NeuroImage 102 (2014) 764–775

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Full-Length Article

The auditory scene: An fMRI study on melody and accompaniment in professional pianists

Danilo Spada [a], Laura Verga [b], Antonella Iadanza [c,d], Marco Tettamanti [d,e,f], Daniela Perani [a,d,e,f,*]

[a] Faculty of Psychology, Vita-Salute San Raffaele University, Milano, Italy
[b] Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[c] Division of Neuroradiology, San Raffaele Scientific Institute, Milano, Italy
[d] C.E.R.M.A.C. (Centro Eccellenza Risonanza Magnetica Alto Campo), San Raffaele Scientific Institute, Vita-Salute San Raffaele University, Milano, Italy
[e] Department of Nuclear Medicine, San Raffaele Scientific Institute, Milano, Italy
[f] Division of Neuroscience, San Raffaele Scientific Institute, Milano, Italy

ARTICLE INFO

Article history:
Accepted 20 August 2014
Available online 28 August 2014

Keywords:
Music
Harmony
Melody
Accompaniment
Salience
Auditory stream segregation

ABSTRACT

The auditory scene is a mental representation of individual sounds extracted from the summed sound waveform reaching the ears of the listeners. Musical contexts represent particularly complex cases of auditory scenes. In such a scenario, melody may be seen as the main object moving on a background represented by the accompaniment. Both melody and accompaniment vary in time according to harmonic rules, forming a typical texture with melody in the most prominent, salient voice. In the present sparse acquisition functional magnetic resonance imaging study, we investigated the interplay between melody and accompaniment in trained pianists, by observing the activation responses elicited by processing: (1) melody placed in the upper and lower texture voices, leading to, respectively, a higher and lower auditory salience; (2) harmonic violations occurring in either the melody, the accompaniment, or both. The results indicated that the neural activation elicited by the processing of polyphonic compositions in expert musicians depends upon the upper versus lower position of the melodic line in the texture, and showed an overall greater activation for the harmonic processing of melody over accompaniment. Both these two predominant effects were characterized by the involvement of the posterior cingulate cortex and precuneus, among other associative brain regions. We discuss the prominent role of the posterior medial cortex in the processing of melodic and harmonic information in the auditory stream, and propose to frame this processing in relation to the cognitive construction of complex multimodal sensory imagery scenes.

Introduction

In the past two decades, there has been a growing interest for music in cognitive neuroscience research, presumably because in musical perception as well as in musical production performed by either naive subjects or professional musicians, both general purpose and highly specialized cognitive functions are at play. In particular, polyphonic music listening, which is common in everyday exposure to Western tonal music of various styles and genres, involves processing both sounds and structural aspects of the acoustic input. The former are referred to with the term texture, and involve the distinction between a prominent, salient stream, called melody, and subordinate streams, collectively called accompaniment. The latter, in Western polyphonic tonal music, are described through the theory of harmony, which deals with the simultaneous combination of notes in chords and chord progressions. Harmony can be viewed as a description of the regularities and rules of the musical tonal system, which people internalize from the first days of life through mere exposure and on the basis of neurobiological constraints. Cognitive neuroscience has elucidated the neural correlates of the violation of such rules both in naive subjects and musicians, but the interplay between melodic salience and harmonic structure has been hitherto largely overlooked.

From the point of view of general purpose cognitive processes, auditory scene analysis (Bregman, 1990, 2007, 2008) mechanisms allow to segregate the auditory flow in streams with different degrees of salience through different kinds of acoustic cues, such as timbre, frequency proximity, spatial distribution, and onset time (Halpern et al., 1996; Hébert et al., 1995; Macherey and Delpierre, 2013; Palmer and Holleran, 1994; Maeder et al., 2001; Walker et al., 2011; Uhlig et al., 2013; Zatorre et al., 1999). With respect to auditory texture, the most salient stream is usually placed in the upper voice and it is referred to – in musical terms – as "melody", whereas the remaining voices in lower streams are collectively defined as "accompaniment". The neural mechanisms underlying polyphonic texture perception have been

* Corresponding author at: Vita-Salute San Raffaele University and Scientific Institute San Raffaele, Via Olgettina 60, I-20132 Milano, Italy. Fax: +39 02 26432717.
E-mail addresses: danilo.spada@harmonio.org (D. Spada), verga@cbs.mpg.de (L. Verga), iadanza.antonella@hsr.it (A. Iadanza), tettamanti.marco@hsr.it (M. Tettamanti), perani.daniela@hsr.it (D. Perani).

http://dx.doi.org/10.1016/j.neuroimage.2014.08.036

investigated particularly in auditory perceptual memory and auditory attention studies. Electro-physiological measurements often focused on the study of mismatch negativities (MMN). The MMN is an auditory Event Related Potential component originating in the auditory cortex around 100–200 ms post stimulus onset and reflecting the processing of infrequent or deviant stimuli (Näätänen et al., 2004, 2007). By investigating the MMN in musicians and in untrained controls, Fujioka et al. (2005) demonstrated that both an upper and a lower melodic voice are represented in the auditory perceptual memory; more importantly, the upper voice appears to be relatively more salient as reflected by larger MMN amplitudes for changes in the upper than in the lower voice in both population groups. Similarly, a study by Lee et al. (2009) reported higher auditory brainstem responses for upper versus lower tones, with trained musicians as experimental subjects, and a recent functional magnetic resonance imaging (fMRI) study by Uhlig et al. (2013) has evidenced the greater salience of melody over accompaniment when listening to complex multi-part musical stimuli. These authors found an increase in the activation of the fronto-parietal attention network when participants were instructed to pay attention to the melody stream as compared to the accompaniment. Finally, a Positron Emission Tomography (PET) study (Satoh et al., 2003) focusing on the contrast between listening to Soprano melodies as opposed to listening to harmony, found activations, notably in the precuneus (Soprano melodies > harmony) and in the anterior cingulate cortex (harmony > Soprano melodies).

