# ARTICLE IN PRESS

### Hearing Research xxx (2013) 1-10



Contents lists available at ScienceDirect

# Hearing Research



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

# Review Thalamocortical mechanisms for integrating musical tone and rhythm

Gabriella Musacchia<sup>a,b,1,2</sup>, Edward W. Large<sup>c,3</sup>, Charles E. Schroeder<sup>d,e,\*</sup>

<sup>a</sup> Communication Sciences and Disorders, Montclair State University, 1515 Broad Street, Bldg. B, Bloomfield, NJ 07003, USA

<sup>b</sup> Center for Molecular & Behavioral Neuroscience, Rutgers University, 197 University Avenue, Newark, NJ 07102, USA

<sup>c</sup> Center for Complex Systems and Brain Sciences, Florida Atlantic University, 777 Glades Rd., BS-12, Room 308, USA

<sup>d</sup> Cognitive Neuroscience and Schizophrenia Program, Nathan S. Kline Institute for Psychiatric Research, 140 Old Orangeburg Rd., Orangeburg,

NY 10962, USA

<sup>e</sup> Department of Psychiatry, Columbia University College of Physicians and Surgeons, 180 Ft. Washington Avenue, New York, NY 10032, USA

#### ARTICLE INFO

Article history: Received 20 July 2013 Received in revised form 21 September 2013 Accepted 26 September 2013 Available online xxx

## ABSTRACT

Studies over several decades have identified many of the neuronal substrates of music perception by pursuing pitch and rhythm perception separately. Here, we address the question of how these mechanisms interact, starting with the observation that the peripheral pathways of the so-called "Core" and "Matrix" thalamocortical system provide the anatomical bases for tone and rhythm channels. We then examine the hypothesis that these specialized inputs integrate acoustic content within rhythm context in auditory cortex using classical types of "driving" and "modulatory" mechanisms. This hypothesis provides a framework for deriving testable predictions about the early stages of music processing. Furthermore, because thalamocortical circuits are shared by speech and music processing, such a model provides concrete implications for how music experience contributes to the development of robust speech encoding mechanisms.

This article is part of a Special Issue entitled < Music: A window into the hearing brain >.

© 2013 Published by Elsevier B.V.

## 1. Introduction

Both speech and music rely on pitch transitions whose production is controlled over time. In everyday listening, the pitch and rhythm of speech and music blend together to create a seamless perception of pitch fluctuations organized in time. The organization of pitch in rhythm aids cognitive processing in both domains (Staples, 1968; Weener, 1971; Shepard and Ascher, 1973; Shields et al., 1974; Quene and Port, 2005; Dilley and Pitt, 2010; Devergie et al., 2010). An extensive body of research has linked music training to enhanced perception and processing of pitch (Pantev et al., 2001, 2003; Munte et al., 2003; Fujioka et al., 2004; Tervaniemi et al., 2005; Besson et al., 2007; Musacchia et al., 2007) and rhythm (Dawe et al., 1995; Bailey and Penhune, 2010), as well as spoken (Magne et al., 2006; Musacchia et al., 2007; Marques et al., 2007; Musacchia et al., 2008; Meinz and Hambrick, 2010; Parbery-Clark et al., 2011) and written (Strait et al., 2011) language. The neuronal mechanisms of pitch and rhythm perception are often investigated separately, and while the interrelationship between the two is a longstanding question, it remains poorly understood. Recent findings show that entrained brain rhythms facilitate neuronal processing and perceptual selection of rhythmic event streams [for review, (Schroeder et al., 2008, 2010]. These findings join with others – specifically relating neuronal entrainment to speech (Luo and Poeppel, 2012) and music (Large and Snyder, 2009) – to prompt renewed interest in the neuronal mechanisms that underlie the brain's ability to form a representation of rhythm and to use this to support the processing of tone. In this manuscript, we advance a hypothesis regarding the how the neuronal mechanisms encoding tone and rhythm cues may interact to produce an integrated representation of music.

