abnormalities in the right Dorso-Lateral Pre-Frontal Cortex and abnormally increased activation in the right Inferior Frontal Gyrus. In contrast, for verbal memory, amusic individuals exhibited brain activations similar to the controls, with both groups recruiting the left Inferior Frontal Gyrus during the retention of verbal information. These findings suggest the existence of separate neural resources for verbal and tonal STM in the amusic brain, with a specific dysfunction for tonal material.

Introduction

Short-term memory (STM) (along with working memory (WM)) is defined as a cognitive ability allowing the retention and the manipulation of information for a short period of time (Baddeley, 2010; Cowan, 2008; D'Esposito, 2007; Logie and D'Esposito, 2007). For the specific case of verbal STM, it has been proposed that the retention of verbal information (e.g., words, syllables or phonemes) is supported by the Phonological loop (Baddeley and Hitch 1974, Valar, 2006), which is supposed to involve at least two mechanisms: 1) a passive storage mechanism (phonological store) that allows the retention of auditory or speech-based information for a short time period (Baddeley, 1992, 2010); and 2) an active rehearsal mechanism in which information is maintained for a longer time period thanks to articulatory rehearsal processes (such as subvocal speech) (Baddeley, 1992, 2003b, 2010, 2012; Baddeley, 1986; Vallar, 2006; Vallar and Baddeley, 1984; Vallar and Papagno, 2002).

When investigating the cerebral correlates of verbal STM, neuroimaging studies (Buchsbaum and D'Esposito, 2008; Buchsbaum et al., 2005) have suggested that rehearsal processes are supported by cortical networks that are similar to those involved in speech perception and production (Schulze and Koelsch, 2012). This hypothesis derives notably from the observation of activations of the Ventro-Lateral-Pre-Frontal Cortex (VLPFC, i.e., Broca's area, Brodmann's area 44) (Awh et al., 1996; Fiez et al., 1996; Gruber and von Cramon, 2003; Paulesu et al., 1993; Ravizza et al., 2004; Schulze and Koelsch, 2012), in particular in the left hemisphere, along with the recruitment of anterior insular areas (Bamiou et al., 2003; Chein and Fiez, 2001; Paulesu et al., 1993) and the cerebellum (Chen and Desmond, 2005; Kirschen et al., 2005; Ravizza et al., 2004; Silveri et al., 1998) during STM tasks for which subvocal rehearsal was the main strategy of maintenance. Furthermore, a role of the Posterior Parietal Cortex (PPC), notably the inferior parietal lobule (IPL) and the Spt (Sylvian–parietal–temporal, left posterior planum temporale) has been hypothesized in supporting the temporary storage (passive phonological store) of verbal information (Buchsbaum and D'Esposito, 2008; Buchsbaum et al., 2005; Hickok et al., 2003). Additionally to the cortical networks supporting passive storage and rehearsal, other pre-frontal regions are recruited by verbal STM, notably the Dorso-Lateral-Prefrontal-

cortex (DLPFC) (Curtis and D'Esposito, 2003; D'Esposito, 2007; Logie and D'Esposito, 2007; Owen, 2000) which is assumed to involve more complex mechanisms, such as, manipulating and/or monitoring of information stored in memory (D'Esposito, 2007; Owen, 2000; Petrides, 1991, 1994).

In contrast to verbal material, the behavioral and cerebral correlates of STM for tonal material have been much less investigated. It has been originally hypothesized that the temporary storage of pitch is supported by a specialized subsystem in the brain (Berz, 1995; Deutsch, 1970b, 1975a; Pechmann and Mohr, 1992; Salame and Baddeley, 1989). However, this hypothesis has been rapidly challenged (Semal et al., 1996), for example by studies suggesting that tonal short-term memory might recruit very similar mechanisms as those involved in the *phonological loop* (Schendel and Palmer, 2007; Williamson et al., 2010a). When investigating the cerebral correlates of pitch memory, a right-lateralized network involving the VLPFC, the DLPFC and the insular cortex, along with the planum temporale, the IPL, the supramarginal gyrus (SMG) and cerebellum (Foster et al., 2013; Foster and Zatorre, 2010; Gaab et al., 2003; Griffiths et al., 1999; Holcomb et al., 1998; Zatorre et al., 1994) has been observed. While the brains regions involved in pitch memory are highly comparable to those recruited for the retention of verbal information, the findings suggested a potential specialization of each hemisphere for different materials (see Peretz and Zatorre 2005). However, when directly comparing the cortical networks involved in auditory STM for verbal and tonal materials either in non-musicians or musicians, such a clear hemispheric specialization failed to be confirmed (Hickok et al., 2003; Koelsch et al., 2009; Schulze and Koelsch, 2012; Schulze et al., 2011b). Based on these findings, it was hypothesized that while verbal and tonal information share several subparts of more general WM and STM brain networks, these two types of information are not necessarily *processed* in the same way (Peretz & Zatorre 2005).

A way to explore this hypothesis is the investigation of individuals exhibiting congenital pitch-related impairments (with other cognitive functions being unaffected) (Tillmann et al., 2013). Such disorders have been referred to as congenital amusia, a lifelong deficit of music processing, that cannot be explained by hearing loss, brain damage, or cognitive deficits (Ayotte et al., 2002; Peretz, 2013; Peretz et al., 2002; Stewart, 2011; Stewart et al., 2006; Tillmann et al., 2013). Interestingly,

amusic individuals' deficits along the pitch dimension have been recently traced down to impairments in pitch memory (Albouy et al., *in preparation;* Albouy et al., 2013a; Albouy et al., 2013b; Gosselin et al., 2009; Tillmann et al., 2009; Williamson et al., 2010b; Williamson and Stewart, 2010). While amusic individuals showed strongly impaired STM performance for pitch sequences (in comparison to their matched controls), they showed similar STM capacities for verbal material (Tillmann et al., 2009; Williamson and Stewart, 2010). The investigation of the cerebral correlates of the pitch memory deficit has revealed that the amusic brain exhibited abnormal and strongly delayed responses in bilateral VLPFC and Superior Temporal Gyrus (STG) during the encoding of the melodies in a pitch STM task, along with right lateralized functional anomalies in the same areas and in the DLPFC and PPC during the retention and the retrieval of the melodic information (Albouy et al., 2013a). These abnormalities were convergent with the previously reported functional and anatomical anomalies observed along the auditory-frontal pathway in the amusic brain (Hyde et al., 2007; Hyde et al., 2006; Hyde et al., 2011; Loui et al., 2009; Mandell et al., 2007) and with the network that was described as supporting pitch perception and memory in typical individuals (Gaab et al., 2003; Hyde et al., 2008; Peretz and Zatorre, 2005; Schulze et al., 2009; Zatorre et al., 1994). Although these findings have improved our understanding of the underpinnings of the disorder, a question remains unresolved: if verbal and tonal STM are supported by the same cortical networks in the brain, how can we explain amusic individuals' preserved abilities for verbal STM?

The present study aims to 1) further our understanding of the relationship between the cortical networks supporting each material and 2) pursue the characterization of the cerebral correlates of memory deficits for pitch in congenital amusia.

During functional Magnetic Resonance Imaging (fMRI) recordings, amusic individuals and matched controls performed two tasks for both verbal and tonal material. For both tasks, two sequences (either words or tone sequences) were presented sequentially with a separation by a silent delay. In the memory task (either Memory Task for tones (MTt) or Memory Task for words (MTw)), participants were required to indicate whether the two sequences were the same or different. In the perception task (Perception Task for tones (PTt) or Perception Task for words (PTw)) participants were required to indicate whether the last two elements of the second sequence were the same or

different (regardless of the first sequence). The PT was designed as a control condition, for which participants listened passively to the same stimuli (i.e., the first sequence) as those used in the MT, but without actively encoding the information in memory. We expected impaired performance in amusic participants relative to controls for the MTt and PTt, but not for the MTw and PTw. We recorded brain activity using a sparse sampling approach (Belin et al., 1999) to minimize any BOLD response or auditory masking due to MRI scanning noise. By contrasting the functional activations in the brain associated to the memorization of the information (MT) with the functional activations related to the perception of the information (PT) for each type of material, we explored the potential differences between verbal and tonal memory in each group and between groups. Based on our previous MEG results (Albouy et al., 2013a), we expected to observe functional abnormalities in right fronto-temporal areas in amusic participants (compared to controls) for tonal memory.

Note: The results and the conclusions presented here are preliminary, and further analyses are planned with the present data.

Material and Methods

Participants and behavioral pre-tests

The data of the present study were acquired in two different fMRI centers, in Lyon (France) and in Montreal (Canada) using identical scanners (3T Philips Achieva, with the same version update R3.2.2) equipped with the same head coil (32 channels), and applying the same imaging parameters at both sites. Similar systems were used for stimulus presentation. Note that all participants were native French speakers. In this preliminary version of the manuscript, we have merged the data from the two centers, without controlling for the potential effects of scanners (that would be minimal here as we used exactly the same systems and parameters). Note that these additional analyses will be done in a future version of the manuscript.

Participants

Eighteen amusic adults and 16 non-musician controls matched for gender, age, handedness, educational background (calculated in years of education), and musical training (calculated in years) participated in the study. The control group was composed of 12 individuals from Lyon and four individuals from Montreal. The amusic group was composed of 13 participants from Lyon and five from Montreal. Note that within each group, participants from Lyon and Montreal did not differ in terms of handedness, years of education and musical training (ps > .66) but differed in terms of age (amusic individuals : (t(16) = 3.02; p = .008; controls : t(14) = 4.65; p < .001), with the participants from Montreal being older than the participants from Lyon in both groups (amusic individuals Lyon mean = 37.07; SD= 13.64 years; controls Lyon mean = 33.16 SD = 10.22 years; amusic individuals Montreal mean = 56.20; SD= 4.27 years; controls Montreal mean = 58.00 SD = 3.83 years). More importantly, there was no difference between amusic and control groups for each city for all demographic characteristics (Lyon all ps >.41; Montreal all ps > .13).

Group characteristics are presented in Table 1. All participants had right-handed laterality (as assessed by Edinburgh handedness questionnaire) and reported no history of neurological or psychiatric disease. They gave their written informed consent, and were paid for their participation.

Moderate (35 dB loss at any frequency) or severe (more than 40 dB loss) peripheral hearing loss at the frequencies of interest (between 250Hz to 1000Hz) was excluded using standard audiometry.

Behavioral pre-tests

All participants had been thoroughly evaluated on previous testing sessions with the Montreal Battery of Evaluation of Amusia both in Lyon and Montreal. Participants were considered as amusic when they obtained an average score on the MBEA below the cut-off score (23.4 on average across the six tasks of the battery, maximum score = 30), the cut-off being two standard deviations below the average of the normal population (Peretz et al., 2003). All amusic participants had scores below the cut-off score, and all controls obtained scores higher than the cut-off score. The average MBEA scores of the amusic group (ranging from 18 to 22.83) differed significantly from those of the control group (ranging from 23.83 to 28.83; see Table 1).

To evaluate Pitch Discrimination Thresholds (PDT), all participants were tested with a two-alternative forced-choice task with an adaptive tracking, using two-down/one-up staircase procedure (see Tillmann et al., 2009 for details). The average PDT of the amusic group (ranging from 0.13 to 2.41 semitones) was higher (worse) than that of the control group (ranging from 0.05 to 0.67 semitones, see Table 1). In agreement with previous findings (Albouy et al., 2013a,b; Foxton et al., 2004; Tillmann et al., 2009), we observed an overlap in pitch discrimination thresholds between amusic and control groups (both in Lyon and in Montreal).

Table 1: Demographic characteristics of participants and data in behavioral pre-tests. Results on the Montreal Battery of Evaluation of Amusia (MBEA) are expressed as number of correct responses (average over the six sub-tests of the battery, maximum score = 30). Pitch Discrimination Thresholds (PDT) are in semitones. Data are reported as a function of group, and groups are compared with t-tests; ''NS'' refers to a non-significant difference (p>.05). Standard deviations are in parentheses.

Characteristics	Controls (n=16)	Amusics (n=18)	t-Test
Age in years	39.37 (14.24)	42.38 (14.60)	*t(32) = 0.60, p = .54 (NS)*
Gender	9F, 7M	11F, 7M	*NA*
Education in years	13.75 (2.76)	15.05 (3.66)	*t(32) = 1.16, p = .25 (NS)*
Musical education in years	0.68 (1.57)	0.83 (1.42)	*t(32) = 0.28, p = .77 (NS)*
Musical battery *(MBEA, Peretz et al., 2003)*	26.55 (1.56)	20.94 (1.53)	*t(32) = 10.86, p < .0001*
Pitch Discrimination Threshold *(PDT, Tillmann et al., 2009)*	0.22 (0.16)	0.90 (0.88)	*t(32) = 3.03, p = .004*

Tasks for tone and word sequences (used in the fMRI recording session)

During the fMRI acquisition, congenital amusic individuals and matched controls performed four tasks: a *Memory Task for tones* (piano tones, MTt), a *Perception Task for tones* (piano tones, PTt), a *Memory Task for words* (monosyllabic words, MTw), and a *Perception Task for words* (monosyllabic words, PTw).

All tasks involved two three-sound (words or tones) sequences (S1, S2), separated by a silent retention period of 9000 ms. Note that as the retention period is very long, and based on reports suggesting that amusic individuals exhibit increasing deficit in short-term memory tasks with increasing delay(Williamson et al., 2010b), we choose such simple material (melodies of 3 tones thus containing simple contour information) to allow amusic participants to perform above chance level.For all sequences, the three sounds had a duration of 250ms and were presented successively without inter-stimulus-interval. For the MTt and MTw, participants were required to indicate whether the S2 sequence was identical or different from the S1 sequence. For the PTt and PTw participants were required to passively listen to the S1 sequence and to be concentrated on the S2 sequence. They

had to indicate, for the S2 sequence, if the two last elements of this sequence were same or different (examples of S2 sequences for tones: G3-*F3-F3* for a same trial, G3-*F3-E3* for a different trial).

Tonal material

120 different three-tone melodies were created using eight piano tones differing in pitch height (Cubase software, Steinberg); all used tones belonged to the key of C Major (C3, D3, E3, F3, G3, A3, B3, C4). For half of the sequences, identical tones were repeated consecutively in the second and third position of the melody (e.g., G3-*F3-F3*); for the second half, the melodies did not contain repetitive tones. S1 melodies (presented as the first sequence) for the MTt and PTt were selected within this set of melodies (60 melodies for the MTt and 60 melodies for the PTt).

Memory Tasks for tones – MTt

During the MTt, participants were asked to indicate whether S1 and S2 melodies, presented in pairs, were the same or different. For different trials, one tone of the S2 melody was different from the S1 melody (in positions 1 to 3, equally distributed across trials) and created a contour-violation in the melody. Note that the pitch interval size between the original tone in S1 and the changed tone in S2 in was controlled in such a way that within the 30 different trials of the MTt, there were 50% of the trials with a changed tone with a small interval size (with equal proportions of 1.5 tone, 2 tones and 2.5 tones, and equally distributed as a function of the position of the changed tone); and 50% of trials with a large interval size (3 tones, 3.5 tones and 4 tones and equally distributed as a function of the position of the changed tone).

There were 60 MTt trials (S1, silence, S2) separated in 30 *same* and 30 *different* trials. For the 30 *same* trials, 15 were composed of melodies that did not contain repetitive tones, and 15 contained tone repetitions in the melodies (as mentioned above, identical tones repeated consecutively at the second and third position of the melody). For the 30 *different* trials, 15 trials were composed of S1 and S2 melodies that did not contain repetitive tones, and 15 trials were divided in two sets so that 7 trials contained repetitive tones only within S1 melodies and 8 trials contained repetitive tones only within S2 melodies.

Perception Task tonal – PTt

Participants had to indicate whether the last two tones of the second melody (S2) were the same or different. Note that a *same/different* task was chosen in order to overcome amusic individuals' difficulties in indicating pitch direction (Foxton et al., 2004b; Liu et al., 2010).

As in the MTt there were 60 PTt trials (S1, silence, S2) separated in 30 *same* and 30 *different* trials. The *Same* trials were divided into 15 trials with S1 and S2 melodies containing identical tones repeated consecutively at the second and third position of the melody and 15 trials containing repetitive tones only within the S2 melody. For the 30 *different* trials, 15 trials were composed of S1 and S2 melodies that did not contain tones repetition, and 15 trials were composed of S1 melodies with repetitive tones only for S1 melodies.
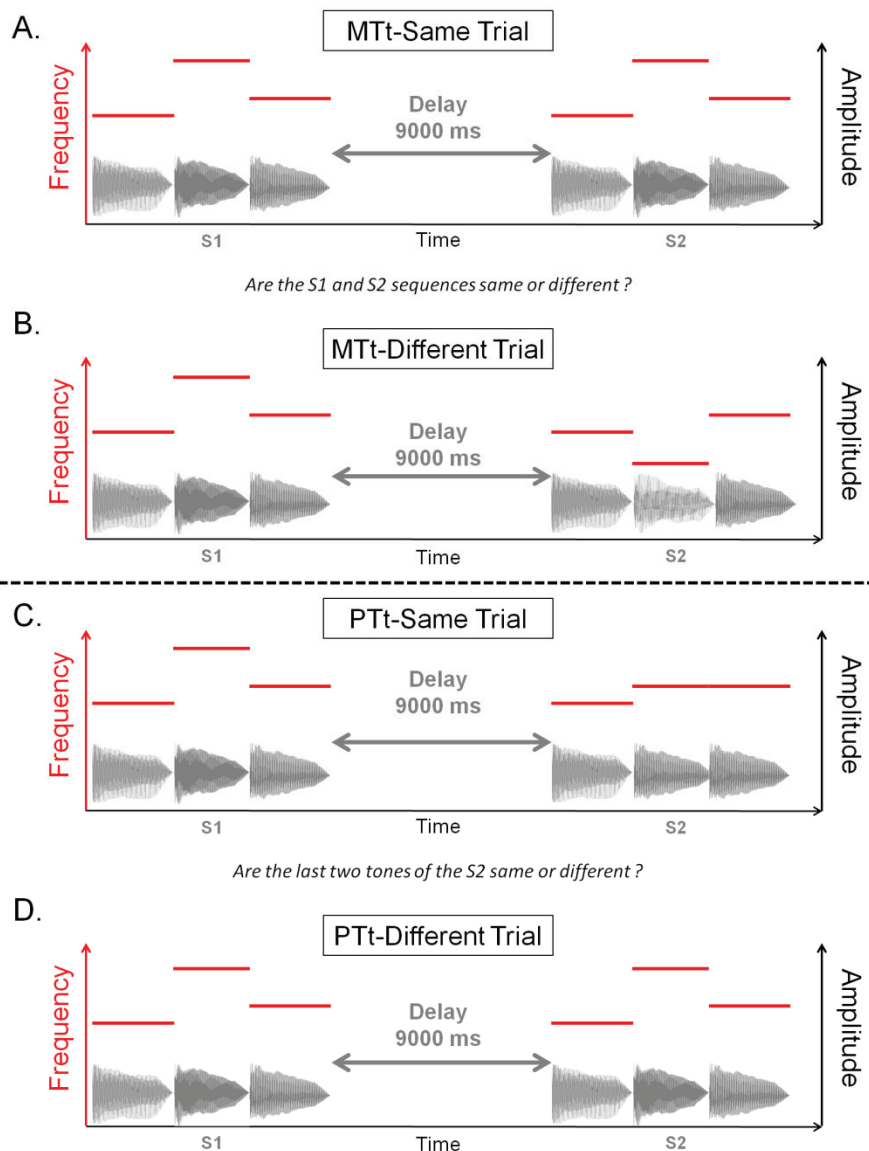
Figure 1: Examples of the stimuli used in the MTt and PTt. A. Example of a same trial for the MTt; tones are represented as spectrograms and their frequencies are illustrated with the red lines. For "same" trials S1 was repeated as the second melody of the pair (S2) after a 9000 ms delay. B. Example of a different trial for the MTt. For "different" trials, the second melody of the pair changed only for one tone (in position 1 to 3) that changed the melodic contour. C. Example of a same trial for the PTt. For "same" trials the two last tones of S2 were identical. D. Example of a different trial for the PTt. For "different" trials, the two last tones of S2 are different.

Verbal material

Six monosyllabic French words (material taken from Tillmann et al. 2009) were used: toux (/tu/ - cough), loup (/lu/ - wolf), boue (/bu/ - dirt), mou (/mu/ - soft), goût (/gu/ - taste) and poux (/pu/ - bug), spoken by a female voice. All recordings were adjusted to the pitch of 230 Hz (within the range of the piano tones used in the MTt and PTt) with STRAIGHT, and equalized in loudness using MATLAB software. The words were selected from a pool of recorded words on the basis of subjective ratings indicating easy intelligibility (i.e., using a subjective scale from 1 (very easy to understand) to 5 (not easy to understand)) by eight native French speakers.

Sixty different sequences were created. For half of these sequences, identical words were repeated consecutively in the second and third position of the sequence; the other half did not contain word repetition within the sequence. S1 sequences for both MTw and PTw were selected within this set of S1 sequences (30 for the MTw and 30 for the PTw).

Memory Task for words - MTw

During the MTw, participants were asked to indicate whether S1 and S2 sequences, presented in pairs, were the same or different. For different trials, one word in the S2 sequence was different from the S1 sequence (in positions 1 to 3, equally distributed across trials). Note that the changed tone was selected from a list of remaining words that were not presented in the S1 sequence.

There were 30 MTw trials (S1, silence, S2) separated in 15 same and 15 different trials. For same trials, 8 were composed of sequences that did not contain identical words repeated consecutively at the second and third position of the sequence, the remaining 7 same trials contained repetitive words

in the sequences. For the 15 different trials, 5 trials were composed of S1 and S2 sequences that did not contain repetitive words, 5 trials contained repetitive words only within S2 sequences and the 5 remaining trials contained repetitive words only in the S1 sequences.

Perception Task - PTw

Participants had to indicate whether the last two words of the second sequence were the same or different (regardless of the S1 sequence). Over the 30 PTw trials (S1, silence, S2), there were 15 same and 15 different trials. Same trials were divided into 8 trials with S1 and S2 sequences containing repetitive words and 7 trials containing repetitive words only in the S2 sequences. For the 15 different trials, 8 were composed of S1 and S2 sequences that did not contain repetitive words and the 7 trials were composed of S1 sequence with words repetition and S2 sequences without words repetition.

Procedure

Amusic individuals and their matched control participants performed the MTt, PTt, MTw and PTw during the fMRI recording. Presentation software (Neurobehavioral systems, Albany, CA, USA) was used to run the experiment and to record button presses. In Lyon, the sound stimulation was performed with MRI-compatible air plugs with foam ear tips (*NordicNeuroLab*). In Montreal, the auditory stimuli were delivered binaurally via a MR-compatible pneumatic system and foam insert earplugs (Etymotic Research, IL, USA). The level of sound presentation was equal for all participants so that the sounds were presented at about 70 dB Sensation Level with a central position (stereo) with respect to the participant's head.

The experiment was divided into six different runs of 8 minutes and 37 seconds each: 4 runs of tonal material, and 2 runs of verbal material. Within a run, MT and PT were presented alternatively (i.e. 15 consecutive trials of MT and 15 consecutive trials of PT) the task order was counterbalanced across runs and participants. In each run, 37 trials (each corresponding to one TR of the sparse sampling protocol) were presented: *1)* 1 trial of visual instruction informing the participant about the task to be performed for the first half of the run, *2)* 5 trials of silence (without auditory stimulation), *3)* 15 trials of either MT or PT as a function of the run, *4)* 1 trial of visual instruction informing the

participant about the task for the second half of the run, and *4)* 15 trials of either PT or MT as a function of the run (e.g., Run 1: *instruction-silence-PT-instruction-MT* and Run 2: *instruction-silence-MT-instruction-PT*).

During the acquisition, participants were required to keep their eyes closed. When they heard a salient tone burst, they were required to open their eyes and look at the screen informing them about the task they had to perform (either the Memory or Perception Task) for the first half of the run. Participants then closed their eyes once they heard the tone burst again, informing them that the task changed for the second part of the run. The runs were separated by 2-3 minutes of break. Note that participants were informed orally by the experimenter about the material (tones or words) and the task order before each run.

Within a run, for each trial, participants listened to the first three-element sequence with a total duration of 750 ms (S1), followed by a silent retention period of 9000 ms, and then the second sequence (S2, 750 ms). They were asked to indicate their answer by pressing one of two keys of a response device with their right hand after the end of S2. They had 2000 ms to respond before the next trial, which occurred 2500 ms to 3000 ms after the end of S2. As a function of the task, trials were presented in a pseudo-randomized order with the constraint that the same trial type (e.g. same, different) could not be repeated more than three times in a row. Before the fMRI acquisition, participants performed a set of 15 practice trials for each task outside the scanner (with a simulation of the scanner noise) with response feedback. No feedback was given during the experiment.

fMRI

fMRI parameters

At the beginning of the MRI session, a high-resolution T1-weighted 3D image was acquired for each participant using a gradient-echo sequence (160 sagittal slices; time to repetition (TR), 8.2 ms; time to echo (TE) 3.8 ms; flip angle (FA), 8°; matrix size, 240x240; field of view (FOV) 240x240mm^2; voxel size, 1 x 1 x 1 mm^3).

The fMRI series were recorded via a gradient-echo EPI pulse sequence to measure whole-brain oxygenation level-dependent (BOLD) signal (47 axial slices acquired in sequential order; TR, 14000 ms (volume acquisition in 3000 ms); TE, 30 ms; FA, 90°; 3 mm slice thickness; no gap; matrix size, 80x79; FOV 240 x 240mm^2; voxel size, 3 x 3 x 3 mm^3). We used a sparse-sampling paradigm to minimize any BOLD response or auditory masking due to MRI scanning noise (Belin et al., 1999). Auditory events were synchronized with fMRI image volume acquisitions using Presentation software (Neurobehavioral Systems), at a rate of one image per trial. Within different blocks, we aimed to capture the hemodynamic response associated with two different processes: (a) the *encoding* of the stimuli (Figure 2A, with fMRI volumes acquired 3500 to 4000 ms after the end of S1, i.e., at the expected peak of the hemodynamic response for processing S1) and (b) the *retention* period of the stimuli (Figure 2B), with fMRI volumes acquired 5500 to 6000 ms after the end of S1 aiming to record the hemodynamic responses associated to the retention of the information in memory (while decreasing the chance of capturing the cerebral activity related to the encoding of S1 stimulus). Note that the encoding scans were performed only for the tonal material (2 runs) in order to explore if amusic individuals' altered responses observed during the encoding of the melodies with MEG (Albouy et al., 2013a) could be observed with another imaging method, and the retention scans were performed for both verbal and tonal materials (2 runs each).
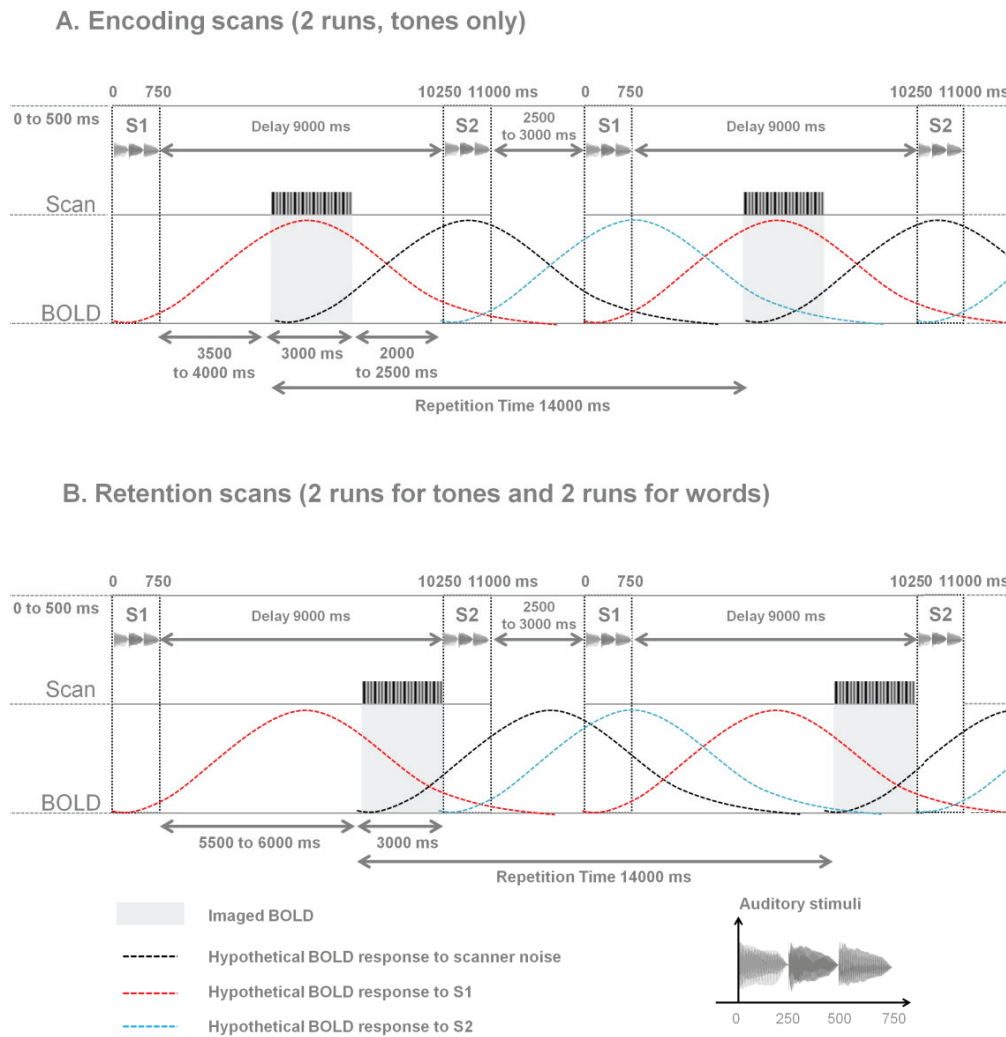
Figure 2: Design for the fMRI experiment and timeline of events during two trials. S1 sequence (tones, words or silence periods) lasted 750 ms and was followed by a constant 9000 ms delay during which occurred 3000 ms of functional data acquisition which was followed by the S2 sequence (750ms). Participants had 2000 ms to respond, the next trial occurring 2500 to 3000 ms after the end of S2. Note that a 0 to 500 ms jitter was added at the beginning of the trial to maximize the detection of the BOLD response to the task. As a function of the run, the acquisition of the whole brain volume was realized at two different time periods. A. Encoding runs (A) (two runs, tones only) acquisition started at 3500 to 4000 ms after the end of S1. B. Retention runs (B)(two runs for tones and two runs for words), the volume acquisition occurred just before S2 (at the end of the delay), the acquisition thus starting at 5500 to 6000 ms after the end of S1. Grey bar, Imaged BOLD, colored dotted lines, hypothetical BOLD signal in response to S1 (red), scanner noise (black) and S2 (blue).

Each run consisted of the acquisition of 5 volumes of silence, 15 volumes of MT (either MTt or MTw as a function of the run), 15 volumes of PT (either PTt or PTw as a function of the run). All together, we recorded 30 volumes of MTt, 30 volumes of PTt and 10 volumes of silence for the *Encoding* runs. For the *Retention* runs, the overall data consisted in 30 volumes of MTt, 30 volumes of PTt and 10 volumes of silence for the tonal runs, and in 30 volumes of MTw, 30 volumes of PTw and 10 volumes of silence for the verbal runs.

Preprocessing

All image preprocessing and statistical analyses were performed using the fMRI functions of SPM8 (Wellcome Trust Centre for Neuroimaging, http://www.fil.ion.ucl.ac.uk/spm/, London, UK). Before preprocessing, all images were checked for artifacts and automatically aligned so that the origin of the coordinate system was located at the anterior commissure. Preprocessing included the realignment of functional images and the co-registration of functional and anatomical data. We then performed a spatial normalization (voxel size, 3x3x3) of the T1 and the EPI images to the Montreal Neurological Institute templates provided with SPM8 (MNI EPI template and T1 template respectively). Finally, images were spatially smoothed (Gaussian kernel, 8 mm full-width at half-maximum (FWHM)).

Statistical analyses

Condition effects were estimated using a general linear model (Friston et al., 1995). Low-frequency drifts were removed using a temporal high-pass filter with a cut-off of 128 s. As a sparse sampling fMRI protocol was used, the data were not convolved with the hemodynamic response function (HRF), and no low-pass filter was applied. At the first level, for each participant, changes in brain regional responses were estimated by a general linear model and the following contrasts were performed: 1) Each task (MTt, PTt, MTw, and PTw, respectively) versus Silence for *Encoding* and *Retentions* scans, 2) *Encoding scans* versus Silence (MTt+PTt vs silence) and *Retention scans* vs Silence (MTt + PTt vs Silence and MTw+PTw vs Silence) 3) MT versus PT for *Encoding* scans for tonal material (MTt vs PTt) and MT versus PT for tonal and verbal material for *Retention* scans (MTt

vs PTt and MTw vs PTw). We then analyzed within and between group effects at the second level. As the analyses presented in the present paper are preliminary, we performed within group analyses in the whole brain with areas of significant BOLD response determined with a lenient significant threshold (either $p < .001$ or $p <. 01$ uncorrected, as a function of the contrasts, see Results). Note that we here reported all the clusters activated in the whole brain for each contrast. Once the areas of activations were determined for each group independently, these areas (either shared activations between groups, or activations in one group only) were used as a Region of Interest (ROI) for the group comparisons ($p < .01$ uncorrected).

Results

__Behavioral data__

Performance

Performance was evaluated in terms of percentages of Hit-FAs (hit rate corresponding to the number of correct responses for different trials / number of different trials*100; and the false alarm rate corresponding to the number of incorrect responses for same trials / number of same trials*100). Behavioral data were analyzed with a 2x2x2 ANOVA with Group as between-participant factor and Material (Tones, Words), and Task (MT, PT) as within-participant factors. The main effect of Group was significant ($F(1,32) = 7.17$, $p = .009$), amusic individuals showing decreased performance in comparison to controls. The main effect of Task was significant ($F(1,32) = 7.17$, $p = .009$), with decreased performance for MT in comparison to PT. The main effect of Material was not significant ($F(1,32) = 2.49$; $p = .12$; but the interaction between Material and Group reached significance ($F(1,32) = 10.44$; $p = .002$). Fischer's LSD post-hoc tests revealed that amusic individuals' performance was decreased in comparison to controls for the Tonal Material ($p = .0001$), but not for the Verbal Material ($p = .80$). For controls, performance for Verbal material was decreased in comparison to Tonal material ($p = .0002$) while amusic individuals exhibited similar performance for both materials ($p=.23$). Finally, note that the interaction between Material, Group and Task did not reach significance ($F(1,32) = 2.88$, $p = .09$). These results are in line with Tillmann et al. (2009) showing preserved short-term memory performance for Verbal Material, and altered performance for Tonal Material in congenital amusia (see Figure 3A).

For the MTt, to assess the potential effect of pitch interval size in the changed tone of S2 (relative to the corresponding tone in S1) on performance, %Hit-FA for different trials were extracted as a function of interval sizes, and analyzed with a 2x2 ANOVA with Group (amusic individuals, controls) as between-participants factor and Interval size (small, large) as within-participant factor. The main effects of Group ($F (1, 32) = 17.00$, $p < .0001$), and of Interval size ($F (1,32) = 17.27$, $p < .0001$) were significant, but the interaction between Group and Interval size ($F (1,32) = 0.15$, $p =.90$)

did not reach significance. This analysis revealed that participants' performance was lower for small interval size in comparison to large interval size (see Figure 3B).
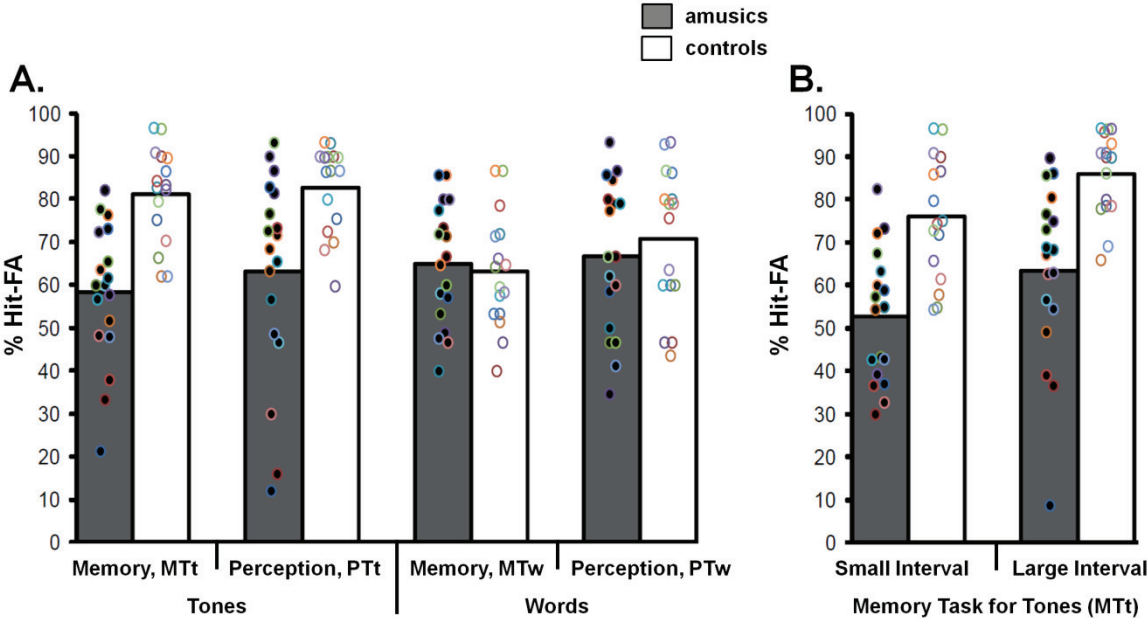
Figure 3 A. Performance of amusic and control groups (white, Controls; Grey, Amusic individuals) in terms of Hits minus FAs, presented as a function of Material (Tonal, Verbal), Task (Memory Task; Perception Task) and Group (amusic individuals n = 18, black circles; controls n = 16, white circles). B. *Performance of amusic and control groups for the Memory task for tones (MTt) in terms of % Hit-FA for different trials as a function of interval size (small, large). Circles: amusic individuals' and matched controls' individual performance.*

Response times (RTs)

We analyzed response times (RTs) for correct responses only. Note that 2 amusic participants and 1 control participant were excluded from this analysis because of average RTs superior to 2000 ms (the time window that was dedicated to respond after the end of S2). Note however that we have included these participants in the performance analysis to check whether they did the task accurately (and respected the instructions). Additionally, note that the analysis of accuracy data without these participants provides very similar results.

RT data was analyzed with a 2x2x2x2 ANOVA with Group as between-participants factor and Material (Tonal, Verbal), Task (Memory Perception) and Trial Type (Same, Different) as within-participant factors. The main effect of Group was marginally significant $F(1,29) = 4.10$; $p = .051$, amusic individuals (mean = 919.33; SD = 240.62) exhibiting slower response times than did controls (mean = 768.57; SD = 270.67) overall. The main effect of Trial Type was significant $F(1,29) = 8.75$; $p = .006$ and revealed that participants responded faster for same trials (mean = 817.37; SD = 268.66) in comparison to different trials (mean = 875.40; SD = 261.21). The main effect of Task was also significant, participant showing shorter RTs for the MT (mean = 810.14; SD = 263.09) than for the PT (mean = 882.63; SD = 265.03).

The main effect of Material was not significant (??), but most interestingly, the interaction between Material and Group was significant $F(1,29) = 7.23$; $p = .01$. Fischer's LSD post hoc tests showed that while controls exhibited shorter RTs for Tonal material (mean = 720.75 ms; SD = 281.23) than for Verbal material (mean = 816.40 ms; SD = 253.04; $p = .015$), amusic individuals did not show this material effect ($p = .23$; mean Tonal = 941.37 ms; SD Tonal = 248.94; mean Verbal = 897.30 ms; SD verbal = 231.86). Furthermore, note that controls' RTs were shorter than amusic individuals' RT for Tonal material ($p = .008$), but not for verbal material ($p = .31$).

Correlation analyses

Two sets of analyses were performed: 1) correlations between the behavioral data recorded during the fMRI acquisition (*Hit-FAs*, *RTs* for each task) and data of the pretest (MBEA, PDT); and 2) correlations between the different measures of the tasks performed during the fMRI acquisition (*MTt, PTt, MTw, PTw*). To explore differences explained not only by the group differences defined via the MBEA or the PDT, we focused on correlations that were significant over all participants *and* for at least one of the two participant groups separately (amusic group or control group).

For the first set of analyses, data from the PDT was negatively correlated with participants' performance in the PTt in amusic individuals ($r(16) = -.52$; $p = .038$) and in all participants ($r(32) = -.60$; $p < .0001$), but not in controls ($r(14) = -.39$; $p = .15$).

For the second set of analyses, we observed a positive correlations between RTs (same and different trials) for correct responses in the PTt and % Hit-FAs of the PTt in amusic individuals (r(14)= .55, p = .027) and in all participants (r(29)= .65, p <.001). Additionally, in both controls and amusic individuals, when comparing tonal and verbal material in the memory and perception tasks, we observed positive correlations between RTs for correct responses in the MTt and the MTw (all ps < .002) and positive correlations between RTs for correct responses in the PTt and PTw (all ps <. 001). Note that these correlations involving RTs data remained significant when performing the same analysis on same and different trials independently. Finally, accuracy data for the MTt and the PTt were positively correlated in amusic individuals (r(16) = .71; p = .002), in all participants (r(32) = .72; p <.0001), but not in controls (r(14) = .35; p =.17).

fMRI

Four main functional sets of analyses were performed. The first set explored the brain activations in amusic individuals and controls when they were listening to melodies independently of the task ([MTt + PTt > Silence], *Encoding scans*) and during *Encoding* of melodies ([MTt > PTt], *Encoding scans*). The second set explored in amusic individuals and controls the potential differences in terms of brain activity for the *Encoding* and the *Retention* of information for melodies in the context of the STM task, ([*Encoding scans* (MTt>PTt) > *Retention scans* (MTt > PTt)] or [*Retention scans* (MTt > PTt) > *Encoding scans* (MTt>PTt)]). The third set of analyses explored the brain activations that were related to the *Retention* of tonal (*Retention scans* [MTt > PTt]) and verbal materials (*Retention scans* [MTw > PTw]), respectively. The fourth set of analyses explored the potential differences between the *Retention* of the two types of materials ([*Retention Tonal* (MTt>PTt) > *Retention Verbal* (MTw > PTw)] and [*Retention Verbal* (MTw > PTw) > *Retention Tonal* (MTt>PTt)]) in the context of the STM task.

Listening to melodies and Encoding of melodies

The first set of analyses was performed on the *Encoding scans*. Within group analysis was performed in the whole brain (p< .001(unc)) and revealed that both groups (see Figure 4,a,b) showed

enhanced activation in bilateral temporal regions encompassing both primary and secondary auditory cortex (when listening to melodies in comparison to silence. The activations extended to the posterior and anterior parts of the STG in both groups. Note however that, in comparison to controls, the amusic group seems to recruit a more diffuse network, involving medial temporal areas, along with occipital, parietal and cerebellar brain regions. The coordinates of activation peaks (and cluster size) are presented in Table 2. Group comparisons were then performed with ROIs defined by the regions that were activated in both groups. This network (as shown by the orange areas in Figure 4c) involved bilateral Heschl's Gyrus (HG), Planum temporale and posterior and anterior portions of the STG. The group comparisons performed in these ROIs did not show any significant differences between the two groups (Figure 4c).

Note additionally that the [MTt > PTt] contrast was then performed for the *Encoding scans* in each group to assess whether supplementary activation could be observed when high-level processing of information is required, notably as listening to melodies in the context of a memory task (active encoding) could be hypothesized to involve more complex processes than simple perception. However, no supra-threshold activation was observed in this analysis, in both amusic and control participants, respectively.
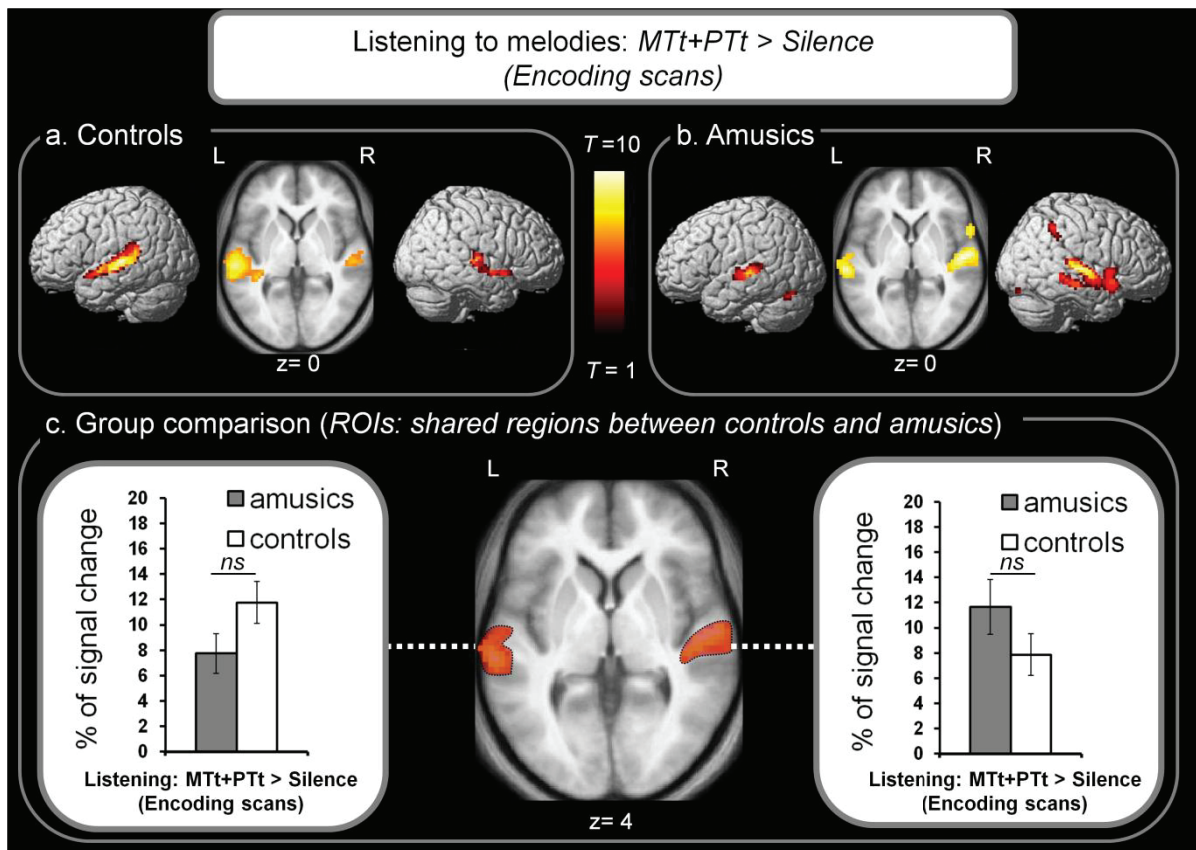
Figure 4. Functional imaging results for the Encoding scans: [(MTt + PTt) > Silence] contrast for controls (a), amusic individuals (b), and for the group comparison (c) performed in ROIs (as indicated with orange areas in the c panel) corresponding to the regions that were activated in both groups for this contrast. Results are displayed at p <0.001 (uncorrected) over the mean structural image of all participants (axial slice) in the MNI space and over the single subject 3D render provided by SPM. The areas of activation are detailed in Table 2. On figure c, the graphics show the average functional data of the ROIs in both amusic individuals and controls in terms of % of signal change (MTt+PTt > Silence). Grey amusic individuals; white, controls. Errors bars represent the SEM.

Table 2. Areas of functional activations for the Encoding scans > silence (MTt + PTt > Silence) contrast in controls and amusic individuals respectively. Coordinates are in the MNI space and correspond to the peak of each cluster. HG, Heschl's Gyrus; IPL, Inferior Parietal Lobule, PT, Planum Temporale; STG, Superior Temporal Gyrus; MTG, Middle Temporal Gyrus; SMG, Supra-Marginal Gyrus; R, Right, L, Left, H, Hemisphere

Contrast	Lobe	Region	H	T	x y z mm	mm3
Amusics MTt+PTt > Silence						
	Temporal	HG, PT, STG	R	6.09	48 -25 4	16764
				5.80	51 -16 1	
				5.72	60 -13 1	
		MTG	L	5.83	-66 -28 4	3510
		PT, STG		4.27	-60 -19 -5	
	Parietal	IPL, SMG	R	5.28	45 -49 49	1863
				4.23	39 -43 37	
	Occipital	Calcarine fissure	L	4.73	9 -76 -17	5508
	Cerebellum	Lobule 6-L		4.66	-30 -67 -26	
				4.28	-6 -79 -14	
Controls MTt+PTt > Silence						
	Temporal	HG, PT, STG	R	4.88	72 -19 7	3834
				4.30	54 11 -8	
				4.30	54 -22 1	
		HG, PT, STG	L	7.65	-63 -28 4	14391
				6.72	-57 -16 -5	
				6.48	-51 -25 -2	
Controls >Amusic individuals: MTt+PTt > Silence						
	*	*	*	*	*	*

Encoding and Retention of melodies in a tonal STM task

The second set of analyses was performed on the *Encoding scans* and *Retention scans* for the tonal material. We explored if the functional brain responses associated to memory processing for tones (MTt> PTt) differed during the *Encoding* and the *Retention* steps of the STM task. Within group analysis was performed in the whole brain (p< .001(unc)) and revealed that the control group (see Figure 5a) showed increased activation in the right DLPFC for the *Retention scans* in comparison to the *Encoding scans* (MTt>PTt). Amusic individuals however, did not exhibit significant differences for *Retention* over *Encoding* of tonal information (MTt > PTt). The coordinates of activation peaks (and cluster size) are presented in Table 3. Group comparisons were then performed using the ROI defined by the regions activated in controls for this contrast. This analysis revealed that within the right DLPFC, amusic individuals showed decreased activation in comparison to controls for the *Retention scans*, but not for the *Encoding scans* (Figure 5b). Note that the reverse contrast ([*Encoding*

scans (MTt>PTt) > *Retention scans* (MTt > PTt)] did not show any significant effect in both amusic individuals and controls.

Figure 5. Functional imaging results for the [Retention scans (MTt > PTt) >Encoding scans (MTt > PTt)] contrast for controls (a), and for the group comparison performed within the ROI (as indicated with an orange area in the (b) panel) corresponding to the region activated in controls for this contrast (b). Results are displayed at p <0.001 (uncorrected) over the mean structural image of all participants (axial slice) in the MNI space and on the single subject render provided by SPM. The areas of activation are detailed in Table 3. On figure b, the graphic shows the average data of the ROI in terms of % of signal change (MTt >PTt) for Encoding and Retention scans. Grey: amusic individuals; white: controls. Errors bars represent the SEM.

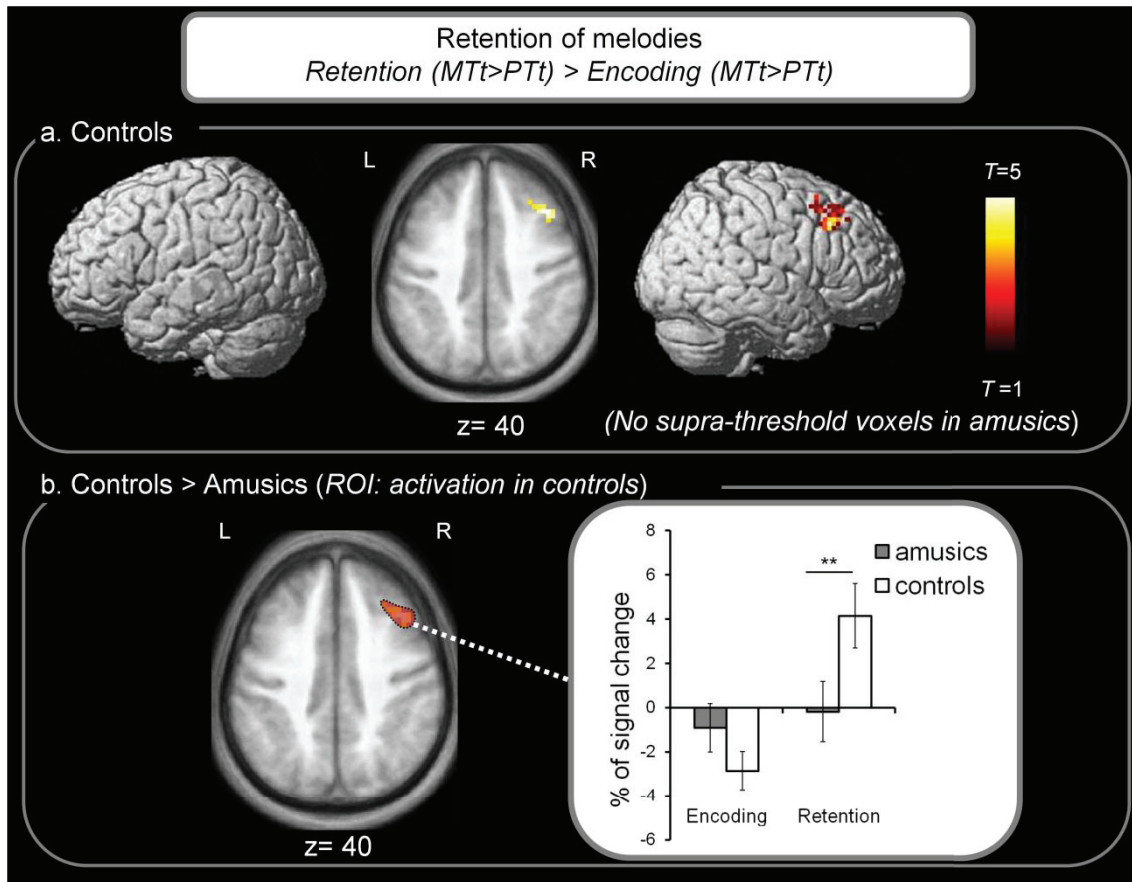*Table 3. Areas of functional activations for the [Retention scans (MTt>PTt) > Encoding scans (MTt > PTt)] contrast for controls only and for the Controls > Amusic individuals comparison. Coordinates are in the MNI space and correspond to the peak of the cluster. DLPFC, Dorso-Lateral Prefrontal Cortex; R, Right, L, Left, H, Hemisphere. * Absence of significant cluster.*

Contrast	Lobe	Region	H	T	x y z mm	mm3
Controls: Retention (MTt > PTt) > Encoding (MTt > PTt)						
	Frontal	DLPFC	R	3.38	42 23 40	4131
				3.17	15 32 40	
				2.94	24 20 46	
Amusisc: Retention (MTt > PTt) > Encoding (MTt > PTt)						
	*	*	*	*	*	*
Controls > Amusics (ROI: activated region in controls)						
	Frontal	DLPFC	R	2.34	42 23 40	513

Retention of tone and word sequences

The third set of analyses was performed on the *Retention scans* for Verbal and Tonal data independently. We explored the cerebral responses associated to memory processing of tones (MTt> PTt) and words (MTw> PTw). For the tonal material, the within-group analysis was performed in the whole brain (p< .01(unc)) and revealed that the control group (see Figure 6a, upper panel) showed increased activation for the MTt in right and left DLPFC in comparison to the PTt, along with increased activations in the left Supplementary Motor Area (SMA), left Middle Frontal Gyrus (MFG), left Pre-central Gyrus and left Temporal pole. In contrast, using the same significance threshold, amusic individuals did not exhibit significant differences between MTt and PTt for the *Retention scans* of the tonal STM task. Group comparisons performed using the ROIs defined by the activated regions in controls revealed that amusic individuals showed significantly decreased activation (p<0.001) only in the right and left DLPFC in comparison to controls (Figure 6A, lower panel).