From the complementary point of view of specialized cognitive functions, studies focusing on musical structure processing established that humans, on the basis of neurophysiological (Koelsch and Jentschke, 2010; Tramo et al., 2001) and of cognitive constraints present from birth (Perani et al., 2010; Trehub, 2001), are able to internalize through mere exposure the harmonic rules of the musical system of their culture. This observation may explain the capacity displayed by naive subjects and – to an even higher degree – by professional musicians to generate precise harmonic expectations during music perception (Carrión and Bly, 2008; Koelsch et al., 2000; Pearce et al., 2010; Tillmann and Bigand, 2001; Tillmann, 2005; Tillmann et al., 2006). The investigation of the neural underpinnings of this process has been extensively carried out with diverse neuroimaging techniques, most often using stimuli characterized by harmonic violations, usually realized through mistuned chords or harmonically incongruent cadences. The former are obtained by shifting the pitch of one or more notes of a chord, which nevertheless maintains its identity and harmonic role in the cadence. The latter are the result of the substitution of one or more notes of the chord totally modifying their harmonic role in the cadence. Such studies consistently reported the involvement of a network of cortical areas encompassing the bilateral superior temporal gyrus and the right inferior frontal cortex (Brattico et al., 2006; Garza Villarreal et al., 2011; Koelsch, 2005, 2011a, 2011b; Koelsch and Siebel, 2005; Maess et al., 2001; Miranda and Ullman, 2007; Ruiz et al., 2009).

While both melodic salience and harmonic structure have been thoroughly explored in musical research, the interplay between these two points of view has in our opinion been largely overlooked until the present. This gap in the literature becomes particularly problematic when considering the increased tendency in more recent studies to use ecological (polyphonic) musical stimuli. In such polyphonic compositions, it should be considered that melody carries in itself harmonic information, in a manner that, at least from the theoretical point of view, is largely independent from the harmonic information characterizing accompaniment. Some experimental hints in favor of this theoretical view are indeed available. Behavioral studies in children have evidenced that the capacity to process melodies develops earlier than the capacity to process accompaniment chords (Trainor and Trehub, 1994; Trehub et al., 1984). Electrophysiological data indicate that melody carries in itself harmonic information (Miranda and Ullman, 2007; Pearce et al., 2010), even pre-attentively (Brattico et al., 2006). One study by Koelsch and Jentschke (2010) directly compared harmonic violations occurring, respectively, in single melodies and harmonized melodies, and showed that melodic information – present both in melodies and in the top voice of chords as well – is processed earlier (N125, 100–150 ms) and more frontally than harmonic information (N180, 160–210 ms) which is processed more broadly over the scalp.

In our experiment, we investigated the neural correlates of the interplay between the processes responsible for melodic salience attribution and the processes involved in harmonic structure analysis. We tackled two crucial experimental questions. First, we investigated the neural correlates of melodic salience, contrasting activations in response to two different polyphonic textures, one with the melody as the uppermost of four voices (Soprano), and the other with the same melody displaced as the lowest voice (Bass). Second, we evaluated how the brain reacts when dealing with unrelated, unexpected musical events occurring in the harmonic context, i.e. harmonic violations, but clearly distinguishing between violations that occur in the melody from those occurring in the accompaniment.

To this aim, we focused on a population of trained musicians, as a previous study investigating harmonic expectations during melodic perception in adults emphasized that the degree of musical training influences the sensitivity to unexpected chords accompanying the harmonically correct melody (Loui and Wessel, 2007). We used sparse acquisition fMRI (Belin et al., 1999; Hall et al., 1999) to present trained musicians with excerpts of original melodies under optimal listening conditions. In order to resolve our first experimental question on the neural correlates of melodic salience, one set of experimental stimuli, all devoid of harmonic violations, was divided according to differing melodic salience (Soprano versus Bass voices). To resolve our second experimental question on the dependence of structural analysis on the harmonic context, we orthogonally compared, in a 2 by 2 factorial design, stimuli (all with Soprano voice) with harmonic violations occurring in either melody, accompaniment, or both melody and accompaniment.

Based on the relevant literature reviewed above, for our effect of melodic salience we expected to find activations correlating with auditory perceptual memory, and possibly also involving the fronto-parietal auditory selective attention network, including the precuneus. With respect to the 2 by 2 factorial design, for the main effect of altered versus correct melody (and possibly also for the main effect of altered versus correct accompaniment), we expected to find activations in the harmonic violation network, including the bilateral superior temporal gyrus and the right inferior frontal cortex. With respect to the crucial interaction between violations occurring in melody and violations occurring in accompaniment, revealing the interplay between the focus on harmonic context and the processing of harmonic violations, we envisaged a possible key role of a subset of brain regions in the perceptual salience and in the harmonic violation networks; more specifically, we expected to observe areas critically involved in the processing of unexpected events (i.e., bilateral superior temporal gyrus, and right inferior frontal gyrus) together with areas deemed important for the processing of salience (i.e., in particular the precuneus).

Materials and methods

Subjects

20 trained musicians (6 men, 14 women) participated in the fMRI study (mean age 30.29 years, SD 7.02, range 22–44 years). They were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971), and had normal hearing and none of them reported history of neurological or psychiatric disorders. All subjects were soloist pianists with a classical curriculum studiorum, recruited from Conservatories. We avoided professional accompanying pianists, who may process a musical excerpt in a bass-salient manner. They all had been playing daily at least since 12 years (mean 21.00 years of practice, SD 6.86, range 12–33 years of practice). Subjects with absolute pitch were excluded. The experiment was approved by the Ethics Committee