There are several compelling reasons to explore the brain's representation of musical structure. One feature that differentiates most music from speech is that it is organized around a beat, a perceived pulse that is temporally regular, or periodic. While periodicity is an ubiquitous aspect of spoken language, it is much less regular and is more often patterned on an alternation of stressed versus unstressed syllables (Patel and Daniele, 2003)

<sup>\*</sup> Corresponding author. Cognitive Neuroscience and Schizophrenia Program, Nathan S. Kline Institute for Psychiatric Research, 140 Old Orangeburg Rd., Orangeburg, NY 10962, USA. Tel.: +1 845 398 6539.

*E-mail addresses:* musacchiag@mail.montclair.edu (G. Musacchia), Cs2388@ columbia.edu, schrod@nki.rfmh.org (C.E. Schroeder).

<sup>&</sup>lt;sup>1</sup> Tel.: +1 973 655 3307; fax: +1 973 655 3406.

<sup>&</sup>lt;sup>2</sup> Tel.: +1 973 353 3277; fax: +1 973 353 1760.

<sup>&</sup>lt;sup>3</sup> Tel.: +1 860 486 2538; fax: +1 860 486 2760.

<sup>0378-5955/\$ -</sup> see front matter © 2013 Published by Elsevier B.V. http://dx.doi.org/10.1016/j.heares.2013.09.017

rather than organized a-priori around a particular rhythm. The beat organization of music is particularly tractable from an experimental point of view because the rhythmic structure can be decomposed into simple (e.g., equally timed or isochronous intervals) or complex (e.g., irregularly timed or syncopated intervals) patterns for understanding how the brain integrates acoustic content at different levels of complexity. Another heuristic feature of music is that the organization of musical rhythm is well matched, and perhaps coupled, to neurophysiology. We believe one basis for this coupling may be related to the hierarchy of neuronal oscillations, which are cyclical fluctuations of activity in single or populations of neurons. In general, neuronal oscillations (or, "brain rhythms") are characterized by peaks of activity in high (e.g.  $> \sim 30$ ), middle (e.g. ~10–30 Hz) and lower (< ~10 Hz) frequency bands. Brain activity in the higher bands is generally coupled to that in lower frequency bands. This coupling, or "nesting" of rhythms, occurs in ongoing activity, both in the absence of stimulation, and when the system is processing auditory input. It is worthwhile to note that the nesting of faster rhythms into slower rhythms in music (e.g., in Habenera, music for a slow Cuban dance, similar to the tango, in duple time) parallels the nesting higher frequencies in lower frequencies in the brain; there is a similar parallel in speech where the formant scale (corresponding to gamma band) is nested within the syllabic scale (theta band) and that in turn is nested within the phrasal scale (delta band) [discussed by Schroeder et al. (2008)]. Thus, investigations of brain rhythm response changes to musical stimulus permutations music may help us to understand fundamental aspects of neuronal information processing in ecologically relevant paradigms. While both music and speech are periodic, the simpler structure of music could act to foster plasticity at a more basic level than the complex periodicity of speech.

In this paper, we advance the hypothesis that the peripheral auditory pathways carry relatively segregated representations of tone and rhythmic features, that feed into separate and specialized thalamocortical projection channels. The so-called thalamic "Core" and "Matrix" systems, which are thalamocortical networks of chemically differentiated projection neurons (Jones, 1998, 2001) that generally correspond to "specific/non-specific" and "lemniscal/ extralemniscal" dichotomies found elsewhere, provide the anatomical bases for these channels. In this model, the classical types of "driving" and "modulatory" input mechanisms (Schroeder and Lakatos, 2009a) are hypothesized to integrate acoustic content within rhythm context in auditory cortex. This hypothesis provides a framework for generating testable predictions about the early stages of music processing in the human brain, and may also be useful for understanding how the neuronal context of rhythm can facilitate music-related plasticity. Although music-related plasticity is not a focus of this paper, our hypothesis generally predicts an overlap of plasticity-related brain changes in speech and music domains, with rhythm-related plasticity over multiple timescales particularly enhanced with music training.