For the verbal material, within-group analysis was performed in the whole brain (p< .001(unc)) and revealed that both amusic individuals and controls showed increased activation in the left VLPFC including the pars opercularis of the IFG (Figure 6B, upper panel) for the MTw in comparison to the PTw. Group comparisons performed in ROIs defined by the shared regions between the two groups did not reveal any significant differences between amusic individuals and controls for

this contrast (Figure 6B, lower panel). The coordinates of activation peaks (and cluster size) are presented in Table 4.
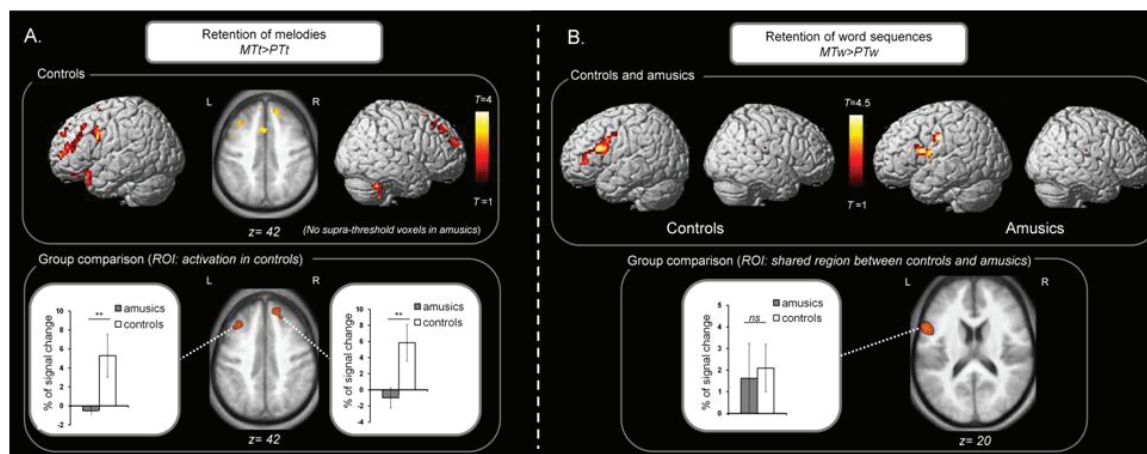
Figure 6. A. Functional imaging results for the Retention scans for (A): tonal material for the MTt > PTt contrast and (B): verbal material for MTw > PTw contrast. Data is presented for controls (A. upper panel), controls and amusic individuals (B. upper panel) and for the group comparison performed in the ROIs corresponding to the regions activated in controls for the MTt > PTt contrast (A, lower panel) and in the ROI corresponding to the regions of activation shared between the two groups (B, lower panel). Results are displayed at p <0.01 (uncorrected) for Tonal material (A) and p <0.001 (uncorrected) for Verbal material over the mean structural image of all participants (axial slice) in the MNI space and on the single subject 3D render provided with SPM. The areas of activation are detailed in Table 4. On the lower panels, the graphics show the average ROIs data in terms of % of signal change (MTt>PTt). Grey, amusic individuals; white, controls. Errors bars represent the SEM.

Table 4 *Areas of functional activations for the Retention scans for verbal and tonal material respectively for the MT > PT contrast (MTt > PTt and MTw > PTw) in controls, amusic individuals and for the Controls vs Amusic individuals contrasts (Note that the reverse contrast: Amusic individuals > Controls did not yield significant clusters). Coordinates are in the MNI space and correspond to the peak of each cluster. DLPFC, Dorso-Lateral Prefrontal Cortex; SMA, Supplementary Motor Area; MFG, Middle Frontal Gyrus; IPL, Inferior Parietal lobule; SMG, Supra-Marginal Gyrus; IFG, Inferior Frontal Gyrus; R, Right, L, Left, H, Hemisphere. * Absence of significant cluster.*

Contrast	Lobe	Region	H	T	x y z mm	mm3
Controls Retention : MTt > PTt						
	Frontal	DLPFC	R	2.85	9 59 25	891
				2.25	21 53 22	1080
				2.53	18 35 46	1080
				2.51	15 44 40	
				2.33	21 26 52	
			L	2.88	-6 50 37	
		SMA	L	3.24	-6 2 52	2619
				3.17	-3 -4 58	
				2.98	-6 11 64	
		MFG	L	3.65	-36 53 7	2160
				3.27	-39 47 13	
				2.50	-36 29 40	
		Pre-central gyrus	L	3.03	-48 8 40	1350
				2.54	-48 8 28	
	Temporal	Temporal pole	L	2.97	-45 20 -20	810
				2.66	-36 20 -32	
	Parietal	IPL, SMG	R	3.40	48 -49 -35	837
				3.15	36 -46 -50	
Controls > Amusics : Retention : MTt > PTt (ROIs: activation in controls)						
	Frontal	DLPFC	R	2.89	18 35 43	540
				2.76	15 47 37	
		MFG, DLPFC	L	3.37	-36 50 10	1188
				3.20	-42 41 16	
				2.97	-33 56 4	
Controls Retention : MTw > PTw						
	Frontal	IFG pars triangularis	L	3.98	-51 20 25	5670
		Precentral gyrus		3.74	-48 8 34	
				2.78	-45 2 43	
		IFG pars triangularis		3.40	-45 38 7	
		IFG pars opercularis		3.43	-45 14 19	
				3.37	-48 11 22	
Amusic individuals Retention : MTw > PTw						
	Frontal	Post central gyrus	L	3.72	-48 -13 34	837
				3.57	-60 -1 13	1944
		IFG pars opercularis		3.35	-60 11 19	
				2.71	-48 8 16	
Controls > Amusics (ROI: activated region in controls)						
		*	*	*	*	*

Comparison between memory for tone and word sequences

The fourth set of analyses explored if the cerebral responses associated to memory processing for tones (MTt> PTt) and words (MTw> PTw) differed in amusic individuals and controls and between the groups (*Retention scans*). Regardless of the contrast ([Tonal > Verbal] or [Verbal > Tonal]) controls did not show any significant differences between tonal and verbal material for these analyses that were performed in the whole brain (p< .001(unc)).

For amusic individuals, the [Tonal > Verbal] contrast revealed that amusic individuals exhibited increased activation in the right pars opercularis of the IFG for memory for tones in comparison to memory for words (see Figure 7B, upper panel). For the reverse contrast ([Verbal > Tonal]), amusic individuals showed increased activation in the left IFG along with activations in temporo-occipital regions (see Figure 7B, upper panel).

Group comparisons performed using the ROIs defined by the activated regions in amusic individuals for these contrasts revealed that, in comparison to controls, amusic individuals showed increased activation in the right VLPFC for tonal material, but were not significantly different from controls in the left IFG for verbal material (Figure 7A, lower panel). The coordinates of activation peaks (and cluster size) are presented in Table 5.

Finally, we performed whole brain correlation analysis with the behavioral data. Interestingly this analysis revealed that amusic individuals showed a negative correlation between performance in the MTt and the % of signal change in the right VLPFC extending to the anterior STG for the contrast MTt > PTt (*Retention scans*) see Figure 7C.
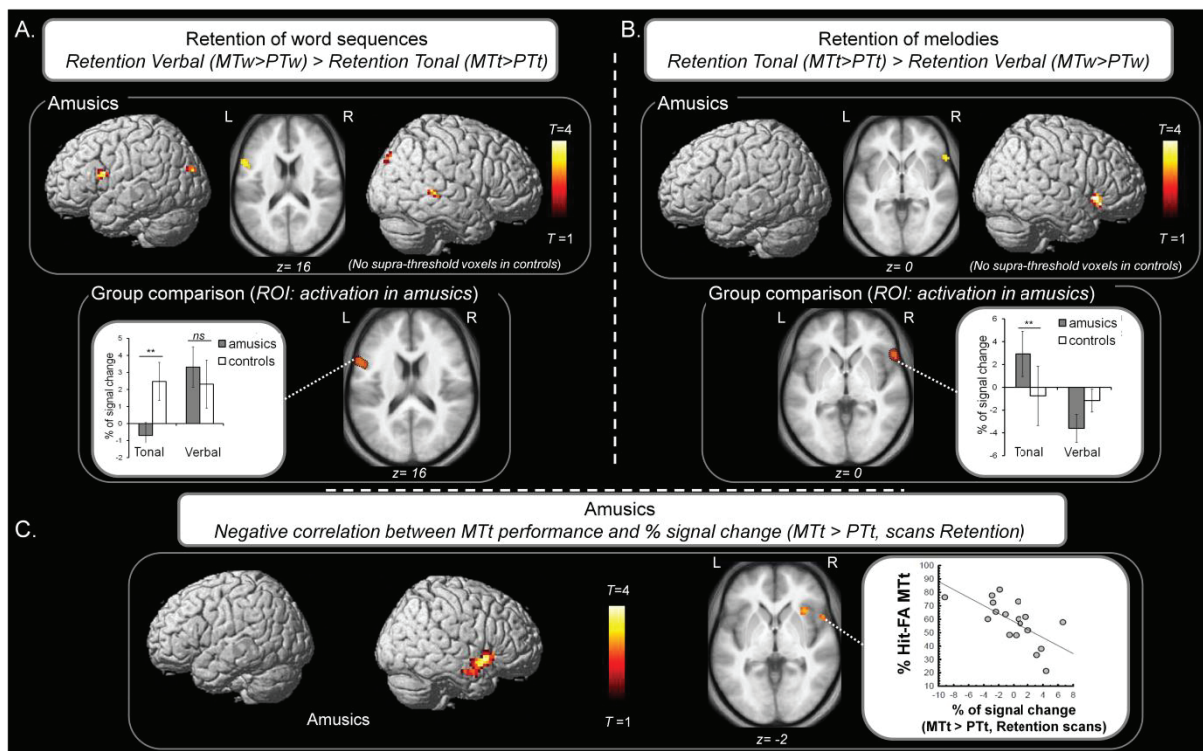
Figure 7A. Functional imaging results for the Retention scans for the contrasts [Verbal Retention (MTw>PTw) > Tonal Retention (MTt > PTt)](A) and [Tonal Retention (MTt>PTt) > Verbal Retention (MTw > PTw)](B), for amusic individuals (A, B, upper panel), and for the group comparison performed in the ROIs (as indicated with orange areas on A lower panel) defined by the regions that were activated in amusic individuals for these contrasts (A,B, lower panel). Results are displayed at p <0.001 (uncorrected) over the mean structural image of all participants (axial view) in the MNI space and on the single subject 3D render provided with SPM. The areas of activation are detailed in Table 5. In lower panels (A, B), the graphics show the average ROIs data in terms of % of signal change (MTt >PTt) for Tonal and Verbal Retention. Errors bars represent the SEM. C. Whole brain correlation analysis between behavioral performance in the MTt and % of signal change for Retention scans in amusic individuals (MTt>PTt). Results are displayed at p <0.001 (uncorrected) over the mean structural image of all participants (axial view) in the MNI space and on the single subject 3D render provided by SPM.

Table 5 *Areas of functional activations for the Retention scans for the comparison of Verbal and Tonal memory (MT > PT) in Amusic individuals and for the Amusic individuals > Controls contrast. Is also presented in the table the coordinates of the negative correlation between behavioral and functional data in amusic individuals. Coordinates are in the MNI space and correspond to the peak of each cluster. IFG, Inferior Frontal Gyrus; MTG, Middle Temporal Gyrus, R, Right, L, Left, H, Hemisphere.*

Contrast	Lobe	Region	H	T	x y z mm	mm3
Amusics Retention: Tonal (MTt > PTt) > Verbal (MTw > PTw)						
	Frontal	IFG pars opercularis	R	2.97	54 20 -5	837
		IFG pars orbitalis		2.58	48 26 -11	
Controls Retention: Tonal (MTt > PTt) > Verbal (MTw > PTw)						
			*	*	*	*
Amusics >Controls Tonal Retention : MTt > PTt (ROIs: activation in amusic individuals)						
	Frontal	IFG pars opercularis	R	2.21	57 17 -2	756
Amusics Retention: Verbal (MTw > PTw) > Tonal (MTt > PTt)						
	Frontal	IFG pars opercularis	L	3.11	-56 14 6	783
	Temporal	MTG	R	3.05	66 -40 1	459
	Occipital	Cuneus	R	3.34	15 -85 37	675
			L	3.18	-4 -70 24	486
Controls Retention: Verbal (MTw > PTw) > Tonal (MTt > PTt)						
			*	*	*	*
Amusics: Negative correlation (%Hit-FA MTt and % signal change (Retention scans (MTt > PTt))						
	Frontal	IFG pars opercularis	R	3.75	51 8 -11	5076
		Insula		2.43	33 20 -2	
	Temporal	MTG	R	2.37	57 -1 -20	
				2.21	57 -7 -20	

Discussion

The present study investigated the cerebral correlates of tonal and verbal STM in congenital amusic individuals and their matched controls. Using fMRI, we explored participants' brain responses associated to 1) the encoding of melodic information in memory, and 2) the retention of verbal and tonal materials in STM. Behavioral results showed that amusic individuals exhibited strongly impaired performance in comparison to controls for tonal STM, but not for verbal STM. The functional results showed that the pitch memory deficits in the amusic brain can be observed in the DLPFC and VLPFC during the retention of tonal material. Interestingly, the functional data revealed that while amusic individuals exhibit altered brain functioning for the retention of tonal material, they show preserved brain activations for the retention of verbal material in STM. This result could be interpreted as suggesting separate neural resources for verbal and tonal STM in the amusic brain. This hypothesis received support in the final analysis, in which we directly compared the fMRI brain responses associated to the retention of verbal and tonal sequences in STM: while amusic individuals recruited differently the same networks for each material, controls showed no significant differences in terms of brain activity between the two materials.

STM deficit for tones but not for words in congenital amusia

In comparison to controls, amusic participants' performance was unimpaired for verbal material (for both the MTw and the PTw), but was impaired for tonal material (MTt and PTt). This observation is in agreement with the behavioral data reported by Tillmann et al. (2009) who suggested that the STM deficit in amusia is pitch-specific. Interestingly, like in Tillmann et al. (2009) study, controls exhibited increased performance for tonal material in comparison to verbal material, while amusic individuals did not, thus suggesting that for controls the MTt is considered as an easier task than the MTtw. Based on this assumption, observing similar performance between the two groups for verbal information but strongly decreased performance for the tonal information in amusic individuals represents evidence for the pitch-related STM deficit in congenital amusia, but not a general STM deficit, which would affect verbal STM too. It might be argued, nevertheless, that amusic individuals' decreased attention, interest or involvement in tonal tasks could explain their decreased performance

for the MTt. However, amusic individuals did not show differences in terms RTs for the verbal and tonal tasks. Furthermore, a positive correlation was observed between RTs data of verbal and tonal tasks in both groups, thus suggesting that similar attentional resources (or implication) are recruited for the two materials in amusic individuals and controls.

In addition, based on two observations of the present study, and in line with Tillmann et al. (2013), we propose that pitch discrimination deficits can be excluded as the origin of the impaired STM performance: 1) several amusic individuals' pitch thresholds (12 out of 18 amusic individuals) were comparable to controls' pitch thresholds (see also Tillmann et al., 2009); and 2) impaired short-term memory for the MTt was observed in amusic individuals (in comparison to controls) even for the changed tones with large interval sizes (i.e., the large changes were superior to all amusic individuals' PDT). These three observations allow us to argue that the deficit in short-term memory performance for the MTt was not merely a consequence of increased PDT in amusic individuals.

It should be noted however, that in contrast to previous studies (Albouy et al., 2013a; Albouy et al., 2013b; Foxton et al., 2004b; Tillmann et al., 2009), we observed an overlap between amusic individuals' and controls' performance for the MTt. Based on a recent study (Albouy et al., *in preparation*), we make the hypothesis that this effect is due to the lack of complexity of melodies used in the present experiment. Indeed, the tones sequences were composed of only 3 tones and half of the pairs contained repetitive tones, thus decreasing the contour complexity of the to-be compared material for the MTt. Thus observation of overlap (or not) in short-term memory tasks might be a function of the complexity of the material or the difficulty of the task: no overlap in terms of performance was observed between controls and amusic individuals for longer sequences (5-tones in Tillmann et al., 2009 and Albouy et al., 2013b; 6-tones in Albouy et al., 2013a, 5 tones in Gosselin et al., 2009), while overlap has been observed for simpler material (as for the pitch discrimination data). Note, however, that we choose such simple material as the delay between the to-be compared melodies in the present experiment was very long (this long delay was need for the sparse sampling acquisition). As it is well known that amusic individuals' STM abilities decreased with increasing retention delay (Williamson et al., 2010), we used simple 3-tone sequences in order to allow amusic participants to exhibit STM performance for tones superior to chance level despite the long delay.

Listening to melodies in congenital amusia

The first set of fMRI analyses investigated whether amusic individuals functional brain responses can differ from that of controls for listening to melodies (*Encoding scans,* [*MTt + PTt > Silence*]*) and* for the active encoding of tonal information in memory (*Encoding scans,* MTt > PTt). This later analysis was done in order to pursue the characterization of the encoding deficits in congenital amusia. While Hyde et al. (2011) have observed near normal cortical functioning in the auditory cortices in the amusic brain during passive listening of melodies, Albouy et al. (2013a) have shown altered and delayed responses in the same areas during the active encoding of melodic information in the context of a STM task.

In the present study, when contrasting the brain responses related to listening to melodies (MTt + PTt > Silence), we observed that, in line with the data from Hyde et al. (2011), both amusic individuals and controls recruited bilateral temporal regions covering both primary and secondary auditory areas. Additionally, when exploring group differences in ROIs corresponding to the common activated areas in the two groups, no differences were observed between amusic individuals and controls. This result suggested that when listening to melodies amusic individuals showed near normal brain activity in bilateral auditory cortices. We then explored whether supplementary activation could be observed when high level auditory processing of information is required, under the assumption that listening to melodies in the context of a memory task (active encoding) involves more complex processes than simple perception ([MTt > PTt] *Encoding* scans only). However, in both amusic individuals and controls participants no significant difference between the two tasks was observed in this analysis of the encoding scans. This result contrasts with the conclusions of Albouy et al (2013a) who proposed that altered brain responses in the auditory cortex support the STM deficits for tones in amusia. However, it should be noted that these discrepant findings could arise from differences in the methodology used in the two studies. The altered responses observed in the MEG study were related to the latency of the brain responses, amusic individuals exhibiting delayed responses (from 20 to 40 ms) in comparison to controls. While these fine temporal differences could be well estimated using the high temporal resolution imaging technique of the MEG, they are not measurable with low temporal resolution imaging methods (such as fMRI).

Retention of melodic information in congenital amusia

To investigate neural correlates of maintenance of tonal information in memory and compare it to the network involved in encoding processes, we compared the functional brain responses recorded during *Encoding* and *Retention scans* of tonal STM (MTt > PTt). While no differences were observed for the *Encoding > Retention* contrast, we observed enhanced brain activation in the right DLPFC in controls for the *Retention > Encoding* comparison. This finding is in agreement with previous data showing a strong implication of the DLPFC in the retention of information in STM (Curtis and D'Esposito, 2003; D'Esposito, 2007; Logie and D'Esposito, 2007; Petrides, 1991, 1995a) and more specifically for pitch memory processing (Peretz and Zatorre, 2005; Zatorre et al., 1994). Interestingly this area was not activated in amusic individuals as revealed by the direct comparison between amusic individuals' and controls' functional data. This result is in agreement with Albouy et al. (2013a) study showing that while controls recruited the right DLPFC during the retention delay of the melodic contour task, as revealed by an analysis of gamma-band activity in the MEG data, (peak at *x = 45 y = 31 z = 25* in Albouy et al. 2013a; *x = 42 y = 23 z = 40* in the present study) amusic individuals did not. In line with this observation, we thus propose that this finding reflects altered high-level representations and/or manipulation of task-relevant information in memory in the amusic brain.

Retention of tone and word sequences in congenital amusia

We first explored the networks activated by the retention of verbal and tonal information independently. To this aim, we examined the MT > PT contrasts for the *Retention scans* for Verbal and Tonal tasks separately. Interestingly, for the verbal material, amusic and control groups showed increased activation in the left VLPFC for the MTw in comparison to the PTw and no significant differences of BOLD signal changes were observed between the two groups in this region. Observing a recruitment of several parts of the left IFG for verbal memory is in line with numerous neuroimaging studies that have hypothesized this region as supporting subvocal rehearsal mechanisms of verbal STM (Awh et al., 1996; Fiez et al., 1996; Gruber and von Cramon, 2003; Paulesu et al., 1993; Ravizza et al., 2004; Schulze and Koelsch, 2012). This result thus comfort the dominant hypothesis suggesting the existence of similarities between the cortical networks supporting STM for words on the one hand and speech perception and production on the other hand. In addition it suggests that STM processing

for verbal material in the amusic brain is preserved and recruits similar networks to that observed in controls. We can thus propose that in congenital amusia, there is not a general STM deficit that could be overcome for verbal material thanks to explicit, verbal strategies (which don't apply for tonal material).

For tonal material, within-group analyses revealed that the control group showed increased activation the right and left DLPFC for the MTt in comparison to the PTt, along with increased activations in the left SMA, the left MFG, the left Pre-central Gyrus and the left Temporal pole. Interestingly, these regions were not activated when contrasting *Retention* and *Encoding* scans in the previous analysis (MTt > PTt) thus suggesting an implication of these brain areas in both mechanisms. This hypothesis is in agreement with neuroimaging studies that have hypothesized a role of this network in manipulating information not only during the retention of information but also during its encoding, even though to a lesser extent (Bidet-Caulet et al., 2009; D'Esposito, 2007; Khan and Muly, 2011). More importantly, amusic individuals did not recruit this network more strongly for MTt than for PTt, and the comparison of amusic individuals and controls brain responses in these regions revealed (as in the previous analysis) that amusic individuals showed decreased activation in comparison to controls only in the right and left DLPFC for tonal material. Observing comparable networks in amusic individuals and controls for verbal material, but altered responses in amusic individuals for tonal material provide the first evidence of the specificity of the STM deficit for pitch in the amusic brain.

Comparison between memory for tone and word sequences in congenital amusia

To explore more precisely this hypothesis, the last analysis of the present study compared STM (MT > PT *Retention scans*) processes in amusic individuals and controls for verbal and tonal materials. In controls, regardless of the contrast (either [Tonal > Verbal] of [Verbal > Tonal]), no differences were observed between the two materials, thus suggesting that the retention of tone and word sequences shares, to some extent, a common network. This hypothesis had previously received support in studies that have compared auditory STM processing for verbal and tonal information in typical listeners (Hickok et al., 2003; Koelsch et al., 2009; Schulze and Koelsch, 2012; Schulze et al.,

2011a) and that have observed a considerable overlap of neural resources underlying STM for both materials.

In contrast to the control group, amusic participants showed significant differences in terms of functional activation for both [Tonal > Verbal] and [Verbal > Tonal] comparisons. For the [Tonal > Verbal] contrast, amusic individuals showed increased activation in the right VLPFC (BA 44). For the reverse contrast [Verbal > Tonal], amusic individuals showed increased activation in the left VLPFC along with activations in temporo-occipital regions. Group comparisons performed in these ROIs revealed that, in comparison to controls, amusic individuals showed increased activation in the right VLPFC for tonal material, but did not exhibit difference in terms % of BOLD signal change in the left VLPFC for verbal material. Additionally, this contrat also suggest altered recruitment of the right VLPFC in amusic individuals in comparison to controls for tonal material- as the contrast used tonal material as a baseline for the comparison. This result further supports that amusic individuals and controls recruit similar cortical networks for the memory of verbal material, but not for tonal material.

The observation of activation in the right VLPFC for tonal memory in amusic individuals could be considered as a marker of the impaired STM abilities for tones in congenital amusia. This hypothesis received support in the observed negative correlation between amusic individuals' performance in the MTt task and the percentages of signal change in the right VLPFC for the MTt > PTt contrast (*Retention scans*). By considering this differential recruitment of frontal areas between amusic individuals and controls for tonal material (amusic individuals recruiting mostly the VLPFC and controls the DLPFC), we propose that additionally to the deficits previously observed in the encoding steps of STM (based on the MEG data of Albouy et al., 2013a), amusic individuals are also impaired in the retention steps of memory for pitch. fMRI studies have indeed suggested a differential role of the DLPFC and VLPFC in STM (D'Esposito, 2007; Logie and D'Esposito, 2007; Owen, 2000). Owen (2000) has proposed that the VLPFC plays a general role in memory, in which it is supposed to trigger active low-level encoding strategies (such as rehearsal), and to initiate explicit retrieval of information from long-term memory. In contrast, the DLPFC has been hypothesized to mediate more complex types of processing and could be considered as a specialized region where stimuli or events, which have been first interpreted and maintained in other association cortical areas, can be re-coded, monitored, and manipulated (D'Esposito, 2007). Observing that for tonal material, amusic individuals

recruit mostly brain areas that are supposed to be more strongly involved in low-level memory processes (encoding, rehearsal for the VLPFC), and observing that altered responses in regions requiring high-level memory (monitoring and manipulation for the DLPFC) mechanisms let us to argue that amusic individuals STM deficit is related to the transformation of tonal information into high level memory representation.

Conclusion

The findings of the present study confirmed the memory deficits in amusic individuals for tonal material and showed for the first time that the brain mechanisms supporting the retention of verbal information seem to be preserved in congenital amusia. However, several questions remain unresolved: while the tonal and verbal memory mechanisms involve the same cortical structures, few is known about the dynamic of the processing and how these regions interact in terms of network activity. Further investigation is thus necessary to understand the dynamic of the retention processes for verbal and tonal memory.

References

Albouy, P., Cousineau, M., Caclin, A., Tillmann, B., and Peretz, I. (in preparation). Building pitch memory traces in congenital amusia.

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

Albouy, P., Schulze, K., Caclin, A., and Tillmann, B. (2013b). Does tonality boost short-term memory in congenital amusia? Brain Res.

Awh, E., Jonides, J., Smith, E.E., Schumacher, E.H., Koeppe, R.A., and Katz, S. (1996). Dissociation of storage and rehearsal in verbal working memory: evidence from positron emission tomography. Psychol Sci 7, 25–31.

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

Baddeley, A. (1992). Working memory. Science 255, 556-559.

Baddeley, A. (2003). Working memory: looking back and looking forward. Nat Rev Neurosci 4, 829-839.

Baddeley, A. (2010). Working memory. Curr Biol 20, R136-140.

Baddeley, A. (2012). Working memory: theories, models, and controversies. Annu Rev Psychol 63, 1-29.

Baddeley, A.D. (1986). Working Memory (Oxford, Clarendon Press).

Bamiou, D.E., Musiek, F.E., and Luxon, L.M. (2003). The insula (Island of Reil) and its role in auditory processing. Literature review. Brain Res Brain Res Rev 42, 143-154.

Belin, P., Zatorre, R.J., Hoge, R., Evans, A.C., and Pike, B. (1999). Event-related fMRI of the auditory cortex. Neuroimage 10, 417-429.

Berz, W.L. (1995). Working Memory in Music: A Theoretical Model. Music Perception: An Interdisciplinary Journal 12, 353-364.

Bidet-Caulet, A., Ye, X.L., Bouchet, P., Guenot, M., Fischer, C., and Bertrand, O. (2009). Non-verbal auditory cognition in patients with temporal epilepsy before and after anterior temporal lobectomy. Front Hum Neurosci 3, 42.

Buchsbaum, B.R., and D'Esposito, M. (2008). The search for the phonological store: from loop to convolution. J Cogn Neurosci 20, 762-778.

Buchsbaum, B.R., Olsen, R.K., Koch, P., and Berman, K.F. (2005). Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. Neuron 48, 687-697.

Chein, J.M., and Fiez, J.A. (2001). Dissociation of verbal working memory system components using a delayed serial recall task. Cereb Cortex 11, 1003-1014.

Chen, S.H., and Desmond, J.E. (2005). Cerebrocerebellar networks during articulatory rehearsal and verbal working memory tasks. Neuroimage 24, 332-338.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? Prog Brain Res 169, 323-338.

Curtis, C.E., and D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. Trends Cogn Sci 7, 415-423.

D'Esposito, M. (2007). From cognitive to neural models of working memory. Philos Trans R Soc Lond B Biol Sci 362, 761-772.

Deutsch, D. (1970). Tones and numbers: specificity of interference in immediate memory. Science 168, 1604-1605.

Deutsch, D. (1975). Auditory memory. Can J Psychol 29, 87-105.

Fiez, J.A., Raife, E.A., Balota, D.A., Schwarz, J.P., Raichle, M.E., and Petersen, S.E. (1996). A positron emission tomography study of the short-term maintenance of verbal information. J Neurosci 16, 808-822.

Foster, N.E., Halpern, A.R., and Zatorre, R.J. (2013). Common parietal activation in musical mental transformations across pitch and time. Neuroimage 75, 27-35.

Foster, N.E., and Zatorre, R.J. (2010). A role for the intraparietal sulcus in transforming musical pitch information. Cereb Cortex 20, 1350-1359.

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

Friston, K., Holmes, A., Worsley, K.J., Poline, J.B., Frith, C.D., and Frackowiak, R.S.J. (1995). Statistical parametric maps in functional imaging: a general linear approach. Hum Brain Mapp 2, 189- 210.

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

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

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

Gruber, O., and von Cramon, D.Y. (2003). The functional neuroanatomy of human working memory revisited. Evidence from 3-T fMRI studies using classical domain-specific interference tasks. Neuroimage 19, 797-809.

Hickok, G., Buchsbaum, B., Humphries, C., and Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. J Cogn Neurosci 15, 673-682.

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

Hyde, K.L., Lerch, J.P., Zatorre, R.J., Griffiths, T.D., Evans, A.C., and Peretz, I. (2007). Cortical thickness in congenital amusia: when less is better than more. J Neurosci 27, 13028-13032.

Hyde, K.L., Peretz, I., and Zatorre, R.J. (2008). Evidence for the role of the right auditory cortex in fine pitch resolution. Neuropsychologia 46, 632-639.

Hyde, K.L., Zatorre, R.J., Griffiths, T.D., Lerch, J.P., and Peretz, I. (2006). Morphometry of the amusic brain: a two-site study. Brain 129, 2562-2570.

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

Khan, Z.U., and Muly, E.C. (2011). Molecular mechanisms of working memory. Behav Brain Res 219, 329-341.

Kirschen, M.P., Chen, S.H., Schraedley-Desmond, P., and Desmond, J.E. (2005). Load- and practice-dependent increases in cerebro-cerebellar activation in verbal working memory: an fMRI study. Neuroimage 24, 462-472.

Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Muller, K., and Gruber, O. (2009). Functional architecture of verbal and tonal working memory: an FMRI study. Hum Brain Mapp 30, 859-873.

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

Logie, R.H., and D'Esposito, M. (2007). Working memory in the brain. Cortex 43, 1-4.

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

Mandell, J., Schulze, K., and Schlaug, G. (2007). Congenital amusia: an auditory-motor feedback disorder? Restor Neurol Neurosci 25, 323-334.

Owen, A.M. (2000). The role of the lateral frontal cortex in mnemonic processing: the contribution of functional neuroimaging. Exp Brain Res 133, 33–43.

Paulesu, E., Frith, C.D., and Frackowiak, R.S. (1993). The neural correlates of the verbal component of working memory. Nature 362, 342-345.

Pechmann, T., and Mohr, G. (1992). Interference in memory for tonal pitch: implications for a working-memory model. Mem Cognit 20, 314-320.

Peretz, I. (2013). The Biological foundations of music: Insights from congenital amusia. In The Psychology of Music, D. Deutsch, ed. (Elsevier), p. in press.

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

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

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

Petrides, M. (1991). Functional specialization within the dorsolateral frontal cortex for serial order memory. Proc Biol Sci 246, 299-306.

Petrides, M. (1994). Frontal lobes and behaviour. Curr Opin Neurobiol 4, 207-211.

Petrides, M. (1995). Functional organization of the human frontal cortex for mnemonic processing. Evidence from neuroimaging studies. Ann N Y Acad Sci 769, 85-96.

Ravizza, S.M., Delgado, M.R., Chein, J.M., Becker, J.T., and Fiez, J.A. (2004). Functional dissociations within the inferior parietal cortex in verbal working memory. Neuroimage 22, 562-573.

Salame, P., and Baddeley, A. (1989). Effects of background music on phonological short-term memory. The Quarterly Journal of Experimental Psychology 41A, 107-122.

Schendel, Z.A., and Palmer, C. (2007). Suppression effects on musical and verbal memory. Mem Cognit 35, 640-650.

Schulze, K., Gaab, N., and Schlaug, G. (2009). Perceiving pitch absolutely: comparing absolute and relative pitch possessors in a pitch memory task. BMC Neurosci 10, 106.

Schulze, K., and Koelsch, S. (2012). Working memory for speech and music. Ann N Y Acad Sci 1252, 229-236.

Schulze, K., Mueller, K., and Koelsch, S. (2011a). Neural correlates of strategy use during auditory working memory in musicians and non-musicians. Eur J Neurosci 33, 189-196.

Schulze, K., Zysset, S., Mueller, K., Friederici, A.D., and Koelsch, S. (2011b). Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. Hum Brain Mapp 32, 771-783.

Semal, C., Demany, L., Ueda, K., and Halle, P.A. (1996). Speech versus nonspeech in pitch memory. J Acoust Soc Am 100, 1132-1140.

Silveri, M.C., Di Betta, A.M., Filippini, V., Leggio, M.G., and Molinari, M. (1998). Verbal short-term store-rehearsal system and the cerebellum. Evidence from a patient with a right cerebellar lesion. Brain 121 (Pt 11), 2175-2187.

Stewart, L. (2011). Characterizing congenital amusia. Q J Exp Psychol (Hove) 64, 625-638.

Stewart, L., von Kriegstein, K., Warren, J.D., and Griffiths, T.D. (2006). Music and the brain: disorders of musical listening. Brain 129, 2533-2553.

Tillmann, B., Albouy, P., and Caclin, A. (2013). Congenital Amusias. In Handbook of Clinical Neurology (Oxford, Elsevier), p. in press.

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

Vallar, G. (2006). Memory systems: The case of phonological short-term memory. A festschrift for Cognitive Neuropsychology. Cogn Neuropsychol 23, 135-155.

Vallar, G., and Baddeley, A.D. (1984). Phonological short-term store, phonological processing and sentence comprehension: A neuropsychological case study. Cognitive Neuropsychology 1, 121-141.

Vallar, G., and Papagno, C. (2002). Neuropsychological impairments of verbal short-term memory. In The Handbook of Memory Disorders, M.D.K.B.A.W. A.D. Baddeley, ed. (Chichester, Wiley), pp. 249–270.

Williamson, V.J., Baddeley, A.D., and Hitch, G.J. (2010a). Musicians' and nonmusicians' short-term memory for verbal and musical sequences: comparing phonological similarity and pitch proximity. Mem Cognit 38, 163-175.

Williamson, V.J., McDonald, C., Deutsch, D., Griffiths, T.D., and Stewart, L. (2010b). Faster decline of pitch memory over time in congenital amusia. Adv Cogn Psychol 6, 15-22.

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

Zatorre, R.J., Evans, A.C., and Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. J Neurosci 14, 1908-1919.

Chapter VI
Building pitch memory traces in congenital amusia

Objectives

Results of <u>Study</u> 2 confirmed the implication of a right fronto-temporal network in amusic individuals' pitch memory deficit. Additionally, it provided further insights in our understanding of the cerebral correlates of the deficits, notably by showing that retention processes differ between verbal and tonal materials in the amusic brain.

While <u>Study 2</u> improved our understanding about the cerebral correlates of the retention deficits, the hypothesis of the altered encoding (<u>Study 1)</u> needs further support. In <u>Study 1</u>, for the encoding of the first melody of the pair, the amusic brain elicited strongly delayed N100m components in bilateral auditory cortices and the inferior frontal gyri, reflecting delayed encoding of the information that might impair the following processing steps. Additionally, the effective connectivity patterns between these two regions were altered in the amusic brain, as assessed by Dynamic Causal Modeling (David et al., 2006). This observation supports the assumption that amusic individuals might show delayed and altered functioning in the early steps of the memory processing (encoding, echoic memory, building of the pitch memory traces) leading to the observed impairments in the pitch dimension (for both pitch discrimination and memory tasks).

Along these lines, <u>Study 3</u> tried to pursue the characterization of amusic individuals' deficits by specifically exploring their pitch encoding abilities. We manipulated sound characteristics that have been previously shown to facilitate the encoding of the information, to assess whether amusic individuals could benefit (or not) of such manipulations in pitch discrimination and short-term memory tasks.

Article 3

Building pitch memory traces in congenital amusia

In preparation

Building pitch memory traces in congenital amusia

Authors: Philippe Albouy [1, 2, 3], Marion Cousineau [4], Anne Caclin [2, 3], Barbara Tillmann [1, 3] and Isabelle Peretz [4]

Affiliations: [1] CNRS, UMR5292; INSERM, U1028; Lyon Neuroscience Research Center, Auditory Cognition and Psychoacoustics team, Lyon, F-69000, France

[2] INSERM, U1028; CNRS, UMR5292; Lyon Neuroscience Research Center, Brain Dynamics and Cognition Team, Lyon, F-69000, France

[3] University Lyon 1, Lyon, F-69000, France

[4] International Laboratory for Brain, Music and Sound Research (BRAMS) and Center for Research on Brain, Language and Music (CRBLM), Department of Psychology, University of Montreal, Montreal, QC, Canada H3C 3J7

Corresponding author: Philippe Albouy, INSERM U1028 - CNRS UMR5292, Brain Dynamics and Cognition Team, Lyon Neuroscience Research Center, Le Vinatier Hospital (Bât. 452); 95, Bd Pinel, 69675 Bron Cedex, France. Phone: +33 4 72 13 89 36. Fax: +33 4 72 13 89 01. E-mail: philippe.albouy@inserm.fr

Abstract

Congenital amusia is a lifelong disorder of music perception and production primarily characterized by impairments along the pitch dimension, notably in pitch discrimination and memory. Previous psychoacoustic research on typical listeners has demonstrated that the discrimination of two short sounds (< 300 ms) differing in pitch is improved when a silent gap is inserted between the to-be-compared sounds (Inter Stimulus Interval). The present study investigated the construction of pitch memory traces in congenital amusia by systematically manipulating sound characteristics (Tone Duration, Inter Stimulus Interval and Stimulus Onset Asynchrony) in single tone comparison and tone sequence comparison tasks. Congenital amusic individuals and matched controls performed two tasks for which they had to indicate whether elements presented in pairs (either single tones or tone-sequences) were the same or different. In Experiment 1, amusic individuals' and controls' abilities for single tone comparison (pitch discrimination with tones being presented without Inter-Stimulus-Interval) were investigated as a function of Tone Duration (100 ms, 350 ms) for two Pitch Interval Sizes (100 cents, 200 cents). In Experiment 2, participants' short-term memory performance for tone sequences were investigated as a function of Tone Duration (100 ms, 350 ms), Inter Stimulus Interval (present or absent) and Stimulus Onset Asynchrony for two different Sequence Lengths (3-tone and 4-tone sequences). For both experiments, when the time given to encode the auditory information was too short, amusic individuals exhibited strongly impaired performance as compared to their matched controls. However, when Tone Duration and/or Inter Stimulus Interval increased, a major improvement in terms of accuracy was observed in both amusic individuals and controls. Most interestingly, when the time dedicated to encode the information was long enough, amusic individuals were able to reach comparable performances as those of controls for the two experiments. Based on these results, we argue that pitch perception and memory deficits in congenital amusia could be mediated by strong impairments in the construction of memory traces for short tones, associated to deficits in the retention of the information.

Key words: *Tone Deafness; short-term memory; single tone comparison; pitch*

1. Introduction

Congenital amusia (Peretz and Hyde, 2003) refers to lifelong deficits in music perception and production. Amusic individuals have difficulties detecting someone singing out-of-tune (including themselves), detecting a wrong or out-of-tune note and recognizing familiar tunes without the help of lyrics (Ayotte et al., 2002; Peretz, 2013; Stewart, 2011; Tillmann et al., 2013). Unlike acquired amusia following brain damage (Nicholson et al., 2002; Nicholson et al., 2003; Patel et al., 1998; Peretz, 1993; Peretz, 2002; Peretz et al., 1997; Steinke et al., 2001; Tillmann et al., 2007), congenital amusia occurs without brain injury and despite normal performance on tests of intelligence and normal cognitive functioning (see Peretz et al., 2002, Foxton et al., 2004, for extensive testing). The most widely investigated hypothesis about the core deficit in congenital amusia is that the musical deficits arise from altered pitch processing, with deficits in pitch discrimination and short-term memory.

Pitch discrimination deficits were observed with numerous behavioral approaches (Ayotte et al., 2002; Foxton et al., 2004b; Hyde and Peretz, 2004; Peretz et al., 2002), showing that amusic individuals exhibit difficulties in detecting pitch changes in simple tone sequences (for changes smaller than one semitone, whereas controls detected changes of a quarter of a semi-tone; Hyde & Peretz, 2004) and show elevated pitch discrimination thresholds in comparison to controls (Foxton et al., 2004b; Tillmann et al., 2009). Additionally, short-term memory deficits for pitch have been reported for amusic individuals: amusic individuals' performance in delayed comparison tasks is more affected than that of controls by increasing durations of the retention interval between single tones (Gosselin et al., 2009; Williamson et al., 2010b) and by increasing length of to-be compared tone sequences (Gosselin et al., 2009). Amusic individuals also exhibit increased sensitivity to interference caused by irrelevant tones presented during the retention interval (Gosselin et al., 2009; Williamson et al., 2010). When investigated with a conventional "span" memory task (Williamson and Stewart, 2010), or with delayed comparison tasks (Tillmann et al., 2009) performance was found to be impaired in congenital amusic individuals for musical material, but not for verbal material.

The first pitch discrimination studies have led to the assumption that congenital amusia arises from a failure to extract pitch attributes with sufficient accuracy. This lack of precision in the first

sensory steps of auditory processing is bound to affect tonal coding and pitch-related processing in general, in turn leading to higher-level deficits such as altered pitch memory abilities (Peretz, 2013; Stewart, 2011; Tillmann et al., 2013). However, this view has been challenged by the observation that some amusic individuals exhibited pitch discrimination thresholds in the range of controls (Foxton et al., 2004; Tillmann et al. 2009; Albouy et al. 2013a,b), suggesting that, to some extent, amusic individuals' are able to extract pitch attributes as well as do controls. In addition, the used memory tasks were constructed with pitch changes that exceeded amusic participants' psychophysically measured pitch difference detection thresholds, thus pointing on deficit unrelated to pitch discrimination impairments. Along these lines, Tillmann et al. (2013) have introduced an alternative hypothesis for the origin of the amusic individuals' deficits that extends beyond perceptual impairments. They proposed that amusic individuals' pitch perception deficits could be traced down to impairments in pitch memory (Albouy et al., 2013a; Albouy et al., 2013b; Gosselin et al., 2009; Tillmann et al., 2009; Williamson et al., 2010b; Williamson and Stewart, 2010). In this *memory* hypothesis, impaired memory for pitch affects performance also in lower-level pitch tasks, such as pitch discrimination measures, as these still require the comparison of two or more events presented sequentially.

This hypothesis calls for further research investigating which step of pitch memory processing might be altered in congenital amusia. Auditory memory has been described as relying on several processing steps in which information has to be 1) extracted by the perceptual systems (extraction of the auditory attributes) (Bregman, 1990), 2) maintained in *echoic memory*, also referred as an internal representation of the sound or pre-perceptual auditory storage (where a first memory trace is maintained for duration of the stimulus for up to about 300 ms (Cowan, 1984; Massaro, 1972)) and 3) stored in auditory short-term memory for several seconds or minutes (Baddeley, 2010).

Under this memory model, it should be noted that pitch discrimination and memory might at least share the first two mechanisms, which is supported by the fact that in the auditory domain (as well as in other perceptual domains), "*discrimination*" and "*memory*" processes are deeply interrelated. Pitch perception in a discrimination paradigm is the relation between the memory traces of a previous stimulus and a stimulus belonging to the present (Demany and Semal, 2007). Similarly, in the context

of tone sequences, each tone of the sequence has to be kept in memory (internal representation) to allow for the construction of an efficient memory trace that can then be actively maintained in short-term memory.

Evidence for relations between "*discrimination*" and "*memory*" processes has been provided in various domains. For example, in a classic memory span task, the length of the sequences that can be correctly recalled varies as a function of how '*discriminable'* the different sounds are. A classic example for verbal material is called the *phonological similarity effect* (Baddeley, 1986). It corresponds to the fact that longer sequences of syllables can be recalled when the syllables are phonologically dissimilar than when they share common phonemes. Similar phenomena have been reported outside the domain of language, notably for pitch and loudness sequences (Cousineau et al., 2009, 2010; McFarland and Cacace, 1992; Williamson et al., 2010a).

This hypothesis of the interrelation between pitch discrimination and short-term memory processes thus requires investigating which step of pitch memory processing might be altered in amusic individuals. A recent study (Albouy et al., 2013a) allowed shedding some light on this issue. In a delayed melody comparison task, thus relying heavily on short-term memory processes, functional abnormalities were observed in the amusic brain for the encoding of melodic information. When encoding the first melody of the pair, the amusic brain elicited strongly delayed N100m components in bilateral auditory cortices and inferior frontal gyri, reflecting delayed encoding of the information that might impair the following processing steps. Additionally, the effective connectivity patterns between these two regions were strongly altered in the amusic brain, as assessed by Dynamic Causal Modeling (David et al., 2006). This observation supports the assumption that amusic individuals might show delayed and altered functioning in the early steps of the memory processing (encoding, echoic memory, building of the pitch memory traces) leading to the observed impairments in the pitch dimension (for both pitch discrimination and memory tasks).

In the present study, we sought to test specifically the contribution of an impaired construction of pitch memory traces to the pitch discrimination and memory impairments in congenital amusia. For short sounds (<300 ms), the amount of time required by normal listeners to construct an appropriate memory trace of that sound can exceed the duration of the sound itself (Demany and Semal, 2007).

Indeed, the detection of frequency changes between successive short tones is facilitated by the introduction of a silent gap between the tones. The increased difficulty without gap is referred to as "backward masking effect" (Massaro and Loftus 1996; Massaro & Idson, 1977; Demany & Semal, 2005) and illustrates the importance of sufficient processing time for the construction of an appropriate internal representation (or memory trace) of the sound.

Based on this observation, the present study aims to test the timing of the elaboration of this internal representation in amusic individuals, and how it affects their performance for both single tone comparison and short-term memory tasks for melodies. Ten amusic individuals and ten matched control participants were required to indicate whether elements (either single tones or melodies) presented in pairs were the same or different. We systematically manipulated the Tone Duration, and the Inter-Stimulus-Interval (ISI) (resulting in the manipulation of Stimulus Onset Asynchrony and the presence or not of a silent gap between the tones) of tones in tone pairs and tone sequences to assess 1) whether such manipulations would affect amusic individuals' performance on single tone comparison and short-term memory for melodies differently or on both processes similarly; 2) whether optimal stimulus parameters could be found in order to allow amusic individuals to reach controls' performance level for both single tone comparison and memory for melodies.

2. Materials and Methods

2.1 Participants

Ten amusic participants and ten non-musician controls matched for age, handedness, educational background (calculated in years of education), and musical training (calculated in years) participated in the study. Amusic group (age ranging from 62 to 72) and control group (age ranging from 61 to 75) were both composed of ten right-handed francophone participants. Moderate (30 dB loss at any frequency) or severe (more than 40 dB loss) peripheral hearing loss in the frequency of interest (between 250 Hz and 1000 Hz) was excluded using standard audiometry, and all participants reported no history of neurological or psychiatric disease. Participants gave their written informed consent, and were paid for their participation. All participants had been thoroughly evaluated in previous testing sessions with the Montreal Battery of Evaluation of Amusia and with the Pitch Change Detection task (PCD, see Table 1 and Figure 3 right panel) from Hyde and Peretz (2004). Participants were considered as amusic if they obtained an average score on the MBEA below the cut-off score (23.4 on average across the six tasks of the battery, maximum score = 30), the cut-off being two standard deviations below the average of the normal population (Peretz et al., 2003). All amusic participants except one had scores below the cut-off score and all controls obtained scores higher than the cut-off score. The amusic participant who obtained an average score superior (23.83) to the cut-off score (23.4) was included as his average score on the pitch subtests of the MBEA (21) was below the average melodic sub-tests' cut-off score (21.6). The average scores of the amusic group (mean = 19.69, SD = 2.56; ranging from 15.33 to 23.83) differed significantly from those of the control group (mean = 26.75, SD = 1.08; ranging from 24.80 to 28; see Table 1).

All participants had previously been tested with the PCD task, in which participants are required to detect a pitch change (from 25 cents (0.25 semitone) to 300 cents (3 semitones)) within the context of a five-tone sequence (isochronous, and otherwise monotonic, see Hyde & Peretz 2004 for details, tones played at a pitch level of C6 (1047 Hz), tone duration = 100ms; ISI = 250ms, SOA = 350 ms).

Table 1: Demographic characteristics of participants and their data in behavioral pretests. Results for the Montreal Battery of Evaluation of Amusia (MBEA) are expressed as number of correct responses (average over the six sub-tests of the battery, maximum score = 30). Pitch Change Detection (PCD) scores were calculated in terms of percentage of Hits (correct response in different trials) minus percentage of false alarms (FA) (incorrect response in same trials) in each group of participants. Data are reported as a function of group, and groups are compared with t-tests; ''NS'' refers to a non-significant difference (p>0.05). Standard deviations are in parentheses.

Characteristics	Controls (n=10)	Amusics (n=10)	t-Test
Age in years	65.10 (4.40)	67.10 (3.5)	*t(18) = 1.12, p = .27 (NS)*
Gender	7 female, 3 male	5 female, 5 male	
Education in years	15 (3.77)	16.5 (2.49)	*t(18) = 1.04, p = .30 (NS)*
Musical education in years	1.60 (1.42)	1.42 (0.96)	*t(18) = 0.36, p = .71 (NS)*
MBEA (cut off score = 23.4) *Peretz et al. (2003)*	26.75 (1.08)	19.69 (2.56)	*t(18) = 8.01, p < .0001*
Pitch Change Detection *Hyde & Peretz (2004)* 25 cents (1/4 semitone)	90.94 (10.80)	31.87 (21.13)	*t(18) = 7.87, p < .0001*
50 cents (1/2 semitone)	94.55 (9.56)	70.49 (22.66)	*t(18) = 3.02, p = .006*
100 cents (1 semitone)	95.37 (9.84)	89.35 (14.30)	*t(18) = 1.09, p = .28 (NS)*
200 cents (2 semitones)	95.37 (10.71)	95.47 (3.64)	*t(18) = 0.02, p = .97 (NS)*
300 cents (3 semitones)	95.38 (9.76)	95.75 (4.68)	*t(18) = 0.10, p = .91 (NS)*

2.2 Set-up

The experiments took place in a sound-attenuated booth and auditory stimuli were presented via SENNHEISER HD 280 pro headphones at participant's "comfortable" sound level. Presentation software (Neurobehavioral systems, Albany, CA, USA) was used to control the presentation of stimuli and record participants' responses given by mouse button presses. Participants performed Experiment 1 first followed by Experiment 2 on the same day of testing, and the testing session lasted approximately one hour.