## 2. Integrating tone and rhythm

The proposed mechanism of tone and rhythm integration entails a specific subset of thalamocortical projections (i.e., the Matrix) that organize activity around a beat and facilitate cortical processing of tones that fall within the rhythmic stream, and are conveyed into cortex by another subset of the thalamocortical projection system (i.e., the Core). This notion is based on five main conclusions stemming from earlier research: 1) neuronal oscillations in thalamocortical systems entrain to repeated stimulation, 2) music and lower frequency neuronal oscillations operate on similar timescales and have a similar hierarchical organization, 3) different ascending projection systems from the thalamus convey acoustic content and modulate cortical rhythmic context; 4) these input systems converge and interact in cortex and 5) neuronal oscillations generating the relevant cortical population rhythms have both optimal and non-optimal excitability phases, during which stimulus processing is generally facilitated or suppressed. In this way, the modulatory influences of oscillations in the Matrix thalamocortical circuit, coupled to the rhythm of a musical stimulus, impose a coordinated "context" for the "content" (e.g. tonal information) carried by networks of Core thalamocortical nuclei, much like interactions between hippocampal interneurons and principal cells (Chrobak and Buzsaki, 1998). Support for these conclusions, along with other important considerations and caveats to our hypothesis, are described below.

## 2.1. Oscillatory entrainment to periodic stimulation

The neuronal oscillations that give rise to the cortical population rhythms (indexed by the electroencephalogram or EEG) are cyclical fluctuations of baseline activity that manifest throughout the brain (for review (Buzsaki and Draguhn, 2004), including both neocortical and thalamic regions (Steriade et al., 1993a; Slezia et al., 2011). Oscillations generally have a 1/f frequency spectrum with peaks of activity at particular frequency bands [(e.g. delta  $\sim 1-3$  Hz, theta ~4–9, alpha ~8–12 Hz, beta ~13–30 Hz, gamma ~30–70 Hz (Berger, 1929); for recent review, see (Buzsaki, 2006)] and interact strongly with stimulus-related responses (Lakatos et al., 2005; Jacobs et al., 2007; Schroeder et al., 2008). For example, when stimuli are presented in a periodic pattern, ambient neuronal oscillations entrain (phase-lock) to the structure of the attended stimulus stream [e.g. (Will and Berg, 2007; Schroeder and Lakatos, 2009b)]. Neuronal entrainment to *salient* periodic stimulation can emerge rapidly without conscious control, and can still facilitate behavioral responses (Stefanics et al., 2010), however, such entrainment can also be attentionally modulated to favor a specific sensory stream when several streams are presented concurrently (Lakatos et al., 2008). Previous work has shown that low-frequency (<1 Hz), periodic sound stimulation promotes regular up/down transitions of activity in non-lemniscal (Matrix) thalamic nuclei (Gao et al., 2009). Interestingly, this entrainment effect emerges rapidly and can persistent for quite some time, modulating the efficacy of synaptic transmission at the entrained interval for some time (at least 10 min) after stimulation is discontinued. These findings suggest that entrainment of neuronal oscillations may also help to capture and retain the contextual structure of rhythmic stimulus intervals.

## 2.2. Mirroring of neuronal and musical rhythms

There are several intriguing parallels between the organization of neuronal rhythms and the organization of musical rhythms. In temporal cortex, as in many brain regions, ongoing neural rhythms display 1/f structure, such that log power decreases as log frequency increases (Freeman et al., 2000). In naturally performed musical rhythms, temporal fluctuations also display a 1/f structure across musical styles (Rankin et al., 2009; Hennig et al., 2011), and such structured temporal fluctuations have been shown to enable temporal prediction (Rankin et al., 2009) and communicate emotional arousal in music (Chapin et al., 2010; Bhatara et al., 2011). Moreover, in auditory cortex, brain rhythms nest hierarchically: delta phase modulates theta amplitude, and theta phase modulates gamma amplitude (Lakatos et al., 2005). In perceptual experiments, pulse occupies the range from about 0.5 to 4 Hz (London, 2004), matching the delta band rather well. Like neuronal rhythm frequencies, musical beats nest hierarchically, such that faster metrical frequencies subdivide pulse frequencies (London,