2.3. Experiment 1: Single Tone Comparison

Participants had to determine whether two piano tones presented without Inter-Stimulus-Interval (ISI) were the same or different.
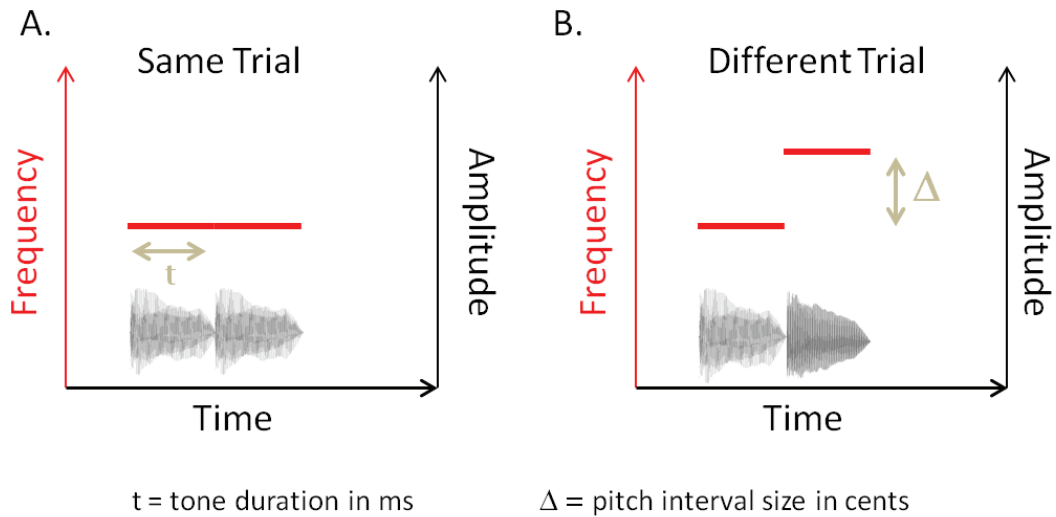

Figure 1: Examples of the stimuli used in the single tone comparison experiment (Experiment 1). Participants made pitch same/different judgments on pairs of tones. A. Example of a same trial; tones are represented as amplitude waveforms and as spectrograms with their frequencies being illustrated by a red line. B. Example of a different trial. The tones had a duration (t) of 100 ms in Conditions 1 and 2, and of 350 ms in Conditions 3 and 4. Δ corresponds to the Pitch Interval Size in different trials, which was either 100 cents (in Conditions 1 and 3) or 200 cents (in Conditions 2 and 4) up or down. Note that the tones were presented without ISI.

2.3.1. Material

The experiment was divided in four conditions, for which we manipulated Tone Duration and Pitch Interval Sizes. For Conditions 1 and 2, each tone had a duration of 100 ms, for Conditions 3 and 4, each tone had a duration of 350 ms (see *t* in Figure 1). There were 64 trials for each Condition (32 same, 32 different). For *different* trials, the second tone was either 100 cents (Condition 1 and Condition 3, see Δ in Figure 1), or 200 cents (Condition 2 and Condition 4) upper or lower than the

first tone[4]. On half (16) of the *different* trials, the pitch changes were presented in the ascending direction, and on the other half (16), in a descending direction.

Tone pairs were created using eight piano tones differing in pitch height (Cubase software, Steinberg); all used tones belonged to the key of C Major (C3, D3, E3, F3, G3, A3, B3, C4). For *same* trials, each of the eight tones of the set was used to create *same* pairs (e.g. *C3-C3*), which were thus presented four times each. For *different* trials, the tone pairs were created as function of the pitch interval sizes. For Condition 1 and 3 ($\Delta = 100$ cents, corresponding to one semitone), among 16 *up*-trials, 8 were composed of *E3- F3* pair, and 8 of *B3-C4* pair, which are the only two one-semitone intervals in the C Major key. For the down condition, among 16 trials, 8 were composed of *F3-E3* pair and the other 8 of the *C4- B3* pair. For Conditions 2 and 4 ($\Delta = 200$ cents, corresponding to one tone), among 16 up-trials, there were 4 trials of each of the following pairs *C3-D3, D3-E3, F3-G3, and G3-A3*. For the down trials, among 16 trials, there were 4 trials of each following pairs *D3-C3, E3-D3, G3-F3, and A3-G3*.

2.2.2. Procedure

Trials were blocked by condition, and the blocks were separated by 2-3 minutes of break. The order of presentation of the conditions was counterbalanced across participants. Participants were informed of the manipulation we made for each condition (i.e., short tone duration, etc.). They were asked to indicate their answers by pressing one of two mouse buttons with their right hand after the end of the auditory stimulation. There was no time limit to respond, and participants pressed the middle mouse button to launch the next trial. No feedback was given during the experiment. Within each block (Condition), the trials were presented in a pseudo-randomized order; the same trial type (i.e. same, different) could not be repeated more than three times in a row and similar tone-pairs (i.e. *E3- F3* for different trials) could not be repeated twice in a row. Before each condition, participants performed a set of 4 practice trials with feedback.

---

[4] The choice of these interval sizes were based on data of the PCD of our participants showing that amusic individuals reached normal pitch detection performance for pitch interval sizes of 100 cents (and more) for tones of 100 ms duration and of 250 ms ISI (resulting in 350ms SOA). The 200 cents pitch change interval was thus chosen as a control material for which we hypothesized normal detection abilities in amusic individuals (even for a short duration).

2.3. Experiment 2: Short-term memory tasks with tone sequences

Participants performed a melodic short-term memory task for which they had to compare two tone sequences (S1 and S2) separated by a silent retention period of 2000 ms. Sequences were composed of either of 3 or 4 piano tones.

2.3.1. Material

There were four Conditions (for each sequence length) based on the manipulation of Tone Duration and Inter-Stimulus-Interval. The duration of the tones was either 100-ms (Conditions 1 and 2) or 350-ms (Conditions 3 and 4), and they were presented either without ISI (Conditions 1 and 3) or with ISI (Conditions 2 and 4), resulting in a range of different Stimulus Onset Asynchronies (SOAs, see Table 2).

Table 2: Characteristics of the short-term memory tasks for tone sequences, note that for the Conditions 2 and 3 the SOA (onset to onset) is equal allowing us to disentangle the contribution of Tone Duration, ISI, or SOA in amusic individuals' short-term memory performance.

Experiment 2: Short Term Memory Tasks	Condition 1	Condition 2	Condition 3	Condition 4
Tone Duration (*t*) in ms	100	100	350	350
Inter Stimulus Interval (ISI) in ms	0	250	0	350
Stimulus Onset Asynchrony (SOA) in ms	100	350	350	700
Overall sequence length 3 Tone sequences	*300*	*800*	*1050*	*1750*
Overall sequence length 4 Tone sequences	*400*	*1150*	*1400*	*2450*

For each sequence length, 128 different melodies (presented as the first sequence, S1) were created using the same piano tones as those used in Experiment 1. For each Condition, there were 32 trials (S1, silence, S2), with 16 *same* and 16 *different* trials. Participants were asked to indicate whether S1 and S2 melodies were the same or different. For *different* trials, one tone in the S2 melody was different from the S1 melody (the change could occur in positions 1 to 3 of the 3-tone sequences,

and in positions 2 or 3 of the 4-tone sequences, equally distributed across trials) and created a contour-violation in the melody[5].

For *different* trials, the Pitch Interval Sizes (see Δ in Figure 2B) were controlled in such a way that no significant differences of interval sizes were observed between all conditions and all sequence lengths used in Experiment 2 (all ps > .09; mean Pitch Interval Size across all conditions = 671 cents (6.71 semitones); SD = 117 cents (1.17 semitones); ranging from 480 cents to 793 cents).
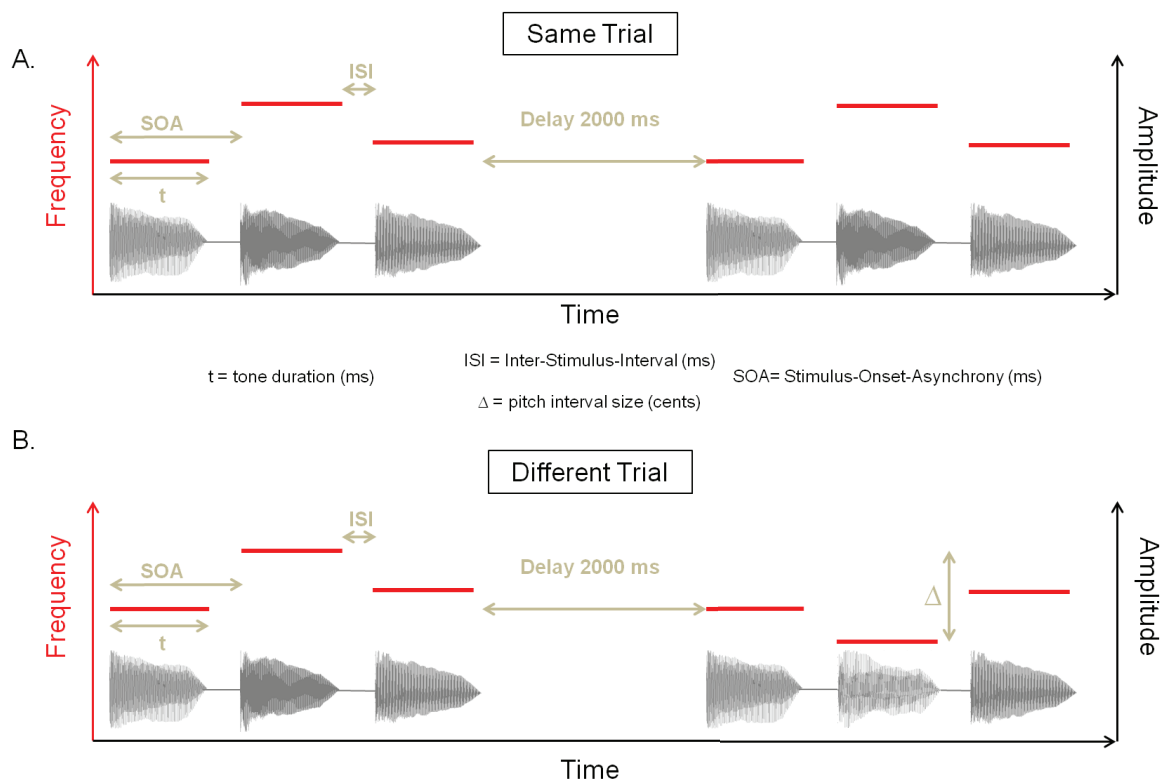

Figure 2: Examples of the musical stimuli used in the short-term memory task for the 3-tone sequence length. A. Example of a same trial; tones are represented as spectrograms and their frequencies are illustrated with the red line. B. Example of a different trial. For "same" trials S1 was repeated as the second melody of the pair (S2) after a 2000 ms delay. For "different" trials, the second melody of the pair changed only for one tone (in position 1 to 3 for the 3-tone sequences and in position 2 and 3 for the 4-tone sequences) that changed the melodic contour.

---

[5] Note that for the 4-tone sequences, the changed tone occurred only in position 2 or 3 to make the task easier, notably by providing participants with two contour change indices, with the preceding tone and the following tone (see Tillmann et al., 2009).

2.3.2 Procedure

Experiment 2 was performed with the same procedure as Experiment 1. Trials were blocked by condition, the order of presentation was counterbalanced between participants, and the blocks were separated by 2-3 minutes of break. Participants were also informed of the stimulus characteristic for each condition (i.e short tone duration, etc.) and were asked to respond by mouse button presses with their right hand after the end of the auditory stimulation. As in Experiment 1, there was no time limit to respond, participants pressed the middle mouse button to launch the next trial and no feedback was given during the experiment. Within each block (condition), the trials were presented in a pseudo-randomized order; with the same constraints as in Experiment 1. Additionally, for *different* trials, trials with a changed tone in S2 in the same position in the sequence could not be repeated twice in a row. As in Experiment 1, participants performed a set of 4 practice trials with feedback before starting each block (condition).

3. Results

3.1. Experiment 1: Single tone comparison

Results are depicted in Figure 3 as percentage of Hit rate (correct response for *different* trials) minus percentage of False Alarm rate (FA, incorrect response for *same* trials). Performance was significantly above chance (i.e., above a Hit-FA value of 0), for all conditions and groups (t-tests, all ps < .0001), except for Condition 1 (t = 100 ms and Δ = 100 cents) in the amusic group (t(9) = 1.52, p = .16).

Hit-FAs were analyzed with a 2x2x2 ANOVA[6] with Group (amusic individuals, controls) as between-participants factor and Tone Duration (t = 100 ms or 350 ms) and Pitch Interval Size (Δ = 100 cents or, 200 cents) as within-participant factors. The main effect of Group was significant: F(1,18) = 44.15 ; p < .0001, with poorer performance on average in amusic individuals (mean = 42.93 ; SD = 35.62) than in controls (mean = 85.63; SD = 18.21). The main effect of Tone Duration was also significant: F(1,18) = 175.22 ; p < .0001, with better mean performance for the long Tone Duration (t = 350 ms; mean = 82.44 ; SD = 19.31) than for the short Tone Duration (t = 100 ms; mean = 45.16 ; SD = 37.54). There was a significant main effect of Pitch Interval Size: F(1,18) = 20.03 ; p < .0001, showing that performance for the large Pitch Interval Size (Δ = 200 cents; mean = 70.63; SD = 34.18) was higher than performance for the small Pitch Interval Size (Δ = 100 cents; mean = 57.97; SD = 35.82).

Interestingly, the Tone Duration by Group interaction was significant F(1,18) = 39.96 ; p < .0001 and further modulated by Pitch Interval Size F(1,18) = 4.37; p = .05. Fischer LSD post hoc tests revealed that while amusic individuals showed higher performance for long tone duration in comparison to short tone duration for both pitch interval sizes (all ps <. 0001) controls, showed the same pattern only for small pitch interval size (Δ = 100 cents) (p= .05) but for large interval size (p =.35) (Δ = 200 cents). Furthermore, while amusic individuals' performance was decreased in comparison to controls for the Conditions 1, 2 and 3 (all ps < .01), they tended to be unimpaired in

---

[6] Note that we tested if the samples (controls and amusic individuals) were derived from a population normally distributed for the measures of interest for both Experiments 1 and 2 using Kolmogorov-Smirnov tests. These analyses revealed that all data were normally distributed (p > .20).

comparison to controls' performance (p = .08, see Figure 3A) for the easiest Condition, as shown by the reached performance level (Condition 4: long Tone Duration (t = 350 ms) and large Pitch Interval (Δ = 200 cents).

Finally, this analysis revealed that in the amusic group, performance was higher for the Condition 3 (t = 350 ms; Δ = 100 cents) in comparison to Condition 2 (t = 100ms; Δ = 200 cents, amusic individuals p <.001; controls p =.05), meaning that they processed better a small pitch change in long tones than a large pitch change in short tones.

To assess the potential differential benefit of Tone Duration and Pitch interval size in both groups and between groups we performed the following subtractions on the accuracy data : 1) Condition 2-Condition 1 allowing us to explore the benefit of Pitch Interval size in short tones; 2) Condition 3- Condition 1 allowing us to investigate the benefit of Tone Duration for small pitch changes. Note that we performed these subtractions as Condition 2 differs from Condition 1 (the most difficult condition) only on the size of the Pitch Interval Size, while Condition 3 differed from Condition 1 only on Tone Duration. The resulting data were analyzed with a 2x2 ANOVA with Group as between participant factor and Type of benefit (Pitch Interval Size, Tone Duration) as within participant factors.

The main effect of group was not significant (F(1,18) = 1.82, p = .19). However the main effect of Type of benefit was significant (F(1,18) = 45.93, p < .0001) and further modulated by Group F(1,18) = 17.76; p < .0001). Fischer LSD post hoc test revealed that while amusic individuals showed stronger benefit of Tone Duration as compared with controls (p = .005), the benefit of Pitch Interval Size was not different between the two groups (p > .58). In addition, amusic individuals benefited more from increasing Tone Duration in comparison to increasing Pitch interval size (p< .0001), while this effect was not significant in controls (p = .08). These results thus suggest a strong involvement of Tone Duration in improving pitch discrimination abilities (or single tone comparison) in amusic individuals rather than Pitch Interval Size.
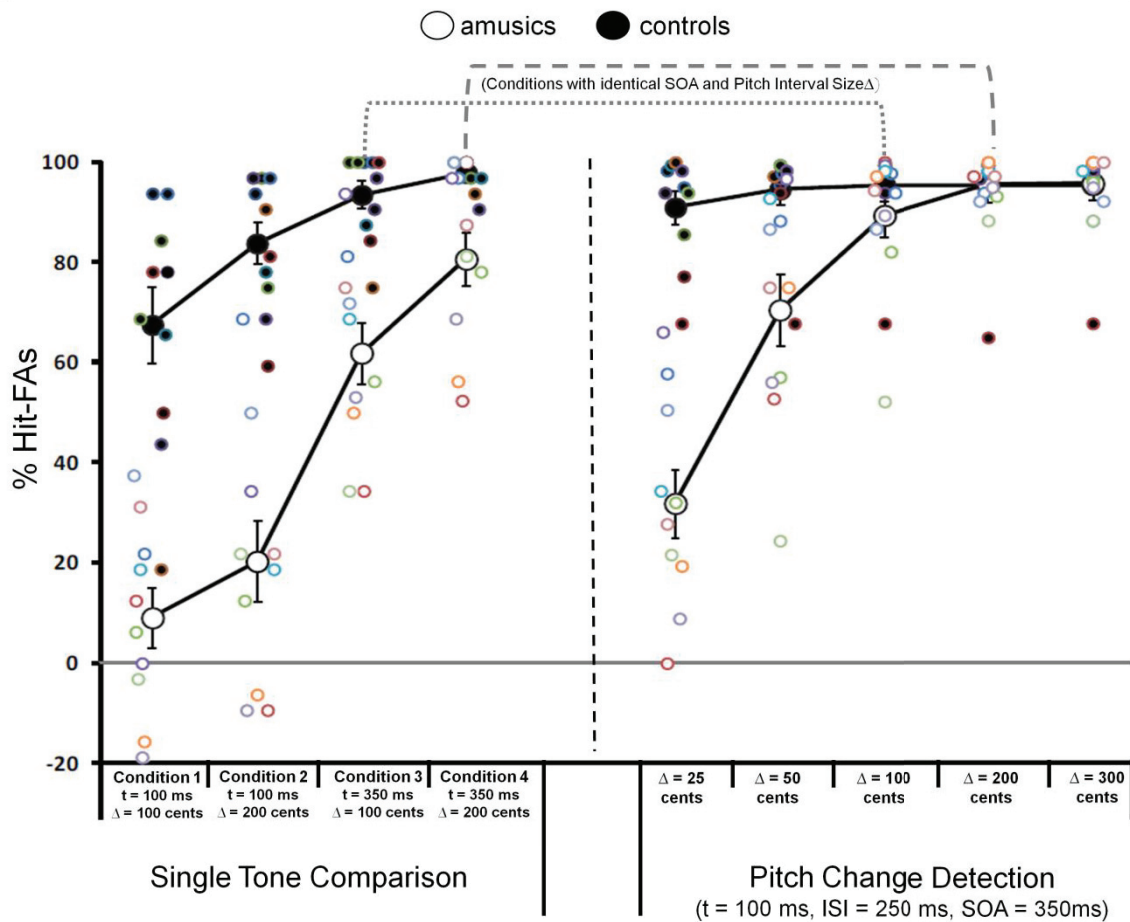
Figure 3. Left Panel: Amusic and controls performance for the Single Tone Comparison Task of Experiment 1 presented as a function of Tone Duration (t) and Pitch Interval Size (Δ).Right Panel Amusic individuals' and controls' performance for the PCD. Black circles, controls; white circles, amusic individuals; colored circles, individual data. Error bars indicate the standard errors of the means; SEM; dotted lines link conditions with similar SOA and Pitch Interval sizes in the two experiments (present Experiment 1 and PCD from Hyde & Peretz, 2004).

Finally, some conditions of Experiment 1 and the PCD share several characteristics (Tone Duration, SOA, and Pitch interval size), even so the tasks were different (either single tone comparison or pitch change detection within a 5 tone monotonic sequence). Conditions 3 and 4 of Experiment 1, the SOA was set to 350 ms as in the PCD (PCD Tone Duration = 100 ms, ISI = 250 ms, SOA = 350 ms; Condition 3 and Condition 4 of Experiment 1 here, Tone Duration t = 350 ms, ITI = 0; SOA = 350 ms). However, the tow experiment differed in terms in of the presence or absence of a

silent gap (ISI) between tones (for the same SOA). It is thus interesting to compare data from the two experiments to assess if the absence of a silent gap between the tones could facilitate amusic individuals' pitch discrimination abilities. Observing higher performance in conditions of Experiment 1 in comparison to the PCD would suggest a benefit of an absence of silent gap on pitch discrimination abilities. In contrast observing higher performance for the PCD in comparison to Experiment 1 would suggest a benefit of the presence of repetitive tones in the PCD task, that offers a stronger reference when the changed tone occurs.

We thus compared the results obtained in the conditions 3 and 4 of this study to previously collected PCD data on the same participants for the two Pitch interval sizes (Δ 100 cents; Δ 200 cents) (Figure 3, right panel). Before presenting this analysis, note, that as previously reported in Hyde & Peretz (2004), amusic individuals' performance was impaired for pitch changes smaller than one semitone (see Table 1 and Figure 3 right panel), while controls showed ceiling performance for changes as small as 0.25 semitones.

Data were analyzed with a 2x2x2 ANOVA with Group as between participant factor and Task (Conditions 3 and 4 of Experiment 1, PCD) and Pitch interval size (Δ 100 cents, Δ 200 cents) as within participant factors. The main effect of group was significant $F(1,18) = 13.53$; $p = .001$ as well as the main effect of Pitch Interval Size ($F(1,18) = 14.79$; $p = .001$. As well as the main effect Task $F(1,18) = 8.64$; $p = .008$

All two-way interactions were significant (all ps < .02), as was the 3-way interaction between Group, Task and Pitch Interval Size ($F(1,18) = 6.07$; $p = .02$). As our question was whether amusic individuals would obtain comparable performance in the two Tasks (Experiment 1, PCD) for both Pitch Interval Sizes ($\Delta = 100$ cents, $\Delta = 200$ cents), we performed planned comparisons to explore this interaction. This analysis revealed that for both Pitch interval sizes, amusic individuals' performance in Experiment 1 were lower than their performance in the PCD (all ps < .0001), while controls' performance in Experiment 1 was only decreased relative to the PCD for the Δ 100 cents Pitch interval Size (p= .03).

3.2. Experiment 2 : Short-term memory tasks for tone sequences

For the short-term memory tasks, Hits-FAs were also computed. Note that one amusic participant was excluded from this analysis as his average performance was 3SD below his group mean for the Condition 4) in both 3-Tones (3.4 % Hits-FA) and 4-Tones sequences (0% Hits-FA). In the remaining participants (9 amusic individuals, 10 controls) performance was significantly above chance, for each condition and each sequence length in each group (t-tests, all ps < .0001). Data were analyzed with a 2x2x4 ANOVA with Group (amusic individuals, controls) as between-participant factor and Sequence Length (3-tones, 4-tones); and Condition (1, 2, 3, 4) as within-participant factors.

The ANOVA revealed that the main effect of Group was significant, $F(1,17) = 4.88$, $p = .04$, with control participants (mean = 80.54; SD = 16.73) reaching higher performance levels than did amusic participants (mean = 67.83; SD = 25.18). The main effect of Sequence Length was significant, $F(1,17) = 38.87$, $p < .0001$, all participants showed higher performance levels for 3-tone sequences (mean = 80.59; SD = 19.50) than for 4-tone sequences (mean = 68.46; SD = 22.84, see Figure 4). The main effect of Condition was also significant $F(1,17) = 33.49$, $p < .0001$, and revealed that both amusic individuals and controls exhibited decreased performances for Condition 1 in comparison to the other Conditions (2, 3, 4; all ps < .0001). Finally, the interaction between Condition and Group $F(3,51) = 4.53$, $p = .006$ was significant.

Fisher LSD post hoc tests were carried out and revealed that amusic individuals were impaired in comparison to the control group for Conditions 1 and 2 (all ps <.04; t = 100 ms) but not for Conditions 3 and 4 (all ps > .11, t = 350 ms).

The last significant interaction observed here did not involve the group factor and was the Sequence Length by Condition interaction $F(3,51) = 5.80$, $p = .034$. Post hoc test revealed that performance was significantly decreased in 4-Tone sequences in comparison to 3-Tone sequences for all conditions (all ps < .005) except for Condition 3 (p = .12). Finally, note that there were no significant differences between Conditions with the same SOA (Conditions 2 and Condition 3 on Figure 4, see Table 2) for the 3-tone sequences (p = .82), and this effect was marginally significant for the 4-tone sequences (p = .07).
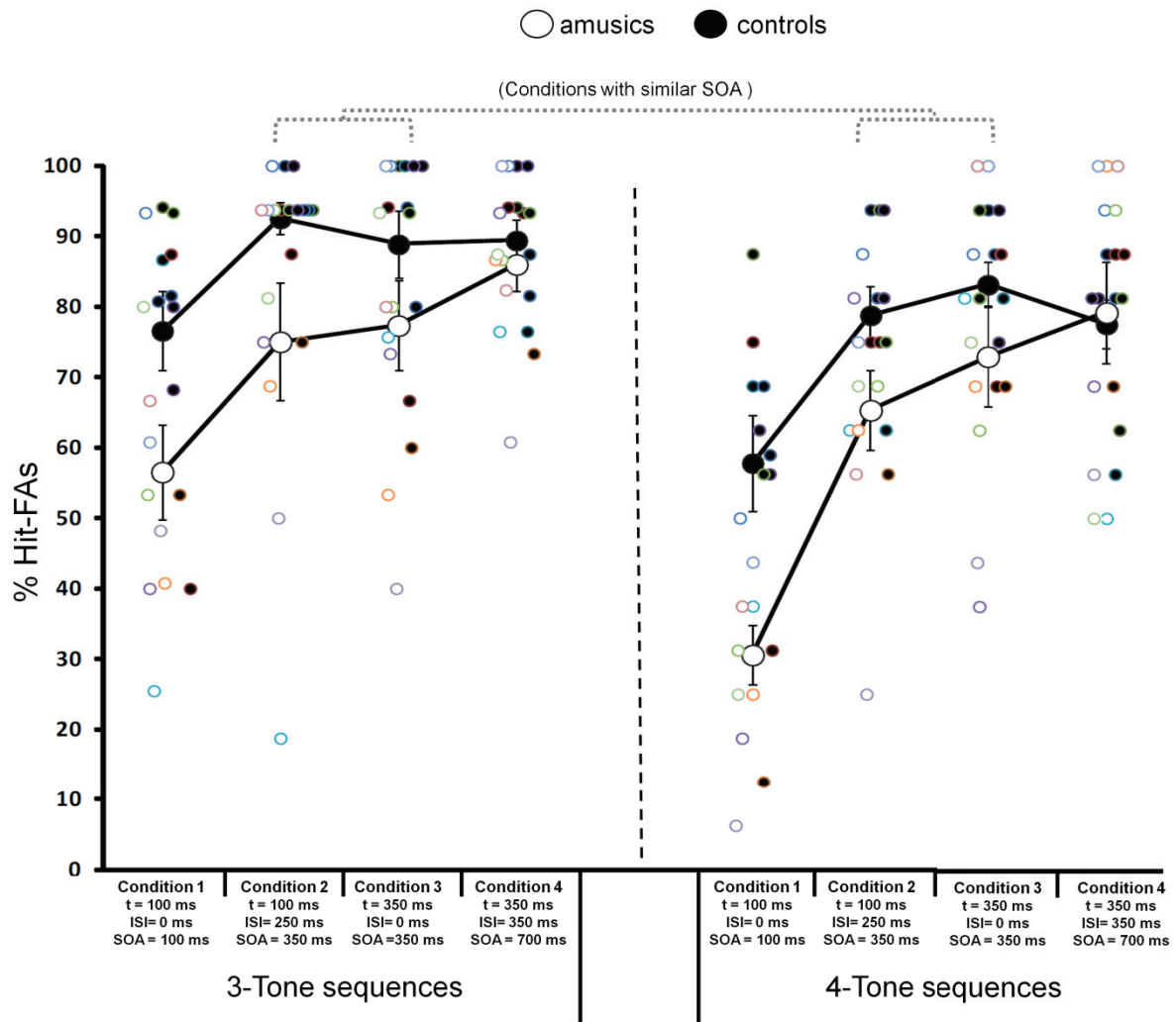
Figure 4: *Amusic and controls performance for the short-term memory tasks (Experiment 2) in terms of percentage of Hits minus FAs presented as a function of Sequence Length (3-Tone Sequences; 4-Tone Sequences), Tone Duration (t =100ms, t =350ms) and ISI (0-250-350ms). Note that the resulting SOA for the different Conditions were as follows (see Table 2): Condition 1 = 100 ms; Conditions 2 and 3 = 350 ms, Condition 4 = 700 ms. Note additionally that the overall sequence lengths varied as function of the manipulations:1) for the 3 Tone Sequences : Condition 1: 300 ms, Condition 2 = 800 ms; Condition 3 = 1050 ms; Condition 4 = 1750 ms. 2) for the 4 Tone Sequences : Condition 1: 400 ms, Condition 2 = 1150 ms; Condition 3 = 1400 ms; Condition 4 = 2450 ms. Black lines, controls; white lines, amusic individuals; circles individual data. Error bars indicate the standard errors of the means, SEM.*

3.3. Correlations

Two sets of analyses were performed: 1) correlations between the behavioral data in Experiments 1 and 2, respectively (*Hit-FA*, for each of the four Conditions in each experiment) and the pretest data (MBEA, PCD); and 2) correlations between data of the Experiments 1 and 2. To explore differences explained not only by the group differences defined via the MBEA or the PCD, we focused on correlations that were significant over all participants *and* for at least one of the two participant groups separately (amusic group or control group).

3.3.1. _MBEA_: For Experiment 1 (single tone comparison), average scores in the MBEA were positively correlated with Condition 2 (Tone Duration t = 100 ms; Δ = 200 cents) in amusic individuals r(8) = .65; p = .03; and all participants r(18) = .93; p < .0001.

For Experiment 2 (short term memory task), average scores in the MBEA were positively correlated with Condition 2 (t = 100 ms, ISI = 250 ms) and Condition 3 (t = 350 ms, ISI = 0 ms) for both sequence lengths (all ps < .024); these correlation were observed in amusic individuals, and over all participants.

3.3.2. _PCD_: For Experiment 1 (single tone comparison), data of the PCD for the most difficult condition (PCD = 25 cents) were positively correlated with the data of Conditions 2, 3 and 4 in amusic individuals and all participants (all ps < .001). For Experiment 2(short term memory task), data of the PCD for the 25 cents condition were positively correlated with data of Condition 2 (t = 100 ms; ISI = 250 ms) for the 3-tone sequences for amusic individuals r(7) = .66; p = .05 and all participants r(17) = .64; p = .002. None of the correlations was significant in controls participants certainly due to a ceiling effect in the data of this group.

3.3.3. Correlations between Experiments 1 and 2: Several correlations were significant; as Tone Duration was the shared stimulus parameter manipulated between Experiments 1 and 2, we present here only correlations between measures for the same Tone Duration.

a)   For t =100 ms: data of Condition 1 (t = 100 ms; ISI = 0 ms) and Condition 2 (t = 100 ms; ISI = 250 ms) of Experiment 1(single tone comparison), were positively correlated with Condition 1 (t = 100 ms; ISI = 0 ms) of 3-tones sequences and Condition 1 (t = 100 ms; ISI = 0 ms) and 2 (t = 100

ms; ISI = 250 ms) of the 4 tones sequences (all ps<.001) of Experiment 2(short term memory task). Note that these correlations were observed in control participants, amusic participants and over all participants.

b)   For t = 350 ms, data of Condition 3 (t = 350 ms; Δ = 100 cents) and Condition 4 (t = 350 ms; Δ = 200 cents) of Experiment 1(single tone comparison) were positively correlated with data of Condition 3 (t = 350 ms; ISI = 0 ms) of 3-tone sequences only (all ps <.001) of Experiment 2(short term memory task). Note that the correlation between Condition 3 in Experiment 1 and Condition 3 in Experiment 2 for the 3-tone sequences was observed in amusic individuals and all participants; and that the correlation between Condition 4 in Experiment 1 and Condition 3 in Experiment 2 was observed in controls and in all participants.

Finally we explored the correlations between tasks comprising the same SOA (Condition 3 and 4 of Experiment 1 (single tone comparison) with Conditions 2 and 3 of Experiment 2 (short-term memory task). This analysis revealed that Condition 3 of Experiment 1 was positively correlated with Condition 2 in Experiment 2 for the 3-Tone sequences and Condition 2 and 3 for the 4-Tone sequences (all ps <.01). Note that these correlations were observed in controls and over all participants (except the Condition 3 for the 4-Tone sequence that was observed also in amusic individuals).

3   **Discussion**

The present study investigated amusic individuals' performance for single tone comparison and short-term memory task for melodies as a function of stimulus characteristics (Tone Duration, ISI, SOA, Sequence Length, Pitch interval size). For the comparison of either single tones or short melodies (pitch-patterns), clear improvements in terms of accuracy were observed in both amusic and controls with increasing Tone Duration and/or ISI. These benefits were observed independently of task difficulty (Pitch Interval size for Experiment 1; Sequence Length for Experiment 2). Most interestingly, we observed that when enough time is given to encode the pitch information, amusic individuals are able to reach performance that is comparable to that of controls in both tasks.

Experiment 1 investigated whether Tone Duration (equal to SOA in this experiment as there was no silent ISI) could affect amusic individuals' single tone comparison (or pitch discrimination) abilities for two different Pitch Interval Sizes ($\Delta = 100$ cents or $\Delta = 200$ cents). Data of Experiment 1 revealed that amusic participants showed altered performance in comparison to controls for single tone comparison, but that these impairments seemed to be strongly related to Tone Duration. While amusic individuals showed impaired performance (as compared to controls) for the short Tone Duration (t = 100 ms), their performance were not different from those of controls for the long Tone Duration (t = 350 ms). By comparing the benefit of Tone Duration and of Pitch Interval size respectively, we observed that the increase of Tone Duration had a more important benefit on amusic individuals' performance than had the increase of Pitch Interval Size.

The benefit of Tone Duration is in line with numerous psychoacoustic studies (Carbotte, 1973; Demany and Semal, 2005; Massaro, 1972; Massaro and Idson, 1977; Small and Campbell, 1962; Sorkin, 1966; Tanner, 1961) showing that auditory discrimination abilities benefit from increasing SOA (or here Tone Duration). This effect can be explained in terms of "backward masking" (Demany and Semal, 2005; Demany and Semal, 2007; Massaro and Idson, 1977; Massaro and Loftus, 1996). When normal listeners are not given the several hundreds of milliseconds that they need to construct the "internal representation" of the pitch of a newly heard sound, the representation of this sound is altered. Massaro (1972; Massaro & Loftus, 1996) proposed that this internal representation constitutes

a "pre-perceptual auditory store" (or echoic memory), where the sound's trace is maintained for a fixed duration independently of the sound itself. If a second sound is presented very soon after the first one, the perceptual analysis of the first sound would be prematurely stopped (the second sound replacing the first one in the pre-perceptual auditory store).

In agreement with these results in normal listeners, Experiment 1 showed that amusic individuals also benefit from long tone durations. The fact that the benefit is stronger for amusic individuals than for control suggests that the time constraints might differ between the two groups, with amusic individuals needing more time than controls to construct a proper 'internal representation' (or pre-perceptual auditory storage) of the sounds. The long tones would allow for such efficient representation to emerge, and would in turn lead to increased discrimination capacities. Finally, it should be noted that amusic individuals' performance for both Pitch Interval Sizes in Experiment 1 (Δ = 100 cents, Δ = 200 cents) even for the long Tone Duration (t = 350 ms) were decreased in comparison to data of the PCD for the same Pitch Interval Sizes. This probably reflects the fact that in the PCD a repetitive sequence of five tones is used (with only the fourth tone being possibly different from the others), instead of a tone pair comparison as required here. In the PCD, the repetitive tones might allow for the construction of an efficient representation of the "standard" tone, and hence facilitate the comparison with the changing tone.

Based on these first conclusions observed with a single tone comparison task (or pitch discrimination task), our study explored whether this benefit of Tone Duration, ISI, and SOA on performance could be transposed to more complex tasks, such as short-term memory processing of pitch patterns. Experiment 2 assessed whether participants' short-term memory performance could vary as a function of Tone Duration (100 ms, 350 ms), Inter Stimulus Interval (present or absent) and Stimulus Onset Asynchrony for two different Sequence Lengths (3-Tones and 4-Tones sequences).

For the short Tone Duration (t = 100 ms), amusic individuals' performance was decreased in comparisons to controls, as previously observed in numerous studies (Albouy et al. 2013a,b; Gosselin et al. 2009; Tillmann et al., 2009; Williamson et al., 2010). Furthermore, this impairment was more important for 4-tone sequences than for 3-tone sequences. This observation is in line with Gosselin et al. (2009) results, showing that amusic individuals were more sensitive to increasing memory load

than controls. The negative impact of increasing memory load also observed in controls is supported by several studies in typical listeners demonstrating that the length of the to-be-remembered information has an impact on short-term memory abilities (Akiva-Kabiri et al., 2009; Demany and Semal, 2007; Grimault et al., 2009; Watson et al., 1990).

For both sequence lengths (3-tones, 4-tones), we demonstrated that Tone Duration and ISI influenced participants' short-term memory performance: amusic and control groups showed better performance for conditions with long Tone Duration and/or with the insertion of a silent ISI (Conditions 2, 3 and 4) in comparison to the condition with short-tone duration and without ISI (Condition 1). Additionally, while amusic individuals showed strongly impaired performance in comparison to controls for the short Tone Duration (Conditions 1 and 2), they were performing as well as controls for the long Tone Duration (Condition 3 and 4). This result is of interest because it allows hypothesizing that the SOA is the principal effect, whit a small benefit (non-significant) of the absence of ISI (all SOA being filled with a tone) This received support in the observation that, while Conditions 2 and Condition 3 have the same SOA (see Table 2), amusic individuals showed impaired performance in comparison to controls for Condition 2 (short tone with ISI), but not for Condition 3 (long tone, without ISI).

As in Experiment 1, these improvements in short-term memory processing suggest that increasing the time given to encode auditory information leads to a better construction of an 'internal representation in a pre-perceptual auditory store' (Massaro, 1972) of each tone of the to-be remembered sequence, leading to an efficient memory of the tone sequence. This result suggests that amusic individuals could strongly improve their short-term memory capacities for tone sequences, when the time dedicated to encode the pitch information is long enough.

Based on the findings in both experiments and on the positive correlations observed between data of Experiments 1 and 2, we argue that congenital amusic individuals' deficits in both single tone comparison (or pitch discrimination) and short-term memory tasks could share the same impairments related to the encoding of pitch in tones of short duration. The hypothesis of a link between pitch discrimination and memory performance has already received support from Demany & Semal (2007) who proposed that pitch discrimination and pitch memory share the same mechanisms as 1) perception

in a discrimination paradigm is based on the relation between a stimulus belonging to the present stimulus and memory traces of previous stimuli, and 2) in the context of a pitch pattern, each tone of the sequence has to be kept in memory to allow for the construction of an efficient memory trace. Here we can thus formulate the hypothesis that amusic individuals' impaired pitch abilities, either for pitch discrimination or pitch memory, could be related to their impaired encoding of short tones.

We will now set this hypothesis in a broader context, notably by considering previous studies that have investigated either pitch discrimination and/or short-term memory processing in congenital amusia. It should be noted that, in these studies (presented in Table 3), a large range of different tone durations, ISI, and SOA have been used (i.e., tone durations from 125 ms to 700 ms, and SOA from 125 to 1200 ms).

Table 3: Sound characteristics in terms of tone duration, ISI and SOA in ms from previous studies in congenital amusia investigating either pitch change detection, pitch discrimination, pitch direction and delayed matching-to-sample tasks (DMST, i.e. short-term memory tasks) for pitch patterns.

Table 3

Study	Task	Tone Duration	ISI	SOA
Peretz et al., 2002				
	Pitch Discrimination	*700*	*0*	*700*
Hyde & Peretz, 2004				
	Pitch Change Detection	*100*	*250*	*350*
Foxton et al., 2004				
	Segmented pitch change detection	*250*	*100*	*350*
	Gliding pitch change detection	*250*	*NA*	
	Pitch direction determination	*250*	*NA*	
	DMST for pitch patterns	*250*	*0*	*250*
	Streaming	*125*	*0*	*125*
Jones et al., 2009				
	Difference Limen for frequency	*250*	*500*	*750*
	DMST for pitch patterns	*200*	*150*	*350*
Liu et al., 2010				
	Pitch change detection	*600*	*600*	*1200*
	Pitch direction discrimination			
Gosselin et al., 2009				
	DMST for pitch patterns	*300*	*0*	*300*
Tillmann et al., 2009				
Albouy et al., 2013b				
	DMST for pitch patterns	*500*	*40*	*540*
	Pitch Discrimination Thresholds	*100*	*150*	*250*
Williamson et al., 2010				
	Single Tone Comparison	*200*	*0 to 15000*	*200 to 15200*
Albouy et al 2013a				
	DMST for pitch patterns	*250*	*0*	*250*

Interestingly, by using such diverse material, these studies have revealed that while some amusic individuals exhibited pitch discrimination threshold in the range of controls, no overlap in terms of performance has been observed between the two groups for short-term memory task for tone sequences (delayed pitch patterns in Table 3).

For tasks requiring pitch discrimination (as well as pitch change detection or pitch direction judgments), the hypothesis of delayed construction of pitch memory traces (internal representation) for short tones could explain that in cases of long SOA (or long tone duration) some amusic individuals are able to reach performance levels that are comparable to those of controls.

However, it might be argued that this hypothesis is not sufficient to interpret the previously reported deficits in pitch memory. Indeed, tone duration and or/SOA used in the studies that have shown memory deficits in amusia were rather long (from 250 to 540ms SOA), thus challenging the main conclusion of the present study. We propose that additional processing steps of auditory short-term memory are impaired in amusic individuals. In the present study, the tones sequences were composed of only 3 or 4 tones thus constituting a rather simple material in terms of contour information and of memory load. In previous studies, the sequences were longer (5 tones in Gosselin et al., 2009; Tillmann et al., 2009 and Albouy et al., 2013b; 6 tones in Albouy et al., 2013a), thus leading to more complex material.

The use of pitch patterns leads to more complex contour information that has to be encoded in memory (Dowling, 1991; Dowling and Fujitani, 1971). According to this view, we could hypothesize that, for complex pitch patterns, amusic individuals are impaired in other processing steps of auditory memory processing. This hypothesis found support in Albouy et al. (2013a) study: during the retention of contour information, the recruitment of a fronto-temporo-parietal network (D'Esposito, 2007; Schulze and Koelsch, 2012), which is well-known for short-term memory is altered in the amusic brain.

Alternatively, it could be hypothesized that increasing tone duration is sufficient to help the encoding of short sequences, but could be not sufficient for longer sequences. According to this hypothesis, also for the short sequences, long tone duration might not fully restore a 'normal pitch memory trace'(which is nevertheless sufficient for encoding less complex contour material).

All together, we thus propose that the present study provides first behavioral evidence of impaired auditory encoding of pitch information for short tones in congenital amusia, and that this impairment contributes to the observed pitch perception and memory deficit. In addition, we proposed that these mechanisms cannot support all the previously reported memory impairments and that pitch perception and memory deficits in congenital amusia could be also mediated by impairments in the short-term retention of pitch information.

References:

Akiva-Kabiri, L., Vecchi, T., Granot, R., Basso, D., and Schon, D. (2009). Memory for tonal pitches: a music-length effect hypothesis. Ann N Y Acad Sci 1169, 266-269.

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

Albouy, P., Schulze, K., Caclin, A., and Tillmann, B. (2013b). Does tonality boost short-term memory in congenital amusia? Brain Res.

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

Baddeley, A. (2010). Working memory. Curr Biol 20, R136-140.

Baddeley, A.D. (1986). Working Memory (Oxford, Clarendon Press).

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

Carbotte, R.M. (1973). Retention of time information in forced-choice duration discrimination. Perception & Psychophysics 14, 440-444.

Cousineau, M., Demany, L., and Pressnitzer, D. (2009). What makes a melody: The perceptual singularity of pitch sequences. J Acoust Soc Am 126, 3179-3187.

Cousineau, M., Demany, L., and Pressnitzer, D. (2010). The role of peripheral resolvability in pitch-sequence processing. J Acoust Soc Am 128, EL236-241.

Cowan, N. (1984). On short and long auditory stores. Psychol Bull 96, 341-370.

D'Esposito, M. (2007). From cognitive to neural models of working memory. Philos Trans R Soc Lond B Biol Sci 362, 761-772.

David, O., Kiebel, S.J., Harrison, L.M., Mattout, J., Kilner, J.M., and Friston, K.J. (2006). Dynamic causal modeling of evoked responses in EEG and MEG. Neuroimage 30, 1255-1272.

Demany, L., and Semal, C. (2005). The slow formation of a pitch percept beyond the ending time of a short tone burst. Percept Psychophys 67, 1376-1383.

Demany, L., and Semal, C. (2007). The Role of Memory in Auditory Perception. In Auditory Perception of Sound Sources, R.R. Fay, and A.N. Popper, eds., pp. 77-113.

Dowling, W.J. (1991). Tonal strength and melody recognition after long and short delays. Percept Psychophys 50, 305-313.

Dowling, W.J., and Fujitani, D.S. (1971). Contour, interval, and pitch recognition in memory for melodies. J Acoust Soc Am 49, Suppl 2:524+.

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

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

Grimault, S., Lefebvre, C., Vachon, F., Peretz, I., Zatorre, R., Robitaille, N., and Jolicoeur, P. (2009). Load-dependent brain activity related to acoustic short-term memory for pitch: magnetoencephalography and fMRI. Ann N Y Acad Sci 1169, 273-277.

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

Massaro, D.W. (1972). Preperceptual images, processing time, and perceptual units in auditory perception. Psychol Rev 79, 124-145.

Massaro, D.W., and Idson, W.L. (1977). Backward recognition masking in relative pitch judgments. Percept Mot Skills 45, 87-97.

Massaro, D.W., and Loftus, G.R. (1996). Sensory and perceptual storage: Data and theory. In Memory, Second Edition, Handbook of Perception and Cognition E.L. Bjork, and R.A. Bjork, eds. (San Diego, Academic Press), pp. 67-99.

McFarland, D.J., and Cacace, A.T. (1992). Aspects of short-term acoustic recognition memory: modality and serial position effects. Audiology 31, 342-352.

Nicholson, K.G., Baum, S., Cuddy, L.L., and Munhall, K.G. (2002). A case of impaired auditory and visual speech prosody perception after right hemisphere damage. Neurocase 8, 314-322.

Nicholson, K.G., Baum, S., Kilgour, A., Koh, C.K., Munhall, K.G., and Cuddy, L.L. (2003). Impaired processing of prosodic and musical patterns after right hemisphere damage. Brain Cogn 52, 382-389.

Patel, A.D., Peretz, I., Tramo, M., and Labreque, R. (1998). Processing prosodic and musical patterns: a neuropsychological investigation. Brain Lang 61, 123-144.

Peretz, I. (1993). Auditory agnosia: a functional analysis. In Thinking in Sound The Cognitive Psychology of Human Audition, S. McAdams, and E. Bigand, eds. (New York, Oxford University Press), pp. 199–230.

Peretz, I. (2002). Brain specialization for music. Neuroscientist 8, 372-380.

Peretz, I. (2013). The Biological foundations of music: Insights from congenital amusia. In The Psychology of Music, D. Deutsch, ed. (Elsevier), p. in press.

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

Peretz, I., Belleville, S., and Fontaine, S. (1997). Dissociations between music and language functions after cerebral resection: A new case of amusia without aphasia. Can J Exp Psychol 51, 354-368.

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

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

Schulze, K., and Koelsch, S. (2012). Working memory for speech and music. Ann N Y Acad Sci 1252, 229-236.

Small, A.M., Jr., and Campbell, R.A. (1962). Temporal differential sensitivity for auditory stimuli. Am J Psychol 75, 401-410.

Sorkin, R.D. (1966). Temporal interference effects in auditory amplitude discrimination. Perception & Psychophysics 1, 55-58.

Steinke, W.R., Cuddy, L.L., and Jakobson, L.S. (2001). Dissociations among functional subsystems governing melody recognition after right-hemisphere damage. Cogn Neuropsychol 18, 411-437.

Stewart, L. (2011). Characterizing congenital amusia. Q J Exp Psychol (Hove) 64, 625-638.

Tanner, W.P. (1961). Physiological implications of psychophysical data. Ann NY Acad Sci 89, 752–765.

Tillmann, B., Albouy, P., and Caclin, A. (2013). Congenital Amusias. In Handbook of Clinical Neurology (Oxford, Elsevier), p. in press.

Tillmann, B., Peretz, I., Bigand, E., and Gosselin, N. (2007). Harmonic priming in an amusic patient: the power of implicit tasks. Cogn Neuropsychol 24, 603-622.

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

Watson, C.S., Foyle, D.C., and Kidd, G.R. (1990). Limits of auditory patterns discrimination for patterns with various durations and numbers of components. J Acoust Soc Am 88, 2631-2638.

Williamson, V.J., Baddeley, A.D., and Hitch, G.J. (2010a). Musicians' and nonmusicians' short-term memory for verbal and musical sequences: comparing phonological similarity and pitch proximity. Mem Cognit 38, 163-175.

Williamson, V.J., McDonald, C., Deutsch, D., Griffiths, T.D., and Stewart, L. (2010b). Faster decline of pitch memory over time in congenital amusia. Adv Cogn Psychol 6, 15-22.

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

Chapter VII Does tonality boost short-term memory in congenital amusia

Objectives

All together the results from studies 1, 2 and 3 have suggested that pitch-related deficits in congenital amusia could be mediated by impaired encoding of pitch information for short tones (Study 1, Study 3) along with impairments in short-term retention of pitch information (Study 1, Study 2). These studies thus suggested that in the context of a short-term memory task, the amusic brain is impaired in all processing steps, that is encoding, retention and retrieval of melodic information (see Study 1).

Study 4 investigated if these short-term memory impairments could be boosted in congenital amusia by exploiting their implicit knowledge about musical structure. It had been shown that non-musicians' implicit knowledge of musical regularities can improve short-term memory for tone information (Schulze et al., 2012). Based on studies reported in Chapter III showing preserved implicit processing in amusic individuals, Study 4 investigated whether it can influence their short-term memory for pitch material.

Article 4

Does tonality boost short-term memory in congenital amusia?

Research Report
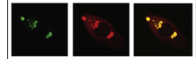
Does tonality boost short-term memory in congenital amusia?

Philippe Albouy[a,b,c,*], Katrin Schulze[d], Anne Caclin[b,c], Barbara Tillmann[a,c]

[a]CNRS, UMR5292, INSERM, U1028, Lyon Neuroscience Research Center, Auditory Cognition and Psychoacoustics Team, Lyon F-69000, France
[b]INSERM, U1028, CNRS, UMR5292, Lyon Neuroscience Research Center, Brain Dynamics and Cognition Team, Lyon F-69000, France
[c]University Lyon 1, Lyon F-69000, France
[d]Developmental Cognitive Neuroscience Unit, UCL Institute of Child Health, London, UK

A B S T R A C T

Congenital amusia is a neuro-developmental disorder of music perception and production. Recent findings have demonstrated that this deficit is linked to an impaired short-term memory for tone sequences. As it has been shown before that non-musicians' implicit knowledge of musical regularities can improve short-term memory for tone information, the present study investigated if this type of implicit knowledge could also influence amusics' short-term memory performance. Congenital amusics and their matched controls, who were non-musicians, had to indicate whether sequences of five tones, presented in pairs, were the same or different; half of the pairs respected musical regularities (tonal sequences) and the other half did not (atonal sequences). As previously reported for non-musician participants, the control participants showed better performance (as measured with d') for tonal sequences than for atonal ones. While this improvement was not observed in amusics, both control and amusic participants showed faster response times for tonal sequences than for atonal sequences. These findings suggest that some implicit processing of tonal structures is potentially preserved in congenital amusia. This observation is encouraging as it strengthens the perspective to exploit implicit knowledge to help reducing pitch perception and memory deficits in amusia.

1. Introduction

Congenital amusia is a lifelong disorder of music perception and production that has been estimated to affect about 4% of the general population (Ayotte et al., 2002; Peretz and Hyde, 2003; Peretz et al., 2002; Stewart, 2006, 2008, 2011). Amusic individuals are unable to recognize a familiar tune (without the help of lyrics) or to detect that someone (including themselves) sings out of tune. This deficit, which affects individuals in their everyday life (McDonald and Stewart, 2008), cannot be explained by peripheral hearing loss, brain lesions, or general cognitive or social impairments (Ayotte et al., 2002).

Seminal studies of congenital amusia (Foxton et al., 2004; Peretz and Hyde, 2003; Peretz et al., 2002) have led to the

*Corresponding author at: INSERM U1028 – CNRS UMR5292, Brain Dynamics and Cognition Team, Lyon Neuroscience Research Center, Le Vinatier Hospital (Bât. 452); 95, Bd Pinel, 69675 Bron Cedex, France. Fax: +33 4 72 13 89 01.
E-mail address: philippe.albouy@inserm.fr (P. Albouy).

assumption that this condition is based on a disorder of fine-grained pitch processing, with amusic individuals exhibiting elevated pitch discrimination thresholds. Pitch discrimination difficulties in amusics were observed in simple tone sequences (for changes smaller than two semi-tones, whereas controls detected changes of a quarter of a semi-tone; Hyde and Peretz, 2004) and with various psychoacoustic approaches (Foxton et al., 2004). Recent studies have extended these findings by showing that amusic individuals may display impaired short-term memory for pitch, even in the absence of elevated pitch discrimination thresholds (Albouy et al., 2013; Gosselin et al., 2009; Tillmann et al., 2009; Williamson et al., 2010; Williamson and Stewart, 2010).

Evidence for a short-term memory deficit in congenital amusia was obtained using pitch comparison tasks both for single tones and for tone sequences. For single tones, a stronger sensitivity to interference (irrelevant tones presented during the silent retention period between to-be-compared tones) in short-term memory tasks was observed in amusic participants in comparison to controls (Gosselin et al., 2009; Williamson et al., 2010). Further support for a short-term memory deficit for single tones was provided by Williamson et al. (2010) who showed that amusic participants had a steeper decline in performance with increasing retention delay between two tones than did controls.

Amusics' short-term memory deficit for pitch was also shown for tone sequences, where amusics showed a greater decrease in performance, relative to controls, with increasing length of the to-be-remembered sequences, thus revealing amusics' sensitivity to memory load (Gosselin et al., 2009). The specificity of the short-term deficit for musical material was shown by Tillmann et al. (2009) who compared short-term memory performance for sequences of tones, timbres, and words. The results indicated that amusic participants performed worse than controls for tone and timbre sequences (see also Marin et al., 2012), but not for words, thus providing evidence that amusics' memory deficit is selective to the musical domain.

Importantly, the pitch changes used in the aforementioned short-term memory studies exceeded the amusics' psychophysically measured pitch discrimination thresholds. The findings therefore point to a memory-based deficit for pitch that extends beyond a perceptual impairment. In addition, it is relevant to note that while some amusics exhibited normal pitch discrimination thresholds (Albouy et al., 2013; Foxton et al., 2004; Tillmann et al., 2009), all amusic participants displayed impairments in pitch memory tasks. These observations have led to the working hypothesis that the core deficit in congenital amusia is an impairment of short-term memory for pitch, and that this impairment may be accompanied by an impairment of fine-grained pitch processing at the perceptual level.

Interestingly, despite explicit processing deficits, recent behavioral data have provided evidence for some spared implicit pitch processing in congenital amusia. Using a priming paradigm, Tillmann et al. (2012a) revealed that some musical knowledge is indeed preserved in congenital amusia, in particular about the syntactic-like functions of chords in the Western musical system (see also Omigie et al., 2012). Similarly, a recent study investigating amusics' perception of

musical emotions has also suggested some spared musical structure processing related to mode (major, minor) (Paquette et al., 2011). Further evidence for implicit tonal processing in amusia has been provided by Tillmann et al. (2012b), showing that amusic individuals can differentiate between tonal and atonal musical pieces. Taken together, these studies reveal that amusics might possess more tonal knowledge than previously thought on the basis of their performance in tasks requiring explicit judgments, such as the detection of a single (out-of-key) tone or the memory-storage of melodies. These data sets thus suggest that congenital amusics might have acquired some knowledge about musical structures by mere exposure to music that obeys to tonal rules, as previously reported for non-musician listeners (Bigand and Poulin-Charronnat, 2006; Tillmann et al., 2000).

In light of the recent evidence for some spared implicit knowledge of tonal structures in congenital amusia, the question arises as to whether amusics' tonal knowledge might also benefit their short-term memory for tone sequences (i.e., for tonal sequences in comparison to atonal sequences). For other materials, such as verbal material or spatial patterns (Bor et al., 2003; Savage et al., 2001), it has been shown that material that is structured leads to better short-term and working memory performance than unstructured material. Similarly, the inherent structure of music, as based on the regularities of the Western tonal system, improved the maintenance of tone sequences in short-term memory for musician and non-musician listeners (in comparison to atonal sequences, Schulze et al., 2012).

The present study investigated amusics' short-term memory for tonal and atonal sequences using the five-tone (tonal and atonal) melodies of Schulze et al. (2012). By comparing the processing of structured and unstructured musical material, we aimed to investigate amusics' implicit tonal knowledge and its impact on short-term memory. Eleven amusic participants and eleven matched controls (all non-musicians) were recruited based on prior testing with the Montreal Battery of Evaluation of Amusia (MBEA, Peretz et al., 2003). Before the short-term memory task, all participants were tested with a two-alternative forced-choice task (using a staircase procedure) to determine their pitch discrimination thresholds (PDT, see Tillmann et al., 2009). Participants' short-term memory abilities for sequences of digits were evaluated using classic forward and backward digit span tests (Wechsler, 1997) to ensure that participants' performance did not differ for non-musical auditory material. For the main musical short-term memory test, we used a delayed matching-to-sample paradigm where participants had to determine whether tonal or atonal melodies (of five tones) presented in pairs were the same or different.

2. Results

2.1. Pre-tests

2.1.1. Montreal battery of Evaluation of amusia (MBEA)
All participants were tested with the MBEA (Peretz et al., 2003). To be considered as amusic, participants had to obtain an average score on the MBEA below the cut-off score (23 on

average across the six tasks of the battery, maximum score=30) that is two standard deviations below the average of the normal population. All amusics obtained scores at or below the cut-off score and all controls obtained scores higher than the cut-off score. The average scores of the amusic group (mean=21.15, SD=1.58; ranging from 18 to 23) differed significantly from the scores of the control group (mean=27.68, SD=.66; ranging from 26.5 to 28.6; $t(20)=12.58$, $p<.0001$).

2.1.2. Pitch discrimination threshold (PDT)

PDTs were determined using a two-alternative forced-choice task with an adaptive tracking, two-down/one-up staircase procedure (see Tillmann et al., 2009, for details). The average PDT of the amusic group (mean=1.32 semitones, SD=1.29; ranging from .13 to 4 semitones) was higher (worse) than that of the control group (mean=.25 semitones, SD=.27; ranging from .06 to .95 semitones, $t(20)=2.68$, $p=.01$). An overlap in the pitch thresholds between amusic and control groups was observed, in agreement with previous findings (Albouy et al., 2013; Foxton et al., 2004; Tillmann et al., 2009).

2.1.3. Digit span

Participants' short-term memory abilities for sequences of digits were investigated with the digit span tests from the Wechsler test (Wechsler, 1997), which measures the maximum number of items that participants can repeat after a single presentation, with the recall being either made in the order of presentation or backwards. For the forward task, controls' mean was 6.54 items (SD=.68) and amusics' mean was 6.36 items (SD=.80). For the backward task, controls' mean was 4.63 items (SD=.80) and amusics' mean was 4.54 items (SD=.93). Digit span data were analyzed with a 2 × 2 ANOVA with Group (amusics, controls) as between-participants factor and Task (forward, backward) as within-participant factor. A significant effect of Task ($F(1, 20)=82.40$, $p<.0001$, MSE=38.20) was observed: performance was lower for the Backward task than for the Forward task. There was no significant effect of Group ($F(1,20)=.23$, $p=.63$, MSE=.20) and no significant interaction between Group and Task ($F(1,20)=.82$, $p=.49$, MSE=.023). These findings thus suggest normal short-term memory abilities in amusics for non-musical material (see also Foxton et al., 2004; Marin et al., 2012; Omigie and Stewart, 2011; Stewart, 2011; Williamson and Stewart, 2010).

2.2. Short-term memory task: performance (d′ and c)

Performance in the short-term memory task was analyzed using signal detection theory by calculating, for each participant and for each condition (tonal or atonal), the discrimination sensitivity (d′) and the response bias (c[1]). For each participant, these analyses were based on hit rate (i.e., number of correct responses for different trials/number of different trials) and false alarm rate (i.e., number of incorrect responses for same trials/number of same trials). Positive

values for c arise when the miss rate (incorrect responses for different trials/number of different trials) exceeds the false alarm rate. Positive values thus indicate a tendency to answer "same", negative values indicate a tendency to answer "different", and c-values around 0 suggest the absence of a response bias. d′ and c were analyzed respectively with a 2 × 2 ANOVA[2] with Group (amusic, control) as the between-participants factor and Tonality (tonal, atonal) as the within-participant factor.

For d′ (Fig. 1A), the main effect of Group was significant, $F(1,20)=37.38$, $p<.0001$, MSE=1.1, with control participants reaching higher performance levels than did amusic participants (controls: mean=2.93, SD=.99; amusics: mean=1.00, SD=.81). The main effect of Tonality was also significant, $F(1,20)=8.00$, $p=.010$, MSE=.34, with higher performance levels for tonal sequences (mean=2.22; SD=1.51) than for atonal sequence (mean=1.22; SD=1.09). Finally, the Tonality-by-Group interaction was also significant, $F(1,20)=8.94$, $p=.007$, MSE=.34: planned comparisons revealed that only control participants showed the advantage in performance for tonal sequences (mean=3.45; SD=1.02) over atonal sequences (mean=2.42; SD=.65; $p=.001$) and not the amusic participants ($p=.91$; tonal sequences: mean=.99, SD=.63; atonal sequences: mean=1.01, $p=.98$). It is important to note that while amusics' performance was decreased in comparison to control participants' performance for both tonal ($F(1,20) = 45.46$; $p<.0001$) and atonal sequences ($F(1,20) = 15.40$; $p<. 0001$), their performance was still above chance level (d′=0) for both conditions (tonal: $t(10)=5.12$, $p<.001$; atonal: $t(10)=3.40$, $p=.007$).

To investigate whether the observed deficit in short-term memory performance was related to amusics' impaired PDT, we performed the same analysis by dividing the amusic group into two sub-groups based on their PDT, with one subgroup composed of the six amusic participants with a PDT in the control range (from .06 to .95 semitones), and the remaining five amusics with elevated PDT composing the other group (ranging from 1.3 to 4 semitones, see stars in Appendix Table 1, Fig. 1B). The d′ data were further analyzed with a 3 × 2 ANOVA with Group (amusics with elevated PDT, amusics with normal PDT, and controls) as the between-participants factor and Tonality (tonal, atonal) as the within-participant factor. The main effect of Group was significant, $F(2,19)=19.73$, $p<.0001$, MSE=42.76, with control participants reaching higher performance levels than did amusic participants of the two groups (amusics with normal PDT: mean=1.25, SD=.37, $p<.0001$; amusics with elevated PDT:= mean=.69, SD=.23, $p<.0001$; controls: mean=2.93, SD=.99). Note that no significant differences were observed between the two subgroups of amusics ($p=.22$). The main effect of Tonality failed to reach significance, $F(1,19)=3.76$, $p=.06$, MSE=1.17, but the interaction between group and tonality was significant, $F(2,19)=6.22$, $p=.008$, MSE=3.88: planned comparisons revealed that only control participants showed the advantage in performance for tonal sequences ($p<.0001$;

---

[1] The correction of d′ and c measures used .01 for cases without false alarms and .99 for the maximum number of hits.

[2] We tested if the samples (controls and amusics) were derived from a population normally distributed for the measures of interest using the Kolmogorov–Smirnov test. These analyses revealed that all data were normally distributed (all $ps>.10$).
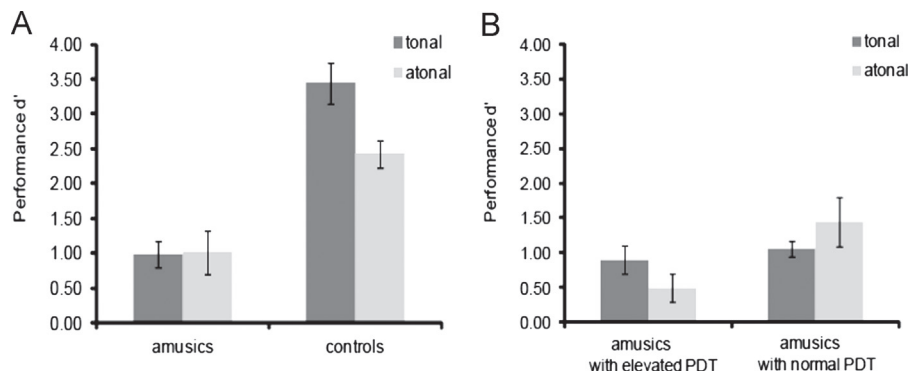
Fig. 1 – (A) Performance of amusic and control groups in terms of d', presented as a function of Tonality (tonal, atonal) and Group (amusics $n=11$; controls $n=11$). (B) Performance of amusic participants, separated in amusics with elevated PDT ($n=5$) and amusics with normal PDT ($n=6$), in terms of d', presented as a function of Tonality (tonal, atonal) and amusic sub-group. Error bars indicate the standard errors of the means; SEM.

mean$=3.45$; SD$=1.02$) over atonal sequences (mean$=2.42$; SD$=.65$; $p=.001$) and not the amusic participants ($p=.24$ for amusics with normal PDT; tonal sequences: mean$=1.05$, SD$=.65$; atonal sequences: mean$=1.44$, SD$=1.16$; $p=.26$ for amusics with elevated PDT; tonal sequences: mean$=.90$, SD$=.68$; atonal sequences: mean$=.49$, SD$=.37$). This additional result thus confirms that the impaired STM performance was not related to amusics' PDT. Moreover, the data pattern showed that only the amusic group with elevated PDTs showed a numerical tendency for increased d' values for tonal sequences in comparison to the atonal sequences.

For c (Fig. 2), we observed that the response bias was significantly superior to zero in each condition for both groups, revealing the participants' general tendency to respond "same" for tonal and atonal sequences (amusics, tonal: $t(10)=5.17$, $p<.001$, mean$=.81$, SD$=.51$; atonal: $t(10)=6.56$, $p<.001$, mean$=.72$, SD$=.36$; and controls, tonal: $t(10)=3.76$, $p=.004$, mean$=.39$, SD$=.34$; atonal: $t(10)=3.72$, $p=.004$, mean$=.64$, SD$=.57$). The 2×2 (Group $\times$ Tonality) ANOVA on c data did not reveal any significant main effect ($p=.14$ for Group and $p=.45$ for Tonality), and the Group-by-Tonality interaction failed to reach significance ($p=.12$)[3].

2.3. Short-term memory task: response times (RTs)

RTs were measured from the end of the second tone sequence of each pair, and thus reflected decision times for the memory task (see Appendix Table 2). For each participant, RTs that were slower than 3SD of the individual participant's mean were excluded (note that only few trials were excluded for each participant; amusics: mean$=.8$SD$=.12$; controls: mean$=1.7$SD$=.48$). For each group, one participant was excluded due to average RTs of 2SD above the group mean. To check that the exclusion of two participants did not change the results for d' and c reported above, we performed

<hr>

[3] For control participants only, the pattern of average bias data was similar to that observed for non-musician listeners in Schulze et al. (2012), revealing that control participants' positive response bias was decreased for tonal sequences compared to atonal sequences. However, the Group-by-Task interaction did not reach significance here ($p=.12$).
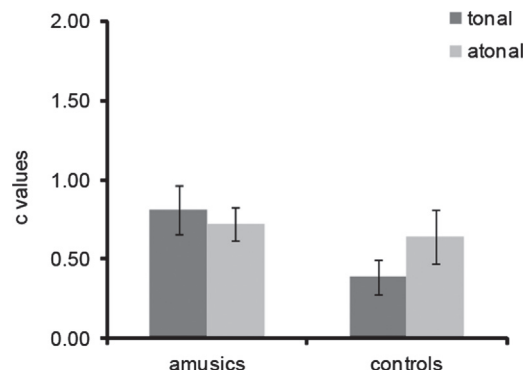

Fig. 2 – Bias data for amusic and control groups in terms of c (error bars indicating SEM), presented as a function of Tonality (tonal, atonal) and Group (amusics $n=11$; controls $n=11$). Note that all biases where positive thus revealing a general tendency to respond "same" in both groups.
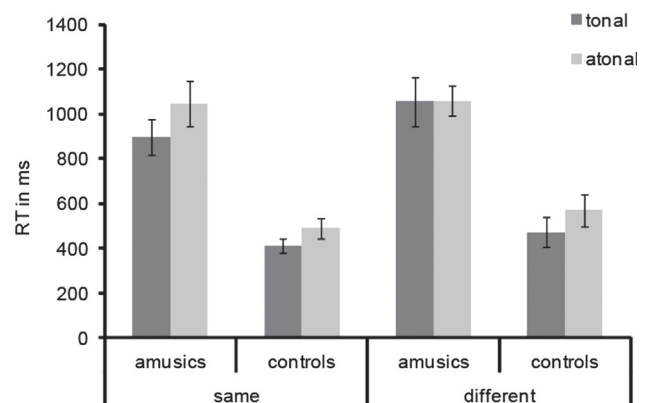
Fig. 3 – RTs (RTs) for all responses (correct and incorrect trials) in amusic and control groups in ms (error bars indicating the standard error of mean; SEM), presented as a function of Tonality (tonal, atonal), Trial type (same, different), and Group (amusics $n=10$; controls $n=10$).

the same analyses as above with the remaining 20 participants: these analyses revealed the same result pattern.

RTs were analyzed with a $2 \times 2 \times 2$ ANOVA with Group (amusics, controls) as the between-participants factor, and

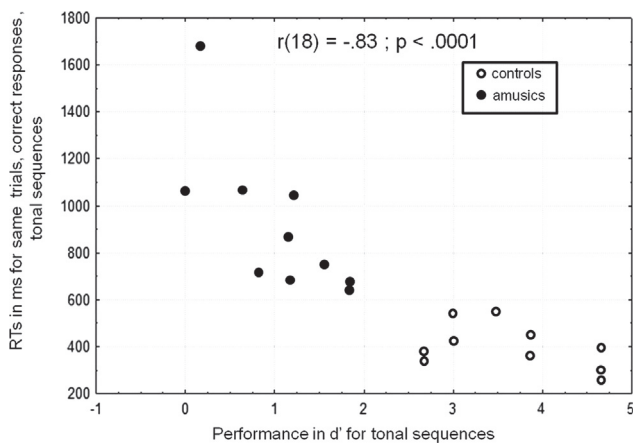**Fig. 4 – Scatter plot of RTs in ms for tonal sequences (same trials, correct responses) against performance (d') for tonal sequences. The correlation between these two variables is significant across all subjects and in the amusic group (see text for details). Note that without the outlier amusic participant (data point at $d' = .17$; RT $= 1677$ ms), this correlation remained significant in all participants ($r(17) = -.88$; $p < .0001$) and in amusics participants only ($r(7) = -.71$; $p = .03$).**

Tonality (tonal, atonal) and Trial type (same, different) as the within-participant factors. For RTs of all trials (i.e., including correct and incorrect responses; Fig. 3), the main effect of Group was significant, $F(1,18) = 33.73$, $p < .0001$, $MSE = 164,812$, revealing overall slower RTs for amusic participants (mean $= 1013$ ms, $SD = 253$) than for control participants (mean $= 485$ ms, $SD = 153$). In addition, the main effect of Trial type was significant, $F(1,18) = 4.93$, $p = .039$, $MSE = 24,479$, with slower RTs for different trials (mean $= 788$ ms, $SD = 64$) than for same trials (mean $= 710$ ms, $SD = 101$). Most interestingly, the main effect of Tonality was significant $F(1,18) = 5.27$, $p = .034$, $MSE = 25,858$, with slower RTs for atonal trials (mean $= 790$ ms, $SD = 71$) than for tonal trials (mean $= 708$ ms, $SD = 103$). None of the interactions were significant (all $p > .112$).

In a second step, we analyzed RTs for correct responses only. For controls, the error rate was 23.13% for different trials and 3.75% for same trials. However, the error rates of amusics were higher than for controls, with an error rate of 58.44% for the different trials and of 13.44% for same trials. This is in agreement with previous data showing that amusics mainly fail to detect differences rather than performing false alarms in short-term memory tasks (Tillmann et al., 2009). Because of the high error rates for different trials in the amusic group, the additional ANOVA on correct RTs was performed for same trials only. A 2×2 ANOVA (Group $\times$ Tonality) confirmed the significant main effect of Group, $F(1,18) = 25.76$, $p < .0001$, $MSE = 105,968$, and Tonality, $F(1,18) = 10.05$, $p = .005$, $MSE = 8569$, and the absence of a significant interaction, $F(1,18) = .03$, $p = .87$, $MSE = 8569$ (for controls, tonal sequences $= 398 \pm 94$ ms, atonal sequences $= 486 \pm 144$ ms; for amusics, tonal sequences $= 916 \pm 315$ ms, atonal sequences $= 1014 \pm 315$ ms). The effect of tonality observed for both groups thus

provided evidence for an advantage of tonal structure in amusics.

2.4. Correlation analyses

Correlation analyses were performed with eleven participants per group, except for correlations with RTs data for which two participants were excluded (one control and one amusic were excluded because of average RTs more than 2SD above their group mean). Two sets of analyses were performed: (1) correlations between the behavioral data in the short-term memory task (d', c, RTs for each condition, difference between tonal and atonal sequences for d', c, and RTs) and data from the pre-tests (MBEA, PDT, and digit span); and (2) correlation between the different measures (e.g., d', RTs) of the behavioral data in the short-term memory task. To explore differences explained not only by the group differences defined via the MBEA, we focus on correlations that were significant over all participants *and* for at least one of the two participant groups separately (amusic group or control group). For the first set of analyses, none of the correlations between data from the pre-tests and the behavioral data in the short-term memory task met these criteria, possibly due to the lack of variance within groups in the pre-tests (see Appendix Table 1).

For the second set of analyses, a negative correlation between d' and RTs was observed for tonal sequences only. For tonal sequences, increasing d' was observed with decreasing RTs for all responses (correct and incorrect, same and different trials). This correlation was significant in amusic participants ($r(8) = -.65$, $p = .044$) and across all participants ($r(18) = -.86$, $p < .0001$). This pattern of negative correlations for tonal sequences was also observed between d' and RTs for same trials only (correct and incorrect responses): it was significant in amusics: $r(8) = .-78$, $p = .007$, and across all participants $r(18) = .-88$, ($p < .0001$). Moreover, this pattern of negative correlations between d' and RTs for tonal sequences also extended when considering RTs of same trials for correct responses only (Fig. 4): amusics: $r(8) = .-74$, $p = .014$; controls $r(8) = .-43$, $p = .21$; all participants $r(18) = .-83$, $p < .0001$) In contrast, for atonal sequences, none of the correlations between RTs and d' were significant (all $p > .44$).

3. Discussion

The present study investigated the effect of tonal structure on short-term memory performance in congenital amusics and non-musician matched controls. More specifically, we explored whether amusics' previously reported implicit knowledge about tonal structure (Gosselin et al., 2011; Omigie et al., 2012; Paquette et al., 2011; Tillmann et al., 2012a,b) could improve short-term memory performance for tone sequences, as previously shown in non-musician listeners by Schulze et al. (2012).

For the non-musician control participants of the present study, short-term memory performance (d') differed for structured (tonal) compared to unstructured (atonal) sequences. This is in agreement with the findings of Schulze et al. (2012), and confirms that the tonal structure helps non-musician

listeners to improve encoding and/or maintaining of pitch information.

In comparison to controls, amusic participants were impaired in the short-term memory task for both tonal and atonal sequences, and the presence of tonal structures did not benefit their short-term memory recognition. This finding is in agreement with previous evidence for amusics' short-term memory deficit for tonal material (Albouy et al., 2013; Gosselin et al., 2009; Tillmann et al., 2009; Williamson et al., 2010, Williamson and Stewart, 2010). In addition, the observation of normal short-term memory abilities in amusics for sequences of digits provides further evidence for the specificity of the deficit to the musical domain (thus supporting the findings of Tillmann et al., 2009, who showed normal short-term memory performance in amusics for sequences of words; see also Williamson and Stewart 2010). As in previous studies (Albouy et al., 2013; Foxton et al., 2004; Tillmann et al., 2009), several amusics' pitch discrimination thresholds (PDT) were comparable to controls' PDT, thus excluding pitch discrimination deficits as the sole origin of the impaired short-term memory performance.

Analysis of the d' results suggested that tonality did not help amusic participants to overcome their short-term memory deficit for melodies. It is relevant to note that previous studies have mostly used tonal material to bring out short-term memory deficits in amusics (Gosselin et al., 2009; Tillmann et al., 2009). Based on this observation, it could be argued that group differences observed in these studies could be linked to the fact that control participants benefit from the presence of tonal structures (based on their "experience/expertise with music"), whereas the amusics do not. However, our study (see also Foxton et al., 2004) demonstrated that even with unstructured material, amusics showed poorer performance in comparison to controls, thus providing new evidence supporting the hypothesis that the core deficit in amusia is related to memory processing of pitch material in general.

However, even though tonality did not improve amusics' short-term memory performance (no benefit in d' and c) in comparison to atonal sequences, some potential benefit of tonality was observed in the RTs data, suggesting that amusics processed the tonal information to some extent. For both amusics and controls, RTs were faster for tonal sequences than for atonal sequences. This observation received further support from the correlation analyses that revealed a negative correlation between performance (d') and RTs only for tonal sequences (i.e., faster responses for higher d') and not for atonal sequences. This type of negative correlation was significant both for all participants and for amusics alone. The observation of this correlation for tonal sequences only can be interpreted as an additional marker of the potential benefit of tonality for amusics. The present findings thus support the hypothesis that amusics process the tonality of the material to some degree, which then influences short-term memory performance for the tone sequences (as revealed in RTs). It remains unclear, however, whether this effect of tonality is specifically related to an improvement of short-term memory abilities. This benefit of tonality could also be a marker of a facilitatory effect for more global cognitive processes, such as encoding or decision

making. Nevertheless, this tonal benefit observed for melodies in controls (and the potential benefit observed in amusics) can be related to results reported by Dowling (1991), suggesting that the various components of a melody (interval, contour, and tonality) are encoded as an integrated whole. This might lead to a stronger memory trace for tonal melodies than for atonal melodies (as observed in Schulze et al., 2012), with atonal melodies missing the benefit of the integration in a tonal framework. This hypothesis thus further predicts that amusics' short-term memory abilities might be improved for more complex tonal material (e.g., chord sequences or harmonized melodies, in contrast to single melodic lines as used in the present study). In a related vein, Tillmann et al. (2012b) used musical pieces (containing melodic and harmonic information) to reveal normal categorization abilities of tonal and atonal pieces in amusics. Using similarly rich material, increased short-term memory accuracy for tonal sequences might also be observed in amusics.

In the present study, differences in the processing of tonal and atonal materials in amusics were revealed in the RTs data, which provides an indirect measure of participants' processing in the memory task. The observation of differences between explicit and implicit processing (as revealed by direct measures (here, accuracy) and indirect measures (here, RTs), respectively) does not imply that different knowledge representations are involved as these differences might be "in degree rather than in kind" (Cleeremans and Jiménez, 2002; Destrebecqz and Cleeremans, 2003). Along these lines, for a case of acquired amusia, Tillmann et al. (2007) proposed that tonal knowledge representations can be conceived as graded representations, with implicit knowledge referring to a representation that is less (or not) available to consciousness, while still influencing perception and behavior.

The present results show that amusics can process tonal material to some degree and provide converging evidence with recent data using indirect measures to show that amusic individuals have acquired some structural knowledge of the Western tonal system (Omigie and Stewart, 2011; Tillmann et al., 2012a) and can process tonality so as to differentiate tonal and atonal structures (Tillmann et al., 2012b) and to distinguish major and minor mode melodies (Gosselin et al., 2011; Paquette et al., 2011). These data can also be related to the unconscious pitch tracking capacities of the amusic brain revealed by EEG (Peretz et al., 2009), showing amusics' near-normal ability to detect pitch irregularities (mistuning) in a musical context, even though amusic individuals cannot detect these irregularities explicitly (see also Moreau et al., 2013). These studies have implications for the development of training and rehabilitation programs of impaired musical functions, notably by encouraging strategies that tap into spared implicit processing resources.

4.   Conclusion

In the present study, we investigated whether previously reported implicit knowledge about tonal structures in amusia (Gosselin et al., 2011; Paquette et al., 2011; Tillmann et al., 2012a,b) could facilitate short-term memory processing.

A benefit was observed in the response time data: both amusics and controls processed the tonal material faster than the atonal material. However tonality facilitated the recognition accuracy only in controls. This data suggests that some implicit processing of tonal structures is preserved in congenital amusia, supporting the perspective to exploit implicit knowledge to help reducing pitch perception and memory deficits in congenital amusia.

5. Experimental procedures

5.1. Participants

Eleven amusic adults and eleven non-musician controls matched for gender, age, handedness, educational background (calculated in years of education starting from the first year of primary school in the French system (i.e., children are then approximately 6 years old)), and musical training (calculated in years) participated in the study. Both the amusic group (seven women; mean age=36.90, SD=10.98 years, ranging from 20 to 54; educational background=14.90, SD=1.44 years; musical training=.95, SD=1.73 years) and the control group (seven women; mean age= 36.18, SD=10.10 years, ranging from 24 to 51; mean education=15, SD=2.60 years; musical education=.36, SD=.80 years) were composed of ten right-handed participants and one left-handed participant. Moderate (25 dB loss at any frequency) or severe (more than 35 dB loss at any frequency) peripheral hearing loss was excluded using standard audiometry, and all participants reported no history of neurological or psychiatric disease. Participants gave their written informed consent, and were paid for their participation. Before performing the main short-term memory task, all participants were tested with the MBEA and with a two-alternative forced-choice task (using a staircase procedure) to evaluate their pitch discrimination thresholds (PDT, see Tillmann et al., 2009).

5.2. Digit span

Participants' short-term memory abilities for non-musical auditory material were estimated using classic forward and backward digit span tests (Wechsler, 1997). These tests evaluate the extent to which an individual can repeat in forward and backward orders a series of numbers, which is presented in increasing length from two to eight digits. Series of digits were presented vocally to the participants by the experimenter with approximately 1 s delay between each element of the series, and the participant was asked to repeat the series in either forward or backward order. Each sequence length was tested twice (but with a different list). If the participant succeeded for a given length, she/he was given a longer list (increased by the addition of one item). The test was stopped following two errors for a given series' length, and the digit span score was defined as the length of the longest list that was remembered correctly. Forward digit span was always tested first and was then followed by the testing of the backward digit span.

5.3. Material for the short-term memory task

Stimuli were the five-tone sequences from Schulze et al. (2012), comprising 16 tonal and 16 atonal sequences: the tonally-structured sequences were created by using a set of tones from the C major scale in equal temperament: C4 (262 Hz), D4, E4, F4, G4, and A4 (440 Hz). All tones belonged to one tonality (C major), and their progression respected musical structures as defined by the Western tonal system. The atonal sequences were created by using another set of eight tones: B3 (247 Hz), C4, C#4, E4, F4, F#4, G4, and G#4 (415 Hz). These tones did not belong to a single tonality and did not have an obvious tonal structure, so that a tonal interpretation in a particular key would be relatively difficult or impossible. The key-finding algorithm proposed by Krumhansl and Schmuckler (cited in Krumhansl, 1990) was used to analyze the strength of tonal centers established by the melodic contexts in both sets of melodies (see Schulze et al., 2012, for details). The probability of occurrence of each note (for tonal and atonal sequences separately) was correlated with the tone profiles of the 12 major and 12 minor keys, respectively (these tone profiles resulted from subjective judgments of listeners on how well a probe tone fits with a preceding tonal context; Krumhansl and Kessler, 1982). The maximum positive correlation provides an indication of the most strongly established key, and the maximum negative correlation indicates the least likely key. For our material, the maximum correlation for the tonal sequences was, as expected, with the C major key: $r(10)=.85$; $p<.001$. The correlations for atonal sequences with all major or minor keys were all inferior to .18 (ns).

The tonally structured sequences and the atonal sequences were matched for various parameters: melodic contour, frequency of occurrence of the tones, range between highest and lowest tones. The stimuli were created with the software Digital performer, Cubase 5.1 (Steinberg) and a Halion Sampler (Steinberg) using an acoustic piano timbre. For both tonal and atonal sequences, each tone had a duration of 500 ms, presented with an inter-stimulus interval of 20 ms. Sequences were presented in pairs, and the sequences of a pair could be either the same or different. For the different pairs, two non-adjacent tones were exchanged so that the melodic contour was preserved (e.g., C A F E G–C A **G** E **F**). The first tone of the second sequence was never changed. Presentation software (Neurobehavioral Systems, Albany, CA, USA) was used to present the stimuli and to record participants' responses.

5.4. Procedure of the short-term memory task

Participants listened to pairs of tone sequences presented with two loudspeakers placed in front of them in a dimly-lit and sound-attenuated testing booth. For each trial, a first sequence (e.g., F G D E C) was presented, and after 3 s of silence, a second sequence was presented with all tones being either in the same order or not (e.g., F G **C** E **D**). Participants were instructed to compare the two sequences and to press one of two mouse buttons to indicate whether the two sequences were the same or different (with "same" being defined as all tones played correctly in the same order).

At the beginning of the experiment, the task was explained with one same trial and one different trial. Error feedback was given only for these practice trials.

The experiment consisted of 64 trials, each melody (16 tonal, 16 atonal) was used as S1 once in a same trial and once in a different trial. A pseudo-randomized presentation was used so that: (1) the same tone sequence was not presented consecutively in a "same" and a "different" trial, and (2) the type of trial (same/different) changed after at most 3 trials (i.e., no more than three consecutive "same" or "different" trials). To choose the order of the sequences, eight different orders were created and presented randomly across participants using a Latin square design. The experiment lasted approximately 15 min, excluding the pre-tests (MBEA, PDT, and digit span) which took place during previous testing sessions for the MBEA and PDT, and on the day of testing for the digit span.

Acknowledgments

Appendix A.   Supplementary materials

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.brainres.2013.09.003.

R E F E R E N C E S

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.

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

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

Bor, D., Duncan, J., Wiseman, R.J., Owen, A.M., 2003. Encoding strategies dissociate prefrontal activity from working memory demand. Neuron 37, 361–367.

Cleeremans, A., Jiménez, L. (Eds.), 2002. Implicit Learning and Consciousness. Psychology Press, Hove, UK.

Destrebecqz, A., Cleeremans, A., 2003. Temporal effects in sequence learning. In: Benjamins (Ed.), Attention and Implicit Learning. Publishing Company, Amsterdam, pp. 181–213.

Dowling, W.J., 1991. Tonal strength and melody recognition after long and short delays. Percept. Psychophys. 50, 305–313.

Foxton, J., Dean, J., Gee, R., Peretz, I., Griffiths, T., 2004. Characterization of deficits in pitch perception underlying 'tone deafness'. Brain 127, 801–810.

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

Gosselin, N., Trainor, L., Belin, P., Peretz, I., 2011. Emotion recognition is spared in amusia. In: Proceedings of the Neurosciences and Music-IV Learning and Memory Meeting, Edinburgh, UK.

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

Krumhansl, C.L., 1990. Cognitive Foundations of Musical Pitch. Oxford University Press, Oxford, UK.

Krumhansl, C.L., Kessler, E.J., 1982. Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. Psychol. Rev. 89, 334–368.

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.

McDonald, C., Stewart, L., 2008. Uses and functions of music in congenital amusia. Music Percept. 25, 345–355.

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

Omigie, D., Pearce, M.T., Stewart, L., 2012. Tracking of pitch probabilities in congenital amusia. Neuropsychologia 50, 1483–1493.

Omigie, D., Stewart, L., 2011. Preserved statistical learning of tonal and linguistic material in congenital amusia. Front. Psychol. 2, 109.

Paquette, S., Gosselin, N., Peretz, I., 2011. Normal use of mode and tempo in emotional judgments by amusic adults. In: Proceedings of the Neurosciences and Music-IV Learning and Memory Meeting, Edinburgh, UK.

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., Hyde, K.L., 2003. What is specific to music processing? Insights from congenital amusia. Trends Cognit. Sci. 7, 362–367.

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

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

Savage, C.R., Deckersbach, T., Heckers, S., Wagner, A.D., Schacter, D.L., Alpert, N.M., Fischman, A.J., Rauch, S.L., 2001. Prefrontal regions supporting spontaneous and directed application of verbal learning strategies: evidence from PET. Brain 124, 219–231.

Schulze, K., Dowling, W.J., Tillmann, B., 2012. Working memory for tonal and atonal sequences during a forward and a backward recognition task. Music Percept. 29, 255–267.

Stewart, L., 2006. Congenital amusia. Curr. Biol. 16, R904–R906.

Stewart, L., 2008. Fractionating the musical mind: insights from congenital amusia. Curr. Opin. Neurobiol. 18, 127–130.

Stewart, L., 2011. Characterizing congenital amusia. Q. J. Exp. Psychol. 64, 625–638.

Tillmann, B., Bharucha, J.J., Bigand, E., 2000. Implicit learning of tonality: a self organizing approach. Psychol. Rev. 107, 885–913.

Tillmann, B., Peretz, I., Bigand, E., Gosselin, N., 2007. Harmonic priming in an amusic patient: the power of implicit tasks. Cognit. Neuropsychol. 24, 603–622.

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

Tillmann, B.,Gosselin, N.,Bigand, E.,Peretz, I., 2012a. Priming paradigm reveals harmonic structure processing in congenital amusia. Cortex 48, 1073–1078.

Tillmann, B., Albouy, P., Lalitte, P., Caclin, A., Bigand, E., 2012b. Perception of tonal and atonal musical excerpts in congenital amusia.. In : 12th International Conference on Music

Perception and Cognition (ICMPC) - 8th Triennial Conference of the European Society for the Cognitive Sciences of Music (ESCOM), Thessaloniki, Greece.

Wechsler, D.S., 1997. Wechsler Adult Intelligence Scale-III (WAIS-III). The Psychological Cooperation, San Antonio.

Williamson, V.J., McDonald, C., Deutsch, D., Griffiths, T.D., Stewart, L., 2010. Faster decline of pitch memory over time in congenital amusia. Adv. Cognit. Psychol. 6, 15–22.

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

Table 1: Amusic individuals and controls' individual data for the pre-tests and for the main short-term memory task. <u>Pre-tests:</u> MBEA: average score of the six subtests of the Montreal Battery of Evaluation of Amusia (max 30); PDT: Pitch Discrimination Threshold (in semitones); Digit span: scores for forward and backward tests (max 8 items). <u>Main short-term memory task</u>: d', accuracy; c, bias (positive values indicate a tendency to answer "same", negative values indicate a tendency to answer "different"), and percent of correct responses as function of Trial Type (Same, Different), presented separately for tonal and atonal sequences.

	Pre-tests		Digit span		Main short-term memory task				% correct responses : Tonal sequences		% correct responses : Atonal sequences	
	MBEA	PDT	forward	backward	d' tonal	d' atonal	c tonal	c atonal	same	different	same	different
A1	20.67	0.2	7	4	0.38	0.26	1.34	1.02	93.75	12.5	87.5	18.75
A2	20.83	4	6	4	0.83	0.36	0.73	0.5	87.5	37.5	75	37.5
A3	21.5	0.14	6	4	1.56	1.64	0.11	0.33	81.25	75	87.5	68.75
A4	22	1.3	8	5	1.85	0.89	0.61	0.44	93.75	62.5	81.25	50
A5	21	0.75	6	7	1.22	0.4	0.93	0.69	93.75	37.5	81.25	31.25
A6	18.83	2.6	6	4	0.65	0	1.21	0.16	93.75	18.75	56.25	43.75
A7	18	0.13	7	4	1.16	2.82	0.09	0.92	75	68.75	100	68.75
A8	21.33	2.41	7	5	1.18	0.38	1.74	1.34	100	12.5	93.75	12.5
A9	23.17	2.18	6	5	0	0.86	1.15	1.1	87.5	12.5	93.75	25
A10	22.5	0.25	5	4	1.85	2.82	0.61	0.92	93.75	62.5	93.75	68.75
A11	22.83	0.64	6	4	0.17	0.73	0.4	0.52	68.75	37.5	81.25	43.75
C1	28.67	0.67	7	5	3.86	2.64	0.4	1	100	93.75	100	62.5
C2	28.17	0.17	7	5	3	2.82	0.83	0.92	100	75	100	68.75
C3	27.83	0.25	6	4	3.86	2.64	0.4	1	100	93.75	100	62.5
C4	27	0.07	7	4	4.65	2.48	0	1.08	100	100	100	56.25
C5	28.17	0.08	7	5	3	2.33	0.83	1.16	100	75	100	50
C6	26.83	0.15	6	5	2.68	2.21	0.19	-0.43	93.75	87.5	73.33	93.75
C7	26.5	0.11	7	5	4.65	2.02	0	0.52	100	100	93.75	68.75
C8	28.17	0.21	5	3	3.48	2.48	0.59	1.08	100	87.5	100	56.25
C9	27.83	0.1	7	6	4.65	3.86	0	-0.4	100	100	93.75	100
C10	27.5	0.9	7	4	1.38	1.15	0.85	0.58	93.75	43.75	87.5	50
C11	27.83	0.12	6	5	2.68	2.02	0.19	0.52	93.75	87.5	93.75	68.75

Table 2: Individual RTs data for amusic individuals and controls for all trials and for correct responses presented as a function of Tonality (tonal. atonal) and Trial Type (same. different). Note that A1 and C10 were excluded from the analysis because of average response times of 2 SD above the group mean.

| | RTs for all responses | | | | RTs for correct responses | |
| | tonal | | atonal | | tonal | atonal |
	tonal	different	same	different	same	same
A2	704.94	1061.35	955.81	904.08	714.26	932.85
A3	703.23	830.99	1148.91	936.45	746.37	1014.14
A4	686.22	969.19	579.01	994.47	646.55	564.68
A5	1021.40	1323.47	1460.38	1301.16	1040.50	1377.00
A6	1021.06	656.53	1400.97	1204.98	1062.23	1171.97
A7	930.79	860.24	1111.13	1031.29	867.38	1111.13
A8	681.19	1084.12	704.81	839.96	681.19	704.81
A9	1104.92	1065.57	925.51	1044.71	1058.88	933.27
A10	669.91	811.04	733.42	839.65	669.91	733.42
A11	1426.56	1883.32	1429.41	1478.71	1677.10	1597.65
C1	358.69	642.49	431.96	469.29	358.69	431.96
C2	539.05	793.03	773.80	1012.17	539.05	773.80
C3	448.00	566.35	425.86	571.43	448.00	425.86
C4	394.05	380.83	636.85	620.47	394.05	636.85
C5	424.38	683.78	524.39	715.54	424.38	524.39
C6	378.02	402.24	465.24	388.39	339.89	416.93
C7	257.33	179.33	298.89	334.41	257.33	310.67
C8	549.74	566.23	591.64	650.31	549.74	591.64
C9	300.01	200.39	331.92	213.25	300.01	331.92
C11	469.68	296.52	423.29	722.70	376.57	423.29

Chapter VIII Boosting pitch encoding with audiovisual interactions in congenital amusia

Objectives

Results of Study 3 (see also Study 1) have suggested that pitch processing deficits in congenital amusia could be mediated by strong impairments in the construction of memory traces for short tones. This hypothesis was supported by the observation of increased performance in amusic individuals in both pitch discrimination and short-term memory tasks for tone sequences when the time dedicated to encode pitch information was long enough. According to these results, we have thus proposed that pitch-related deficits in congenital amusia could be mediated by impaired encoding of pitch information for short tones (Study 1, Study 3) along with impairments in the short-term retention of pitch information (Study 1, Study 2, Study 4, Tillmann et al., 2013).

Along this line we aimed to explore how and whether congenital amusic individuals' pitch encoding abilities could be boosted by using a new approach in the domain that consists in facilitating amusic individuals' pitch encoding with multisensory integration. In Study 5, we thus explored whether the first processing steps could be boosted by presenting simultaneous uninformative visual information during the encoding of the pitch information This study was based on previous studies showing that the combination of sensory information across senses can enhance perceptual abilities (Ernst and Bulthoff, 2004; Stein and Meredith, 1993).

Article 5
Boosting pitch encoding with audiovisual interaction in congenital amusia

In preparation

Boosting pitch encoding with audiovisual interaction in congenital amusia

Authors: Philippe Albouy [1, 2, 3,], Yohana Lévêque[1, 2, 3], Krista L. Hyde [4], Patrick Bouchet [1, 3], Barbara Tillmann [2, 3] and Anne Caclin [1, 3]

Affiliations:

[1] Lyon Neuroscience Research Center, Brain Dynamics and Cognition Team, CRNL, INSERM U1028, CNRS UMR5292, Lyon, F-69000, France

[2] Lyon Neuroscience Research Center, Auditory Cognition and Psychoacoustics Team, CRNL, INSERM U1028, CNRS UMR5292, Lyon, F-69000, France

[3] University Lyon 1, Lyon, F-69000, France;

[4] Montreal Neurological Institute, McGill University, 3801 University Street, Montreal, QC H3A 2B4 Canada

Corresponding author: Philippe Albouy, Montreal Neurological Institute, McGill University, 3801 University Street, Montreal, QC H3A 2B4 Canada. Phone: XXX Fax: XXX E-mail: phv.albouy@gmail.com

Abstract

Previous research has shown that multisensory facilitatory interaction is maximally effective when responses to unisensory modalities are weak. In congenital amusia, deficits in music perception are currently understood as being mediated by the impaired encoding of pitch. The present study investigated if audiovisual facilitation could be observed in congenital amusia and boost amusic individuals' pitch encoding abilities in a pitch change detection task. Amusic individuals and their matched controls, who were non-musicians, performed two Tasks. In Task 1, they were required to detect auditory, visual or audiovisual stimuli as rapidly as possible. In Task 2, they were required to detect (as accurately and as rapidly as possible) a pitch change (with different pitch interval sizes) within a sequence of five identical tones that was presented either only auditorily (A condition), or with simultaneous uninformative visual signal presented on the screen (AV condition). Results of Task 1 revealed that, as previously reported for typical listeners, both amusic and control participants exhibited audiovisual facilitation (shorter response times for audiovisual stimuli than for either auditory or visual stimuli). Most interestingly, no differences were observed between the two participant groups in terms of bimodal facilitation, thus showing for the first time that amusic individuals have preserved visual and audiovisual integration capacities. Results of Task 2 revealed that for the pitch detection task, both amusic individuals and controls benefited from simultaneous visual information in terms of accuracy (better performance for the AV condition in comparison to the A condition). Moreover, participants of both groups exhibited shorter response times in the AV condition than in the A condition. These audiovisual benefits in terms of response times were observed for different ranges of difficulty for each group (that is, for different pitch interval sizes of the changed tone). These findings suggest that amusic individuals can benefit from multisensory integration to improve their pitch processing abilities and that this benefit varied as a function of the range of difficulty. This study constitutes the first step towards the perspective to exploit multisensory paradigms to help reducing pitch-related deficits in congenital amusia. In addition this study demonstrated for the first time that uninformative visual stimulation can boost pitch encoding in typical listeners, thus providing a new evidence of facilitative effect of audiovisual integration upon pitch perception.

1. Introduction

It is well established that the combination of sensory information across senses can modify the qualitative perceptual experience (Calvert and Thesen, 2004; Ernst and Bulthoff, 2004; Jousmaki and Forss, 1998; McGurk and MacDonald, 1976; Shams et al., 2000; Stein and Meredith, 1993; von Kriegstein, 2012), even by disturbing (Bertelson and Radeau, 1981) or by enhancing perception (von Kriegstein, 2012). For the latter point, benefits of multisensory integration upon perception were observed in various processes, from very simple mechanisms (such as speeding reaction times (Gielen et al., 1983; Hershenson, 1962; Posner et al., 1976)) to more complex processes (such as facilitating learning (Seitz et al., 2006)). In the specific case of audiovisual interaction, the simultaneous presentation of auditory information has been shown to improve visual performance, such as discriminating briefly flashed visual patterns (Vroomen and de Gelder, 2000), detecting a dimly flashed light (Frassinetti et al., 2002a; Frassinetti et al., 2002b; Teder-Salejarvi et al., 2005) or boosting visual detection thresholds (Caclin et al., 2011). In a related vein, it has been shown that congruent visual stimuli could improve auditory performance (von Kriegstein, 2012), by improving speech intelligibility in the presence of other distracting sounds (Ross et al., 2007; Sumby and Pollack, 1954; von Kriegstein, 2012) or by improving sound localization (Shelton and Searle, 1980).

Interestingly, recent research has shown that these intersensory effects depend on the participants' performance when the modalities are presented in isolation. For example, an improvement of visual localization performance by concurrent auditory stimulation has been observed only under conditions of induced myopia (Hairston et al., 2003). Similarly, it has been proposed that cochlear-implanted deaf participants integrate audiovisual information better than normal-hearing participants (Rouger et al., 2007). Recently, Caclin et al. (2011) observed an audiovisual benefit upon visual detection thresholds only in myopic participants with a poor visual-only performance. These effects were associated to the *inverse effectiveness principle* of multisensory integration (Stein and Meredith, 1993), which states that such interactions are maximally effective when the responses to the unimodal constituents are weak. Further supports for this hypothesis could be found in the study by Giard and Peronnet (1999) in which the behavioral gain in processing bimodal (relative to unimodal)

stimuli was associated with increased activity (neural interactions) in the cortex of the weaker unisensory modality (Giard and Peronnet, 1999).

Based on these principles, the present study aimed to explore the effectiveness of facilitatory audiovisual interaction in a condition of impaired auditory processing that has been referred to as congenital amusia (Peretz, 2013; Peretz and Hyde, 2003; Stewart, 2011; Tillmann et al., 2013). Congenital amusia is a lifelong disorder of music perception and production. Amusic individuals are unable to recognize a familiar tune (without the help of lyrics) or to detect that someone (including themselves) sings out of tune. This disorder cannot be explained by peripheral hearing loss, brain lesions, or general cognitive impairments (Ayotte et al., 2002; Peretz, 2003; Peretz et al., 2002; Peretz and Hyde, 2003). Hyde and Peretz (2004) have reported pitch discrimination difficulties in simple tone sequences in congenital amusia for pitch changes smaller than 1/2 tone, whereas controls could detect changes of a 1/8 of a tone (Hyde and Peretz, 2004). Converging evidence for a deficit in fine-grained pitch processing was obtained with an independent sample of amusic participants and with various psychoacoustic approaches (Foxton et al., 2004b). These studies led to the assumption that congenital amusia arises from a failure to extract pitch information with sufficient accuracy. This lack of precision in the first processing steps of auditory perception is bound to affect tonal coding and pitch-related processing in general, in turn leading to higher-level deficits, such as altered pitch memory abilities (Peretz, 2013; Stewart, 2011; Tillmann et al., 2013).

Interestingly, recent findings went beyond this view notably by proposing that pitch perception impairments in this disorder could be mediated by altered construction of pitch memory traces (Albouy et al., *in preparation;* Albouy et al., 2013a; Tillmann et al., 2013). This hypothesis was supported by the observation of increased performance in amusic individuals in both pitch discrimination and short-term memory tasks for tone sequences when the time dedicated to encode pitch information was sufficiently long. According to these results, and based on previously reported functional abnormalities in the amusic brain (Albouy et al., 2013a; Hyde et al., 2011), pitch-related deficits in congenital amusia could be mediated by impaired auditory encoding of pitch information (Albouy et al., *in preparation;* Tillmann et al., 2013).

In the present study, we explored for the first time the potential effectiveness of multisensory interactions in improving congenital amusic individuals' abilities in the auditory encoding of pitch information. We hypothesize that appropriate audiovisual information might boost pitch perception in individuals suffering from this lifelong disorder. Indeed, audiovisual facilitative interactions involving pitch have already been reported in the literature. For example, auditory pitch and visual vertical positions have been described as synthetically corresponding dimensions (Melara and O'Brien, 1987), which means that judging pitch height is easier when low and high tones are paired with visual stimuli of corresponding lower or higher positions along a vertical axis (Lidji et al., 2007; Rusconi et al., 2006). However, it is also possible to obtain audiovisual facilitation even with uninformative stimuli, particularly for participants with impaired performance in the modality of interest: For example, Caclin et al. (2011) observed that uninformative sounds (i.e., sounds conveying no information about the answer to be given in the visual task) can improve visual detection thresholds in myopic participants with poor visual-only performance, – a facilitation effect that was not observed in normal-seeing participants and myopic participants with good visual-only performance. Along these lines, the main objective of the present study was to assess whether audiovisual facilitatory interactions (improving auditory performance with uninformative visual stimulation) could boost amusic individuals' performance in a pitch change detection task. It is relevant to note that in addition to investigate for the first time the potential benefit of multisensory integration in congenital amusia, this study also examines the potential benefit of uninformative visual stimulation upon pitch perception in typical listeners (our non-musician matched controls), question which, to our knowledge, has not been investigated so far. Finally, in addition to the investigation of pitch processing in relation to multi sensory integration, and as a preliminary step to guide the interpretation of the data, we also assessed for the first time whether audiovisual integration capacities in a simple detection task were preserved or altered in congenital amusic individuals as compared with controls.

Participants performed two tasks. In Task 1, we tested amusic individuals' and controls' abilities in detecting auditory (A), visual (V), and Audiovisual (AV) stimuli with a speeded detection paradigm. In Task 2, we tested the effect of uninformative visual stimuli upon pitch change detection within five tone sequences, in amusic individuals and matched controls. The auditory task was adapted

from Hyde and Peretz (2004)'s study of pitch change detection in congenital amusia (see above). The focus of interest is the facilitatory influence of uninformative visual stimuli on auditory pitch processing for the amusic individuals (impaired in the auditory modality, that is on pitch processing), thus representing the mirror situation of previous work from our group in visually-impaired participants (Caclin et al., 2011).