G. Musacchia et al. / Hearing Research xxx (2013) 1-10



Fig. 1. A model for integration content and context cues in music based on "Core" and "Matrix" divisions of the auditory thalamocortical projection pathways. Parvalbuminpositive (Parv. +) "Core" thalamocortical neurons (blue) lying mostly in the ventral division of the Medial Geniculate body (MGv) receive frequency-tuned inputs from the tonotopically organized internal nuclear portions of the inferior colliculus. These neurons terminate in the granular cortical layers of Core regions of primary auditory cortex (A1), giving rise to the relatively precise tonotopic representation observed there (Pandya et al., 1994; Hashikawa et al., 1995; Kimura et al., 2003). In contrast, calbindin positive (Calbin. +) "Matrix" thalamocortical neurons (red) lying mainly in the Magnocellular (mc) and dorsal (d) Medial Geniculate divisions receive broadlytuned (diffuse) inputs input from the non-tonotopic external nuclear portions of the inferior colliculus. These inputs project broadly to secondary/tertiary (e.g., auditory belt and parabelt), as well as primary regions, terminating in the superficial cortical layers (Pandya et al., 1994; Hashikawa et al., 1995; Kimura et al., 2003). As detailed in the text, we propose that that activity conveyed by thalamocortical Matrix projections, and entrained to the accent/beat pattern of a musical rhythm, dynamically reset and "modulate" the phase of ongoing cortical rhythms, so that the cortical rhythms themselves synchronize with the rhythmic "contextual" pattern of the music. Because oscillatory phase translates to excitability state in a neuronal ensemble, a potential consequence of the modulatory phase resetting that underlies synchrony induction and maintenance is that responses "driven" by specific "content" that is associated with contextual musical accents (e.g., notes or words) will be amplified, relative to content that occurs off the beat. Abbreviations: IPS: Intraparietal Sulcus; 3b: Area 3b, Primary somatosensory cortex; SII: Secondary somatosensory cortex; STG: Superior Temporal Gyrus.

2004). Perceptual and behavioral experiments put the upper limit of pulse frequencies at about 2.5–4 Hz (Large, 2008) and the upper limit for fast metrical frequencies at about 8–10 Hz (Repp, 2005). Note also that the onset of acoustic events within rhythmic context evokes nested gamma band responses (Pantev, 1995; Snyder and Large, 2005). Critically auditory cortical rhythms synchronize to acoustic stimulation in the range of rhythmic frequencies (Snyder and Large, 2005; Will and Berg, 2007; Lakatos et al., 2008; Stefanics et al., 2010; Fujioka et al., 2012), and delta and theta rhythms typically emerge with engagement of motor cortex (Saleh et al., 2010) which is a central nexus of rhythm-related brain activity (Besle et al., 2011; Saleh et al., 2010) (Chen et al., 2008). Experiments have confirmed that synchronization with auditory rhythms matches predictions of entrained oscillations (Large, 2008) and have demonstrated qualitative shifts in coordination mode at the transition from pulse frequencies and fast metrical frequencies (Kelso et al., 1990).

# 2.3. Transmission of tonal and rhythmic information via separate thalamocortical circuits

The mechanism we propose for integrating pitch and rhythm processing involves the contrasting physiology and anatomy of the so-called "Core" and "Matrix" systems described by Jones (Jones, 1998, 2001) (Fig. 1). Note that the Core/Matrix dichotomy

corresponds generally to the "specific/non-specific" and "lemniscal/extralemniscal" terms found elsewhere in the literature (Kaas and Hackett, 1999; Kraus and Nicol, 2005; Lomber and Malhotra, 2008; Abrams et al., 2011). In this model, Core and Matrix thalamocortical projection neurons are chemically differentiated by the calcium binding proteins they express, with Core neurons expressing parvalbumin (Pa) and Matrix neurons expressing calbindin (Cb).

In the auditory system, Pa-positive Core thalamic neurons lie mostly within the ventral division of the medial geniculate nucleus (de Venecia et al., 1995; Jones, 2003), receive ascending inputs from the internal nucleus of the inferior colliculus (Oliver and Huerta, 1992; Winer and Schreiner, 2005), and exhibit narrow frequency tuning (Calford, 1983; Anderson et al., 2007). They project mainly to Layer 4 and deep layer 3 of the auditory cortical "Core" regions including A1, and the more rostral regions (R and RT) (McMullen and de Venecia, 1993; Pandya et al., 1994; Hashikawa et al., 1995; Kimura et al., 2003), where their pure tone preferences give rise to the well-known tonotopic maps of these regions. Cb-positive Matrix thalamic neurons comprise much of the dorsal and Magnocellular divisions of the medial geniculate complex, but are also scattered throughout the ventral division of the MGB (de Venecia et al., 1995; Jones, 2003). They receive input from the external nuclear division of inferior colliculus (Oliver and Huerta, 1992; Winer and Schreiner, 2005), and thus, exhibit broad frequency tuning, as well as strong sensitivity to acoustical transients (Calford, 1983; Anderson et al., 2007). Matrix neurons project broadly to Layer 2 of all the auditory cortices, including both Core (A1, R and Rt), as well as the surrounding belt regions (Jones, 1998, 2001, 2003). Their notably poor frequency tuning is consistent with the tuning properties of neurons in auditory belt regions, though it probably does not completely account for these properties.