2. Material and Methods

2.1. Participants

Sixteen amusic adults and sixteen non-musician controls matched for gender, age, educational background (calculated in years of education starting from the first year of primary school in the French system, about 6 years old), and musical training (calculated in years) participated in the study. The amusic group was composed of fourteen right-handed participants and two left-handed participants and the control group was composed of fifteen right-handed participants and one left-handed participant. Moderate (25 dB losses at any frequency) or severe (more than 35 dB loss at any frequency) peripheral hearing losses were excluded using standard audiometry, and all participants reported no history of neurological or psychiatric disease. Participants gave their written informed consent, and were paid for their participation. In a previous testing session, all participants were tested with the Montreal Battery of Evaluation of Amusia (Peretz et al., 2003) and with a two-alternative forced-choice task (using a staircase procedure) to evaluate their pitch discrimination thresholds (Tillmann et al., 2009). Participants' demographic characteristics and data from the pre-tests are presented in Table 1.

2.2. Montreal Battery of Evaluation of Amusia (MBEA)

All participants were tested with the MBEA (Peretz et al., 2003). To be considered as amusic, they had to obtain an average score on the MBEA below the cut-off score (23.4 on average across the six tasks of the battery, maximum score = 30), the cut-off being two standard deviations below the average of the normal population (Peretz et al., 2003). All amusic individuals obtained scores below the cut-off score and all controls obtained scores higher than the cut-off score. The average scores of the amusic group (ranging from 18 to 23) differed significantly from the scores of the control group (ranging from 23.83 to 28.83).

2.3. Pitch Discrimination Threshold (PDT)

PDTs were determined using a two-alternative forced-choice task with an adaptive tracking, two-down/one-up staircase procedure (see Tillmann et al., 2009, for details). The average PDT of the

amusic group (ranging from 0.13 to 4 semitones) was higher (worse) than that of the control group (ranging from 0.05 to 0.95 semitones). In agreement with previous findings, we observed an overlap in pitch thresholds between amusic and control groups, (Albouy et al., 2013a; Foxton et al., 2004b; Jones et al., 2009b; Liu et al., 2010; Tillmann et al., 2013; Tillmann et al., 2009).

Table 1: Demographic characteristics of participants and data in behavioral pretests. Results on the Montreal Battery of Evaluation of Amusia (MBEA) are expressed as number of correct responses (average over the six sub-tests of the battery). Pitch Discrimination Threshold (PDT) scores are reported semitones. Data is reported as a function of group, significance levels on corresponding t-tests; ''NS'' refers to a non-significant difference (p> .05). Standard deviations are in parentheses.

Characteristics	Controls (n=16)	Amusics (n=16)	t-Test
Age in years	34.75 (10.69)	34.87 (12.49)	*NS*
Gender	10 female, 6 male	10 female, 6 male	
Education in years	14.65 (2.41)	13.87 (2.06)	*NS*
Musical education in years	0.56 (1.36)	0.71 (1.48)	*NS*
MBEA (Peretz et al., 2003)			
Mean score (cut off score = 23.4)	26.75 (1.08)	21.17 (1.51)	*t(30) = 12.32, p < .0001*
Pitch Discrimination Threshold (Tillmann et al., 2009)			
Threshold in semitones	0.24 (0.22)	1.12 (1.11)	*t(30) = 3.06, p = .004*

2.4.   Experimental set-up for Tasks 1 and 2

The experiment took place in a sound-attenuated booth. Presentation software (Neurobehavioral systems, Albany, CA, USA) was used to control the presentation of stimuli and record participants' responses given by mouse button presses. The mouse was directly connected to the parallel port of the computer (interrupt signals) thus providing precise RTs measurements. A 36 × 27 cm^2 CRT computer screen (1024 × 768 pixels) was located 95cm in front of the participant. The screen background color was grey and had a luminance of 20 cd/m^2. The visual stimuli were white disks (150-pixel diameter) of 60 cd/m^2 luminance for the two Tasks. Auditory stimuli were piano tones (see below) presented using two speakers located at each side of the screen at a level of 70 dB SPL A, measured at the participants' head using a Brüel & Kjær type 2239 sonometer.

2.5. Task 1: speeded detection of auditory, visual and audiovisual stimuli

Congenital amusics and matched controls performed one block of a simple detection task comprising visual, audio, and audiovisual stimuli. Participants were required to press the left mouse button as rapidly as possible when a stimulus appeared (V, A, or AV) independently of the nature of the stimulation. They responded with their right hand by clicking on the left mouse button with their index finger (note that for the three left-handed participants, they were free to use their left hand, but all were right-handed mouse users). The sound stimulation was a piano tone of 100 ms duration, played at the pitch level of C6 (1047 Hz), and synthesized using a Roland SC 50 sound canvas (Roland Corporation, Los Angeles, California). The visual stimulation was a white disk presented at the center of the screen (diameter = 150 pixels) for a duration of 100 ms. There were 90 trials in total with 30 trials per condition (A, V, AV), the trials were presented in a pseudo-randomized order with the constraint that the same trial type (i.e. A, V, AV), could not be repeated more than three times in a row. Note that for the AV condition, auditory and visual stimuli were presented simultaneously. Participants were given 1000 ms to respond after the onset of any stimulus (disk, sound, disk and sound simultaneously). Inter-trial-intervals varied from 1000 to 2000 ms randomly. Before the main task, participants performed a set of 15 practice trials. We measured the Response Times (RTs) for each trial, calculated from the onset of the stimulus to explore in both groups if a potential benefit of AV integration (shorter RTs) could be observed in comparison to unimodal stimulations (A,V).

It is widely admitted that observing shorter RTs (on average) for a bimodal condition (i.e. AV) in comparison to unimodal conditions (i.e. A, V) does not necessarily reflect cross-modal integration per se. It had been argued that this multisensory benefit could result instead from *statistical facilitation* (Raab, 1962). *Statistical facilitation* reflects the fact that, for each trial, participants could benefit from the fastest modality between the two unimodal stimulations. As a participant responds to whichever of two stimulus modalities (here A and V) that is processed first, the RT in the combined-stimulus situations is based on the distribution of the *minima* between RTs of the two stimulus modalities (Figure 1).
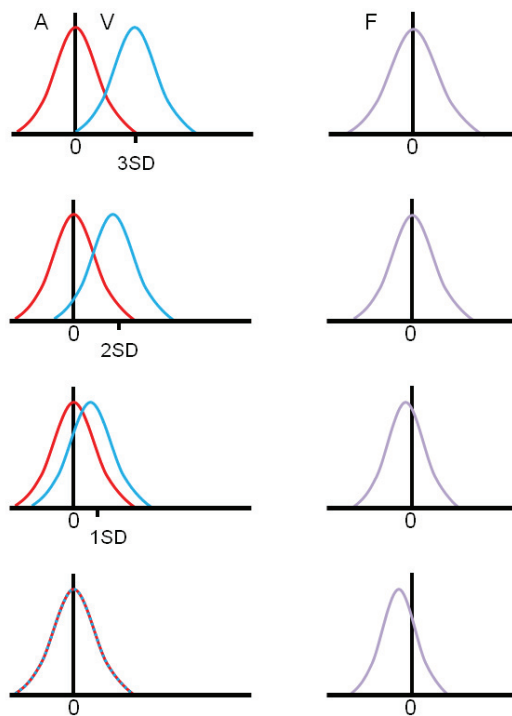
Figure 1. Statistical facilitation from Raab's model (modified from Raab, 1962). Distribution of RTs of two stimulus modalities A(t) and V(t), are shown on the left panel. The distributions of minima, F(t), (right panel), are calculated as: $F(t) = A(t) \int_t^\infty V(t)dt + V(t) \int_t^\infty A(t)dt$. When A(t) and V(t) do not overlap (upper panel), the mean of the distribution of F(t) is equal to that of A(t), providing no statistical facilitation. When the two stimulus distributions overlap more, for example when the mean of V(t) is about 1 SD greater than that of A(t), the mean of F(t) is shortened. When the two distributions overlap completely, the greatest shortening of the mean of F(t) is observed, (lower panel), this effect being referred to as statistical facilitation.

This effect has lead to the *Race Model* hypothesis for simple RTs, stating that a *race* between the two unimodal stimulations could explain bimodal facilitation: when investigating the RTs to either a V signal or an A signal, mean RT is shortened when both signals are presented simultaneously relative to when only one is presented.

To avoid this problem and to quantify a proper multisensory effect on RTs, Miller (1982) proposed that combined stimulations (audiovisual in the present case) have to produce RTs that are shorter than those predicted by race models. In this view, RTs distribution must be as

$$P \ (RTAV <t) > [P \ (RTA <t) + P \ (RTV <t)]$$

where P (RT <t) represent the density function of cumulative probabilities of the RTs (Miller's inequality, 1982).

For the present Task, we thus tested whether RTs distributions would satisfy this inequality. It is important to note that the inequality can be violated only when the value of t is relatively small. As t gets larger, the *P (RTAV <t)* goes to a maximum of 1 while the *[P (RTA <t) + P (RTV <t)]* goes to a maximum of 2. Thus, for large t the inequality is sure to hold, but in the opposite direction: *P (RTAV <t) < [P (RTA <t) + P (RTV <t)]* (Miller, 1982). Therefore, in the present study, the tests of interest focus on small t values.

Furthermore, it should be noted that the violation of Miller's inequality is only applicable when investigating bimodal facilitation in a single participant. For group studies, it does not guarantee that the bimodal facilitation is not due to variability of the distribution of participants' RTs (sampling bias). To statistically evaluate audiovisual facilitation in a group of participants and to avoid sampling bias issues, we assumed that as the application of Miller' inequality uses RTs distributions, and not averaged data, the distributions of different participants have to be merged without losing information about their auditory, visual and audiovisual distributions. For this aim, we used a well-known clustering method for distributions (Vincent, 1912), known as *Vincentization*, that allows testing Miller's inequality in a group of participants. This method involves calculating a given fractile limit of the group distributions by performing an average of the data of this fractile limit for all participants of the group. This method allows avoiding variability associated to absolute differences in terms of RTs between participants and to highlight distributions differences in all participants of the group. We thus have calculated, for each participant, the sum of the RTs distributions of the unimodal data (A+V, the cumulative sum of this sum of distributions is two). The first half of the resulting sum of distributions (i.e., up to the point where the cumulative sum reached one) was then divided in 20 equal Fractiles

(providing 19 values corresponding to RT limits between the 20 Fractiles) and performed the same clustering for the distribution of audiovisual RTs. We were thus able to statistically test the differences between audiovisual data and data resulting from of the sum of unimodal distributions of RTs (A+V).

2.6 Task 2 : Pitch Change Detection Task for Audio and Audiovisual material

The Pitch Change Detection task (PCD) was an adaptation of the pitch task used in Hyde & Peretz (2004, stimuli by courtesy of the authors). Participants performed the PCD in two different conditions (either Audio only, A, or with uninformative visual stimuli: Audiovisual, AV). In both conditions, they had to indicate within a sequence of 5 isochronous tones, whether the fourth tone was changed in pitch in comparison to the other tones of the sequence. The stimuli were 100 ms piano tones played at C6 (1047 Hz, same piano tone as the one used in Task 1) presented with a 250 ms Inter-Stimulus-Interval (that is, an SOA of 350 ms), leading to a total sequence length of 1500 ms. For different trials, the fourth tone was displaced by either a 1/4 tone, a 1/8 tone or a 1/16 tone (12.5, 25, and 50 cents respectively, with 100 cents corresponding to a 1/2 tone) upward or downward from C6. Participants were asked to press as rapidly as possible the right mouse button when they detected a change in the sequence and the left mouse button when they were unable to detect a change. They responded with their right hand by clicking on the left mouse button with their index finger and on the right mouse button with their middle finger. In the first condition (A), participants performed the PCD with the auditory stimulation only. In the second condition, they performed the PCD with simultaneously presented uninformative visual stimuli displayed on the screen (AV). The visual stimulation consisted in 5 disks (with the same characteristics as the disk used in Task 1) appearing consecutively from the left to the right (x1 = -384, x2 = -171, x3 = 0, x4 = -171, x5 = -384, in pixels relative to the screen center) of the screen at the middle of its vertical axis. Each visual disk appeared for a duration of 100 ms simultaneously with one of the auditory stimuli. It is of importance to note that the visual stimulation is uninformative with respect to whether or not an auditory change was present in the sequence (for a similar procedure, see Caclin et al., 2011).

For each condition (A, AV), there were 192 trials, with 96 trials without changed tone and 96 trials with a changed tone (with equal proportion (16 trials) of each of the 6 types of pitch changes). There

were two blocks of 96 trials each (48 trials without a changed tone, 48 trials with a changed tone with equal proportion of all types of pitch changes) for each condition (A, AV). Conditions were presented in alternating blocks with the starting condition being counterbalanced across participants (A, AV, A, AV or AV, A, AV, A). Note that matched participants (a given amusic with his matched control) performed the tasks with the same block order. The blocks were separated by 2-3 minutes of break. Within a block, trials were presented with several constraints: the same trial type (i.e. trial with a changed tone, trial without a changed tone), could not be repeated more than three times in a row, and trials with a changed tone with the same pitch interval size (either 1/4 tone, 1/8 tone or 1/16 tone) could not be repeated twice in a row.

For each trial, after the auditory stimulation (that lasted 1500 ms) a question mark ("?") appeared on the screen, and participants had 2000 ms to respond before the start of the next trial. Participants were informed about the condition order and were asked to respond as accurately and as rapidly as possible once the question mark appeared. Before the task, participants performed one practice block of each condition with a set of 15 trials with feedback.

3. Results

3.1. Speeded detection of Auditory, Visual and Audiovisual stimuli (Task 1)

RTs data was transformed using natural logarithm for each trial (in order to normalize the distribution). RTs analyses were performed on the mean of the RTs for each condition and participant. Note that data from one control participant were excluded from this analysis as her average RTs (mean = 5.94 (~379.93 ms)) was 2SD above the mean of her group (5.82 (~336.97 ms)) for the A condition. RTs were analyzed with a 2x3 ANOVA with Group (amusics, controls) as between-participants factor and conditions (A, V, AV) as within-participant factor. The ANOVA revealed a significant main effect of condition ($F_{(2,58)}$= 114.60, $p < .0001$), with shorter RTs for the AV condition in comparison to the A and the V conditions (all $ps < .0001$). Additionally, note that the RTs for the A condition were significantly shorter than for the V conditions ($p = .01$). No significant Group differences ($F_{(1,29)} = 0.80$; $p= .37$) or interaction between Group and Condition ($F_{(2,58)} = 0.38$; $p= .68$) were observed (see Figure 2).
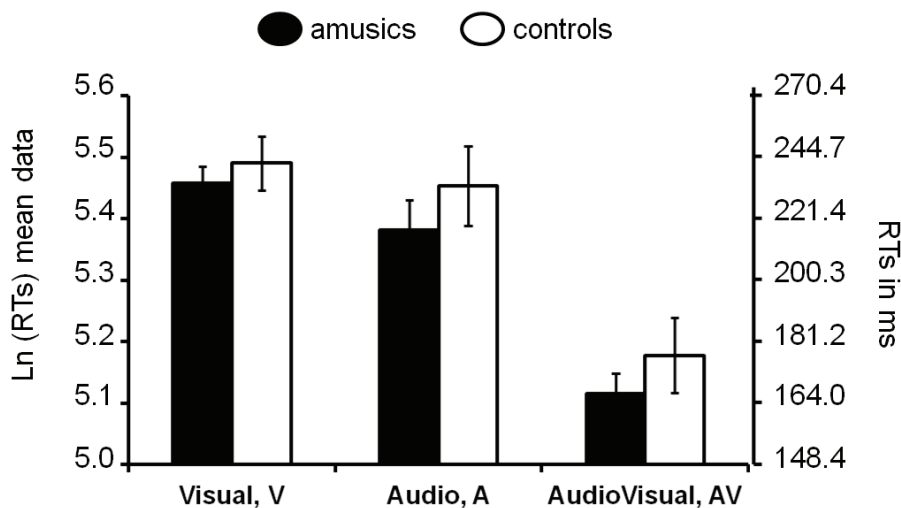

Figure 2: Reaction times of amusic and control groups (black, amusics; white, controls) as a function of the three conditions (Visual, V; Audio, A; Audiovisual; AV).Left y axis corresponds to the RTs transforms on which the statistics were performed; Left y axis corresponds to the RTs data in ms. Errors bars represent SEM.

To assess whether the benefit of the AV condition was related to multisensory integration processes, we tested Miller's inequality with a 2x2x19 ANOVA with Group (amusics, controls) as between-participants factor and conditions (Audio+Visual, A+V and Audiovisual, AV) and Fractiles (19 values corresponding to RT limits between the 20 Fractiles) as within-participant factors. The ANOVA revealed 1) a significant main effect of condition ($F(1,29)=43.73$, $p < .0001$), with shorter RTs in AV in comparison to A+V; 2) an expected significant effect of Fractiles ($F(18,522) = 392.95$, $p < .0001$); and 3) a significant interaction between Condition and Fractiles ($F(18,522) = 28.73$, $p < .0001$). To analyze this interaction, Fischer LSD post hoc tests were performed and revealed that RTs for AV were shorter in comparison to A+V for fractiles 1 to 17 (all ps <. 0001) and were slower for fractiles 18 and 19 (all ps <.0001) (Note that these differences for the longer fractiles are a non-interesting and unavoidable side-effect of the procedure used to create the A+V distribution, see Methods). No significant effect of Group was observed ($p = .29$), nor any interactions involving the Group factor (all ps > .62) (see Figure 3).

In a second step, we analyzed RTs for unimodal data (V and A) to assess whether RT distributions for A and V conditions could be different between amusic individuals and controls. Similarly to what has been done above for the AV and A+V RT distributions, A and V RT distributions were vincentized in 20 fractiles. RTs for unimodal stimulations (A or V) were analyzed with a 2x2x19 ANOVA with Group as between-participants factor and Condition (A, V) and Fractiles as within-participant factors. The main effect of condition was significant ($F(1,29) = 7.31$, $p = .049$) with shorter RTs for the A condition in comparison to the V condition. The interaction between Fractiles and Condition (A, V) was significant ($F(18,552) = 17.05$, $p < .0001$), and was further modulated by Group, as indicated in the three-way interaction between Group, Fractiles and Condition ($F (18,522) = 3.95$; $p< .0001$). Post-hoc tests revealed that amusic individuals presented shorter RTs in the A condition in comparison to the V condition for fractiles 1 to 11 (all p values < .001) and the inverse pattern (RTV< RTA) for fractiles 16 to 19 (all ps <.001). For controls, shorter RTs in the A condition (in comparison to the V condition) was observed for the fractiles 1 to 7 (all p values < .001) and no significant differences were observed between RTs of A and V conditions for fractiles 8 to 19 (all p > .09). None of the direct comparisons between groups reached significance (all ps >.21).
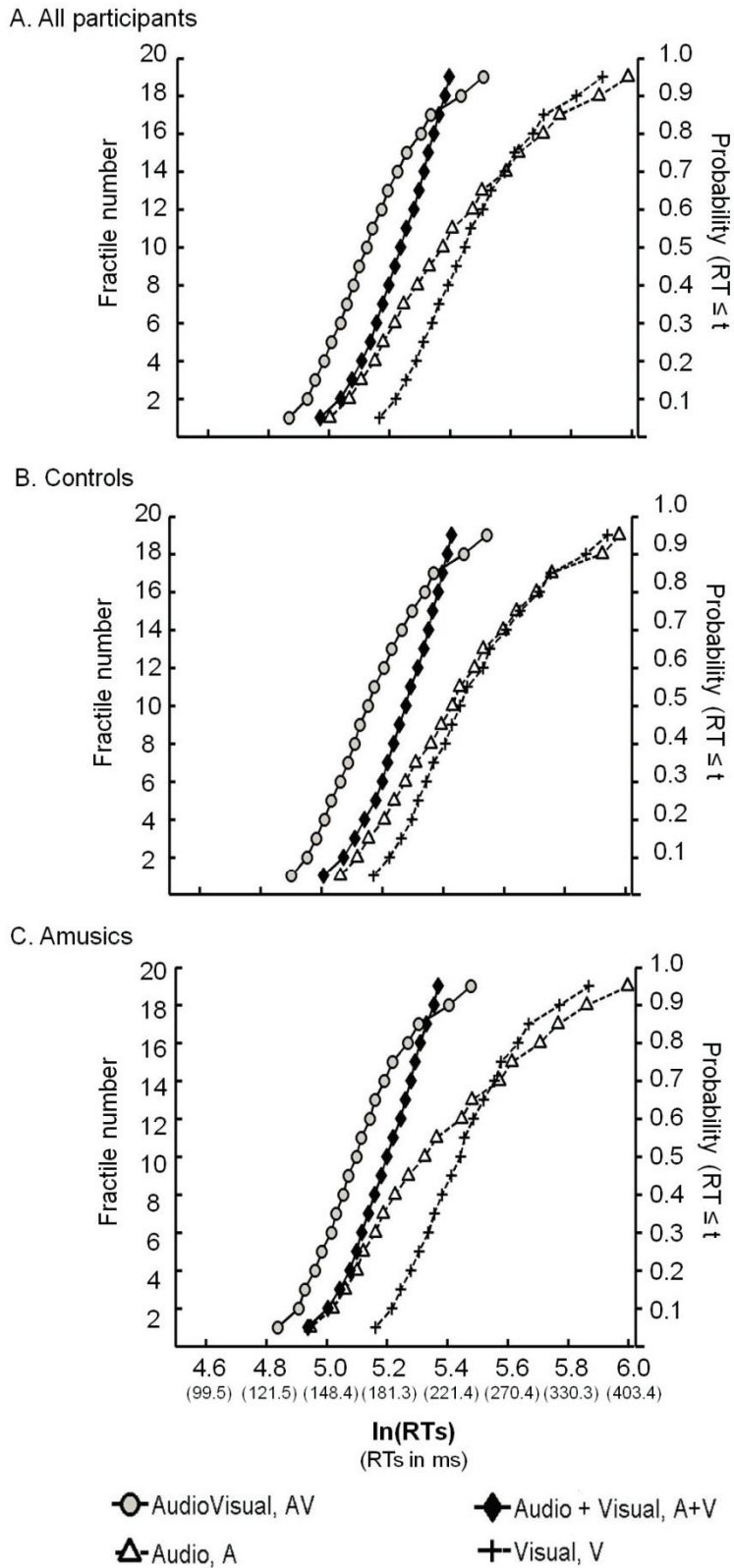
Figure 3: Density functions of cumulative probabilities for RTs in audio, A; visual,V; Auditory + Visual, A + V; and AudioVisual, AV conditions. A. Density functions for all participants. B. Density functions for control participants. C. Density functions for amusic participants.

3.2. Pitch Change Detection Task for Audio and Audiovisual material

3.2.1. Performance

One control participant did not understand the PCD task (she responded every two trials) and was thus excluded from the analysis. Performance in the PCD task was analyzed using signal detection theory by calculating, for each participant, each Condition (A, AV) and each Interval Size (1/4 tone, 1/8 tone, 1/16 tone), discrimination sensitivity with d' and the response bias c[7]. For each participant, these analyses were based on z-values of the hit rate (i.e., number of correct responses for trials with a changed tone / number of trials with a changed tone) and the false alarm rate (i.e., number of incorrect responses for trials without a changed tone / number of trials without a changed tone). Positive values for c arise when the miss rate (incorrect responses for trials with a changed tone / number of trials with a changed tone) exceeds the false alarm rate. Positive values thus indicate a tendency to answer "no change detected", negative values indicate a tendency to answer "change detected", and c-values around 0 suggest the absence of a response bias.

d' and c were analyzed separately with a 2x2x3 ANOVA with Group (amusics, controls) as between-participants factor and Condition (A, AV) and Interval Size (1/4 tone, 1/8 tone, 1/16 tone) as within-participant factors.

Performance was significantly superior to chance level (d'=0) in each condition for each group (all ps < .001, see Figure 4A). The main effect of Group was significant, ($F(1,29) = 4.53$, p = .04), and revealed that control participants reached higher performance levels than did amusic individuals. The main effect of condition was significant ($F(1,29) = 5.13$, p = .03) revealing that participants reached higher performance level for the AV condition in comparison to the A condition. The main effect of Pitch Interval Size ($F(2,58) = 38.96$, p < .0001) was also significant and further modulated by Group ($F(2,58) = 4.01$, p = .02). Fischer LSD post hoc tests revealed that amusic individuals showed decreased performance with decreasing Pitch interval size (all ps < .0001), while controls reached similar performance for 1/4 and 1/8 tone intervals (p = .34), but decreased performance for 1/16 tone

---

[7] The correction of d' and c measures used .01 for cases without false alarms and .99 for the maximum number of hits.

interval (all ps < .001). Moreover, note that amusic individuals showed decreased performance in comparison to controls for 1/8 (p = .03), and 1/16 (p = .002) (i.e., the most difficult) interval sizes, but not for the 1/4 tone interval (p = .37). Note that no other significant effects or interactions were observed (all ps > .69).

For c (Figure 4B), the response bias was significantly superior to zero in each condition for both groups, revealing participants' general tendency to respond "no change detected" for both A and AV conditions (all ps < .01). However, note that the main effect of Condition was significant (F(1, 29) = 5.68, p = .02) with higher positive c values for the AV condition than for the A condition. The main effect of Pitch Interval size (F(2,58) = 38.80, p <.0001) was significant as and was further modulated by Group (F(2,58) = 3.96, p =.024). Post hoc test revealed that amusic individuals showed increased bias with decreasing Pitch interval size (all ps < .001), while controls showed similar bias data for 1/4 and 1/8 tone intervals (p = .33) and increased bias for 1/16 tone interval (p < .001). Moreover, note that amusic participants' and controls' bias data did not differ for all pitch interval sizes (all ps > .13). No other significant effects or interactions were observed (all ps > .37).

A. Performance

B. Bias
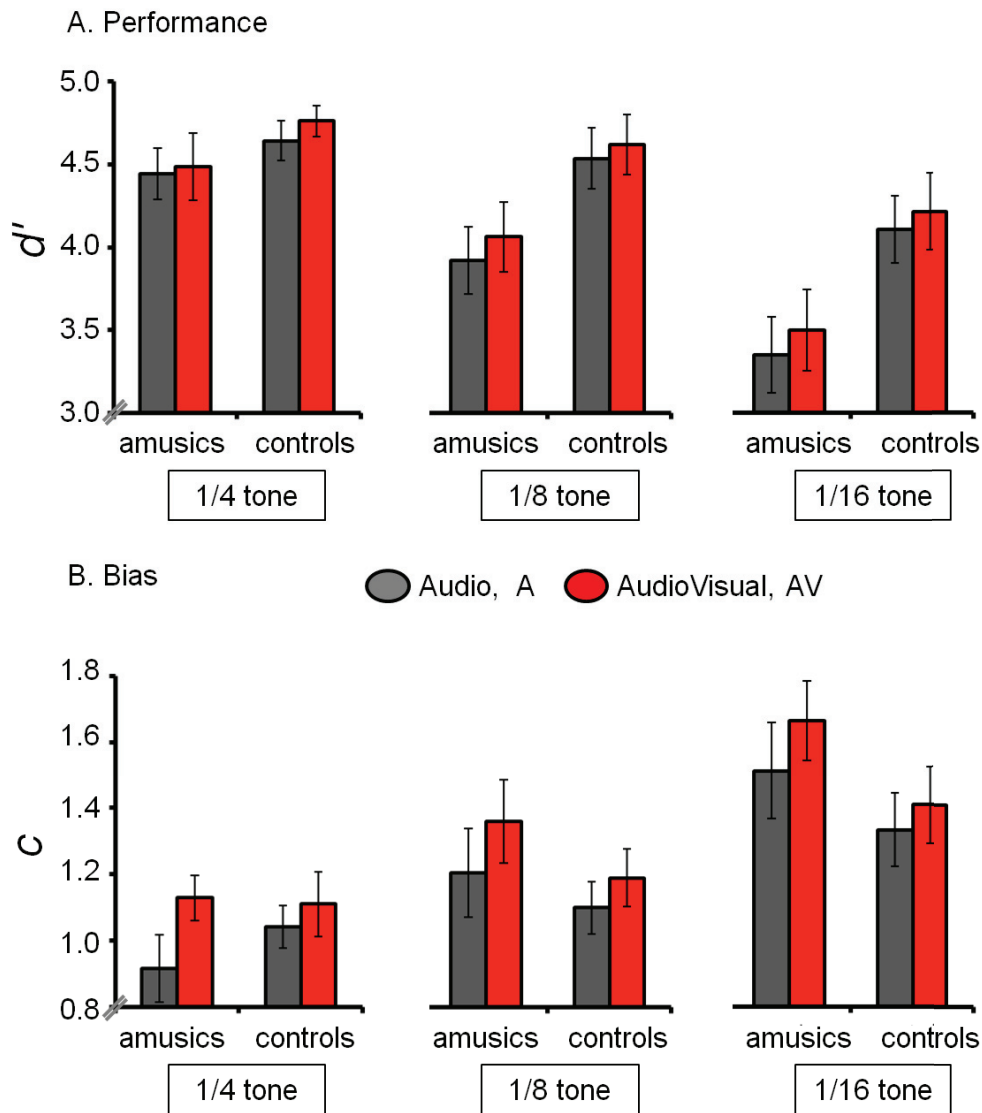
○ Audio, A ● AudioVisual, AV

Figure 4. A. Participants performance in terms of d' presented as a function of Group (amusics, controls), Pitch Interval Size (1/4 tone, 1/8 tone, 1/16 tone) and Condition (A, AV). B. Bias data for amusic and control groups in terms of c, presented as a function of Pitch Interval Size (1/4 tone, 1/8 tone, 1/16 tone) and Condition (A, AV). Note that all biases where positive thus revealing a general tendency to respond "no change" in both groups.

Additional analyses were performed by investigating the effect of the Trial Type (with a changed tone, without changed tone) in participants' performance. We analyzed the percentage of correct responses with a 2x2x2 ANOVA, with Group as between-participants factor and Condition (A, AV), and Trial Type (with changed note, without changed note) as within-participant factors.

In addition to the main effect of Group ($F(1,29)= 4.84$; $p=.03$), the effect of Trial Type was significant ($F(1,29)= 16.79$; $p < .0001$) revealing that participants' performance was decreased for trials with a changed note in comparison to trials without a change note. Note that the Group by Trial Type interaction was marginally significant ($F(1,29)= 3.08$; $p = .06$). Observing that both amusic and controls showed decreased performance for trials with changed tone in comparison to trials without change suggests that participants missed pitch changes rather than hearing non-existing changes (thus fitting with other data in the literature reporting increased miss-rates rather than FA-rates in congenital amusia, Tillmann et al., 2009; Albouy et al., 2013b).

3.2.2. *Response Times (RTs)*

As in Task 1, analyses were performed on the transformed RTs. We analyzed RTs for correct responses only to assess whether a facilitative effect of AV stimulation could be observed. In a first step, we investigated the effect of the Trial Type, and then explored the RTs as a function of the Pitch Interval size in the different trials to assess if a potential benefit of the AV condition could be related to task difficulty.

RTs data for correct responses were analyzed with a 2x2x2 ANOVA with Group as between participant factor and Condition (A, AV), and Trial Type (with changed note or without changed note) as within-participant factors. Analyses revealed a significant main effect of Group ($F(1,29)= 16.48$; $p<.0001$), with shorter RT in controls than in amusic individuals. The main effect of Trial Type was significant ($F(1,29)= 38.58$; $p < .0001$), showing shorter RTs for trials with a changed tone in comparison to trials without change. Finally, the main effect of Condition was significant ($F(1,29)= 19.09$; $p <. 001$) and revealed that participants showed shorter RTs for the AV condition in comparison to the A condition (Figure 5). No other effects or interactions were significant (all ps $> .25$).
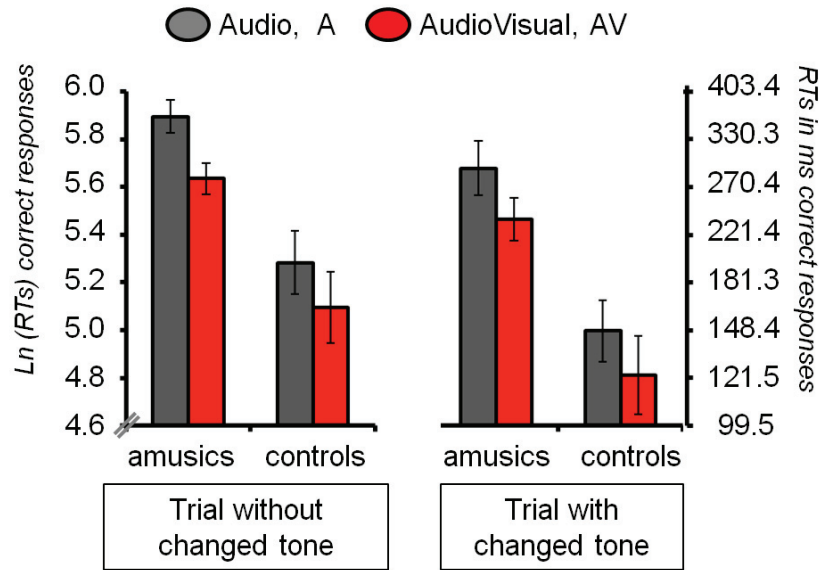
Figure 5: *RTs for correct responses presented as function of Trial Type, Group, and Condition.*

To assess if the observed benefit of AV condition on the RTs could be related to task difficulty, we analyzed the RTs data for trials with a changed tone as function of the pitch interval size. As in the previous analysis, we investigated RTs for correct responses only.

RTs data were analyzed with a 2x2x3 ANOVA with Group as between participant factor and Condition (A, AV) and Pitch Interval Size (1/4 tone, 1/8 tone, 1/16 tone) as within participant factors. The ANOVA revealed that additionally to the main effect of Group ($F(1,29)= 15.76$; $p < .0001$), the main effect of Pitch Interval Size (($F(2,58)= 39.32$; $p < .0001$) was significant revealing that RTs for the larger (i.e., easier) Pitch Interval Size (1/4 tone) were shorter than the other Pitch Interval Sizes (1/8 tone, 1/16 tone, all ps $< .0001$) and that RT 1/8 tone interval size were shorter than for 1/16 tone interval size ($p < .001$).

The main effect Condition was significant ($F(1,29)= 11.93$; $p < .001$) revealing shorter RTs for the AV condition in comparison to the A condition. More interestingly, the three-way interaction between Group, Condition and Pitch Interval Size ($F(2,58) = 4.51$; $p = .01$) was significant. Fisher LSD post-hoc tests were carried out and revealed that the benefit of the AV condition (shorter RTs) compared to the A condition was observed for different Pitch Interval Sizes depending on the group.

For amusic individuals, shorter RTs in the AV condition in comparison to the A condition were observed for 1/4 tone and 1/8 tone (all ps < .001), but not for the 1/16 (p= .16) while controls showed this benefit of AV condition for 1/8 tone (p= .005) and 1/16 tone (p < .0001) pitch intervals, but not for the 1/4 tone interval (p = .15, see Figure 6).
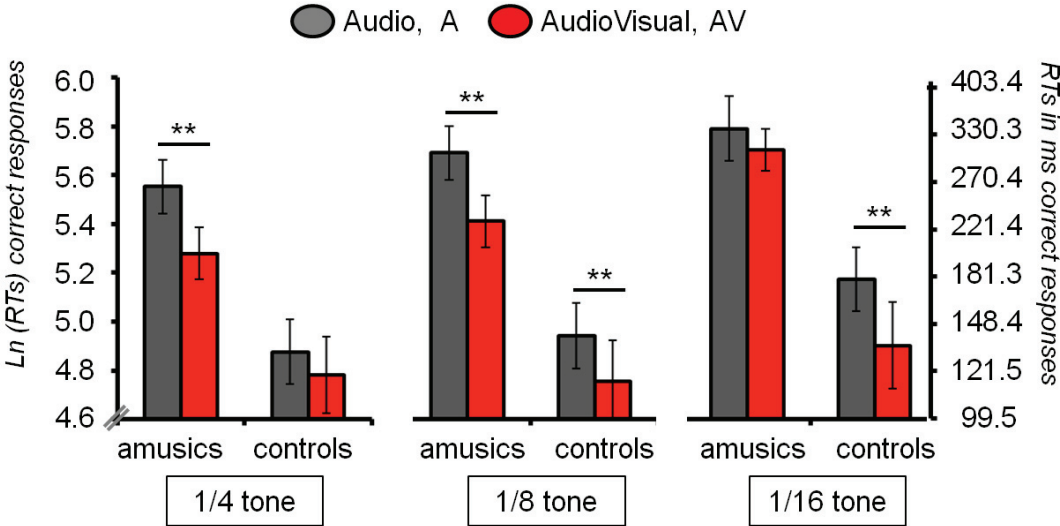

Figure 6. RTs for correct responses for trials with changed tone presented as a function of Group, Condition for each Pitch Interval size (1/4 tone; 1/8 tone; 1/16 tone).

3.3. Correlations

Three sets of analyses were performed: 1) correlations between data from the pre-tests (MBEA, PDT) and data from Tasks 1 and 2 (RTs, magnitude of the AV benefit in Task 1, *d'*, *c*, RTs, difference between AV and V conditions for *d'*, *c*, and RTs); 2) correlations between the different measures (e.g., *d'*, RTs) of the behavioral data in Tasks 1 and 2 and; 3) correlations between the different measures (e.g., *d'*, RTs) of Task 2. To explore differences explained not only by the group differences defined via the MBEA or the PDT, we focus on correlations that were significant over all participants *and* for at least one of the two participant groups separately (amusic group or control group).

Correlation analyses were performed with all participants but note that one control participant was excluded from Task 1 and one other control participant was excluded from Task 2 (see Results). For the first set of analyses two correlations were observed:

1) data of the MBEA was positively correlated with performance on Task 2 for both AV and A conditions. Note that these correlations were observed in controls (A condition ($r(13) = .63$; $p = .01$); AV condition ($r(13) = .52$; $p = .04$)) and over all participants for A condition $r(29) = .54$; $p = .002$; for AV condition ($r(13) = .41$; $p = .02$)) but not in amusic participants (all ps $> .15$) .

2) Data of the PDT was negatively correlated with performance data of Task 2 for both AV and A conditions. These correlations were observed in controls (A condition ($r(13) = -.52$; $p = .04$); AV condition ($r(13) = .57$; $p = .02$) and over all participants (A condition ($r(29) = .44$; $p = .01$); AV condition ($r(13) = .45$; $p = .01$)) but not in amusic participants (all ps $> .20$).

For the second set of analyses, none of the correlations between data from Task 1 and Task 2 were significant.

For the third set of analyses, a negative correlation between the AV benefits in terms of accuracy (difference between AV and A conditions for d', a positive value corresponding to AV>A performance) and RTs benefits (difference between AV and A conditions for RTs, a negative value corresponding to faster RTs for the AV in comparison to the A condition) was observed. This correlation was significant in amusic participants ($r(14) = -.55$, $p = .02$) and across all participants ($r(29) = -.51$; $p = .004$) and was only marginally significant for controls ($r(13) = -.44$, $p = .09$) (see Figure 7). In addition, note that for both groups, data for the AV and A condition were positively correlated (all ps $< .001$) for both performance and RTs data (all ps $<. 0001$). Finally performance data and RTs for each condition and each interval size were negatively correlated (all ps $< .001$).
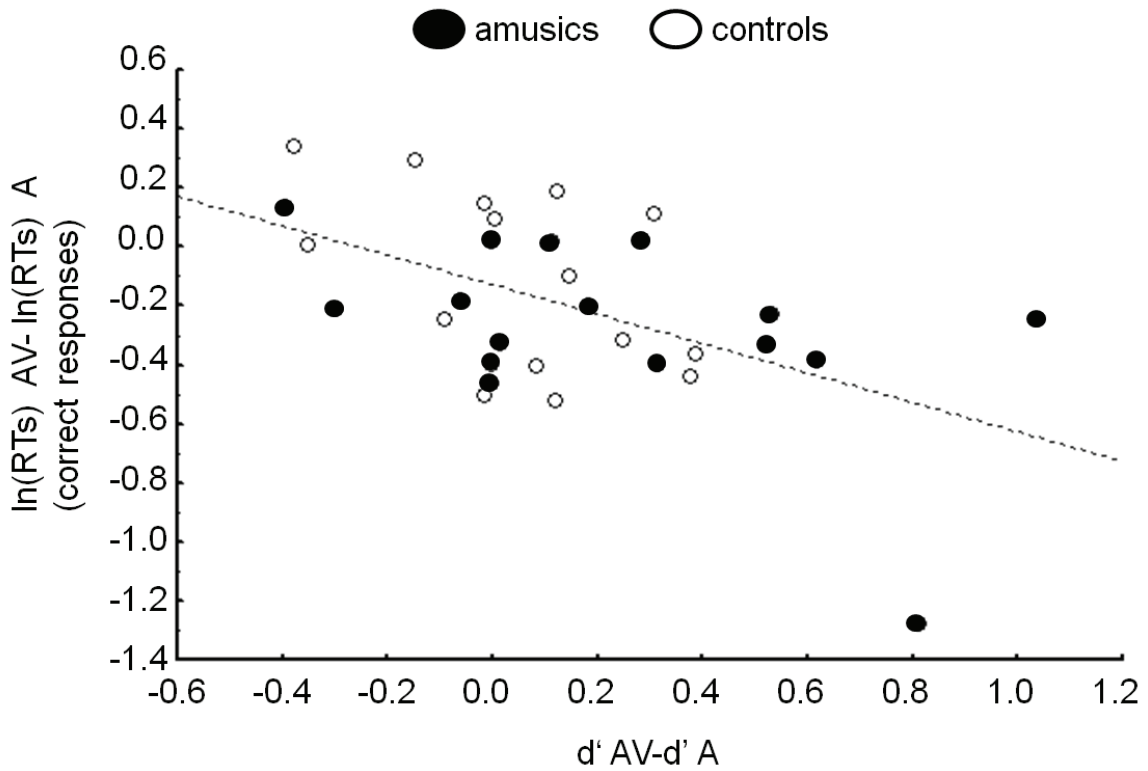
Figure 7. Scatter plot of the benefit in terms of RTs for the AV condition in comparison to the A condition (ln(RTs) AV- ln(RTs) A, correct responses) against the benefit in terms of performance for the AV condition in comparison to the A condition (d' AV- d' A). The correlation between these two variables is significant across all participants and in the amusic group (see results for details). Note that without the two outliers (amusic participants, data points at x1 = 0.81; y1 = -1.27 and x2 = 1.03; y2 = -0.24), this correlation remained significant in all participants (r(27) = -.43; p = .17) but only marginally significant in amusic participants only (r(12) = -.45; p = .09).

4. Discussion

The present study investigated whether audiovisual integration could improve pitch processing abilities in congenital amusia. Results of Task 1 revealed that both amusic individuals and controls exhibited audiovisual facilitation (shorter response times for audiovisual stimuli than for either auditory or visual stimuli) and no differences were observed between the two participant groups in terms of facilitation. Results of Task 2 revealed that for the pitch detection task, participants of both groups exhibited higher performance and shorter response times in the AV condition than in the A condition. For RTs, these audiovisual benefits were observed for different ranges of difficulty for each group (that is, for different pitch interval sizes of the changed tone) thus suggesting that optimal multisensory facilitation could be observed only within an appropriate range of unimodal performance.

Task 1 investigated for the first time if simple auditory and visual processing are preserved in congenital amusia and if AV facilitation could be observed in this deficit. By using a speeded detection task for visual, audio and audiovisual stimuli, we observed that both amusic individuals and controls exhibited shorter RTs for AV stimuli in comparison to the unimodal stimuli (A, V). More interestingly, by exploring Miller's inequality for RTs distributions, we observed that this benefit was related to multisensory integration processes (rather than statistical facilitation (Raab, 1962)) as RTs were shorter for the AV distribution in comparison to the A+V distribution in all participants and in each group independently.

These results are in line with numerous studies showing that the combination of auditory and visual information enhances perceptual abilities as revealed by response times (Ernst and Bulthoff, 2004; Gielen et al., 1983; Hershenson, 1962; Posner et al., 1976; Stein and Meredith, 1993) in typical listeners. This effect is known since a long time and has been historically interpreted by the *energy-summation hypothesis* (Nickerson, 1973), which assumed that the *energy* from the two stimuli, in the context of a bimodal presentation, is combined and leads to more rapid processing in early stages of the perceptual process (such as sensory encoding), as suggested by the shorter detection times. This hypothesis has received support from the observation of cross-connections between early sensory areas, as for example demonstrated for connections from the auditory to the visual cortex (Falchier et

al., 2002; Rockland and Ojima, 2003). Moreover, this hypothesis is in line with the observation that the superior colliculus, a subcortical structure involved in orienting to external stimuli, contains many neurons responding to bi- or trimodal stimuli, and could mediate cross-modal improvements in simple detection tasks (Stein and Meredith, 1993). Furthermore, in a study using the speeded detection task together with recordings of Event-Related Potentials (ERPs) (Fort et al., 2002), it had been evidenced that audiovisual interactions take place both in sub-cortical and cortical-structures. These studies thus suggest that the multisensory integration is performed in the first steps of the perceptual process, thus facilitating the early encoding of either visual or auditory information. Note, however, that an alternative explanation was the *preparation-enhancement* hypothesis, which was also proposed by Nickerson (1973), who assumed that facilitation could also concern more later stages of processing: one unimodal stimulus (referred to as 'accessory') *enhances* and *prepares* the processing of the other 'obligatory' stimulus. The differentiation between the role of early and later steps of multisensory integration is still under debate, but one might assume that multimodal facilitation could rather be considered as a combination of both processes (Bauer, 2008), involving local and long distance synchronization and communication between brain areas.

Interestingly for the purpose of the present study, we observed comparable AV facilitation in amusic individuals and controls. This observation let us to argue that multisensory integrating systems (either in the early or later stages of the processing) for visual and auditory information are preserved in the amusic brain. Moreover, the direct comparison between A and V conditions showed that, as for control participants, amusic individuals showed shorter detection times for the A stimuli in comparison to the V stimuli. This result is convergent with the hypothesis suggesting that amusic individuals present altered encoding of tones only when those have to be memorized (altered memory traces of tones leading to impairments in pitch discrimination and memory, see Albouy et al., in preparation and Albouy et al., 2013a) but present preserved sensory extraction of sound information in a simple context. Taken together, results from Task 1 allow us to conclude that amusic individuals processes A, V and most importantly AV information as well as typical listeners, thus permitting the exploration of bimodal benefits on amusic individuals' abilities in processing pitch material.

To this aim, in Task 2, we used a pitch change detection task for which amusic individuals were previously shown to be impaired (Albouy et al., *in preparation;* Hyde and Peretz, 2004) in comparison to controls. As expected, amusic individuals showed decreased performance (as measured with *d'*) in comparison to controls for the small Pitch interval sizes (1/8 tone; 1/16 tone), but not for the large interval size (1/4 tone). While this finding is in agreement with previous evidence for amusic individuals' pitch perception deficit (Albouy et al., *in preparation;* Albouy et al., 2013b; Ayotte et al., 2002; Foxton et al., 2004b; Jones et al., 2009b; Liu et al., 2010; Peretz et al., 2002; Tillmann et al., 2011c; Tillmann et al., 2009), it differs in some points with previous studies using the same PCD task and showing that amusic individuals are impaired in detecting changes smaller than 1/2 tone (see Hyde and Peretz, 2004 and Peretz and Hyde 2003). In the present study, amusic individuals were able to perform the task as well as controls for the 1/4 tone interval thus differing from Hyde and Peretz (2004) study. Based on the observation that performance in the PCD task was related to participants' pitch discrimination thresholds in the present study (see the negative correlation between *d'* in Task 2 and PDT), this inter-study difference might be due to the better pitch thresholds in the amusic group from Lyon in comparison to the amusic group from Montreal (as revealed by the overlap in PDT values between amusic individuals and controls in the present study, see Methods). In addition to these differences between amusic and controls in terms of performance, both groups exhibited decreased performance with increasing task difficulty (decreasing Pitch Interval Size). This effect was also reflected in the response bias data *c* showing participants' increasing tendency to respond '*no change detected'* in the sequences with increasing task difficulty.

Interestingly, a benefit in terms of performance (*d'*) was observed for the bimodal (AV) in comparison to the unimodal (A) condition. This result demonstrates for the first time that visual synchronous stimulation can help amusic participants to improve their pitch processing abilities. The present finding confirms the hypothesis that amusic individuals process the visual stimuli as well as controls, and are able to use it to enhance their pitch perception performance for the tone sequences. Moreover, thanks to controls' data, we demonstrated here that uninformative visual stimulation can boost typical listeners' pitch discrimination (or encoding) abilities, an effect which, to our knowledge, has never been demonstrated before.

In addition to this benefit of the multisensory integration in terms of accuracy, RTs data confirmed that amusic individuals and controls used the visual information in the auditory pitch change detection. For both amusic individuals and controls, RTs were faster for the AV condition than for the A condition. Moreover, this benefit of multisensory integration on RTs was related to the benefit observed on performance as revealed by a negative correlation between d' benefit (d' AV-d'A) and RTs benefit (RTs AV-RTs A) (see Figure 7).

When exploring RTs, we observed that the AV benefit was related to task difficulty and varied as a function of pitch interval sizes in both participant groups. We observed that amusic individuals and controls did not show the benefit on RTs of the AV condition in comparison to the A condition for the same pitch interval sizes (reflecting task difficulty): while controls had shorter RTs for 1/8 tone and 1/16 tone pitch changes, amusic individuals exhibited shorter RTs for the 1/4 tone and 1/8 tone interval sizes, but not for the 1/16 tone pitch changes, that is the most difficult task. This finding is line with several pieces of evidences suggesting that multisensory interactions depend on the participant's performance when the modalities are presented in isolation (Caclin et al., 2011; Hairston et al., 2003; Rouger et al., 2007; Serino et al., 2007) and was observed in particular when the unisensory performance is poor (difficult task). As proposed by Caclin et al. (2011), this could be understood as a law of inverse effectiveness (Stein and Meredith, 1993), but at the population or participant scale. This observation found support in the data of controls which showed AV benefit only for the most difficult tasks, but not for the easiest task. Similarly, amusic individuals did not benefit from the AV condition for the smallest pitch interval size, that is the most difficult task. This result suggests that to observe a benefit of bimodal stimulation, participants have to be in an appropriate range of performance. These observations suggest that to observing AV improvement, the range of efficiency is related to participants' performance, and in consequence, to the difficulty of the task. Moreover, as for the simple detection Task 1, it should be noted that audiovisual interactions have the same influence on both participant groups, as revealed by the RTs data in the same range of unimodal performance.

The present study thus demonstrates that amusic individuals are able to process and benefit from multisensory information to speed up detection as well as improving pitch processing. Although the cerebral dynamics of these processes remain to be investigated, it might be argued that uninformative

visual stimuli influence either early or late levels of cortical processing, by potentially amplifying and boosting the brain responses to auditory signals. The results presented here concerned both accuracy and RT data, that is encouraging in further investigating the implication of multisensory integration for the development of training and rehabilitation programs of impaired musical functions in this lifelong disorder. Finally, note that the modulation of effectiveness of the AV benefits with task difficulty was observed only in RTs, which could be considered as an indirect measure. Further work is thus needed to explore whether this benefit in terms of performance could vary with difficulty for both low level (sensory encoding) and high-level processes (such as memory) in congenital amusia.

5. References

Albouy, P., Cousineau, M., Caclin, A., Tillmann, B., and Peretz, I. (in preparation). Building pitch memory traces in congenital amusia.

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

Albouy, P., Schulze, K., Caclin, A., and Tillmann, B. (2013b). Does tonality boost short-term memory in congenital amusia? Brain Res.

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

Bauer, M. (2008). Multisensory integration: a functional role for inter-area synchronization? Curr Biol 18, R709-710.

Bertelson, P., and Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. Percept Psychophys 29, 578-584.

Caclin, A., Bouchet, P., Djoulah, F., Pirat, E., Pernier, J., and Giard, M.H. (2011). Auditory enhancement of visual perception at threshold depends on visual abilities. Brain Res 1396, 35-44.

Calvert, G.A., and Thesen, T. (2004). Multisensory integration: methodological approaches and emerging principles in the human brain. J Physiol Paris 98, 191-205.

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

Falchier, A., Clavagnier, S., Barone, P., and Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. J Neurosci 22, 5749-5759.

Fort, A., Delpuech, C., Pernier, J., and Giard, M.H. (2002). Dynamics of cortico-subcortical cross-modal operations involved in audiovisual object detection in humans. Cereb Cortex 12, 1031-1039.

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

Frassinetti, F., Bolognini, N., and Ladavas, E. (2002a). Enhancement of visual perception by crossmodal visuo-auditory interaction. Experimental Brain Research 147, 332-343.

Frassinetti, F., Pavani, F., and Ladavas, E. (2002b). Acoustical vision of neglected stimuli: interaction among spatially converging audiovisual inputs in neglect patients. J Cogn Neurosci 14, 62-69.

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

Gielen, S.C., Schmidt, R.A., and Van den Heuvel, P.J. (1983). On the nature of intersensory facilitation of reaction time. Percept Psychophys 34, 161-168.

Hairston, W.D., Laurienti, P.J., Mishra, G., Burdette, J.H., and Wallace, M.T. (2003). Multisensory enhancement of localization under conditions of induced myopia. Exp Brain Res 152, 404-408.

Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. J Exp Psychol 63, 289-293.

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

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

Jones, J.L., Zalewski, C., Brewer, C., Lucker, J., and Drayna, D. (2009). Widespread auditory deficits in tune deafness. Ear Hear 30, 63-72.

Jousmaki, V., and Forss, N. (1998). Effects of stimulus intensity on signals from human somatosensory cortices. Neuroreport 9, 3427-3431.

Lidji, P., Kolinsky, R., Lochy, A., and Morais, J. (2007). Spatial associations for musical stimuli: a piano in the head? J Exp Psychol Hum Percept Perform 33, 1189-1207.

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

McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices. Nature 264, 746-748.

Melara, R.D., and O'Brien, T.P. (1987). Interaction between synesthetically corresponding dimensions. Journal of Experimental Psychology: General 116, 323-336

Miller, J. (1982). Divided attention: evidence for coactivation with redundant signals. Cogn Psychol 14, 247-279.

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

Peretz, I. (2003). Brain specialization for music: New evidence from congenital amusia. In The Cognitive Neuroscience of Music, I.P.I.a.Z. RJ, ed. (Oxford, Oxford University Press), pp. 192-203.

Peretz, I. (2013). The Biological foundations of music: Insights from congenital amusia. In The Psychology of Music, D. Deutsch, ed. (Elsevier), p. in press.

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

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

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

Posner, M.I., Nissen, M.J., and Klein, R.M. (1976). Visual dominance: an information-processing account of its origins and significance. Psychol Rev 83, 157-171.

Raab, D.H. (1962). Statistical facilitation of simple reaction times. Trans N Y Acad Sci 24, 574-590.

Rockland, K.S., and Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. Int J Psychophysiol 50, 19-26.

Ross, L.A., Saint-Amour, D., Leavitt, V.M., Javitt, D.C., and Foxe, J.J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. Cereb Cortex 17, 1147-1153.

Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., and Barone, P. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. Proc Natl Acad Sci U S A 104, 7295-7300.

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

Seitz, A.R., Kim, R., and Shams, L. (2006). Sound facilitates visual learning. Curr Biol 16, 1422-1427.

Serino, A., Farne, A., Rinaldesi, M.L., Haggard, P., and Ladavas, E. (2007). Can vision of the body ameliorate impaired somatosensory function? Neuropsychologia 45, 1101-1107.

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

Shelton, B.R., and Searle, C.L. (1980). The influence of vision on the absolute identification of sound-source position. Percept Psychophys 28, 589-596.

Stein, B.E., and Meredith, M.A. (1993). Merging of the senses (Cambridge, MIT Press).

Stewart, L. (2011). Characterizing congenital amusia. Q J Exp Psychol (Hove) 64, 625-638.

Sumby, W.H., and Pollack, I. (1954). Visual contribution to speech intelligibility in noise. J Acoust Soc Am 26, 212–215.

Teder-Salejarvi, W.A., Di Russo, F., McDonald, J.J., and Hillyard, S.A. (2005). Effects of spatial congruity on audiovisual multimodal integration. J Cogn Neurosci 17, 1396-1409.

Tillmann, B., Albouy, P., and Caclin, A. (2013). Congenital Amusias. In Handbook of Clinical Neurology (Oxford, Elsevier), p. in press.

Tillmann, B., Rusconi, E., Traube, C., Butterworth, B., Umilta, C., and Peretz, I. (2011). Fine-grained pitch processing of music and speech in congenital amusia. J Acoust Soc Am 130, 4089-4096.

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

Vincent, S.B. (1912). The function of the vibrissae in the behavior of the white rat. Behavior Monographs 1.

von Kriegstein, K. (2012). A Multisensory Perspective on Human Auditory Communication.

Vroomen, J., and de Gelder, B. (2000). Sound enhances visual perception: cross-modal effects of auditory organization on vision. J Exp Psychol Hum Percept Perform 26, 1583-1590.

III. General Discussion

Chapter IX Discussion and Perspectives

For the last decade, research in neuroscience has increasingly used music as a tool to explore human cognition and brain functioning. This research area is considerably expanding and raises questions about neural correlates underlying the processing of specific aspects of music (e.g., perception of pitch, rhythm, timbre etc.) and more general cognitive functions (working memory, mental imagery, auditory processing etc.). Within this general framework, the present PhD work aimed to improve our understanding about how humans perceive and memorize complex sound structures through two main approaches:

Axis 1) by investigating the behavioral and neural correlates of auditory perception and memory for tonal and verbal sequences in congenital amusia as compared to the normally functioning brain and;

Axis 2) by investigating how and whether congenital amusic individuals' pitch perception and memory abilities could be improved.

These investigations have revealed anatomical and functional abnormalities in the amusic brain and have provided the first steps toward the perspective to develop tools of rehabilitation for this developmental disorder. The discussion sections of the Chapters IV to VIII of this thesis have already pointed out the major implications of each of the studies presented.

The aim of the present section is twofold: 1) presenting a short summary of the general conclusions that could be made on the basis of this PhD work by linking the studies together and 2) by setting this work into a broader context, notably by presenting how our research could improve the understanding of auditory perception and short-term memory in the normally functioning brain.

9.1. General discussion

9.1.1. *Axis 1: Behavioral and neural correlates of auditory perception and memory for tonal and verbal sequences*

Behavior

Axis 1 contained three studies (Study 1, 2 and 3) using behavioral approaches together with multimodal neuroimaging methods (VBM, MEG, fMRI). Behavioral results have confirmed the short-term memory deficit for tones in congenital amusia and have explored the potential link between the pitch perception and pitch memory deficits in this developmental disorder.

First, by investigating amusic individuals' performance in a short-term memory task for either verbal or tonal material (Study 2), we observed similar performance between amusic and control groups for verbal information, but strongly decreased performance for the tonal information in amusic individuals. This result let us argue that amusic individuals present a pitch-related short-term memory deficit, but not a general short-term memory deficit, which would affect verbal memory too (Tillmann et al., 2009; Williamson and Stewart, 2010).

Second, by combining short-term memory tasks for tone sequences with pitch discrimination (Study 1 and 2) or pitch change detection (Study 3) measurements in the same participants, we investigated whether, as previously hypothesized (Hyde and Peretz, 2004; Peretz and Hyde, 2003), the abnormalities in pitch perception underlie the phenomenon of congenital amusia (see Chapter III). We observed that amusic individuals were strongly impaired in short-term memory performance for tone sequences in comparison to controls, without overlap between the two groups (except in Study 2 for which we used very simple material). In contrast, measures of pitch discrimination thresholds showed large overlaps between amusic and control participants with numerous amusic participants exhibiting pitch thresholds in the range of those of the controls. Based on this observation, we argued that, as suggested by Tillmann et al. (2013), the pitch memory deficit in amusic individuals could not be related solely to impaired discrimination of tones. Moreover, data from Study 3 allowed us to propose that the pitch perception deficit in congenital amusia could be related to an impaired construction of pitch memory traces of short-tones. When the time dedicated to encode information was too short, amusic individuals exhibited strongly impaired performance in comparison to controls for both pitch discrimination and pitch memory tasks. In contrast, when the time dedicated to encode the information was long enough, amusic individuals were able to reach performance comparable to that of controls in both tasks. Observing these group differences only for short tone duration suggests that amusic individuals exhibit an altered construction of memory traces of short tones (encoding) that affect their pitch perception and memory abilities. In this view, it might be argued that pitch discrimination, pitch direction and pitch memory deficits constitute combined elements in congenital amusia (Tillmann et al., 2013), these pitch-related deficits being related to an impaired encoding of short tones (Study 3). All together, we thus propose that congenital amusia could be supported by impaired auditory encoding of pitch information, and that this impairment contributes to some extent to the observed pitch perception and memory deficit. To further examine the short-term memory deficit for tone sequences in congenital amusia, we have used neuroimaging approaches (Study 1, Studyv2). These studies have investigated whether altered brain functioning could support the deficits in congenital amusia as a function of the three processing steps of the short-term memory task:

1) *Encoding*, the processes whereby information is registered (Cowan, 1984); 2) *Storage* or *Retention*, corresponding to the maintenance and manipulation of information over time; (Baddeley, 1992, 2000, 2003a, 2010, 2012; Baddeley, 1986; Baddeley and Hitch, 1974; Champod and Petrides, 2007, 2010; D'Esposito, 2007; D'Esposito et al., 1999; Vallar and Baddeley, 1984); and 3) *Retrieval*, which refers to the accessing of the information stored in memory (Baddeley, 1986).

Neural Correlates

Study 1 combined MEG and structural MRI recordings to test the hypothesis of functional and anatomical abnormalities in amusic individuals' brains related to their pitch perception and memory deficits. During the encoding of the information, amusic individuals and control participants recruited both primary and secondary auditory cortices as well as inferior frontal areas. These results were provided converging evidence with numerous studies (Curtis and D'Esposito, 2003; D'Esposito, 2007; Khan and Muly, 2011; Owen, 2000), suggesting that in addition to the involvement of primary sensory areas, the VLPFC plays a general role in memory by triggering active low-level encoding strategies. By comparing amusic individuals and control participants' brain responses, we observed that, in line with the anatomical abnormalities observed in the right fronto-temporal network (Hyde et al., 2007; Hyde et al., 2006; Mandell et al., 2007) in terms of white and grey matter concentrations (see also Study 1), the amusic brain exhibited abnormal and strongly delayed responses in bilateral VLPFC and STG during the encoding of melodies in the short-term memory task. As revealed by Dynamic Causal Modeling (David et al., 2006), these abnormal responses were supported by dysfunctional connectivity between and within several brain areas, the amusic brain notably exhibiting 1) increased lateral connectivity between left and right auditory cortices; 2) decreased intrinsic connectivity in both auditory cortices, and 3) decreased right fronto-temporal backward connectivity (Hyde et al., 2011). These observations allow us to make the assumption that amusic individuals' deficits in short-term memory are supported by delayed and altered brain functioning during the first steps of memory processing (encoding, echoic memory, building of the pitch memory traces), participating to impairments in the following steps of short-term memory (Retention, Retrieval). Moreover, they suggest that, in the normal functioning brain, dynamic fronto-temporal interactions are necessary to efficiently construct an appropriate memory trace of the auditory information in the context of a short-term memory task.

Along these lines, Study 3 tried to pursue the characterization of controls and amusic individuals' pitch encoding abilities with a behavioral approach. By manipulating sound characteristics that have been previously shown to facilitate the encoding of the information (Demany and Semal, 2008), we observed that for pitch discrimination of single tones and for short-term memory of tone sequences, when the time given to encode the auditory information was too short, amusic individuals exhibited strongly impaired performance as compared to their matched controls. This result is fully in line with the conclusions made in Study 1 and suggests that in comparison to controls, amusic individuals showed altered encoding of short sounds. Interestingly, however, when tone duration and/or Inter Tone Interval (or Inter Stimulus Interval, ISI) increased within the tone sequences, a major improvement in terms of accuracy was observed particularly in amusic individuals but also in controls participants. In line with Study 1, it could be hypothesized that this improved encoding abilities could be associated to more a efficient connectivity between the right VLPFC and STG allowing the construction of an appropriate memory trace of the sounds (due to long-enough presentation of the to-be-encoded tone) in the amusic brain (see Figure 23).

The following steps of the processing of information in memory (retention and retrieval) were investigated in Study 1 and 2. In Study 1, by analyzing oscillatory synchronization in the gamma frequency during the retention period of the short-term memory task, we observed enhanced gamma-power in the right DLPFC in control participants in comparison to amusic participants. This finding was in agreement with previous data showing that high-level representations of task-relevant information are reflected in the gamma-frequency (Jensen et al., 2007), and that increased gamma-frequency power in fronto-temporal areas is related with short-term memory processing (Kaiser et al., 2003). Moreover, this finding is in line with previous data showing a strong implication of the DLPFC in the retention of information in short-term memory (Curtis and D'Esposito, 2003; D'Esposito, 2007; Logie and D'Esposito, 2007; Petrides, 1991, 1995a) and more specifically for pitch memory (Peretz and Zatorre, 2005; Zatorre et al., 1994). This recruitment of the right DLPFC during retention of tonal information was less observed in amusic individuals (in comparison to controls), thus providing another evidence of altered processing of melodic information in the amusic brain. This altered recruitment was also observed in Study 2 that used fMRI to investigate the cerebral correlates of auditory short-term memory for tonal and verbal materials. When contrasting brain activations related to the retention of tonal information as compared to the encoding of this information, we observed enhanced brain activation in the right DLPFC in controls. However, this area was not activated in amusic individuals as revealed by the direct comparison between amusic participants' and control participants' functional data, reflecting altered high-level representations and/or altered manipulation of task-relevant information in memory in the amusic brain.

Interestingly, when investigating brain activation related to the retention of verbal information as compared to tonal information, we observed that both amusic and control participants showed increased activation in the left VLPFC. This result is in line with numerous neuroimaging studies that have hypothesized this region as supporting subvocal rehearsal mechanisms of verbal short-term memory (Awh et al., 1996; Fiez et al., 1996; Gruber and von Cramon, 2003; Paulesu et al., 1993; Ravizza et al., 2004; Schulze and Koelsch, 2012). Moreover, it suggests that, as observed behaviorally, short-term processing for verbal material in the amusic brain is preserved and recruits similar networks to those observed in controls. Moreover, in a broader context, observing similar brain activations between amusic individuals and control participants for verbal material but not for tonal material suggests that memory for tones and words do not recruit the same networks.

Finally, when directly contrasting brain activity related to tonal and verbal memory (Tonal > Verbal), we observed that amusic participants activated the right VLPFC: the activation in this region was negatively correlated with amusic individuals' performance in the memory task for tones. This observation suggests that the recruitment of the rVLPFC during the retention of tonal information could be considered as a marker of the impaired memory abilities for tones in congenital amusia. By considering this differential recruitment of frontal areas between amusic and control participants for tonal material (amusic individuals recruiting mostly the VLPFC and controls the DLPFC), we have proposed that additionally to the deficits previously observed in the encoding steps of short-term

memory (based on the MEG data of Study 1), amusic individuals are also impaired in the retention steps of memory for pitch. This hypothesis is supported by numerous fMRI studies that have suggested a differential role of the DLPFC and VLPFC in short-term memory (D'Esposito, 2007; Logie and D'Esposito, 2007; Owen, 2000). While the VLPFC is known to play an active role in low-level encoding strategies, the DLPFC has been hypothesized to mediate more complex types of processing such as the retention and manipulation of complex material (D'Esposito, 2007). Observing that for tonal material, amusic individuals recruit mostly brain areas that are supposed to be more strongly involved in low-level memory processes (encoding for the VLPFC), and observing that altered responses in regions requiring high-level memory (monitoring and manipulation for the DLPFC) mechanisms let us argue that amusic individuals' short-term memory deficit is related to the transformation of tonal information into high level memory representation (see Figure 23).

Finally, the retrieval of the information from memory was studied in Study 1 by exploring the source reconstruction of the event-related fields elicited by the changed tone in the second melody of the short-term memory task. We observed activity in bilateral STG and in the VLPFC in the control group (note that this recruitment was less extended in amusic individuals). Observing that the same network is recruited during encoding and retrieval of melodic information is in line with numerous research showing the role of these areas in music perception and memory (Gaab et al., 2003; Griffiths et al., 1999; Griffiths et al., 2000b; Janata et al., 2002a; Janata et al., 2002b; Peretz and Zatorre, 2005; Tillmann et al., 2003; Zatorre et al., 2002; Zatorre et al., 1994). The data from Study 1 revealed an abnormal recruitment of frontal (VLPFC) and temporal (STG) regions during the retrieval of melodic information in the amusic brain, - data that further reflect the functional correlates of the deficit in short-term memory of melodic information in congenital amusia. All together, Studies 1, 2 and 3 have improved our understanding about the behavioral and neural correlates of the pitch perception and short-term memory deficits in congenital amusia. These studies suggested that in the context of a short-term memory task, the amusic brain is impaired in all processing steps, that is encoding, retention and retrieval of melodic information. In the second Axis of this PhD, we wanted to benefit from these novel findings to explore, with two different approaches, the potential tools of rehabilitation of this developmental disorder.

9.1.2. Axis 2: Boosting pitch perception and memory in congenital amusia

Exploiting amusic individuals' implicit knowledge of the musical system

Study 4 investigated whether amusic individuals' short-term memory impairments could be attenuated by exploiting their implicit knowledge about musical structures. Based on a previous study (Schulze et al., 2012) showing that non-musicians' implicit knowledge of musical regularities can improve their short-term memory for tone information (Schulze et al., 2012) and based on studies reported in Chapter III showing some preserved implicit musical processing in amusic individuals, we investigated whether implicit knowledge of musical structures can influence amusic individuals' short-term memory for pitch material. This was done by investigating amusic and control participants'

short-term memory abilities for tone sequences that respected musical regularities (tonal sequences) as compared to tone sequences that did not respect these regularities (atonal sequences). As previously reported for non-musician participants, the control participants showed better performance for tonal sequences than for atonal ones. However, this was not observed for the amusic participants. These results suggested that tonality did not help amusic participants to overcome their short-term memory deficit for melodies.

However, even though tonality did not improve amusic individuals' short-term memory performance in comparison to atonal sequences, some potential benefit of tonality was observed in their response times data, suggesting that amusic individuals processed the tonal information at least to some extent. For both amusic and control participants, response times were faster for tonal sequences than for atonal sequences. We interpreted this finding as supporting the hypothesis that amusic individuals process the tonality of the material to some degree, which then influences short-term memory performance for the tone sequences (as revealed in response times). We have assumed that this tonal benefit observed for melodies in controls (and the potential benefit observed in amusic individuals) can be related to results reported by Dowling (1991), suggesting that the various components of a melody (interval, contour, and tonality) are encoded as an integrated whole (Dowling, 1991). This might lead to a stronger memory trace for tonal melodies than for atonal melodies (as observed in Schulze et al. 2012), with atonal melodies missing the benefit of the integration in a tonal framework. This hypothesis thus further predicts that amusic individuals' short-term memory abilities might be improved for more complex tonal material (e.g., chord sequences or harmonized melodies, in contrast to single melodic lines as used in the present study). However, whether this effect of long-term knowledge impacted either encoding, retention or retrieval of the information is not determined yet and will need future investigations.

Investigating the benefit of multisensory integration

Finally, the last study of the present PhD work investigated whether amusic individuals' pitch encoding abilities could be boosted via multisensory integration. As presented above, results of Study 1 and 3 have suggested that pitch processing deficits in congenital amusia could be mediated by strong impairments in the construction of memory traces for short tones (encoding). Along these lines, Study 5 aimed to explore how and whether congenital amusic individuals' pitch abilities could be boosted by using a new approach in the domain that consists in facilitating amusic individuals' pitch encoding with multisensory integration. By using a simple speeded detection task for auditory, visual or audio-visual cues, we observed that both amusic individuals and controls exhibited audio-visual facilitation (shorter response times for audio-visual stimuli than for either auditory or visual stimuli) and no differences were observed between the two participant groups in terms of facilitation. This effect suggested that both groups of participants are able to encode better multimodal information than uni-modal ones during the first steps of the perceptual process, thus facilitating the early processing of either visual or auditory information. As we observed comparable audio-visual facilitation in amusic

individuals and control participants, we argued that multisensory integrating systems for visual and auditory information are preserved in the amusic brain and thus permit the exploration of such bimodal benefits on amusic individuals' abilities in processing pitch material. To this aim, amusic and control participants performed a pitch change detection task for which amusic individuals were previously shown to be impaired (Hyde and Peretz, 2004) in comparison to control participants. Amusic and control participants were required to detect a pitch change (with different pitch interval sizes) within a sequence of five identical tones that was presented either only auditorily (A condition), or with simultaneous uninformative visual information presented on the screen (AV condition). As expected, amusic individuals showed decreased performance in comparison to controls for the small pitch interval sizes, but not for the large interval size and both groups exhibited decreased performance with increasing task difficulty (decreasing Pitch Interval Size). However, no differences in terms of performance were observed between the unimodal and bimodal conditions in terms of performance. This suggests that visual uninformative synchronous stimulation did not help amusic participants to improve their pitch processing abilities.

Interestingly, even though the AV condition did not improve amusic individuals' performance for the pitch detection task, in comparison to the A condition, some benefit of the multisensory condition was observed in the response times data, suggesting that amusic individuals and controls used the visual information in the auditory pitch change detection task to some extent. Furthermore, this effect was related to task difficulty and varied as a function of the participant group: control participants showed the AV benefits for the difficult (small) interval size and amusic participants showed the AV benefit for the easy (large) interval size. This finding was line with several pieces of evidences suggesting that multisensory interactions depend on the participant's performance when the modalities are presented in isolation (Caclin et al., 2011; Hairston et al., 2003; Rouger et al., 2007; Serino et al., 2007) and was observed in particular when the uni-sensory performance (for the task-relevant dimension, here pitch) is poor (difficult task). This result suggests that to observe a benefit of bimodal stimulation, participants have to be in an appropriate range of performance. However, the results presented in Study 5 concerned only reaction times data, and further work is thus needed to explore whether this benefit could be transposed to amusic individuals' performance for both low level (sensory encoding) and high-level processes (such as memory processing) in congenital amusia. Observing such improvements would have implications for the development of training and rehabilitation programs of impaired musical functions in this disorder.
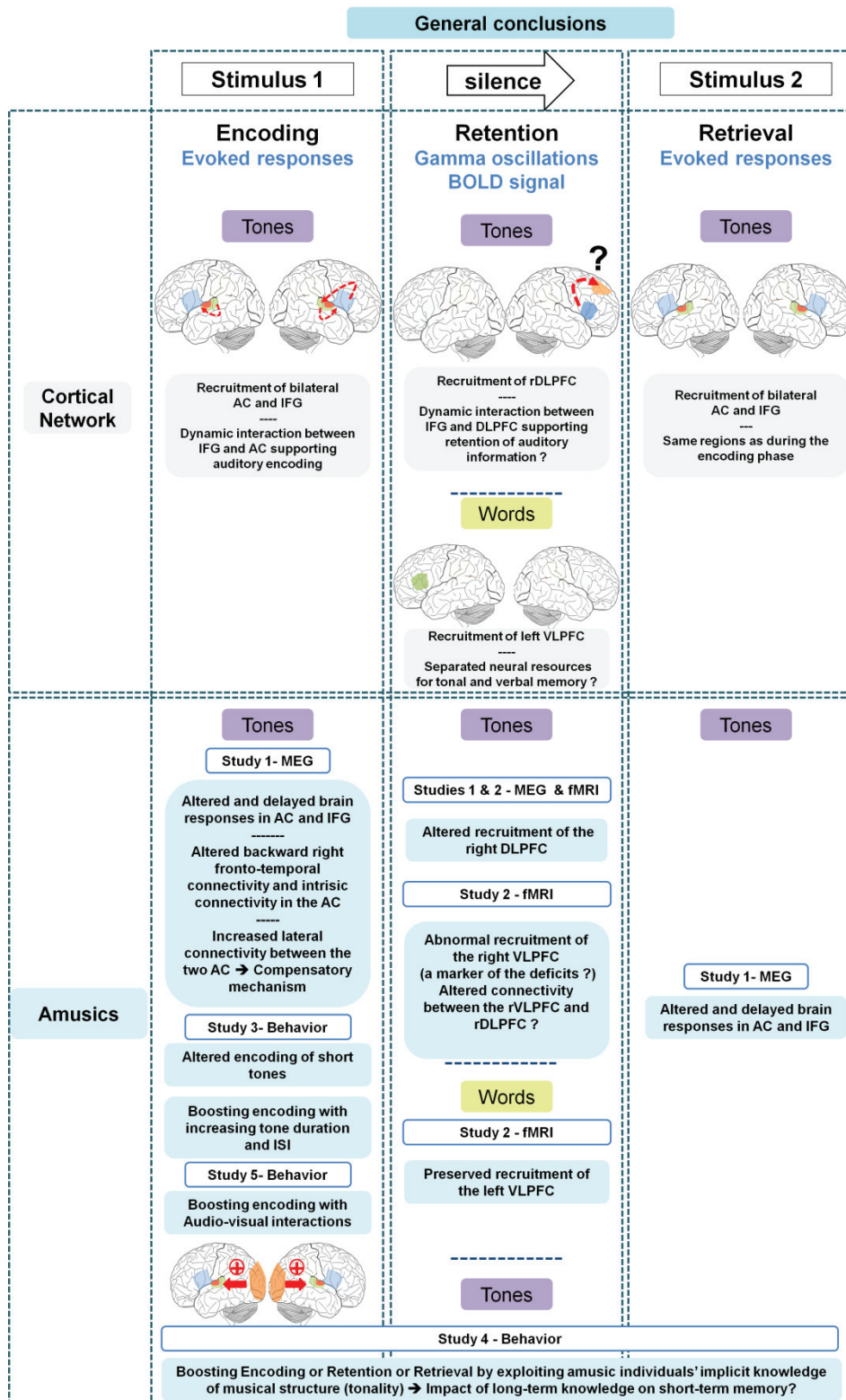
Figure 23: Summary of the general conclusions of studies 1 to 5 presented as a function of encoding, retention and retrieval phases of the short-term memory task. Grey Boxes: conclusions that could be made for the normal functioning brain; Blue boxes: conclusions that could be made for the amusic brain. The question marks "?" illustrate hypotheses that have emerged from our results.

All together, the five studies presented in this PhD work have improved our understanding of the behavioral and cerebral correlates of congenital amusia and have paved the way towards some rehabilitation strategies of this disorder. These general conclusions are summarized in Figure 23.

Moreover, our work has allowed making several conclusions concerning more general questions about short-term memory. In the following section, we will thus elaborate on these conclusions and propose perspectives for the study of each processing step underlying auditory short-term memory (Encoding, Retention, Retrieval) in order to 1) improve our understanding of behavioral and functional underpinnings of the disorder of congenital amusia, 2) further our comprehension of auditory perception and memory in general 3) help the design of rehabilitation tools for this disorder that could further be transposed to other pathologies that share deficits on the processing of the pitch dimension in the auditory signal (such as cochlear hearing loss).

9.2. Perspectives: Encoding of pitch information

9.2.1. *Pitch encoding and fronto-temporal connectivity*

Data from Studies 1 and 3 suggest that altered encoding of short tones might lead to the deficits observed in congenital amusia. In Study 1 (see Chapter IV), using a classic short-term memory task with melodies, we observed functional abnormalities during the encoding of melodic information: the amusic brain elicited abnormal (decreased and delayed) N100m components in bilateral Inferior Frontal Gyrus (pars opercularis) and Superior Temporal Gyri. Furthermore, these abnormal responses were associated with connectivity as revealed by Dynamic Causal Modeling (David et al., 2006). This observation supports the assumption that amusic individuals might show delayed and altered functioning in the early steps of memory processing (encoding, echoic memory, building of the pitch memory traces, see Study 3), leading to the observed impairments for the processing of the pitch dimension (for both pitch discrimination and memory tasks).

Our data suggest that *causal* interactions between the right inferior frontal cortex and the right auditory cortex support the encoding of pitch patterns. This observation is in agreement with the contemporary hypothesis in neuroscience suggesting that both perception and memory processes are supported by dynamic interactions between different brain areas (Friston et al., 2003; Garrido et al., 2007; Rauschecker and Scott, 2009; Varela et al., 2001).

For the case of auditory perception, our results are in line with Garrido et al. (2007) study who proposed that auditory evoked brain responses are mediated by interactions between fronto-temporal cortical areas and suggested that late event-related components (> 100 ms) are especially driven by backward connections (as observed in Study 1). They proposed that forward and backward connections have a differential role in auditory perception: backward connections are supposed to play a critical role in providing top-down predictions of bottom-up sensory input (Friston, 2003), while forward connections rather support only bottom up processes (Penny et al., 2004). This hypothesis was notably supported by studies suggesting that cortical feedback is necessary in recognition tasks (Lamme and Roelfsema, 2000) or discrimination tasks of visual stimuli (Hupe et al., 1998) that require the comparison of a previously presented stimulus to a new stimulus arriving in the perceptual system.

By considering the role of forward and backward connections, our work thus suggests that the encoding of melodic contour patterns, which need to keep a trace of a previously presented stimulus in memory in order to compare it (top-down) to a sound occurring later (bottom up), are supported by backward connections between inferior frontal regions and auditory cortices.

In Study 1, we have proposed that this altered functioning of the backward connections in the amusic brain might lead to an impairment in the construction of short-term memory traces of tones, and could explain other pitch-related impairments in congenital amusia, such as their pitch discrimination deficit. Discrimination and memory processes are deeply interrelated as both processes require the construction of a relation between the memory traces of a previous stimuli and a stimulus belonging to the present. Study 3 tested the contribution of impaired pitch encoding to the pitch discrimination and memory impairments observed in congenital amusia. By manipulating stimulus characteristics, Study 3 has demonstrated that amusic individuals exhibited strongly impaired performance (as compared to their matched controls) for pitch discrimination and short-term memory tasks when the time dedicated to encode the information was short. However, amusic individuals were able to reach performance comparable to that of controls for the two experiments when the time dedicated to encode the information was long enough (here from 350 to 700 ms). Based on these results, we argue that pitch perception and memory deficits in congenital amusia could be mediated by strong impairments in the construction of memory traces for short tones, - a phenomenon that reflected in delayed functioning of the auditory cortex and altered connectivity between frontal and temporal regions.

However, it should be noted that no differences between amusic individuals and controls were observed when analyzing the listening (without task) and encoding (in a short-term memory task) of melodies in Study 2 (i.e., the fist melody of the pair). When contrasting the brain responses related to the encoding of melodies, both groups recruited bilateral temporal regions covering both primary and secondary auditory areas. While we have discussed in Study 2 that these discrepant findings could arise from the different methodologies used in Studies 1 and 2, an alternative hypothesis needs to be considered as in Study 2 up to now, we have not performed connectivity analyses yet. In Hyde et al.'s (2011) study, altered fronto-temporal connectivity in the amusic individuals' brain was observed during passive listening. Although this study did not explore causal interactions, it suggests that altered functioning in terms of network communication exists in the amusic brain, even for low-level processing of pitch information.

This hypothesis can be investigated with a high temporal resolution method (such as MEG), notably by comparing the cerebral correlates of passive listening for pitch in amusic individuals and controls to high-level processing of the same stimuli (i.e., for memorization). We have started such investigation with the data of Study 1 (unpublished data, see Figure 24). We briefly present here data from one amusic participant (age, 28 years; education, 18 years; musical education 0 year) and one control participant (age, 35 years; education, 15 years, musical education 0 year) who have passively listened to a train of monotone and isochronous tones (using G3, tone duration 250 ms, the tone in the

middle of the tone set of the first sequence in Study 1; separated by a 2s-ISI). This bloc of passive listening was added after the acquisition of the memory task data (note recordings of passive listening will be also acquired after a memory task).
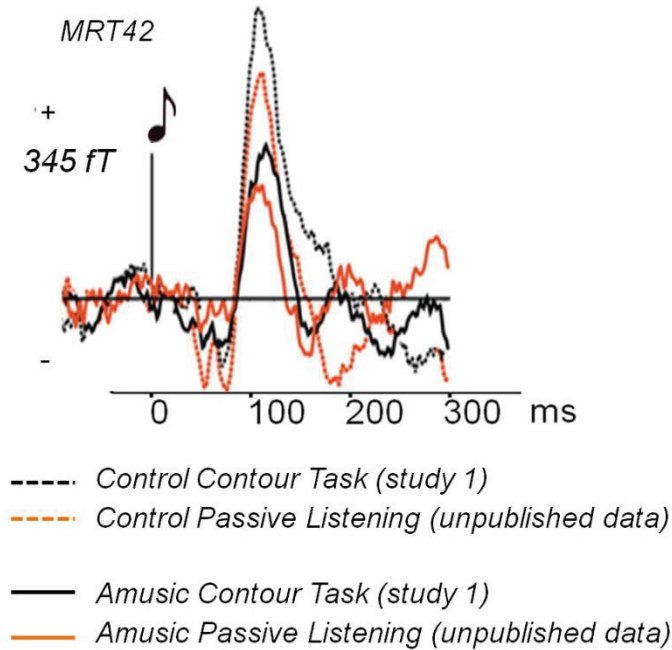

Figure 24. ERF at a right hemisphere MEG sensor (MRT 42, 'MEG Right Temporal) for the first tone of the first sequence in the Contour task of Study 1 and the passive listening of single tones for one control and one amusic participant.

In these two participants, we observed a modulation of the N1m amplitude as a function of the task: the N1m amplitude was larger for the *Contour Task* than for *passive listening*. This effect is well known and could be associated to the involvement of additional components underlying attentional processes in the Contour task in comparison to passive listening (Giard et al., 1988). Moreover, we observed that the amusic's brain response was decreased in comparison to the control's responses for both signals, suggesting that differences between the two groups could be observed even during passive listening of pitch information, although the between-subject difference is less pronounced in the passive listening condition. These data are preliminary, and call, of course, for further investigation in a group of amusic individuals and their matched controls. By performing distributed source modeling of these signals and by investigating the potential differences in terms of connectivity between high-level processing (memory) and low-level processing (passive listening), we should be able to determine if altered and delayed responses associated to altered fronto-temporal connectivity could be observed even during passive listening of pitch in congenital amusia. Note that in this future study, we should also use the same material (a single tone comparison for the memory task and a

single tone as passive listening) just to make sure that if a difference is observed, it is not due to the fact that listeners know that there will be a full sequence following the first tone (as is Study 1).

Finally, it should be noted that to pursue the characterization of the deficits in this developmental disorder, further research investigating the encoding of other types of information (such as verbal material) are needed, notably by exploring if the delayed responses observed in Study 1 for tones, could be considered as a general impairment of auditory encoding or are specific for pitch.

9.2.2. *Pitch encoding and hemispheric specialization*

In addition to the description of the implication of dynamic interactions between distant cortical regions in pitch encoding, several results of our research suggest that the encoding of musical information in the brain is lateralized (hemispheric asymmetry while both are involved). Notably, in Study 1, the altered functional responses, along with the anatomical abnormalities, mainly involved the right hemisphere. As described in the introduction (see Chapter I, II, III), numerous studies have hypothesized a specific role of the right hemisphere in the processing of musical information. This was notably based on the observed musical deficits occurring after right hemispheric brain lesions (see Chapter III, section 3.1 pages 80 to 82), and suggested by neuroimaging studies in typical listeners showing that the right hemisphere supports the processing of pitch information (Griffiths, 1999; Griffiths et al., 1999; Hyde et al., 2008; Janata et al., 2002a; Janata et al., 2002b; Koelsch et al., 2005; Koelsch et al., 2009; Peretz and Zatorre, 2005; Schulze et al., 2009; Schulze et al., 2011a; Schulze et al., 2011b; Stewart et al., 2008; Tillmann et al., 2003; Tillmann et al., 2006; Zatorre et al., 2002; Zatorre et al., 1994).

It the present discussion section, we propose another approach that allows the investigation of the relative involvement of each hemisphere in the encoding of pitch information, notably by analyzing which brain regions of the involved cerebral network are disengaged in the processing. This can be investigated by analyzing alpha oscillations (8-13 Hz) in neurophysiological data (such as EEG or MEG). Recent findings have suggested that the amplitude of alpha-oscillations reflects the level of cortical inhibition (or the level of cortical excitability) (Klimesch et al., 2007). The alpha inhibition hypothesis proposes that a small-amplitude of alpha-oscillations is a signature of regions of active neuronal processing, whereas large-amplitude alpha-oscillations reflect the inhibition and disengagement of task-irrelevant cortical areas. The core phenomenon supporting this hypothesis is the alpha-amplitude suppression, also known as event-related desynchronization. Event-related desynchronization appears in sensory areas that are processing external stimuli and reflects a state of high cortical excitability (Palva and Palva, 2007). Event-related synchronization (increased amplitude) is observed in areas that are not related to the processing of the stimulus and reflects a state of inhibition (comparatively low excitability) (Meyer et al., 2013; Obleser and Weisz, 2012; Obleser et al., 2012). The basic principles underlying this inhibition hypothesis can be illustrated by considering the phase of oscillatory activity together with the level of excitation of single neurons. In this view, alpha-oscillatory activity is induced by inhibitory cells and reflects rhythmic changes between phases

of maximal and minimal inhibition. When investigating short-term memory for pitch, alpha oscillations have been shown to increase in the left hemisphere during the retention of information (van Dijk et al., 2010), thus suggesting that the left hemisphere is to some extent less involved in the processing of pitch than the right hemisphere (see also Obleser et al., 2012).

Along these lines, by investigating alpha oscillations during the first melody of a pair in a short-term memory task, we thus would be able to pursue the characterization of hemispheric specialization for pitch encoding. We have performed a first step of this investigation, by performing additional analyses of the data of Study 1 (poster, Albouy P, Caclin A and Tillmann B: *Altered alpha-activity related to impaired short-term memory of tone sequences in congenital amusia. Presented at BIOMAG 2012 conference, Paris, France, 26-30 August 2012, Figure 25*). To assess the relative involvement of each hemisphere in the encoding of pitch information in the short term memory task, we have explored if enhanced alpha-oscillations (synchrony) could be observed in cortical areas specialized in processing unrelated to our task (e.g., visual cortical areas), thus allowing participants to focus on auditory processing. In this additional analysis, we did not make the distinction between evoked and induced oscillations (in contrast to the analyses we have performed for the gamma oscillations, see Study 1) and assessed whether event-related alpha synchronization could differ between the two hemispheres as well as between tasks and groups in the data of Study 1.

We have performed the same preprocessing steps and analyses as described in Study 1 for gamma oscillations, but here we investigated a different frequency band (8-11Hz) and used a different time window of interest (i.e., to study the encoding of melodies 500 to 1500 ms after the onset of the first tone of the first sequence, see Study 1). To localize the neural sources in the frequency band of interest and explore the potential differences between tasks and groups, we applied an adaptive spatial filtering or beamforming technique (Dynamic Imaging of Coherent Sources, DICS, see Study 1).

For each trial, we estimated the power in the 9-11 Hz frequency band at each volumetric element for both active (encoding) and baseline periods (-1000 to 0) before the onset of the first tone of the first sequence), and we computed the log ratio between these two power values to determine the extent of the emerging alpha oscillations. In this preliminary analysis, after normalization of each brain in the MNI space, statistical comparisons were performed to compare groups for each task, with two sided t-tests, corrected for multiple comparisons using cluster-level statistics, as implemented in Fieldtrip.
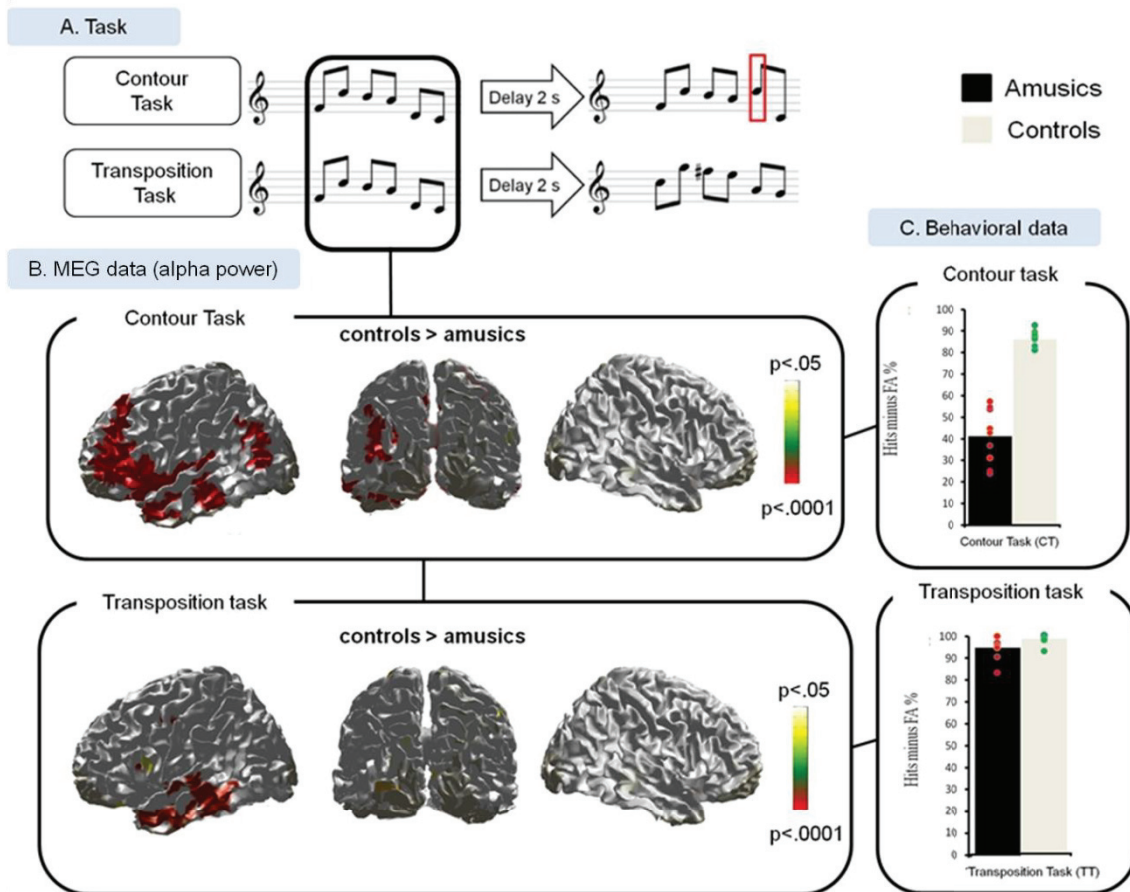
Figure 25: A. Examples of the musical stimuli used for different trials of the Contour Task (CT) and the Transposition Task (TT) of Study 1. The second melody of the pair changed by only one tone in CT (indicated by the red square) and was a transposition of the first melody in TT (example of the pitch-up condition). B. Cortical meshes showing regions where alpha-power was significantly superior in controls in comparison to amusic individuals for the first sequence of the pair. Red regions represent significant p-values for the comparison controls > amusic individuals (two sided t-tests corrected using cluster-based multiple comparison) in controls (upper panel) and amusic individuals (lower panel). Note that data were baseline-corrected using a -1000 to 0 ms time window before the onset of the first tone of the first melody. C. Behavioral performance for the contour Task (upper panel) and Transposition Task (lower panel) in terms of Hits minus FAs (see Study 1), presented as a function of the group (controls, amusic individuals); data of Study 1. Black: average performance of amusic individuals; Grey: average performance of controls; circles: participants' individual performance.

For the contour task of Study 1, controls showed increased alpha synchronization (relative to baseline) in the left DLPFC and VLPFC (BA9/46, BA 44, BA45) in comparison to amusic participants along with increased alpha synchronization in the left auditory cortex, left temporal lobe, left Middle Temporal Gyrus and left supramarginal gyrus. Interestingly, no differences were observed between the two groups in the right hemisphere. For the Transposition Task, which was the easier control task of Study 1, controls showed increased alpha synchronization only in the left temporal pole and in the

MTG in comparison to amusic individuals. Note that no differences were observed for the amusic individuals > controls contrast.

These results, interpreted in the context of the alpha inhibition hypothesis, suggest that during the encoding of the difficult contour task (CT) controls (in comparison to amusic individuals) exhibit reduced excitability in several regions of the left hemisphere that could be thus considered as unrelated to the task (left auditory-fronto-parietal areas). This group difference also suggests that in amusic individuals, the left hemisphere could be involved in the processing of pitch information, by playing a role in compensatory mechanisms to the altered mechanisms in the right hemisphere (see Study 1). Interestingly, these group differences were less extended for the easier task, thus suggesting that the degree of inhibition of the left hemisphere is related to task difficulty.

These preliminary results on alpha-oscillation are encouraging because they are consistent with the results presented in Study 1, notably showing increased gamma oscillations in the right hemisphere during the retention of the contour information in controls. Thus observing increased alpha oscillation (reduced excitability) during the encoding of melodies and increased gamma oscillations (increased excitability) in the right hemisphere during the delay in controls suggest that mechanisms underlying pitch processing mainly involve the contribution of right cortical areas. In addition, this analysis adds a new piece of evidence for the understanding of congenital amusia. It suggests that the neurophysiological mechanisms underlying the encoding processing of pitch material differ between the two groups, in particular for the relative involvement of the two hemispheres. This conclusion is also in line with the DCM analysis showing altered connectivity in amusic individuals in the right hemisphere.

9.3. Perspectives: Retention of pitch and verbal information

To explore the cerebral correlates of the deficits in pitch processing in congenital amusia, Studies 1 and 2 have investigated the retention mechanisms of pitch information with two different methods. Interestingly, both studies have highlighted the role of the right DLPFC during the storage (or manipulation) of pitch information.

9.3.1. Cortical networks supporting the retention of pitch and verbal information

Study 2 investigated short-term memory for tonal material and verbal material in congenital amusic individuals and their matched controls with fMRI. In addition to investigating pitch memory, this study thus assessed whether amusic individuals' cerebral network underlying normal-level performance with verbal material is similar to the network engaged by matched control participants.

For tonal memory, amusic individuals exhibited abnormally decreased brain responses in the right DLPFC and abnormally increased activation in the right VLPFC. In contrast, for verbal memory, amusic individuals showed brain activations similar to the brain activations of controls, with both groups recruiting the left inferior frontal regions during the retention of verbal information.

Chapter IX: Perspectives: Retention of Pitch and Verbal Information

These findings suggest separate neural resources for verbal and tonal short-term memory in the amusic brain, with a specific dysfunction for tonal material.

For controls however, the dissociation between verbal and tonal memory were less conclusive (see also Schulze and Koelsch, 2012). While the investigation of the cerebral correlates of the retention of each material revealed lateralization specificity (right DLPFC for tonal material and left DLPFC and VLPFC for verbal material), the direct comparison between the two materials revealed no activation differences. One might thus conclude that the maintenance of verbal and tonal information share the same network. However, several questions remain unresolved: while verbal and tonal information seem to involve the same cortical structures, few is known about the dynamic of the processing and how these regions interact in terms of network activity. Further investigation is thus necessary to understand the dynamic of the retention processes for verbal and tonal memory. To this aim, we are currently performing a new study with high-resolution technique and high temporal resolution, that is intracranial electroencephalography.

To date, we have recorded 19 drug-resistant epileptic patients, candidate for surgery with either left lateralized, right lateralized or bilateral implantations in the auditory cortex and in the DLPFC, VLPFC and PPC. These patients performed short-term memory tasks that contained either tone sequences, word sequences or timbre sequences (material from Study 1 and 2, and from Tillmann et al., 2009). We added a passive listening session of the tone sequences to assess the memory-specificity of the effect for this material. Additionally, for all materials, we manipulated the strength of the memory load (number of elements in the to-be-encoded sequence from 3 elements to 6 elements) as well as the duration of the silent retention period between the to-be-compared sequences (from 2000 ms to 8000 ms).

The data of this study are still under analysis but will permit, by the comparison of each material, to precisely define the dynamic of the retention processes of tonal information, and compare it to the dynamic of the retention processes for other types of material (verbal, timbre). Interestingly, thanks to the quality of the iEEG signal (Lachaux et al., 2012), we will be able to link the functional data for each trial with the behavioral performance, thus allowing us a posteriori to determine functional markers that could predict performance in our tasks of interest (see Figure 26). Note that in addition to the cerebral correlates of retention of auditory information, the data of this study will also allow a more precise investigation of the cerebral correlates of encoding and retrieval of the auditory information.
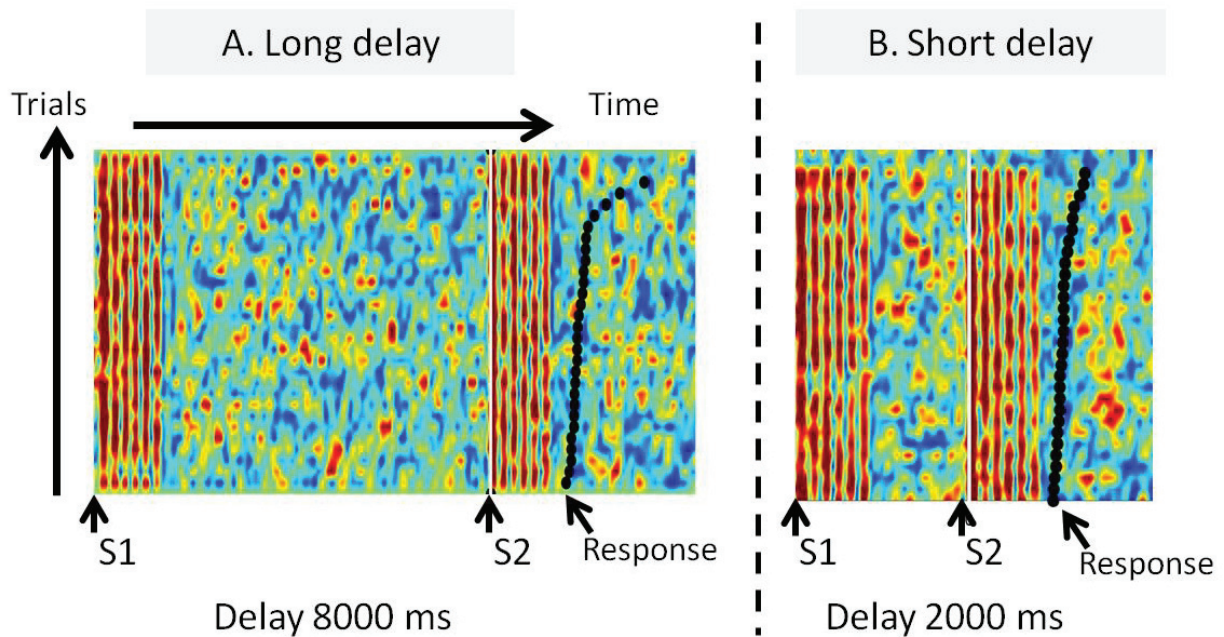
Figure 26: Illustration of raw data obtained in a participant with an electrode in the right Heschl's Gyrus. Data are represented for two short-term memory tasks for two 6-tone sequences separated by a silent delay (same characteristic as stimuli used in the Contour Task in Study 1) for (A) a long retention period and (B) a short retention period. The time lines or trials are shown with black arrows: S1, onset of the first tone of the first sequence; S2, onset of the first tone of the second sequence; Response, Participant response times. Plots correspond to single-trial gamma (30-40Hz) power with trials sorted according to RTs (fastest to slowest response times). Black circles show response times for each trial (i.e., latency of button press after the end of the S2 sequence). As observed here, this method allows us to explore the link between brain oscillations (as well as event-related responses) with behavior (performance, responses times) at the level of each trial.

9.3.2. Linking BOLD and electrophysiological signals

Comparing the results of Studies 1 and 2 is also interesting from a methodological point of view as the comparison allows hypothesize that the BOLD signal and the electrophysiological signal (in the present case gamma oscillations) measure the same cerebral mechanisms. It has been suggested that the fMRI method (with high spatial, but poor temporal resolution) and electrophysiological measures (with poor spatial, but high temporal resolution) are complementary (Lachaux et al., 2007). In the fMRI signal, the activity of a small neural population is measured as hemodynamic values during a relatively long time window (over seconds), in contrast to electrophysiological approaches, where the signal corresponds to a composite high-temporal resolution signal with a complex organization in time and frequency (Lachaux et al., 2012). Over the last decade, several advances in the linkage between the two methods came from animal studies combining simultaneous hemodynamic and spikes/local field potential measurements (Kayser et al., 2004; Kim et al., 2004; Logothetis et al., 2001; Niessing et al., 2005).

These studies have shown a close correspondence between the BOLD signal and the gamma band component of the Local Field Potential, in both monkey and cat visual cortex. In humans, Lachaux et al. (2007) combined fMRI and intra-cranial EEG recordings of the same epileptic patients during a semantic decision task. They observed a close spatial correspondence between the regions of fMRI activations and the recording sites showing EEG energy modulations in the gamma range. Based on these results, the observed convergence between the data of Study 1 (MEG) and Study 2 (fMRI) strongly supports that the increase of local gamma activity is related to an increase of the BOLD signal (Lachaux et al., 2012; Lachaux et al., 2007; Niessing et al., 2005).

9.4. Perspectives: Retrieval of pitch information

In the present PhD work, only Study 1 has investigated the cerebral correlates of the retrieval of pitch information. Source reconstruction of the Event Related Fields elicited by the changed tone in the second sequence of the contour task (the difference wave between different trials and same trials) allowed us to observe activity in bilateral STG and pars opercularis of the IFG (BA 44) in the control group. This result suggests that the same network is recruited during encoding and retrieval of melodic information, and is in line with numerous studies showing the role of these areas in music perception and memory (Owen, 2000; Gaab et al. 2003; Griffiths, 2001; Zatorre et al., 1994, 2002; Peretz & Zatorre 2005). Additionally, this data revealed an abnormal recruitment of frontal (IFG) and temporal (STG) regions during the retrieval of melodic information in the amusic brain, - data that further reflect the functional correlates of the deficit in the STM of melodic information in congenital amusia.

While this study furthers our understanding about the cerebral correlates underlying the retrieval of melodic information, it did not inform us about the dynamic of this network. Further work is thus needed to understand whether frontal and temporal cortices causally interact in retrieval.

To explore this question, we have performed with the data of Study 1 a new DCM analysis based on the retrieval of melodic information (poster, Albouy P, Mattout J, Sanchez G, Tillmann B and Caclin A (2013) -*Altered encoding and retrieval of melodies in congenital amusia: Dynamic Causal Modeling of MEG data* presented at the 19th Annual Meeting of the Organization for Human Brain Mapping June 16-20, 2013 Seattle, WA, USA). To this aim, we have explored the group differences in the event-related responses associated to the changed tone in the second sequence of the contour task. As in Study 1, to characterize the two participant groups with a high signal-to-noise ratio and to investigate the putative differences in effective connectivity between the groups, we compared and modeled the difference between the grand average data of the controls (average over the nine participants) and the grand average data of the amusic individuals (average over the nine participants). Note that this analysis was realized only on correct different trials with the same number of trials in each group (adjusted for each matched participant). The models were based on the same network architecture as used for the encoding of the melodic information (see Study 1). We assumed four sources, modeled as equivalent current dipoles (ECDs), over left and right primary auditory cortices

(A1), left and right pars opercularis of the inferior frontal gyrus (IFG). Using these sources, we constructed the following DCM: An extrinsic input entered bilaterally to the primary auditory cortices (A1), which were connected to their ipsilateral IFG. Inter-hemispheric (lateral) connections were placed between left and right A1. All connections were reciprocal (i.e., connected with forward and backward connections or with bilateral connections) (see Figure 27).