Based on their proclivities for pure tones and acoustic transients, the Core and Matrix provide potential substrates for encoding and central projection of music's pitch- and rhythmrelated cues, respectively. This proposed division of labor, and in particular, the role of the Matrix in using entrained neuronal oscillations to represent and convey rhythm-related influences is more intriguing in light of the proposition that these oscillations represent the "context" for the processing of specific sensory "content" (Buzsaki, 2006); pitch components of music would equate to specific tonal "content," while rhythmic components would equate to "context." It is important to note here that these designations are not absolute, as pitch cues can give contextual meaning or emotional context in certain cases, but rather broad acoustic delineations that are useful to describe and examine the general structure of music.

## 2.4. Convergence and interaction of Core and Matrix circuits

The separation of thalamocortical music processing into tonal and rhythmic components mediated by the Core and Matrix systems would also have predictable downstream effects in the neocortex. As mentioned above, Pa-positive thalamic Core inputs target the middle layers of primary cortical regions (A1, R and Rt), whereas Cb-positive thalamic Matrix afferents project over a much wider range of territory, targeting the most superficial layers of these areas, as well as the surrounding belt cortical regions (Jones, 1998, 2001), as illustrated in Fig. 1. Core thalamic inputs clearly drive action potentials in the target neurons, and thus are reasonably considered as "driving" inputs, (e.g., (Sherman and Guillery, 2002) in auditory cortex). Matrix inputs, through their upper layer terminations, are in a position to function as "modulatory" inputs, which would not elicit action potentials directly, but rather, control the likelihood that appropriate driving inputs would do

# ARTICLE IN PRESS

G. Musacchia et al. / Hearing Research xxx (2013) 1-10

this. This idea is admittedly speculative, but fits with the notion that Matrix afferents control cortical synchrony in general (Jones, 2001). One area where this has been extensively explored is in multisensory investigations. To make the proposition more

concrete, we illustrate the contrasting physiology of driving and modulatory inputs on the laminar activity profile in A1 (Fig. 2), by comparing the effects of auditory (driving) versus somatosensory (modulatory) inputs into A1 (Lakatos et al., 2007).



B.Grouping of neuronal firing according to cortical state in cat parietal area



## C. Phase-coupled neuronal excitation in macaque primary auditory cortex



**Fig. 2.** Oscillations control neuronal excitability. Panel A. On the left, four aspects of neural activity are shown: the low pass filtered local field potential (LFP) (A1), the amplitude activity (B1), phase activity (C1) and multiunit activity (MUA) (D1) (Slezia et al., 2011). A2–D2 shows the same indices for an oscillation that increases in frequency. Panel B. MUA bursting of cat parietal corticothalamic neurons (top trace) occurs in the Delta frequency range and co-occurs with slower (0.3–0.4 Hz) LFP rhythms (middle trace) and scalp recorded electroencephalographic (EEG) activity (bottom trace) (Steriade et al., 1993b). Panel C. In macaque (Lakatos et al., 2005), a current source density (CSD) profile for the supragranular layers of auditory cortex (i) sampled using a multielectrode (left) shows Theta-band (5–9 Hz) oscillatory activity (superimposed black line). Drop lines from ii shows coincidence of simultaneously recorded MUA activity. Reproduced, with permission, from [33] (A), [38] (B), [77] (C).