Given this network architecture, we used a factorial design and a family inference approach to assess various types of connection modulations to explain the group differences in auditory evoked responses (control participants defined the baseline). Factor 1 refers to the modulation of intrinsic connectivity in bilateral auditory cortices (and includes two families (or levels), corresponding to models where these intrinsic connections were modulated (or not) between the two groups). Factor 2 refers to the modulation of lateral connections between the two auditory cortices (two families: models which included modulation or not). Factor 3 refers to the type of connections between auditory and frontal areas that are modulated (namely forward, backward, or both forward and backward connections) or not (four families). Factor 4 refers to the hemispheric location of the above modulated connections, either in the right hemisphere, the left hemisphere, or both (3 families). We thus fitted and compared 48 models. Assuming equal prior probabilities over models, we used Bayesian model selection (BMS) to compare model families (Penny et al., 2010). This rests upon the free energy (or approximate marginal likelihood or evidence) for each model, and yields a posterior probability associated with each model family.
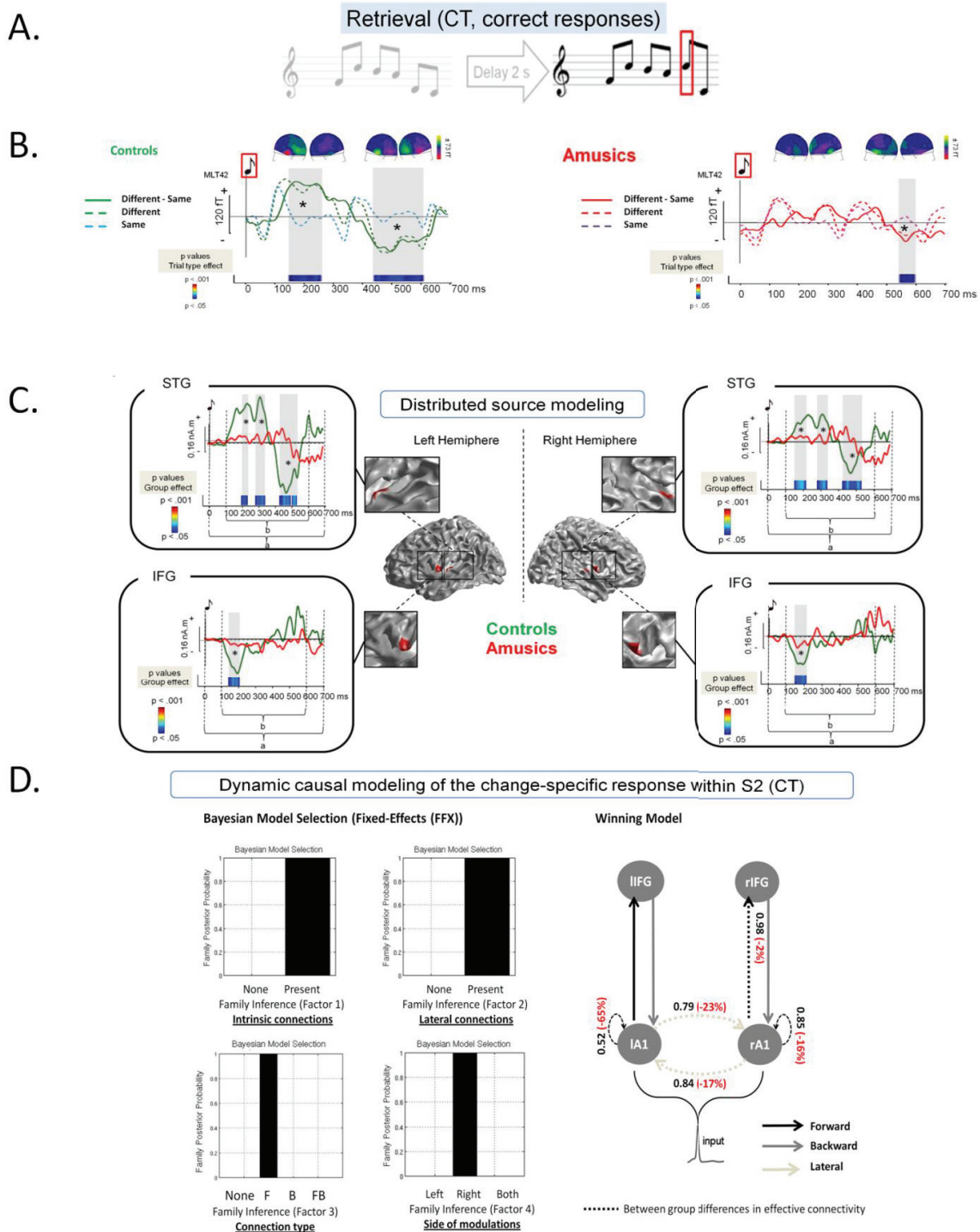
Figure 27 A. Example of the musical stimuli used for different trials of the Contour Task (CT) of Study 1. The second melody of the pair changed by only one tone in CT (indicated by the red square) B. Grand average of a left temporal MEG sensor (MLT42) for a 0 to 700 ms time window after the onset of the changed tone in S2 for the Contour Task for each group and each type of trial. For controls (left). Green dotted line: Different Trials, correct responses; blue dotted line: Same Trials, correct responses; green plain line: Difference Wave (Different trials – Same Trials for correct responses). For amusic individuals (right). Red dotted line: Different Trials, correct responses; purple dotted line:

Same Trials, correct responses; red plain line: Difference Wave (Different Trials – Same Trials for correct responses). Two-sample t-tests were performed at each time sample on sensor amplitudes in the 0 to 700 time window in the two groups of participants. p-values are reported across time in the lower panel with blue for p < .05; green for p < .01; and red for p < .001. Note that only effects lasting longer than 15 ms were reported (ref Guthrie & Buchwald).

C. Source reconstruction of the brain responses specifically evoked by the changed tone in S2. Cortical meshes show bilateral regions that were significantly different from baseline (as indicated by the brown areas). These regions were the bilateral STG as well as the opercular part of the Inferior Frontal Gyrus (see Table 2). The surrounding panels correspond to the grand average of source data for each region and for the time window where the inversion was performed (0 to 700 ms after the changed tone onset, as indicated by a) for the control group (green) and the amusic group (red). Two sample t-tests were performed at each time sample and for each region on source amplitude in the 100 to 600 ms time window (as indicated by b) in the two groups of participants. p-values are reported across time below the source amplitudes with blue for p < .05; green for p < .01; and red for p < .001. Note that only effects lasting longer than 15 ms were reported.

D. DCM results. Results of the four family-wise inferences. For each inference, the posterior probability of each model family is depicted. For each comparison, the family associated with a high posterior probability (p > .99) could be retained as the winning family. As each comparison yields a clear winning family, we could identify a winning model.

Winning model. Dashed arrows indicate modulated connections (i.e., connections that differ between groups) and solid arrows indicate fixed connections. Significant changes in effective coupling between controls and amusic individuals are specified (in black: amount of coupling change between groups; in red: corresponding relative coupling with amusic individuals coupling expressed in % of control coupling).

Posterior estimates obtained with the winning model enabled us to conclude that, compared to controls, amusic participants showed an abnormally decreased lateral connectivity between the two auditory cortices, decreased intrinsic modulations in both auditory cortices, and decreased forward connectivity between the right inferior frontal gyrus and the right auditory cortex. Based on this result, we thus argue that while encoding and retrieval recruit the same brain regions, the dynamic of interaction between these regions is different. Based on Garrido et al. (2007), we suggest that encoding of auditory information is supported by top-down processes requiring feedback connections between the inferior frontal gyrus and the auditory cortex, while retrieval (which involve also online comparison with the preceding sound) could be considered more as a cooperation of bottom up and top-down processes involving interactions between the two auditory cortices as well as forward connection between the auditory cortex and the inferior frontal gyrus.

9.5. Conclusions

9.5.1. Functional and anatomical data

All together, this PhD work suggests that amusic individuals' pitch deficits are related to impairments in the construction of short-term memory traces and in altered retention of pitch information for a short time period. These deficits are associated altered functioning of the auditory cortex along with altered recruitment of fronto-parietal regions in each step of the processing of short-term memory: that is encoding, retention and retrieval of information. Additionally, our findings have also allowed improving our understanding of the normally functioning brain. An overview of the results discussed in the present general discussion section is depicted in Figure 28.
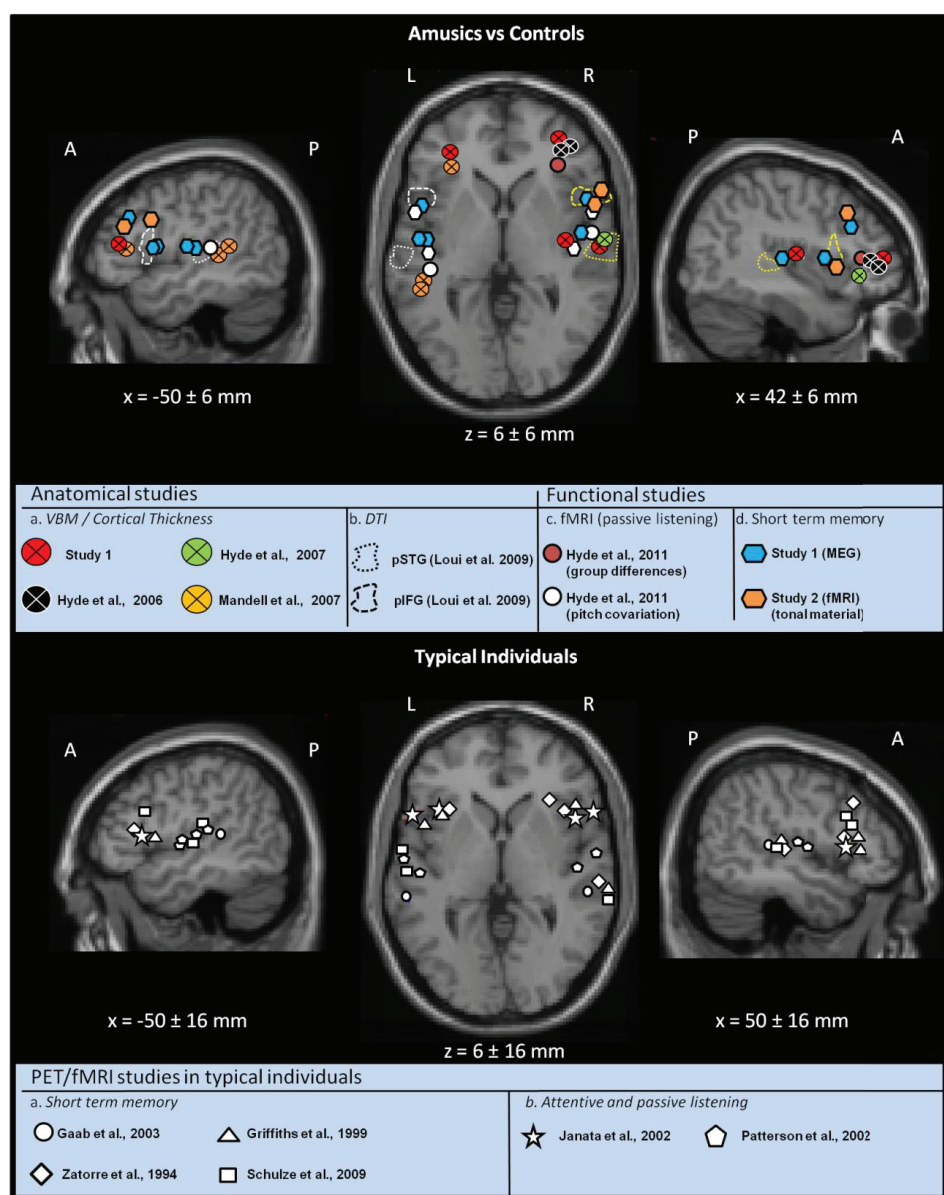

Figure 28. Upper panel: *Brain regions observed in anatomical and functional studies comparing control participants and amusic individuals (white: no group difference; color: group difference). a)*

Coordinates of the anatomical group differences observed with VBM (Hyde et al., 2006 (black); Mandell et al., 2007 (orange); Study 1) and Cortical thickness (Hyde et al., 2007 (green)) approaches. b) Approximate location of the posterior inferior frontal gyrus (pIFG) and posterior superior temporal gyrus (pSTG) seed regions used in the tractography study of the arcuate fasciculus by Loui et al. (2009) (bold dotted lines and plain dotted line, respectively). Note that these regions were manually selected by a neurologist who was blind concerning the protocol and previously reported activations. Group differences were observed only in the right hemisphere (yellow color). c) Coordinates of the functional group differences observed with fMRI in the right inferior frontal gyrus (Hyde et al., 2011, purple circle), whereas in the auditory cortices both amusic individuals and control exhibit similar BOLD activity to pitch variations (Hyde et al., 2011, white circles). d) Coordinates of the functional group differences during a short-term memory task (Study 1, MEG, blue hexagons, Study 2, fMRI, orange hexagones).

Lower panel: *activation peaks observed in studies using PET (Positron Emission Tomography) or functional MRI to investigate music processing in typical individuals. Coordinates of activations obtained with pitch material for a) auditory short-term memory tasks (Zatorre et al., 1994, diamonds; Griffths et al., 1999, triangles; Gaab et al., 2003, circles; Schulze et al., 2009, squares); b) attentive listening (Janata et al., 2002, stars) and passive listening (Patterson et al., 2002, pentagons).*

All activations are displayed on the single subject T1 template provided by SPM8. The figure displays activations with coordinates that were 6 mm up or down the central slice and 6 mm left or right of the lateral slices for the upper panel, and ± 16 mm for the lower panel. R = right; L = left; A = anterior; P = posterior).

9.5.2. *Perspective of rehabilitation*

Beyond the understanding of amusic individuals' perceptual and memory processes and their cerebral underpinnings, the present PhD work has also intended to pave the way towards rehabilitation tools for amusic individuals. Studies 4 and 5 have tried to boost pitch short-term memory and pitch discrimination, respectively, with two approaches that are different than approaches started to be used in other teams (singing lessons (Anderson et al., 2012), or passive listening in children (Mignault Goulet et al., 2012)).

By exploring implicit knowledge related to musical structure (that requires knowledge about the Western tonal system), in a short-term memory task (Study 4), we have shown that some implicit knowledge is preserved in congenital amusia and could be potentially used to improve their short-term memory abilities. However, we were not able to assess which step of the processing benefited from this knowledge, and this could constitute a next step of investigation with neurophysiological methods. Moreover, in the task used in Study 4, amusic individuals' performance was strongly impaired in comparison to controls, thus suggesting that the task was too difficult to observe benefit in terms of accuracy. Based on the conclusions made in Study 3, in would be interesting to test the same paradigm

with a more simple material (shorter sequence, slower timing etc.) that amusic individuals can encode efficiently.

Finally, Study 5 tried to boost pitch processing in congenital amusia with multisensory interactions. While amusic individuals (and controls) did not benefit from simultaneous visual information in terms of accuracy, we observed a benefit of audio-visual integration in terms of response times. More interestingly, this benefit was modulated as a function of task difficulty. These results are thus encouraging and call for future research that will lead to design a tool for congenital amusia rehabilitation.

Such rehabilitation programs could be based on a perceptual learning procedure that could be boosted using facilitatory multisensory interactions. Perceptual learning consists in an improvement of the discrimination of a perceptual attribute, after several training sessions with a particular perceptual task. It has been observed in all sensory modalities, for a variety of attributes (Gilbert et al., 2001). For example, and interestingly, training with tasks requiring pitch sequence comparison leads to improvements in pitch-contour (i.e. melodic) perception (Foxton et al., 2004a). Neurophysiological studies have revealed that perceptual training can lead to plastic changes up to the level of primary sensory cortices, even in adult subjects: converging evidence has been obtained from animal (Recanzone et al., 1993; Schoups et al., 2001) and human (Schiltz et al., 1999) studies. Such perceptual training procedures have been found efficient for the rehabilitation of perceptual deficits, for example using a task that is training contrast sensitivity in the case of amblyopia (Polat et al., 2004; Polat et al., 2009). Other successful approaches include auditory-based training procedures for the rehabilitation of language-learning impairment (Tallal, 2004). Finding similarly efficient procedures in congenital amusia could then be further transposed to pathologies that share deficits for the processing of the pitch dimension in the auditory signal, such as cochlear hearing loss (in particular for people receiving cochlear implants (Cousineau et al., 2010) or individuals with dyslexia (Kujala et al., 2003).

List of Publications

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., Schulze, K., Caclin, A., Tillmann, B., 2013. Does tonality boost short-term memory in congenital amusia? ***Brain Research,*** 2013 Sep 13. doi:pii: S0006-8993(13)01238-9. 10.1016

Albouy, P. Lévêque, Y., Hyde, K., Bouchet, P, Tillmann, B, Caclin, A*.,* 2013. Boosting pitch encoding with audio-visual interactions in congenital amusia. In preparation.

Albouy, P.*,* Cousineau, M., Caclin, A, Tillmann, B., Peretz, I.,. 2013. Building pitch memory traces in congenital amusia. In preparation.

Albouy, P.*,* Tillmann, B., Bermudez, P, Zatorre, R.J., Peretz, I., Caclin, A. 2013. When memory for tones differs from memory for words: evidence from separated neural resources for tonal and verbal material in congenital amusia. In preparation.

Albouy, P., Mattout, J., Sanchez, G., Tillmann, B., Caclin, A., 2013. Altered retrieval of melodic information in congenital amusia: a Dynamic Causal Modeling study of MEG data. In preparation.

Tillmann, B., **Albouy, P**., Lalitte, P., Caclin, A., Bigand, E., 2013. Perception of tonal and atonal musical excerpts in congenital amusia. In preparation.

Tillmann, B., **Albouy, P**., Caclin, A., Bigand, E., 2013. Feeling of familiarity despite congenital amusia: Evidence from a gating paradigm with real musical pieces. In preparation.

Torrecillos, F., **Albouy, P**., Brochier, T., Malfait, N., 2013. Does the processing of sensory and reward-prediction errors involve common neural resources? Evidence from a fronto-central negative potential modulated by movement execution errors. ***The Journal of Neuroscience***, in press

Book chapter:

Tillman B, **Albouy P**, Caclin A., 2013. Congenital Amusias in "The human auditory system: fundamental organization and clinical disorders -Handbook of Clinical Neurology", Eds Celesia G.G. & Hickok GS.

References

Ahveninen, J., Jaaskelainen, I.P., Raij, T., Bonmassar, G., Devore, S., Hamalainen, M., Levanen, S., Lin, F.H., Sams, M., Shinn-Cunningham, B.G., et al. (2006). Task-modulated "what" and "where" pathways in human auditory cortex. Proc Natl Acad Sci U S A 103, 14608-14613.

Akiva-Kabiri, L., Vecchi, T., Granot, R., Basso, D., and Schon, D. (2009). Memory for tonal pitches: a music-length effect hypothesis. Ann N Y Acad Sci 1169, 266-269.

Alain, C., Arnott, S.R., Hevenor, S., Graham, S., and Grady, C.L. (2001). "What" and "where" in the human auditory system. Proc Natl Acad Sci U S A 98, 12301-12306.

Albouy, P., Cousineau, M., Caclin, A., Tillmann, B., and Peretz, I. (in preparation). Building pitch memory traces in congenital amusia.

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

Albouy, P., Schulze, K., Caclin, A., and Tillmann, B. (2013b). Does tonality boost short-term memory in congenital amusia? Brain Res.

Allen, G. (1878). Note deafness. Mind 3, 157-167.

Altmann, C.F., Bledowski, C., Wibral, M., and Kaiser, J. (2007). Processing of location and pattern changes of natural sounds in the human auditory cortex. Neuroimage 35, 1192-1200.

Anderson, S., Himonides, E., Wise, K., Welch, G., and Stewart, L. (2012). Is there potential for learning in amusia? A study of the effect of singing intervention in congenital amusia. Ann N Y Acad Sci 1252, 345-353.

Anourova, I., Nikouline, V.V., Ilmoniemi, R.J., Hotta, J., Aronen, H.J., and Carlson, S. (2001). Evidence for dissociation of spatial and nonspatial auditory information processing. Neuroimage 14, 1268-1277.

Atkinson, R.C., and Shiffrin, R.M. (1968). Human memory: a proposed system and its control processes. In The Psychology of Learning and Motivation: Advances in Research and Theory, K.W. Spence, ed. (New York, Academic Press), pp. 89–195.

Awh, E., and Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. Trends Cogn Sci 5, 119-126.

Awh, E., Jonides, J., Smith, E.E., Schumacher, E.H., Koeppe, R.A., and Katz, S. (1996). Dissociation of storage and rehearsal in verbal working memory: evidence from positron emission tomography. Psychol Sci 7, 25–31.

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

Baddeley, A. (1992). Working memory. Science 255, 556-559.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? Trends Cogn Sci 4, 417-423.

Baddeley, A. (2003a). Working memory and language: an overview. J Commun Disord 36, 189-208.

Baddeley, A. (2003b). Working memory: looking back and looking forward. Nat Rev Neurosci 4, 829-839.

Baddeley, A. (2010). Working memory. Curr Biol 20, R136-140.

Baddeley, A. (2012). Working memory: theories, models, and controversies. Annu Rev Psychol 63, 1-29.

Baddeley, A.D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. Q J Exp Psychol 18, 362-365.

Baddeley, A.D. (1976). The Psychology of Memory (New York, Basic Books).

Baddeley, A.D. (1986). Working Memory (Oxford, Clarendon Press).

Baddeley, A.D. (1996). Exploring the Central Executive. Quarterly Journal of Experimental Psychology 49, 5–28.

Baddeley, A.D. (1997). Human memory: Theory and practice (East Sussex, Psychology Press).

Baddeley, A.D., Grant, S., Wight, E., and Thomson, N. (1975a). Imagery and visual working memory. In Attention and performance, P.M.A. Rabbitt, and S. Domic, eds. (London, Academic Press), pp. 205-217.

Baddeley, A.D., and Hitch, G. (1974). Working memory. In The Psychology of Learning and Motivation, B. G.H., ed. (New York, Academic Press), pp. 47–89.

Baddeley, A.D., and Lieberman, K. (1980). Spatial working memory. In Attention and performance, R. Nickerson, ed. (Hillsdale, Lawrence Erlbaum Associates Inc), pp. 521-539.

Baddeley, A.D., Thomson, N., and Buchanan, M. (1975b). Word length and the structure of short-term memory. Journal of Verbal Learning and Verbal Behavior 14, 575-589.

Baddeley, A.D., Vallar, G., and Wilson, B. (1987). Sentence comprehension and phonological memory: some neurological evidence. Attention and Performance 12, 509-529.

Baier, B., Karnath, H.O., Dieterich, M., Birklein, F., Heinze, C., and Muller, N.G. (2010). Keeping memory clear and stable--the contribution of human basal ganglia and prefrontal cortex to working memory. J Neurosci 30, 9788-9792.

Baker, S.N., Philbin, N., Spinks, R., Pinches, E.M., Wolpert, D.M., MacManus, D.G., Pauluis, Q., and Lemon, R.N. (1999). Multiple single unit recording in the cortex of monkeys using independently moveable microelectrodes. J Neurosci Methods 94, 5-17.

Bamiou, D.E., Musiek, F.E., and Luxon, L.M. (2003). The insula (Island of Reil) and its role in auditory processing. Literature review. Brain Res Brain Res Rev 42, 143-154.

Barrett, D.J., and Hall, D.A. (2006). Response preferences for "what" and "where" in human non-primary auditory cortex. Neuroimage 32, 968-977.

Bartlett, J.C., and Dowling, W.J. (1980). Recognition of transposed melodies: a key-distance effect in developmental perspective. J Exp Psychol Hum Percept Perform 6, 501-515.

Bauer, M. (2008). Multisensory integration: a functional role for inter-area synchronization? Curr Biol 18, R709-710.

Baumann, S., Meyer, M., and Jancke, L. (2008). Enhancement of auditory-evoked potentials in musicians reflects an influence of expertise but not selective attention. J Cogn Neurosci 20, 2238-2249.

Baumgart, F., Gaschler-Markefski, B., Woldorff, M.G., Heinze, H.J., and Scheich, H. (1999). A movement-sensitive area in auditory cortex. Nature 400, 724-726.

Belin, P., and Zatorre, R.J. (2000). 'What', 'where' and 'how' in auditory cortex. Nat Neurosci 3, 965-966.

Belin, P., Zatorre, R.J., Hoge, R., Evans, A.C., and Pike, B. (1999). Event-related fMRI of the auditory cortex. Neuroimage 10, 417-429.

Bendor, D., and Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. Nature 436, 1161-1165.

Berkowska, M., and Dalla Bella, S. (2009). Acquired and congenital disorders of sung performance: A review. Adv Cogn Psychol 5, 69-83.

Bernstein, J.G., and Oxenham, A.J. (2005). An autocorrelation model with place dependence to account for the effect of harmonic number on fundamental frequency discrimination. J Acoust Soc Am 117, 3816-3831.

Bertelson, P., and Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. Percept Psychophys 29, 578-584.

Bertrand, O., Perrin, F., and Pernier, J. (1991). Evidence for a tonotopic organization of the auditory cortex observed with auditory evoked potentials. Acta Otolaryngol Suppl 491, 116-122; discussion 123.

Berz, W.L. (1995). Working Memory in Music: A Theoretical Model. Music Perception: An Interdisciplinary Journal 12, 353-364.

Besle, J., Fischer, C., Bidet-Caulet, A., Lecaignard, F., Bertrand, O., and Giard, M.H. (2008). Visual activation and audiovisual interactions in the auditory cortex during speech perception: intracranial recordings in humans. J Neurosci 28, 14301-14310.

Bidet-Caulet, A., and Bertrand, O. (2005). Dynamics of a temporo-fronto-parietal network during sustained spatial or spectral auditory processing. J Cogn Neurosci 17, 1691-1703.

Bidet-Caulet, A., Fischer, C., Besle, J., Aguera, P.E., Giard, M.H., and Bertrand, O. (2007). Effects of selective attention on the electrophysiological representation of concurrent sounds in the human auditory cortex. J Neurosci 27, 9252-9261.

Bidet-Caulet, A., Ye, X.L., Bouchet, P., Guenot, M., Fischer, C., and Bertrand, O. (2009). Non-verbal auditory cognition in patients with temporal epilepsy before and after anterior temporal lobectomy. Front Hum Neurosci 3, 42.

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

Bigand, E., and Tillmann, B. (2005). Effect of Context on the Perception of Pitch Structures. In Pitch Perception, C.J. Plack, and A.J. Oxenham, eds. (Verlag, Springer), pp. 306-351.

Blum, R.A. (1952). Effect of subtotal lesions of frontal granular cortex on delayed reaction in monkeys. Archives of Neurology and Psychiatry Res 67, 375-386.

Bor, D., Duncan, J., Wiseman, R.J., and Owen, A.M. (2003). Encoding strategies dissociate prefrontal activity from working memory demand. Neuron 37, 361-367.

Bourassa, D., and Besner, D. (1994). Semantics at a glance: is the right hemisphere special? Can J Exp Psychol 48, 543-552.

Brandimonte, M.A., Hitch, G.J., and Bishop, D.V. (1992). Verbal recoding of visual stimuli impairs mental image transformations. Mem Cognit 20, 449-455.

Braun, A., McArdle, J., Jones, J., Nechaev, V., Zalewski, C., Brewer, C., and Drayna, D. (2008). Tune deafness: processing melodic errors outside of conscious awareness as reflected by components of the auditory ERP. PLoS One 3, e2349.

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

Brochard, R., Dufour, A., and Despres, O. (2004). Effect of musical expertise on visuospatial abilities: evidence from reaction times and mental imagery. Brain Cogn 54, 103-109.

Brodmann, K. (1909). Vergleichende Loakalisationslehre der Grosshirnrinde (Leipzig, JA Barth).

Brugge, J.F., Volkov, I.O., Garell, P.C., Reale, R.A., and Howard, M.A., 3rd (2003). Functional connections between auditory cortex on Heschl's gyrus and on the lateral superior temporal gyrus in humans. J Neurophysiol 90, 3750-3763.

Brugge, J.F., Volkov, I.O., Oya, H., Kawasaki, H., Reale, R.A., Fenoy, A., Steinschneider, M., and Howard, M.A., 3rd (2008). Functional localization of auditory cortical fields of human: click-train stimulation. Hear Res 238, 12-24.

Brugge, J.F., Volkov, I.O., Reale, R.A., Garell, P.C., Kawasaki, H., Oya, H., Li, Q., and Howard, M.A. (2005). The posteriolateral superior temporal auditory field in humans. Functional organization and connectivity. In The Auditory Cortex – Toward a Synthesis of Human and Animal Research, R. Konig, P. Heil, E. Budinger, and H. Scheich, eds. (Mahwah, NJ, Erlbaum), pp. 145–162.

Buchsbaum, B.R., and D'Esposito, M. (2008). The search for the phonological store: from loop to convolution. J Cogn Neurosci 20, 762-778.

Buchsbaum, B.R., Olsen, R.K., Koch, P., and Berman, K.F. (2005). Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. Neuron 48, 687-697.

Bushara, K.O., Weeks, R.A., Ishii, K., Catalan, M.J., Tian, B., Rauschecker, J.P., and Hallett, M. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. Nat Neurosci 2, 759-766.

Cabeza, R., and Nyberg, L. (2000). Imaging cognition II: An empirical review of 275 PET and fMRI studies. J Cogn Neurosci 12, 1-47.

Caclin, A., Bouchet, P., Djoulah, F., Pirat, E., Pernier, J., and Giard, M.H. (2011). Auditory enhancement of visual perception at threshold depends on visual abilities. Brain Res 1396, 35-44.

Calvert, G.A., and Thesen, T. (2004). Multisensory integration: methodological approaches and emerging principles in the human brain. J Physiol Paris 98, 191-205.

Cantor, J., Engle, R.W., and Hamilton, G. (1991). Short-term memory, working memory, and verbal abilities: How do they relate? Intelligence 15, 29-246.

Carbotte, R.M. (1973). Retention of time information in forced-choice duration discrimination. Perception & Psychophysics 14, 440-444.

Catani, M., and Mesulam, M. (2008). The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. Cortex 44, 953-961.

Catani, M., and Thiebaut de Schotten, M. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. Cortex 44, 1105-1132.

Celesia, G.G. (1976). Organization of auditory cortical areas in man. Brain 99, 403-414.

Champod, A.S., and Petrides, M. (2007). Dissociable roles of the posterior parietal and the prefrontal cortex in manipulation and monitoring processes. Proc Natl Acad Sci U S A 104, 14837-14842.

Champod, A.S., and Petrides, M. (2010). Dissociation within the frontoparietal network in verbal working memory: a parametric functional magnetic resonance imaging study. J Neurosci 30, 3849-3856.

Chan, A.S., Ho, Y.C., and Cheung, M.C. (1998). Music training improves verbal memory. Nature 396, 128.

Chein, J.M., and Fiez, J.A. (2001). Dissociation of verbal working memory system components using a delayed serial recall task. Cereb Cortex 11, 1003-1014.

Chen, S.H., and Desmond, J.E. (2005). Cerebrocerebellar networks during articulatory rehearsal and verbal working memory tasks. Neuroimage 24, 332-338.

Clarke, S., Adriani, M., and Tardif, E. (2005). "What" and "Where" in human audition: Evidence from anatomical, activation and lesion studies. In Auditory cortex towards a synthesis of human and animal research, R. Konig, P. Heil, E. Budinger, and H. Scheich, eds. (Hillsdale, Lawrence Erlbaum Associates), pp. 77–94.

Clarke, S., Bellmann, A., Meuli, R.A., Assal, G., and Steck, A.J. (2000). Auditory agnosia and auditory spatial deficits following left hemispheric lesions: evidence for distinct processing pathways. Neuropsychologia 38, 797-807.

Clarke, S., Bellmann Thiran, A., Maeder, P., Adriani, M., Vernet, O., Regli, L., Cuisenaire, O., and Thiran, J.P. (2002). What and Where in human audition: Selective deficits following focal hemispheric lesions. Experimental Brain Research 147, 8–15.

Clarke, S., and Thiran, A.B. (2004). Auditory neglect: what and where in auditory space. Cortex 40, 291-300.

Cohen, N.J., and Squire, L.R. (1980). Preserved learning and retention of pattern-analyzing skill in amnesia: dissociation of knowing how and knowing that. Science 210, 207-210.

Colle, H.A., and Welsh, A. (1976). Acoustic masking in primary memory. Journal of Verbal Learning and Verbal Behaviour 15, 17-32.

Corbetta, M., and Shulman, G.L. (2002). Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 3, 201-215.

Courtney, S.M., Petit, L., Maisog, J.M., Ungerleider, L.G., and Haxby, J.V. (1998). An area specialized for spatial working memory in human frontal cortex. Science 279, 1347-1351.

Cousineau, M., Demany, L., and Pressnitzer, D. (2009). What makes a melody: The perceptual singularity of pitch sequences. J Acoust Soc Am 126, 3179-3187.

Cousineau, M., Demany, L., and Pressnitzer, D. (2010). The role of peripheral resolvability in pitch-sequence processing. J Acoust Soc Am 128, EL236-241.

Cousineau, M., McDermott, J.H., and Peretz, I. (2012). The basis of musical consonance as revealed by congenital amusia. Proc Natl Acad Sci U S A 109, 19858-19863.

Cowan, N. (1984). On short and long auditory stores. Psychol Bull 96, 341-370.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. Psychol Bull 104, 163-191.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? Prog Brain Res 169, 323-338.

Cowan, N., Wood, N.L., Wood, P.K., Keller, T.A., Nugent, L.D., and Keller, C.V. (1998). Two separate verbal processing rates contributing to short-term memory span. J Exp Psychol Gen 127, 141-160.

Craik, F.I.M., and Lockhart, R.S. (1972). Levels of processing: a framework for memory research. Journal of Verbal Learning and Verbal Behavior 11, 671–684.

Crone, N.E., Boatman, D., Gordon, B., and Hao, L. (2001). Induced electrocorticographic gamma activity during auditory perception. Brazier Award-winning article, 2001. Clin Neurophysiol 112, 565-582.

Croxson, P.L., Johansen-Berg, H., Behrens, T.E., Robson, M.D., Pinsk, M.A., Gross, C.G., Richter, W., Richter, M.C., Kastner, S., and Rushworth, M.F. (2005). Quantitative investigation of connections of the prefrontal cortex in the human and macaque using probabilistic diffusion tractography. J Neurosci 25, 8854-8866.

Cuddy, L.L. (1971). Absolute judgment of musically-related pure tones. Can J Psychol 25, 42-55.

Cuddy, L.L., Balkwill, L.L., Peretz, I., and Holden, R.R. (2005). Musical difficulties are rare: a study of "tone deafness" among university students. Ann N Y Acad Sci 1060, 311-324.

Curtis, C.E., and D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. Trends Cogn Sci 7, 415-423.

D'Esposito, M. (2007). From cognitive to neural models of working memory. Philos Trans R Soc Lond B Biol Sci 362, 761-772.

D'Esposito, M., Aguirre, G.K., Zarahn, E., Ballard, D., Shin, R.K., and Lease, J. (1998a). Functional MRI studies of spatial and nonspatial working memory. Brain Res Cogn Brain Res 7, 1-13.

D'Esposito, M., Ballard, D., Aguirre, G.K., and Zarahn, E. (1998b). Human prefrontal cortex is not specific for working memory: a functional MRI study. Neuroimage 8, 274-282.

D'Esposito, M., and Postle, B.R. (1999). The dependence of span and delayed-response performance on prefrontal cortex. Neuropsychologia 37, 1303-1315.

D'Esposito, M., Postle, B.R., Ballard, D., and Lease, J. (1999). Maintenance versus manipulation of information held in working memory: an event-related fMRI study. Brain Cogn 41, 66-86.

Da Costa, S., van der Zwaag, W., Marques, J.P., Frackowiak, R.S.J., Clarke, S., and Saenz, M. (2011). Human Primary Auditory Cortex Follows the Shape of Heschl's Gyrus. The Journal of Neuroscience 31, 14067–14075.

Dalla Bella, S., Giguere, J.F., and Peretz, I. (2007). Singing proficiency in the general population. J Acoust Soc Am 121, 1182-1189.

Dalla Bella, S., Giguere, J.F., and Peretz, I. (2009). Singing in congenital amusia. J Acoust Soc Am 126, 414-424.

Dalla Bella, S., and Peretz, I. (2003). Congenital amusia interferes with the ability to synchronize with music. Ann N Y Acad Sci 999, 166-169.

Dalla Bella, S., Tremblay-Champoux, A., Berkowska, M., and Peretz, I. (2012). Memory disorders and vocal performance. Ann N Y Acad Sci 1252, 338-344.

David, O., and Friston, K.J. (2003). A neural mass model for MEG/EEG: coupling and neuronal dynamics. Neuroimage 20, 1743-1755.

David, O., Kiebel, S.J., Harrison, L.M., Mattout, J., Kilner, J.M., and Friston, K.J. (2006). Dynamic causal modeling of evoked responses in EEG and MEG. Neuroimage 30, 1255-1272.

de Cheveigné, A. (2010). Pitch Perception Models. In Neural Coding and Perception, C.J. Plack, A.J. Oxenham, R.R. Fay, and A.N. Popper, eds. (Springer), pp. 71–104.

Demany, L., and Semal, C. (2005). The slow formation of a pitch percept beyond the ending time of a short tone burst. Percept Psychophys 67, 1376-1383.

Demany, L., and Semal, C. (2007). The Role of Memory in Auditory Perception. In Auditory Perception of Sound Sources, R.R. Fay, and A.N. Popper, eds., pp. 77-113.

Demany, L., and Semal, C. (2008). The Role of Memory in Auditory Perception. In Springer Handbook of Auditory Research (Springer), pp. 77-113.

Deutsch, D. (1970a). Dislocation of tones in a musical sequence: a memory illusion. Nature 226, 286.

Deutsch, D. (1970b). Tones and numbers: specificity of interference in immediate memory. Science 168, 1604-1605.

Deutsch, D. (1972). Effect of repetition of standard and comparison tones on recognition memory for pitch. J Exp Psychol 93, 156-162.

Deutsch, D. (1974). Generality of interference by tonal stimuli in recognition memory for tonal pitch. Q J Exp Psychol 26, 229–234.

Deutsch, D. (1975a). Auditory memory. Can J Psychol 29, 87-105.

Deutsch, D. (1975b). Facilitation by repetition in recognition memory for tonal pitch. Mem Cognit 3, 263-266.

Deutsch, D. (1977). Memory and attention in music. In Music and the brain, M. Critchley, and R.A. Henson, eds. (London, William Heinemann Medical Books).

Dewar, K.M., Cuddy, L.L., and Mewhort, D.J. (1977). Recognition memory for single tones with and without context. J Exp Psychol Hum Learn 3, 60-67.

di Pellegrino, G., and Wise, S.P. (1991). A neurophysiological comparison of three distinct regions of the primate frontal lobe. Brain 114 (Pt 2), 951-978.

Divenyi, P.L., and Robinson, A.J. (1989). Nonlinguistic auditory capabilities in aphasia. Brain Lang 37, 290-326.

Douglas, K.M., and Bilkey, D.K. (2007). Amusia is associated with deficits in spatial processing. Nat Neurosci 10, 915-921.

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

Dowling, W.J. (1982). Melodic information processing and its development. In The psychology of music, D. Deutsch, ed. (New York, New York Academic Press), pp. 413–429.

Dowling, W.J. (1991). Tonal strength and melody recognition after long and short delays. Percept Psychophys 50, 305-313.

Dowling, W.J., and Fujitani, D.S. (1971). Contour, interval, and pitch recognition in memory for melodies. J Acoust Soc Am 49, Suppl 2:524+.

Dowling, W.J., and Harwood, D.L. (1986). Music Cognition (Academic Press).

Drachman, D.A., and Arbit, J. (1966). Memory and the hippocampal complex. II. Is memory a multiple process? Arch Neurol 15, 52-61.

Drayna, D., Manichaikul, A., de Lange, M., Snieder, H., and Spector, T. (2001). Genetic correlates of musical pitch recognition in humans. Science 291, 1969-1972.

Ducommun, C.Y., Murray, M.M., Thut, G., Bellmann, A., Viaud-Delmon, I., Clarke, S., and Michel, C.M. (2002). Segregated processing of auditory motion and auditory location: an ERP mapping study. Neuroimage 16, 76-88.

Durstewitz, D., Seamans, J.K., and Sejnowski, T.J. (2000). Neurocomputational models of working memory. Nat Neurosci 3 Suppl, 1184-1191.

Duvernoy, H.M. (1999). The Human Brain: Surface, Three-dimensional Sectional Anatomywith MRI, and Vascularization (Wien, Springer-Verlag).

Eacott, M.J., Gaffan, D., and Murray, E.A. (1994). Preserved recognition memory for small sets, and impaired stimulus identification for large sets, following rhinal cortex ablations in monkeys. Eur J Neurosci 6, 1466-1478.

Ebbinghaus, H. (1885/1913). Memory: a contribution to experimental psychology (Columbia University Teachers College).

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

Ehret, G., and Romand, R. (1997). The central auditory system (New York, Oxford University Press).

Engel, A.K., Fries, P., and Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. Nat Rev Neurosci 2, 704-716.

Engle, R.W. (2002). Working memory capacity as executive attention. Curr Dir Psychol Sci 11, 19–23.

Engle, R.W., Tuholski, S.W., Laughlin, J.E., and Conway, A.R.A. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. J Exp Psychol Gen 128, 309–331.

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

Falchier, A., Clavagnier, S., Barone, P., and Kennedy, H. (2002). Anatomical evidence of multimodal integration in primate striate cortex. J Neurosci 22, 5749-5759.

Fiez, J.A. (2001). Bridging the gap between neuroimaging and neuropsychology: using working memory as a case-study. J Clin Exp Neuropsychol 23, 19-31.

Fiez, J.A., Raife, E.A., Balota, D.A., Schwarz, J.P., Raichle, M.E., and Petersen, S.E. (1996). A positron emission tomography study of the short-term maintenance of verbal information. J Neurosci 16, 808-822.

Formisano, E., Kim, D.S., Di Salle, F., van de Moortele, P.F., Ugurbil, K., and Goebel, R. (2003). Mirror-symmetric tonotopic maps in human primary auditory cortex. Neuron 40, 859-869.

Fort, A., Delpuech, C., Pernier, J., and Giard, M.H. (2002). Dynamics of cortico-subcortical cross-modal operations involved in audio-visual object detection in humans. Cereb Cortex 12, 1031-1039.

Foster, N.E., Halpern, A.R., and Zatorre, R.J. (2013). Common parietal activation in musical mental transformations across pitch and time. Neuroimage 75, 27-35.

Foster, N.E., and Zatorre, R.J. (2010). A role for the intraparietal sulcus in transforming musical pitch information. Cereb Cortex 20, 1350-1359.

Foxton, J.M., Brown, A.C., Chambers, S., and Griffiths, T.D. (2004a). Training improves acoustic pattern perception. Curr Biol 14, 322-325.

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

Foxton, J.M., Nandy, R.K., and Griffiths, T.D. (2006). Rhythm deficits in 'tone deafness'. Brain Cogn 62, 24-29.

Francès, R., ed. (1958). La perception de la musique (Paris, Librairie Philosophique J. Vrin.).

Frassinetti, F., Bolognini, N., and Ladavas, E. (2002a). Enhancement of visual perception by crossmodal visuo-auditory interaction. Experimental Brain Research 147, 332-343.

Frassinetti, F., Pavani, F., and Ladavas, E. (2002b). Acoustical vision of neglected stimuli: interaction among spatially converging audiovisual inputs in neglect patients. J Cogn Neurosci 14, 62-69.

Frey, S., Campbell, J.S., Pike, G.B., and Petrides, M. (2008). Dissociating the human language pathways with high angular resolution diffusion fiber tractography. J Neurosci 28, 11435-11444.

Friston, K. (2003). Learning and inference in the brain. Neural Netw 16, 1325-1352.

Friston, K., Holmes, A., Worsley, K.J., Poline, J.B., Frith, C.D., and Frackowiak, R.S.J. (1995). Statistical parametric maps in functional imaging: a general linear approach. Hum Brain Mapp 2, 189- 210.

Friston, K.J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. Neuroimage 19, 1273-1302.

Fullerton, B.C., and Pandya, D.N. (2007). Architectonic analysis of the auditory-related areas of the superior temporal region in human brain. J Comp Neurol 504, 470-498.

Funahashi, S., Bruce, C.J., and Goldman-Rakic, P.S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. J Neurophysiol 61, 331-349.

Fuster, J. (1989). The Prefrontal Cortex: Anatomy, Physiology, and Neuropsychology of the Frontal Lobes (New York, Raven Press).

Fuster, J.M., and Alexander, G.E. (1971). Neuron activity related to short-term memory. Science 173, 652-654.

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

Galaburda, A., and Sanides, F. (1980). Cytoarchitectonic organization of the human auditory cortex. J Comp Neurol 190, 597-610.

Garell, P.C., Bakken, H., Greenlee, J.D., Volkov, I., Reale, R.A., Oya, H., Kawasaki, H., Howard, M.A., and Brugge, J.F. (2012). Functional Connection Between Posterior Superior Temporal Gyrus and Ventrolateral Prefrontal Cortex in Human. Cereb Cortex.

Garrido, M.I., Kilner, J.M., Kiebel, S.J., and Friston, K.J. (2007). Evoked brain responses are generated by feedback loops. Proc Natl Acad Sci U S A 104, 20961-20966.

Gathercole, S.E., and Baddeley, A.D. (1993). Working Memory and Language (Hillsdale, Erlbaum).

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

Giard, M.H., Perrin, F., Pernier, J., and Peronnet, F. (1988). Several attention-related wave forms in auditory areas: a topographic study. Electroencephalogr Clin Neurophysiol 69, 371-384.

Gibson, J.J. (1966). The senses considered as perceptual systems, Houghton Mifflin Co. edn (Boston).

Gielen, S.C., Schmidt, R.A., and Van den Heuvel, P.J. (1983). On the nature of intersensory facilitation of reaction time. Percept Psychophys 34, 161-168.

Gilbert, C.D., Sigman, M., and Crist, R.E. (2001). The neural basis of perceptual learning. Neuron 31, 681-697.

Goldman-Rakic, P.S. (1988). Topography of cognition: parallel distributed networks in primate association cortex. Annu Rev Neurosci 11, 137-156.

Goldstein, J.L. (1973). An optimum processor theory for the central formation of the pitch of complex tones. J Acoust Soc Am 54, 1496-1516.

Golubock, J.L., and Janata, P. (2013). Keeping timbre in mind: working memory for complex sounds that can't be verbalized. J Exp Psychol Hum Percept Perform 39, 399-412.

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

Gosselin, N., Trainor, L., Belin, P., and Peretz, I. (2011). Emotion recognition is spared in amusia. In The Neurosciences and Music-IV Learning and Memory Meeting, Edinburgh, UK.

Griffiths, T.D. (1999). Human complex sound analysis. Clin Sci (Lond) 96, 231-234.

Griffiths, T.D., and Green, G.G. (1999). Cortical activation during perception of a rotating wide-field acoustic stimulus. Neuroimage 10, 84-90.

Griffiths, T.D., Green, G.G., Rees, A., and Rees, G. (2000a). Human brain areas involved in the analysis of auditory movement. Hum Brain Mapp 9, 72-80.

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

Griffiths, T.D., Kumar, S., Sedley, W., Nourski, K.V., Kawasaki, H., Oya, H., Patterson, R.D., Brugge, J.F., and Howard, M.A. (2010). Direct recordings of pitch responses from human auditory cortex. Curr Biol 20, 1128-1132.

Griffiths, T.D., Penhune, V., Peretz, I., Dean, J.L., Patterson, R.D., and Green, G.G. (2000b). Frontal processing and auditory perception. Neuroreport 11, 919-922.

Grimault, S., Lefebvre, C., Vachon, F., Peretz, I., Zatorre, R., Robitaille, N., and Jolicoeur, P. (2009). Load-dependent brain activity related to acoustic short-term memory for pitch: magnetoencephalography and fMRI. Ann N Y Acad Sci 1169, 273-277.

Gross, J., Baillet, S., Barnes, G.R., Henson, R.N., Hillebrand, A., Jensen, O., Jerbi, K., Litvak, V., Maess, B., Oostenveld, R., et al. (2013). Good practice for conducting and reporting MEG research. Neuroimage 65, 349-363.

Gruber, O., and von Cramon, D.Y. (2003). The functional neuroanatomy of human working memory revisited. Evidence from 3-T fMRI studies using classical domain-specific interference tasks. Neuroimage 19, 797-809.

Gueguin, M., Le Bouquin-Jeannes, R., Faucon, G., Chauvel, P., and Liegeois-Chauvel, C. (2007). Evidence of functional connectivity between auditory cortical areas revealed by amplitude modulation sound processing. Cereb Cortex 17, 304-313.

Gutschalk, A., Mase, R., Roth, R., Ille, N., Rupp, A., Hahnel, S., Picton, T.W., and Scherg, M. (1999). Deconvolution of 40 Hz steady-state fields reveals two overlapping source activities of the human auditory cortex. Clin Neurophysiol 110, 856-868.

Hackett, T.A. (2003). The comparative anatomy of the primate auditory cortex. In Primate audition: ethology and neurobiology A.A. Ghazanfar, ed. (Boca Raton, CRC Press), pp. 199–219.

Hackett, T.A. (2007). Organization and correspondence of the auditory cortex of humans and nonhuman primates. . In Evolution of the nervous system, J.H. Kaas, ed. (Oxford, Elsevier), pp. 109–119.

Hackett, T.A. (2008). Anatomical organization of the auditory cortex. J Am Acad Audiol 19, 774-779.

Hackett, T.A., Preuss, T.M., and Kaas, J.H. (2001). Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. J Comp Neurol 441, 197-222.

Hackett, T.A., Stepniewska, I., and Kaas, J.H. (1999). Prefrontal connections of the parabelt auditory cortex in macaque monkeys. Brain Res 817, 45-58.

Hairston, W.D., Laurienti, P.J., Mishra, G., Burdette, J.H., and Wallace, M.T. (2003). Multisensory enhancement of localization under conditions of induced myopia. Exp Brain Res 152, 404-408.

Hannon, E.E., and Trehub, S.E. (2005). Tuning in to musical rhythms: infants learn more readily than adults. Proc Natl Acad Sci U S A 102, 12639-12643.

Helmholtz, H. (1866/1962). Concerning the perceptions in general, Treatise on Physiological Optics, J. Southall, Trans., III. edn (Dover, New York.).

Helmholtz, H.L.F. (1863). Die Lehre von den Tonempfindungen als physiologische Grundlage fur die Theorie der Musik" published in translation (1954) "On the sensations of tone as a physiological basis for the theory of music" (New York, Dover).

Henry, M., and McAuley, J.D. (2010). On the prevalence of congenital amusia. Music Perception 27, 413-418.

Herdener, M., Esposito, F., Scheffler, K., Schneider, P., Logothetis, N.K., Uludag, K., and Kayser, C. (2013). Spatial representations of temporal and spectral sound cues in human auditory cortex. Cortex.

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

Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. J Exp Psychol 63, 289-293.

Hickok, G. (2009). The functional neuroanatomy of language. Phys Life Rev 6, 121-143.

Hickok, G., Buchsbaum, B., Humphries, C., and Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. J Cogn Neurosci 15, 673-682.

Hickok, G., Houde, J., and Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. Neuron 69, 407-422.

References

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. Nat Rev Neurosci 8, 393-402.

Hitch, G.J., Burgess, N., Towse, J., and Culpin, V. (1996). Temporal grouping effects in immediate recall: a working memory analysis. Quarterly Journal of Experimental Psychology 49A, 116-139.

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

Howard, M.A., 3rd, Volkov, I.O., Abbas, P.J., Damasio, H., Ollendieck, M.C., and Granner, M.A. (1996). A chronic microelectrode investigation of the tonotopic organization of human auditory cortex. Brain Res 724, 260-264.

Howard, M.A., Volkov, I.O., Mirsky, R., Garell, P.C., Noh, M.D., Granner, M., Damasio, H., Steinschneider, M., Reale, R.A., Hind, J.E., et al. (2000). Auditory cortex on the human posterior superior temporal gyrus. J Comp Neurol 416, 79-92.

Howard, M.W., Rizzuto, D.S., Caplan, J.B., Madsen, J.R., Lisman, J., Aschenbrenner-Scheibe, R., Schulze-Bonhage, A., and Kahana, M.J. (2003). Gamma oscillations correlate with working memory load in humans. Cereb Cortex 13, 1369-1374.

Huettel, S.A., Song, A.W., and McCarthy, G. (2009). Functional Magnetic Resonance Imaging (Massachusetts, Sinauer).

Hulme, C., Maughan, S., and Brown, G.D.A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to shortterm memory span. Journal of Memory and Language 30, 685-701.

Hulme, C., Roodenrys, S., Schweickert, R., Brown, G.D., Martin, M., and Stuart, G. (1997). Word-frequency effects on short-term memory tasks: evidence for a redintegration process in immediate serial recall. J Exp Psychol Learn Mem Cogn 23, 1217-1232.

Hupe, J.M., James, A.C., Payne, B.R., Lomber, S.G., Girard, P., and Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. Nature 394, 784-787.

Hutchins, S., Gosselin, N., and Peretz, I. (2010a). Identification of Changes along a Continuum of Speech Intonation is Impaired in Congenital Amusia. Front Psychol 1, 236.

Hutchins, S., and Peretz, I. (2012a). Amusics can imitate what they cannot discriminate. Brain Lang 123, 234-239.

Hutchins, S., and Peretz, I. (2013). Vocal pitch shift in congenital amusia (pitch deafness). Brain Lang 125, 106-117.

Hutchins, S., Zarate, J.M., Zatorre, R.J., and Peretz, I. (2010b). An acoustical study of vocal pitch matching in congenital amusia. J Acoust Soc Am 127, 504-512.

Hutchins, S.M., and Peretz, I. (2012b). A frog in your throat or in your ear? Searching for the causes of poor singing. J Exp Psychol Gen 141, 76-97.

Hyde, K.L., Lerch, J.P., Zatorre, R.J., Griffiths, T.D., Evans, A.C., and Peretz, I. (2007). Cortical thickness in congenital amusia: when less is better than more. J Neurosci 27, 13028-13032.

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

Hyde, K.L., Peretz, I., and Zatorre, R.J. (2008). Evidence for the role of the right auditory cortex in fine pitch resolution. Neuropsychologia 46, 632-639.

Hyde, K.L., Zatorre, R.J., Griffiths, T.D., Lerch, J.P., and Peretz, I. (2006). Morphometry of the amusic brain: a two-site study. Brain 129, 2562-2570.

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

Inoue, M., Mikami, A., Ando, I., and Tsukada, H. (2004). Functional brain mapping of the macaque related to spatial working memory as revealed by PET. Cereb Cortex 14, 106-119.

Inui, K., Okamoto, H., Miki, K., Gunji, A., and Kakigi, R. (2006). Serial and parallel processing in the human auditory cortex: a magnetoencephalographic study. Cereb Cortex 16, 18-30.

Isseroff, A., Rosvold, H.E., Galkin, T.W., and Goldman-Rakic, P.S. (1982). Spatial memory impairments following damage to the mediodorsal nucleus of the thalamus in rhesus monkeys. Brain Res 232, 97-113.

Jacobsen, C.F. (1935). Function of frontal association areas in primates. Arch Neurol Psychiatry 33, 558–560.

Jacobsen, C.F. (1936). Studies of cerebral function in primates. I. The functions of the frontal association areas in monkeys. Comp Psychol Monogr 13, 1-60.

James, W. (1890). The Principles of Psychology (New York).

Janata, P., Birk, J.L., Van Horn, J.D., Leman, M., Tillmann, B., and Bharucha, J.J. (2002a). The cortical topography of tonal structures underlying Western music. Science 298, 2167-2170.

Janata, P., Tillmann, B., and Bharucha, J.J. (2002b). Listening to polyphonic music recruits domain-general attention and working memory circuits. Cogn Affect Behav Neurosci 2, 121-140.

Jansen, B.H., and Rit, V.G. (1995). Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. Biol Cybern 73, 357-366.

Jeneson, A., and Squire, L.R. (2012). Working memory, long-term memory, and medial temporal lobe function. Learn Mem 19, 15-25.

Jensen, O., Kaiser, J., and Lachaux, J.P. (2007). Human gamma-frequency oscillations associated with attention and memory. Trends Neurosci 30, 317-324.

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

Jiang, C., Hamm, J.P., Lim, V.K., Kirk, I.J., and Yang, Y. (2012). Impaired categorical perception of lexical tones in Mandarin-speaking congenital amusics. Mem Cognit 40, 1109-1121.

Johnsrude, I.S., Penhune, V.B., and Zatorre, R.J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. Brain 123 (Pt 1), 155-163.

Jones, J.L., Lucker, J., Zalewski, C., Brewer, C., and Drayna, D. (2009a). Phonological processing in adults with deficits in musical pitch recognition. J Commun Disord 42, 226-234.

Jones, J.L., Zalewski, C., Brewer, C., Lucker, J., and Drayna, D. (2009b). Widespread auditory deficits in tune deafness. Ear Hear 30, 63-72.

Jones, M.R. (1976). ime, our lost dimension: Toward a new theory of perception, attention, and memory. Psychological Review 83, 323-355.

Jones, M.R., and Boltz, M. (1989). Dynamic attending and responses to time. Psychological Review 96, 459-491.

Jones, M.R., and Yee, W. (1993). Attending to auditory events : the role of temporal organization. In Thinking in sound : the cognitive psychology of human audition S. McAdams, and E. Bigand, eds. (Oxford, Oxford University Press), pp. 69-112.

Jonides, J., Schumacher, E.H., Smith, E.E., Koeppe, R.A., Awh, E., Reuter-Lorenz, P.A., Marshuetz, C., and Willis, C.R. (1998). The role of parietal cortex in verbal working memory. J Neurosci 18, 5026-5034.

Jousmaki, V., and Forss, N. (1998). Effects of stimulus intensity on signals from human somatosensory cortices. Neuroreport 9, 3427-3431.

Kaas, J.H., and Hackett, T.A. (1999). 'What' and 'where' processing in auditory cortex. Nat Neurosci 2, 1045-1047.

Kaas, J.H., and Hackett, T.A. (2000). Subdivisions of auditory cortex and processing streams in primates. Proc Natl Acad Sci U S A 97, 11793-11799.

Kail, R., and Hall, L.K. (2001). Distinguishing short-term memory from working memory. Memory and Cognition 29, 1-9.

Kaiser, J., and Bertrand, O. (2003). Dynamics of working memory for moving sounds: an event-related potential and scalp current density study. Neuroimage 19, 1427-1438.

Kaiser, J., Ripper, B., Birbaumer, N., and Lutzenberger, W. (2003). Dynamics of gamma-band activity in human magnetoencephalogram during auditory pattern working memory. Neuroimage 20, 816-827.

Kalmus, H., and Fry, D.B. (1980). On tune deafness (dysmelodia): frequency, development, genetics and musical background. Ann Hum Genet 43, 369-382.

Kane, M.J., Hambrick, D.Z., Tuholski, S.W., Wilhelm, O., Payne, T.W., and Engle, R.W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. J Exp Psychol Gen 133, 189-217.

Kawagoe, R., Takikawa, Y., and Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. Nat Neurosci 1, 411-416.

Kayser, C., Kim, M., Ugurbil, K., Kim, D.S., and Konig, P. (2004). A comparison of hemodynamic and neural responses in cat visual cortex using complex stimuli. Cereb Cortex 14, 881-891.

References

Kelly, C., Uddin, L.Q., Shehzad, Z., Margulies, D.S., Castellanos, F.X., Milham, M.P., and Petrides, M. (2010). Broca's region: linking human brain functional connectivity data and non-human primate tracing anatomy studies. Eur J Neurosci 32, 383-398.

Khan, Z.U., and Muly, E.C. (2011). Molecular mechanisms of working memory. Behav Brain Res 219, 329-341.

Kim, D.S., Ronen, I., Olman, C., Kim, S.G., Ugurbil, K., and Toth, L.J. (2004). Spatial relationship between neuronal activity and BOLD functional MRI. Neuroimage 21, 876-885.

Kirschen, M.P., Chen, S.H., Schraedley-Desmond, P., and Desmond, J.E. (2005). Load- and practice-dependent increases in cerebro-cerebellar activation in verbal working memory: an fMRI study. Neuroimage 24, 462-472.

Klauer, K.C., and Zhao, Z. (2004). Double dissociations in visual and spatial short-term memory. J Exp Psychol Gen 133, 355-381.

Klimesch, W., Sauseng, P., and Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition-timing hypothesis. Brain Res Rev 53, 63-88.

Koelsch, S., Gunter, T.C., Wittfoth, M., and Sammler, D. (2005). Interaction between syntax processing in language and in music: an ERP Study. J Cogn Neurosci 17, 1565-1577.

Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Muller, K., and Gruber, O. (2009). Functional architecture of verbal and tonal working memory: an FMRI study. Hum Brain Mapp 30, 859-873.

Koski, L., and Paus, T. (2000). Functional connectivity of the anterior cingulate cortex within the human frontal lobe: a brain-mapping meta-analysis. Exp Brain Res 133, 55-65.

Kujala, T., Belitz, S., Tervaniemi, M., and Naatanen, R. (2003). Auditory sensory memory disorder in dyslexic adults as indexed by the mismatch negativity. Eur J Neurosci 17, 1323-1327.

Lachaux, J.P., Axmacher, N., Mormann, F., Halgren, E., and Crone, N.E. (2012). High-frequency neural activity and human cognition: past, present and possible future of intracranial EEG research. Prog Neurobiol 98, 279-301.

Lachaux, J.P., Fonlupt, P., Kahane, P., Minotti, L., Hoffmann, D., Bertrand, O., and Baciu, M. (2007). Relationship between task-related gamma oscillations and BOLD signal: new insights from combined fMRI and intracranial EEG. Hum Brain Mapp 28, 1368-1375.

Lamme, V.A., and Roelfsema, P.R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. Trends Neurosci 23, 571-579.

Langers, D.R., Backes, W.H., and van Dijk, P. (2007). Representation of lateralization and tonotopy in primary versus secondary human auditory cortex. Neuroimage 34, 264-273.

Langers, D.R., and van Dijk, P. (2012). Mapping the tonotopic organization in human auditory cortex with minimally salient acoustic stimulation. Cereb Cortex 22, 2024-2038.

Large, E., and Jones, M. (1999). The Dynamics of Attending: How People Track Time-Varying Events. Psychological Review 106, 119-159.

Lebrun, M.A., Moreau, P., McNally-Gagnon, A., Mignault Goulet, G., and Peretz, I. (2012). Congenital amusia in childhood: a case study. Cortex 48, 683-688.

Levy, R., and Goldman-Rakic, P.S. (2000). Segregation of working memory functions within the dorsolateral prefrontal cortex. Exp Brain Res 133, 23-32.

Lewis, J.W., Beauchamp, M.S., and DeYoe, E.A. (2000). A comparison of visual and auditory motion processing in human cerebral cortex. Cereb Cortex 10, 873-888.

Lewis, S.J., Dove, A., Robbins, T.W., Barker, R.A., and Owen, A.M. (2004). Striatal contributions to working memory: a functional magnetic resonance imaging study in humans. Eur J Neurosci 19, 755-760.

Licklider, J.C. (1951). A duplex theory of pitch perception. Experientia 7, 128-134.

Lidji, P., Kolinsky, R., Lochy, A., and Morais, J. (2007). Spatial associations for musical stimuli: a piano in the head? J Exp Psychol Hum Percept Perform 33, 1189-1207.

Liegeois-Chauvel, C., Musolino, A., Badier, J.M., Marquis, P., and Chauvel, P. (1994). Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. Electroencephalogr Clin Neurophysiol 92, 204-214.

Liegeois-Chauvel, C., Musolino, A., and Chauvel, P. (1991). Localization of the primary auditory area in man. Brain 114 (Pt 1A), 139-151.

Liegeois-Chauvel, C., Peretz, I., Babai, M., Laguitton, V., and Chauvel, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. Brain 121 (Pt 10), 1853-1867.

Liu, F., Jiang, C., Thompson, W.F., Xu, Y., Yang, Y., and Stewart, L. (2012a). The mechanism of speech processing in congenital amusia: evidence from Mandarin speakers. PLoS One 7, e30374.

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

Liu, F., Xu, Y., Patel, A.D., Francart, T., and Jiang, C. (2012b). Differential recognition of pitch patterns in discrete and gliding stimuli in congenital amusia: evidence from Mandarin speakers. Brain Cogn 79, 209-215.

Lockhart, R.S. (2000). Methods of memory research. In The Oxford Handbook of Memory, E. Tulving, and F.I.M. Craik, eds. (Oxford, Oxford University Press), pp. 45–57.

Logie, R.H., and D'Esposito, M. (2007). Working memory in the brain. Cortex 43, 1-4.

Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. Nature 412, 150-157.

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

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

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

Loui, P., and Schlaug, G. (2012). Impaired learning of event frequencies in tone deafness. Ann N Y Acad Sci 1252, 354-360.

Luria, A., Tsvetkova, L., and Futer, J. (1965). Aphasia in a composer. J Neurol Sci, 288–292.

Lutzenberger, W., Ripper, B., Busse, L., Birbaumer, N., and Kaiser, J. (2002). Dynamics of gamma-band activity during an audiospatial working memory task in humans. J Neurosci 22, 5630-5638.

Maeder, P.P., Meuli, R.A., Adriani, M., Bellmann, A., Fornari, E., Thiran, J.P., Pittet, A., and Clarke, S. (2001). Distinct pathways involved in sound recognition and localization: a human fMRI study. Neuroimage 14, 802-816.

Maess, B., Koelsch, S., Gunter, T.C., and Friederici, A.D. (2001). Musical syntax is processed in Broca's area: an MEG study. Nat Neurosci 4, 540-545.

Mainy, N., Kahane, P., Minotti, L., Hoffmann, D., Bertrand, O., and Lachaux, J.P. (2007). Neural correlates of consolidation in working memory. Hum Brain Mapp 28, 183-193.

Mandell, J., Schulze, K., and Schlaug, G. (2007). Congenital amusia: an auditory-motor feedback disorder? Restor Neurol Neurosci 25, 323-334.

Marklund, P., Fransson, P., Cabeza, R., Petersson, K.M., Ingvar, M., and Nyberg, L. (2007). Sustained and transient neural modulations in prefrontal cortex related to declarative long-term memory, working memory, and attention. Cortex 43, 22-37.

Marmel, F., Tillmann, B., and Dowling, W.J. (2008). Tonal expectations influence pitch perception. Percept Psychophys 70, 841-852.

Massaro, D.W. (1972). Preperceptual images, processing time, and perceptual units in auditory perception. Psychol Rev 79, 124-145.

Massaro, D.W., and Idson, W.L. (1977). Backward recognition masking in relative pitch judgments. Percept Mot Skills 45, 87-97.

Massaro, D.W., and Loftus, G.R. (1996). Sensory and perceptual storage: Data and theory. In Memory, Second Edition, Handbook of Perception and Cognition E.L. Bjork, and R.A. Bjork, eds. (San Diego, Academic Press), pp. 67-99.

McAdams, S. (1989). Segregation of concurrent sounds. I: Effects of frequency modulation coherence. J Acoust Soc Am 86, 2148-2159.

McAdams, S. (1993). Recognition of sound sources and events In Thinking in sound : the cognitive psychology of human audition, S. McAdams, and E. Bigand, eds. (Oxford, Oxford University Press), pp. 146-198.

McAdams, S., and Bigand, E. (1993). Thinking in sound : the cognitive psychology of human audition, Oxford University Press edn (Oxford).

McDonald, C., and Stewart, L. (2008). Uses and functions of music in congenital amusia. Music Perception 25, 345-355.

McFarland, D.J., and Cacace, A.T. (1992). Aspects of short-term acoustic recognition memory: modality and serial position effects. Audiology 31, 342-352.

McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices. Nature 264, 746-748.

McKeown, D., and Wellsted, D. (2009). Auditory memory for timbre. J Exp Psychol Hum Percept Perform 35, 855-875.

Meddis, R., and Hewitt, M.J. (1991a). Virtual pitch and phase sensitivity of a computer model of the auditory periphery. i: Pitch identification. J Acoust Soc Am 89, 2866–2882.

Meddis, R., and Hewitt, M.J. (1991b). Virtual pitch and phase sensitivity of a computer model of the auditory periphery. ii: Phase sensitivity. J Acoust Soc Am 89.

Meddis, R., and O'Mard, L. (1997). A unitary model of pitch perception. J Acoust Soc Am 102, 1811-1820.

Melara, R.D., and O'Brien, T.P. (1987). Interaction between synesthetically corresponding dimensions. Journal of Experimental Psychology: General 116, 323-336

Mercer, T., and McKeown, D. (2010). Updating and feature overwriting in short-term memory for timbre. Atten Percept Psychophys 72, 2289-2303.

Meyer, L., Obleser, J., and Friederici, A.D. (2013). Left parietal alpha enhancement during working memory-intensive sentence processing. Cortex 49, 711-721.

Mignault Goulet, G., Moreau, P., Robitaille, N., and Peretz, I. (2012). Congenital amusia persists in the developing brain after daily music listening. PLoS One 7, e36860.

Miller, G.A., Galanter, E., and Pribram, K.H. (1960). Plans and the structure of behavior (New York, Holt, Rinehart and Winston, Inc.).

Miller, J. (1982). Divided attention: evidence for coactivation with redundant signals. Cogn Psychol 14, 247-279.

Milner, B. (1966). Amnesia following operation on the temporal lobes. In Amnesia, C.W.M. Whitty, and O.L. Zangwill, eds. (London, Butterworth).

Mishkin, M., Ungerleider, L.G., and Macko, K.A. (1983). Object vision and spatial vision: two cortical pathways. Trends Neurosci 6, 414–417.

Moore, B.C.J. (2003). An introduction to the psychology of hearing, Academic Press edn (San Diego, USA.).

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

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., and Zilles, K. (2001). Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. Neuroimage 13, 684-701.

Muller, N.G., and Knight, R.T. (2006). The functional neuroanatomy of working memory: contributions of human brain lesion studies. Neuroscience 139, 51-58.

Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. Biol Cybern 66, 241-251.

Näätänen, R., and 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.

Nairne, J.S. (1992). The loss of positional certainty in long-term memory. Psychological Science 3, 199-202.

Nan, Y., Sun, Y., and Peretz, I. (2010). Congenital amusia in speakers of a tone language: association with lexical tone agnosia. Brain 133, 2635-2642.

Neath, I. (1998). Human memory: An introduction to research, data, and theory (Pacific Grove, Brooks/Cole).

Nguyen, S., Tillmann, B., Gosselin, N., and Peretz, I. (2009). Tonal language processing in congenital amusia. Ann N Y Acad Sci 1169, 490-493.

Nicholson, K.G., Baum, S., Cuddy, L.L., and Munhall, K.G. (2002). A case of impaired auditory and visual speech prosody perception after right hemisphere damage. Neurocase 8, 314-322.

Nicholson, K.G., Baum, S., Kilgour, A., Koh, C.K., Munhall, K.G., and Cuddy, L.L. (2003). Impaired processing of prosodic and musical patterns after right hemisphere damage. Brain Cogn 52, 382-389.

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

Niessing, J., Ebisch, B., Schmidt, K.E., Niessing, M., Singer, W., and Galuske, R.A. (2005). Hemodynamic signals correlate tightly with synchronized gamma oscillations. Science 309, 948-951.

Norman, D.A., and Shallice, T. (1986). Attention to action: willed and automatic control of behaviour. In Consciousness and Self-regulation Advances

in Research and Theory, R.J. Davidson, G.E. Schwarts, and D. Shapiro, eds. (New York, Plenum), pp. 1–18.

Nystrom, L.E., Braver, T.S., Sabb, F.W., Delgado, M.R., Noll, D.C., and Cohen, J.D. (2000). Working memory for letters, shapes, and locations: fMRI evidence against stimulus-based regional organization in human prefrontal cortex. Neuroimage 11, 424-446.

Obleser, J., and Weisz, N. (2012). Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. Cereb Cortex 22, 2466-2477.

Obleser, J., Wostmann, M., Hellbernd, N., Wilsch, A., and Maess, B. (2012). Adverse listening conditions and memory load drive a common alpha oscillatory network. J Neurosci 32, 12376-12383.

Obrig, H., and Villringer, A. (1997). Near-infrared spectroscopy in functional activation studies. Can NIRS demonstrate cortical activation? Adv Exp Med Biol 413, 113-127.

Ohm, G.S. (1843). Über die Definition des Tones, nebst daran geknüfter Theorie der Sirene and ähnlicher tonbildender Vorichtungen. Ann Phys Chem 59, 513-565.

Ollinger, J.M. (1987). Estimation algorithms for dynamic tracer studies using positron-emission tomography. IEEE Trans Med Imaging 6, 115-125.

Olson, I.R., and Berryhill, M. (2009). Some surprising findings on the involvement of the parietal lobe in human memory. Neurobiol Learn Mem 91, 155-165.

Omigie, D., Müllensiefen, D., and Stewart, L. (2012a). The experience of music in congenital amusia. Music Perception 30, 1-18.

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

Omigie, D., Pearce, M.T., Williamson, V.J., and Stewart, L. (2013). Electrophysiological correlates of melodic processing in congenital amusia. Neuropsychologia 51, 1749-1762.

Omigie, D., and Stewart, L. (2011). Preserved statistical learning of tonal and linguistic material in congenital amusia. Front Psychol 2, 109.

Otto, T., and Eichenbaum, H. (1992). Complementary roles of the orbital prefrontal cortex and the perirhinal-entorhinal cortices in an odor-guided delayed-nonmatching-to-sample task. Behav Neurosci 106, 762-775.

Overath, T., Cusack, R., Kumar, S., von Kriegstein, K., Warren, J.D., Grube, M., Carlyon, R.P., and Griffiths, T.D. (2007). An information theoretic characterisation of auditory encoding. PLoS Biol 5, e288.

Owen, A.M. (2000). The role of the lateral frontal cortex in mnemonic processing: the contribution of functional neuroimaging. Exp Brain Res 133, 33–43.

Owen, A.M., Evans, A.C., and Petrides, M. (1996). Evidence for a two-stage model of spatial working memory processing within the lateral frontal cortex: a positron emission tomography study. Cereb Cortex 6, 31-38.

Palva, S., and Palva, J.M. (2007). New vistas for alpha-frequency band oscillations. Trends Neurosci 30, 150-158.

Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G., and Elbert, T. (1995). Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. Electroencephalogr Clin Neurophysiol 94, 26-40.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L.E., and Hoke, M. (1998). Increased auditory cortical representation in musicians. Nature 392, 811-814.

Pantev, C., Roberts, L.E., Elbert, T., Ross, B., and Wienbruch, C. (1996). Tonotopic organization of the sources of human auditory steady-state responses. Hear Res 101, 62-74.

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

Paquette, S., Gosselin, N., and Peretz, I. (2011). Normal use of mode and tempo in emotional judgments by amusic adults. In The Neurosciences and Music-IV Learning and Memory Meeting, Edinburgh, UK.

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

Patel, A.D., Peretz, I., Tramo, M., and Labreque, R. (1998). Processing prosodic and musical patterns: a neuropsychological investigation. Brain Lang 61, 123-144.

Patel, A.D., Wong, M., Foxton, J., Lochy, A., and Peretz, I. (2008). Speech Intonation Perception Deficits in Musical Tone Deafness (congenital amusia). Music Perception 25, 357-368.

Patterson, R.D., Uppenkamp, S., Johnsrude, I.S., and Griffiths, T.D. (2002). The processing of temporal pitch and melody information in auditory cortex. Neuron 36, 767-776.

Paulesu, E., Frith, C.D., and Frackowiak, R.S. (1993). The neural correlates of the verbal component of working memory. Nature 362, 342-345.

Paus, T., Petrides, M., Evans, A.C., and Meyer, E. (1993). Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: a positron emission tomography study. J Neurophysiol 70, 453-469.

Pechmann, T., and Mohr, G. (1992). Interference in memory for tonal pitch: implications for a working-memory model. Mem Cognit 20, 314-320.

Penagos, H., Melcher, J.R., and Oxenham, A.J. (2004). A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. J Neurosci 24, 6810-6815.

Penhune, V.B., Zatorre, R.J., and Feindel, W.H. (1999). The role of auditory cortex in retention of rhythmic patterns as studied in patients with temporal lobe removals including Heschl's gyrus. Neuropsychologia 37, 315–331.

Penny, W.D., Stephan, K.E., Mechelli, A., and Friston, K.J. (2004). Modelling functional integration: a comparison of structural equation and dynamic causal models. Neuroimage 23 Suppl 1, S264-274.

Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. Brain 113 (Pt 4), 1185-1205.

Peretz, I. (1993). Auditory agnosia: a functional analysis. In Thinking in Sound The Cognitive Psychology of Human Audition, S. McAdams, and E. Bigand, eds. (New York, Oxford University Press), pp. 199–230.

Peretz, I. (1996). Can we lose memory for music? A case of music agnosia in a nonmusician. J Cogn Neurosci 8, 481-496.

Peretz, I. (2001). Brain specialization for music. New evidence from congenital amusia. Ann N Y Acad Sci 930, 153-165.

Peretz, I. (2002). Brain specialization for music. Neuroscientist 8, 372-380.

Peretz, I. (2003). Brain specialization for music: New evidence from congenital amusia. In The Cognitive Neuroscience of Music, I.P.I.a.Z. RJ, ed. (Oxford, Oxford University Press), pp. 192-203.

Peretz, I. (2008). Musical Disorders: From Behavior to Genes. Current Directions in Psychological Science 17, 329-333.

Peretz, I. (2013). The Biological foundations of music: Insights from congenital amusia. In The Psychology of Music, D. Deutsch, ed. (Elsevier), p. in press.

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

Peretz, I., Belleville, S., and Fontaine, S. (1997). Dissociations between music and language functions after cerebral resection: A new case of amusia without aphasia. Can J Exp Psychol 51, 354-368.

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

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

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

Peretz, I., and Coltheart, M. (2003). Modularity of music processing. Nat Neurosci 6, 688-691.

Peretz, I., Cummings, S., and Dube, M.P. (2007). The genetics of congenital amusia (tone deafness): a family-aggregation study. Am J Hum Genet 81, 582-588.

Peretz, I., Gosselin, N., Tillmann, B., Cuddy, L.L., Gagnon, B., Trimmer, C., Paquette, S., and Bouchard, B. (2008). On-line Identification of Congenital Amusia. Music Perception 25, 331-343.

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

Peretz, I., and Kolinsky, R. (1993). Boundaries of separability between melody and rhythm in music discrimination: a neuropsychological perspective. Q J Exp Psychol A 46, 301-325.

Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., and Belleville, S. (1994). Functional dissociations following bilateral lesions of auditory cortex. Brain 117 (Pt 6), 1283-1301.

Peretz, I., Saffran, J., Schon, D., and Gosselin, N. (2012). Statistical learning of speech, not music, in congenital amusia. Ann N Y Acad Sci 1252, 361-367.

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

Pernier, J. (2007). Electro et magnéto encéphalographie biophysique, techniques et méthodes (Paris, Hermes Science,Lavoisier).

Petkov, C.I., Kayser, C., Augath, M., and Logothetis, N.K. (2006). Functional imaging reveals numerous fields in the monkey auditory cortex. PLoS Biol 4, e215.

Petrides, M. (1991). Functional specialization within the dorsolateral frontal cortex for serial order memory. Proc Biol Sci 246, 299-306.

Petrides, M. (1994). Frontal lobes and behaviour. Curr Opin Neurobiol 4, 207-211.

Petrides, M. (1995a). Functional organization of the human frontal cortex for mnemonic processing. Evidence from neuroimaging studies. Ann N Y Acad Sci 769, 85-96.

Petrides, M. (1995b). Impairments on nonspatial self-ordered and externally ordered working memory tasks after lesions of the mid-dorsal part of the lateral frontal cortex in the monkey. J Neurosci 15, 359-375.

Petrides, M., and Pandya, D.N. (1999). Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. Eur J Neurosci 11, 1011-1036.

Petrides, M., and Pandya, D.N. (2002). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. Eur J Neurosci 16, 291-310.

Petrides, M., Tomaiuolo, F., Yeterian, E.H., and Pandya, D.N. (2012). The prefrontal cortex: comparative architectonic organization in the human and the macaque monkey brains. Cortex 48, 46-57.

Pfeuty, M., and Peretz, I. (2010). Abnormal pitch--time interference in congenital amusia: evidence from an implicit test. Atten Percept Psychophys 72, 763-774.

Pfordresher, P.Q., and Brown, S. (2007). Poor-pitch singing in the absence of tone deafness. Music Perception 25, 95-115.

Phillips-Silver, J., Toiviainen, P., Gosselin, N., Piche, O., Nozaradan, S., Palmer, C., and Peretz, I. (2011). Born to dance but beat deaf: a new form of congenital amusia. Neuropsychologia 49, 961-969.

Piccirilli, M., Sciarma, T., and Luzzi, S. (2000). Modularity of music: evidence from a case of pure amusia. J Neurol Neurosurg Psychiatry 69, 541-545.

Pickering, S.J., Gathercole, S.E., Hall, M., and Lloyd, S.A. (2001). Development of memory for pattern and path: further evidence for the fractionation of visuo-spatial memory. Q J Exp Psychol A 54, 397-420.

Pickles, J.O. (1988). An introduction to the physiology of hearing (London, Academic Press).

Picton, T.W., Hillyard, S.A., Krausz, H.I., and Galambos, R. (1974). Human auditory evoked potentials. I. Evaluation of components. Electroencephalogr Clin Neurophysiol 36, 179-190.

Plomp (1964). The ear as a frequency analyzer. J Acoust Soc Am 36, 1628–1636.

Polat, U., Ma-Naim, T., Belkin, M., and Sagi, D. (2004). Improving vision in adult amblyopia by perceptual learning. Proc Natl Acad Sci U S A 101, 6692-6697.

Polat, U., Ma-Naim, T., and Spierer, A. (2009). Treatment of children with amblyopia by perceptual learning. Vision Res 49, 2599-2603.

Polster, M.R., and Rose, S.B. (1998). Disorders of auditory processing: evidence for modularity in audition. Cortex 34, 47-65.

Poremba, A., Saunders, R.C., Crane, A.M., Cook, M., Sokoloff, L., and Mishkin, M. (2003). Functional mapping of the primate auditory system. Science 299, 568-572.

Posner, M.I., Nissen, M.J., and Klein, R.M. (1976). Visual dominance: an information-processing account of its origins and significance. Psychol Rev 83, 157-171.

Postle, B.R., Berger, J.S., and D'Esposito, M. (1999). Functional neuroanatomical double dissociation of mnemonic and executive control processes contributing to working memory performance. Proc Natl Acad Sci U S A 96, 12959-12964.

Postle, B.R., and D'Esposito, M. (1999). "What"-Then-Where" in visual working memory: an event-related fMRI study. J Cogn Neurosci 11, 585-597.

Postle, B.R., Stern, C.E., Rosen, B.R., and Corkin, S. (2000a). An fMRI investigation of cortical contributions to spatial and nonspatial visual working memory. Neuroimage 11, 409-423.

Postle, B.R., Zarahn, E., and D'Esposito, M. (2000b). Using event-related fMRI to assess delay-period activity during performance of spatial and nonspatial working memory tasks. Brain Res Brain Res Protoc 5, 57-66.

Raab, D.H. (1962). Statistical facilitation of simple reaction times. Trans N Y Acad Sci 24, 574-590.

Ranganath, C., and Blumenfeld, R.S. (2005). Doubts about double dissociations between short- and long-term memory. Trends Cogn Sci 9, 374-380.

Rao, R.P., and Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat Neurosci 2, 79-87.

Rauschecker, J.P. (1998). Cortical processing of complex sounds. Curr Opin Neurobiol 8, 516-521.

Rauschecker, J.P., and Scott, S.K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat Neurosci 12, 718-724.

Rauschecker, J.P., and Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc Natl Acad Sci U S A 97, 11800-11806.

Rauschecker, J.P., and Tian, B. (2004). Processing of band-passed noise in the lateral auditory belt cortex of the rhesus monkey. J Neurophysiol 91, 2578-2589.

Rauschecker, J.P., Tian, B., and Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. Science 268, 111-114.

Rauschecker, J.P., Tian, B., Pons, T., and Mishkin, M. (1997). Serial and parallel processing in rhesus monkey auditory cortex. J Comp Neurol 382, 89-103.

Ravizza, S.M., Delgado, M.R., Chein, J.M., Becker, J.T., and Fiez, J.A. (2004). Functional dissociations within the inferior parietal cortex in verbal working memory. Neuroimage 22, 562-573.

Reale, R.A., Calvert, G.A., Thesen, T., Jenison, R.L., Kawasaki, H., Oya, H., Howard, M.A., and Brugge, J.F. (2007). Auditory-visual processing represented in the human superior temporal gyrus. Neuroscience 145, 162-184.

Recanzone, G.H. (2011). Perception of auditory signals. Ann N Y Acad Sci 1224, 96-108.

Recanzone, G.H., Schreiner, C.E., and Merzenich, M.M. (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. J Neurosci 13, 87-103.

Reid, T., and Hamilton, W. (1850). Essays on the Intellectual Powers of Man abridged with notes and illustrations from Sir William and others (Cambridge, Bartlett, J.).

Rivier, F., and Clarke, S. (1997). Cytochrome oxidase, acetylcholinesterase, and NADPH-diaphorase staining in human supratemporal and insular cortex: evidence for multiple auditory areas. Neuroimage 6, 288-304.

Rockland, K.S., and Ojima, H. (2003). Multisensory convergence in calcarine visual areas in macaque monkey. Int J Psychophysiol 50, 19-26.

Romani, G.L. (1986). Tonotopic organization of the human auditory cortex revealed by steady state neuromagnetic measurements. Acta Otolaryngol Suppl 432, 33-34.

Romani, G.L., Williamson, S.J., and Kaufman, L. (1982). Tonotopic organization of the human auditory cortex. Science 216, 1339-1340.

Romanski, L.M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P.S., and Rauschecker, J.P. (1999). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat Neurosci 2, 1131-1136.

Ross, L.A., Saint-Amour, D., Leavitt, V.M., Javitt, D.C., and Foxe, J.J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. Cereb Cortex 17, 1147-1153.

References

Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., and Barone, P. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. Proc Natl Acad Sci U S A 104, 7295-7300.

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

Saenz, M., and Langers, D.R. (2013). Tonotopic mapping of human auditory cortex. Hear Res.

Saffran, J.R. (2003). Musical learning and language development. Ann N Y Acad Sci 999, 397-401.

Sala, J.B., and Courtney, S.M. (2007). Binding of what and where during working memory maintenance. Cortex 43, 5-21.

Sala, J.B., Rama, P., and Courtney, S.M. (2003). Functional topography of a distributed neural system for spatial and nonspatial information maintenance in working memory. Neuropsychologia 41, 341-356.

Salame, P., and Baddeley, A. (1989). Effects of background music on phonological short-term memory. The Quarterly Journal of Experimental Psychology 41A, 107-122.

Salame, P., and Baddeley, A.D. (1987). Noise, unattended speech and short-term memory. Ergonomics 30, 1185-1193.

Samson, S., and Zatorre, R.J. (1992). Learning and retention of melodic and verbal information after unilateral temporal lobectomy. Neuropsychologia 30, 815-826.

Sarkisov, S.A., Filimonoff, I.N., Kononowa, I.P., Preobrazeanskaja, N.S., and Kukuewa, L.A. (1955). Atlas of the cytoarchitectonics of the human cerebral cortex (in Russian) (Moscow, Medgiz).

Savage, C.R., Deckersbach, T., Heckers, S., Wagner, A.D., Schacter, D.L., Alpert, N.M., Fischman, A.J., and Rauch, S.L. (2001). Prefrontal regions supporting spontaneous and directed application of verbal learning strategies: evidence from PET. Brain 124, 219-231.

Scheich, H., Brechmann, A., Brosch, M., Budinger, E., and Ohl, F.W. (2007). The cognitive auditory cortex: task-specificity of stimulus representations. Hear Res 229, 213-224.

Scheich, H., Heil, P., and Langner, G. (1993). Functional organization of auditory cortex in the mongolian gerbil (Meriones unguiculatus). II. Tonotopic 2-deoxyglucose. Eur J Neurosci 5, 898-914.

Schendel, Z.A., and Palmer, C. (2007). Suppression effects on musical and verbal memory. Mem Cognit 35, 640-650.

Schiltz, C., Bodart, J.M., Dubois, S., Dejardin, S., Michel, C., Roucoux, A., Crommelinck, M., and Orban, G.A. (1999). Neuronal mechanisms of perceptual learning: changes in human brain activity with training in orientation discrimination. Neuroimage 9, 46-62.

Schonwiesner, M., von Cramon, D.Y., and Rubsamen, R. (2002). Is it tonotopy after all? Neuroimage 17, 1144-1161.

Schonwiesner, M., and Zatorre, R.J. (2008). Depth electrode recordings show double dissociation between pitch processing in lateral Heschl's gyrus and sound onset processing in medial Heschl's gyrus. Exp Brain Res 187, 97-105.

Schoups, A., Vogels, R., Qian, N., and Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. Nature 412, 549-553.

Schulze, K., Dowling, W.J., and Tillmann, B. (2012). Working Memory for Tonal and Atonal Sequences during a Forward and a Backward Recognition Task. Music Perception 29, 255-267.

Schulze, K., Gaab, N., and Schlaug, G. (2009). Perceiving pitch absolutely: comparing absolute and relative pitch possessors in a pitch memory task. BMC Neurosci 10, 106.

Schulze, K., and Koelsch, S. (2012). Working memory for speech and music. Ann N Y Acad Sci 1252, 229-236.

Schulze, K., Mueller, K., and Koelsch, S. (2011a). Neural correlates of strategy use during auditory working memory in musicians and non-musicians. Eur J Neurosci 33, 189-196.

Schulze, K., and Tillmann, B. (2013). Working memory for pitch, timbre, and words. Memory 21, 377-395.

Schulze, K., Zysset, S., Mueller, K., Friederici, A.D., and Koelsch, S. (2011b). Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. Hum Brain Mapp 32, 771-783.

Sedley, W., Teki, S., Kumar, S., Overath, T., Barnes, G.R., and Griffiths, T.D. (2012). Gamma band pitch responses in human auditory cortex measured with magnetoencephalography. Neuroimage 59, 1904-1911.

Seitz, A.R., Kim, R., and Shams, L. (2006). Sound facilitates visual learning. Curr Biol 16, 1422-1427.

Semal, C., Demany, L., Ueda, K., and Halle, P.A. (1996). Speech versus nonspeech in pitch memory. J Acoust Soc Am 100, 1132-1140.

Serino, A., Farne, A., Rinaldesi, M.L., Haggard, P., and Ladavas, E. (2007). Can vision of the body ameliorate impaired somatosensory function? Neuropsychologia 45, 1101-1107.

Shallice, T., and Warrington, E.K. (1970). Independent functioning of verbal memory stores: a neuropsychological study. Q J Exp Psychol 22, 261-273.

Shamma, S., and Klein, D. (2000). The case of the missing pitch templates: how harmonic templates emerge in the early auditory system. J Acoust Soc Am 107, 2631-2644.

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

Shelton, B.R., and Searle, C.L. (1980). The influence of vision on the absolute identification of sound-source position. Percept Psychophys 28, 589-596.

Silveri, M.C., Di Betta, A.M., Filippini, V., Leggio, M.G., and Molinari, M. (1998). Verbal short-term store-rehearsal system and the cerebellum. Evidence from a patient with a right cerebellar lesion. Brain 121 (Pt 11), 2175-2187.

Singer, W. (1999). Neuronal synchrony: a versatile code for the definition of relations? Neuron 24, 49-65, 111-125.

Singer, W., and Gray, C.M. (1995). Visual feature integration and the temporal correlation hypothesis. Annu Rev Neurosci 18, 555-586.

Sloboda, J.A., and Parker, D.H.H. (1985). Immediate recall of melodies. In Musical structure and cognition, P. Howell, I. Cross, and R. West, eds. (London, Academic Press).

Sloboda, J.A., Wise, K.J., and Peretz, I. (2005). Quantifying tone deafness in the general population. Ann N Y Acad Sci 1060, 255-261.

Small, A.M., Jr., and Campbell, R.A. (1962). Temporal differential sensitivity for auditory stimuli. Am J Psychol 75, 401-410.

Smith, E.E., and Jonides, J. (1998). Neuroimaging analyses of human working memory. Proc Natl Acad Sci U S A 95, 12061-12068.

Smith, R.L. (1977). Short-term adaptation in single auditory nerve fibers: some poststimulatory effects. J Neurophysiol 40, 1098-1111.

Smyth, M.M., and Pendleton, L.R. (1989). Working memory for movements. Q J Exp Psychol A 41, 235-250.

Sorkin, R.D. (1966). Temporal interference effects in auditory amplitude discrimination. Perception & Psychophysics 1, 55-58.

Sporns, O. (2013). The human connectome: Origins and challenges. Neuroimage 80, 53-61.

Stein, B.E., and Meredith, M.A. (1993). Merging of the senses (Cambridge, MIT Press).

Steinke, W.R., Cuddy, L.L., and Jakobson, L.S. (2001). Dissociations among functional subsystems governing melody recognition after right-hemisphere damage. Cogn Neuropsychol 18, 411-437.

Steinschneider, M., Nourski, K.V., Kawasaki, H., Oya, H., Brugge, J.F., and Howard, M.A., 3rd (2011). Intracranial study of speech-elicited activity on the human posterolateral superior temporal gyrus. Cereb Cortex 21, 2332-2347.

Stewart, L. (2011). Characterizing congenital amusia. Q J Exp Psychol (Hove) 64, 625-638.

Stewart, L., Overath, T., Warren, J.D., Foxton, J.M., and Griffiths, T.D. (2008). fMRI evidence for a cortical hierarchy of pitch pattern processing. PLoS One 3, e1470.

Stewart, L., von Kriegstein, K., Warren, J.D., and Griffiths, T.D. (2006). Music and the brain: disorders of musical listening. Brain 129, 2533-2553.

Stewart, L., and Walsh, V. (2007). Music perception: sounds lost in space. Curr Biol 17, R892-893.

Sumby, W.H., and Pollack, I. (1954). Visual contribution to speech intelligibility in noise. J Acoust Soc Am 26, 212–215.

Sweet, R.A., Dorph-Petersen, K.A., and Lewis, D.A. (2005). Mapping auditory core, lateral belt, and parabelt cortices in the human superior temporal gyrus. J Comp Neurol 491, 270-289.

Szentagothai, J. (1978). The Ferrier Lecture, 1977. The neuron network of the cerebral cortex: a functional interpretation. Proc R Soc Lond B Biol Sci 201, 219-248.

Talavage, T.M., Sereno, M.I., Melcher, J.R., Ledden, P.J., Rosen, B.R., and Dale, A.M. (2004). Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity. J Neurophysiol 91, 1282-1296.

Tallal, P. (2004). Improving language and literacy is a matter of time. Nat Rev Neurosci 5, 721-728.

Tallon-Baudry, C., and Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. Trends Cogn Sci 3, 151-162.

Tanibuchi, I., and Goldman-Rakic, P.S. (2003). Dissociation of spatial-, object-, and sound-coding neurons in the mediodorsal nucleus of the primate thalamus. J Neurophysiol 89, 1067-1077.

Tanner, W.P. (1961). Physiological implications of psychophysical data. Ann NY Acad Sci 89, 752–765.

Tardif, E., Spierer, L., Clarke, S., and Murray, M.M. (2008). Interactions between auditory 'what' and 'where' pathways revealed by enhanced near-threshold discrimination of frequency and position. Neuropsychologia 46, 958-966.

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

Terhardt, E. (1974). Pitch, consonance, and harmony. J Acoust Soc Am 55, 1061-1069.

Thompson, W.F. (2007). Exploring variants of amusia: Tone deafness, rhythm impairment, and intonation insensitivity. In The inaugural International Conference on Music Communication Science 5-7 December 2007 (Sydney Australia http://marcs.uws.edu.au/links/ICoMusic).

Thompson, W.F., Marin, M.M., and 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.

Thorn, A.S., Gathercole, S.E., and Frankish, C.R. (2005). Redintegration and the benefits of long-term knowledge in verbal short-term memory: an evaluation of Schweickert's (1993) multinomial processing tree model. Cogn Psychol 50, 133-158.

Tian, B., and Rauschecker, J.P. (2004). Processing of frequency-modulated sounds in the lateral auditory belt cortex of the rhesus monkey. J Neurophysiol 92, 2993-3013.

Tillmann, B. (2005). Implicit investigations of tonal knowledge in nonmusician listeners. Ann N Y Acad Sci 1060, 100-110.

Tillmann, B., Albouy, P., and Caclin, A. (2013). Congenital Amusias. In Handbook of Clinical Neurology (Oxford, Elsevier), p. in press.

Tillmann, B., Bharucha, J.J., and Bigand, E. (2000). Implicit learning of tonality: a self-organizing approach. Psychol Rev 107, 885-913.

Tillmann, B., Burnham, D., Nguyen, S., Grimault, N., Gosselin, N., and Peretz, I. (2011a). Congenital Amusia (or Tone-Deafness) Interferes with Pitch Processing in Tone Languages. Front Psychol 2, 120.

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

Tillmann, B., Janata, P., and Bharucha, J.J. (2003). Activation of the inferior frontal cortex in musical priming. Brain Res Cogn Brain Res 16, 145-161.

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

Tillmann, B., Koelsch, S., Escoffier, N., Bigand, E., Lalitte, P., Friederici, A.D., and von Cramon, D.Y. (2006). Cognitive priming in sung and instrumental music: activation of inferior frontal cortex. Neuroimage 31, 1771-1782.

Tillmann, B., Peretz, I., Bigand, E., and Gosselin, N. (2007). Harmonic priming in an amusic patient: the power of implicit tasks. Cogn Neuropsychol 24, 603-622.

Tillmann, B., Peretz, I., and Samson, S. (2011b). Neurocognitive Approaches to Memory in Music: Music is Memory. In The Memory Process: Neuroscientific and Humanistic Perspective, S. Nalbatian, P. Matthews, and J.L. McClelland, eds. (MIT press), pp. 377-394.

Tillmann, B., Rusconi, E., Traube, C., Butterworth, B., Umilta, C., and Peretz, I. (2011c). Fine-grained pitch processing of music and speech in congenital amusia. J Acoust Soc Am 130, 4089-4096.

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

Tremblay-Champoux, A., Dalla Bella, S., Phillips-Silver, J., Lebrun, M.A., and Peretz, I. (2010). Singing proficiency in congenital amusia: imitation helps. Cogn Neuropsychol 27, 463-476.

Tulving, E. (1987). Multiple memory systems and consciousness. Hum Neurobiol 6, 67-80.

Upadhyay, J., Ducros, M., Knaus, T.A., Lindgren, K.A., Silver, A., Tager-Flusberg, H., and Kim, D.S. (2007). Function and connectivity in human primary auditory cortex: a combined fMRI and DTI study at 3 Tesla. Cereb Cortex 17, 2420-2432.

Vallar, G. (2006). Memory systems: The case of phonological short-term memory. A festschrift for Cognitive Neuropsychology. Cogn Neuropsychol 23, 135-155.

Vallar, G., and Baddeley, A.D. (1984). Phonological short-term store, phonological processing and sentence comprehension: A neuropsychological case study. Cognitive Neuropsychology 1, 121-141.

Vallar, G., Di Betta, A.M., and Silveri, M.C. (1997). The phonological short-term store-rehearsal system: patterns of impairment and neural correlates. Neuropsychologia 35, 795-812.

Vallar, G., and Papagno, C. (2002). Neuropsychological impairments of verbal short-term memory. In The Handbook of Memory Disorders, M.D.K.B.A.W. A.D. Baddeley, ed. (Chichester, Wiley), pp. 249–270.

van Dijk, H., Nieuwenhuis, I.L., and Jensen, O. (2010). Left temporal alpha band activity increases during working memory retention of pitches. Eur J Neurosci 31, 1701-1707.

Varela, F., Lachaux, J.P., Rodriguez, E., and Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. Nat Rev Neurosci 2, 229-239.

Vincent, S.B. (1912). The function of the vibrissae in the behavior of the white rat. Behavior Monographs 1.

von Economo, C., and Horn, L. (1930). Über Windungsrelief, Maße und Rindenarchitektonik des Supratemporalfläche, Ihre individuellen und Ihre Seitenunterschiede. Z Neurol Psychiat 130, 678–757.

von Economo, C., and Koskinas, G.H. (1925). Die Cytoarchitektonik der Hirnrinde des erwachsenen Menschen (Wien-Berlin, Springer).

von Kriegstein, K. (2012). A Multisensory Perspective on Human Auditory Communication.

Vos, P.G., and Troost, J.M. (1989). Ascending and descending melodic intervals: statistical findings and their perceptual relevance. Music Perception 6, 383–396.

Vroomen, J., and de Gelder, B. (2000). Sound enhances visual perception: cross-modal effects of auditory organization on vision. J Exp Psychol Hum Percept Perform 26, 1583-1590.

Walker, K.M., Bizley, J.K., King, A.J., and Schnupp, J.W. (2011). Cortical encoding of pitch: recent results and open questions. Hear Res 271, 74-87.

Wallace, M.N., Johnston, P.W., and Palmer, A.R. (2002). Histochemical identification of cortical areas in the auditory region of the human brain. Exp Brain Res 143, 499-508.

Wallis, J.D., and Miller, E.K. (2003). From rule to response: neuronal processes in the premotor and prefrontal cortex. J Neurophysiol 90, 1790-1806.

Wang, X.J. (2001). Synaptic reverberation underlying mnemonic persistent activity. Trends Neurosci 24, 455-463.

Warren, J.D., Zielinski, B.A., Green, G.G., Rauschecker, J.P., and Griffiths, T.D. (2002). Perception of sound-source motion by the human brain. Neuron 34, 139-148.

Warren, R.M. (2008). Auditory Perception. An Analysis and Synthesis third edn (Cambridge, Cambridge University Press).

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

Watanabe, Y., and Funahashi, S. (2004a). Neuronal activity throughout the primate mediodorsal nucleus of the thalamus during oculomotor delayed-responses. I. Cue-, delay-, and response-period activity. J Neurophysiol 92, 1738-1755.

Watanabe, Y., and Funahashi, S. (2004b). Neuronal activity throughout the primate mediodorsal nucleus of the thalamus during oculomotor delayed-responses. II. Activity encoding visual versus motor signal. J Neurophysiol 92, 1756-1769.

Watkins, K.E., Vargha-Khadem, F., Ashburner, J., Passingham, R.E., Connelly, A., Friston, K.J., Frackowiak, R.S., Mishkin, M., and Gadian, D.G. (2002). MRI analysis of an inherited speech and language disorder: structural brain abnormalities. Brain 125, 465-478.

Watson, C.S., Foyle, D.C., and Kidd, G.R. (1990). Limits of auditory patterns discrimination for patterns with various durations and numbers of components. J Acoust Soc Am 88, 2631-2638.

Wendelken, C., Bunge, S.A., and Carter, C.S. (2008). Maintaining structured information: an investigation into functions of parietal and lateral prefrontal cortices. Neuropsychologia 46, 665-678.

Wessinger, C.M., VanMeter, J., Tian, B., Van Lare, J., Pekar, J., and Rauschecker, J.P. (2001). Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. J Cogn Neurosci 13, 1-7.

Williamson, V.J., Baddeley, A.D., and Hitch, G.J. (2010a). Musicians' and nonmusicians' short-term memory for verbal and musical sequences: comparing phonological similarity and pitch proximity. Mem Cognit 38, 163-175.

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

Williamson, V.J., Liu, F., Peryer, G., Grierson, M., and Stewart, L. (2012). Perception and action de-coupling in congenital amusia: sensitivity to task demands. Neuropsychologia 50, 172-180.

Williamson, V.J., McDonald, C., Deutsch, D., Griffiths, T.D., and Stewart, L. (2010b). Faster decline of pitch memory over time in congenital amusia. Adv Cogn Psychol 6, 15-22.

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

Winkler, I., and Cowan, N. (2005). From sensory to long-term memory: evidence from auditory memory reactivation studies. Exp Psychol 52, 3-20.

Winkler, I., Paavilainen, P., and Naatanen, R. (1992). Can echoic memory store two traces simultaneously? A study of event-related brain potentials. Psychophysiology 29, 337-349.

Wong, P.C., Ciocca, V., Chan, A.H., Ha, L.Y., Tan, L.H., and Peretz, I. (2012). Effects of culture on musical pitch perception. PLoS One 7, e33424.

Wood, C.C., and Wolpaw, J.R. (1982). Scalp distribution of human auditory evoked potentials. II. Evidence for overlapping sources and involvement of auditory cortex. Electroencephalogr Clin Neurophysiol 54, 25-38.

Wood, N., and Cowan, N. (1995). The cocktail party phenomenon revisited: how frequent are attention shifts to one's name in an irrelevant auditory channel? J Exp Psychol Learn Mem Cogn 21, 255-260.

Woods, D.L., Stecker, G.C., Rinne, T., Herron, T.J., Cate, A.D., Yund, E.W., Liao, I., and Kang, X. (2009). Functional maps of human auditory cortex: effects of acoustic features and attention. PLoS One 4, e5183.

Xiang, J., Chuang, S., Wilson, D., Otsubo, H., Pang, E., Holowka, S., Sharma, R., Ochi, A., and Chitoku, S. (2002). Sound motion evoked magnetic fields. Clin Neurophysiol 113, 1-9.

Yang, Y., Engelien, A., Engelien, W., Xu, S., Stern, E., and Silbersweig, D.A. (2000). A silent event-related functional MRI technique for brain activation studies without interference of scanner acoustic noise. Magn Reson Med 43, 185-190.

Yvert, B., Crouzeix, A., Bertrand, O., Seither-Preisler, A., and Pantev, C. (2001). Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. Cereb Cortex 11, 411-423.

Zacks, J.M. (2008). Neuroimaging studies of mental rotation: a meta-analysis and review. J Cogn Neurosci 20, 1-19.

Zatorre, R.J. (1985). Discrimination and recognition of tonal melodies after unilateral cerebral excisions. Neuropsychologia 23, 31-41.

Zatorre, R.J., Belin, P., and Penhune, V.B. (2002). Structure and function of auditory cortex: music and speech. Trends Cogn Sci 6, 37-46.

Zatorre, R.J., Bouffard, M., and Belin, P. (2004). Sensitivity to auditory object features in human temporal neocortex. J Neurosci 24, 3637-3642.

Zatorre, R.J., Evans, A.C., and Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. J Neurosci 14, 1908-1919.

Zatorre, R.J., and Samson, S. (1991). Role of the right temporal neocortex in retention of pitch in auditory short-term memory. Brain 114 (Pt 6), 2403-2417.