G. Musacchia et al. / Hearing Research xxx (2013) 1-10



**Fig. 3.** Contrasting laminar profiles and physiology of "driving" and "modulatory" inputs. A) A schematic of the multi-contact electrode and positioned astride the cortical laminae is shown on the right. Box-plots show quantified onset latencies of the best-frequency pure tone (blue) and somatosensory stimulus (red) related current source density (CSD) response in supragranular (S), granular (G) and infragranular (I) layers across experiments. Lines in the boxes depict the lower quartile, median, and upper quartile values. Notches in boxes depict 95% confidence interval for medians of each distribution. Brackets indicate the significant differences (Games–Howell post-hoc test, p < 0.01). B) Quantified multi-unit activity (MUA) amplitudes (over 38 cases) from the representative supragranular, granular and infragranular channels (S, G, and I) over the 15–60 m time interval for the same conditions as in A, along with a bimodal stimulation condition. Brackets indicate the significant differences (Games–Howell, p < 0.01). C) left, Quantified (n = 38) post-/pre-stimulus single trial oscillatory amplitude ratios (0-250 ms/–500 to -250 m) for different frequency bands (different colors) of the auditory, somatosensory and bimodal supragranular responses. Stars denote the amplitude ratios significantly different from 1 (one-sample *t*-tests, p < 0.01). right. Quantified (n = 38) intertrial coherence (ITC) expressed as a vector quantity (mean resultant length), measured at 15 ms post-stimulus (i.e., at the initial peak response). For somatosensory events, increase in phase concentration only occurs in the low-delta (1-2.2 Hz), theta (4.8-9.3 Hz) and gamma (25-49 Hz) bands, indicated by colored arrows on the right. Adapted, with permission, from (Lakatos et al., 2007).

In this illustration, laminar current source density (CSD) onset latency profiles (Fig. 2A) index net local synaptic current flow. Patterns of activity show that auditory input initiates a classic feedforward synaptic activation profile, beginning in Layer 4, while synaptic activation stemming from somatosensory input begins in the supragranular layers. Ouantified multiunit activity amplitude (MUA - Fig. 2C) for the same laminar locations show robust neuronal firing with auditory input, but no detectable change in local neuronal firing with somatosensory input. The combination of synaptic currents (CSD) with neuronal firing (MUA), plus the initiation of the auditory response profile in Layer 4, identifies it as a 'driving' (Core) input, while the CSD input profile beginning outside of Layer 4, in the absence of a MUA identifies somatosensory as a subthreshold 'modulatory' input (Lakatos et al., 2007). Importantly, both the extremely short latency of the input and its high supragranular target zone suggest that it represents a Matrix input. In the tone-rhythm integration mechanism we propose, acoustic content carried by Core afferents would form the driving input to Layer 4, while rhythmic context information borne by Matrix afferents would form the modulatory input to the supragranular site.

How would the hypothesized tone - rhythm interaction look in this system? Notice first that interaction of the modulatory and driving inputs during bimodal stimulation in our example produces significant enhancement of neuronal firing (MUA, Fig. 2C). We predict that modulatory rhythmic input would similarly enhance the neural representation of tonal inputs. It is further noteworthy that the modulatory somatosensory inputs appear to use phase reset of ongoing neuronal oscillations as an instrument for enhancement of auditory processing. Fig. 2C (left) depicts pre- to post-stimulus amplitude ratios for single trial oscillatory activity in six frequency bands extending from delta to gamma. As expected of a pure modulatory input (Shah et al., 2004; Makeig et al., 2004), somatosensory input triggers no detectable change on the power of ongoing oscillations, but produces phase concentration (reset) and coherence over trials (intertrial coherence or ITC) in the low delta, theta and gamma bands (Fig. 2C-right). This is how we predict that rhythmic inputs will behave. In contrast, the auditory (driving) input provokes broad band increase in both oscillatory power and ITC across the spectrum, as expected for an 'evoked' type of response, and this is how we predict tonal inputs will function. Thus rhythmic input would impact mainly by resetting the phase of

ongoing oscillations and 'sculpting' the temporal structure of ambient oscillatory activity. Rhythm would 'modulate' the dynamics of activity in A1, but would not 'drive' auditory neurons to fire to any great extent.

# *2.5.* Thalamocortical mediation of cortical excitability control at rhythmic intervals

Several lines of research provide a conceptually and empirically appealing explanation for the way that entrained, rhythmic input to the cortex modulates cortical excitability at beat intervals (Large and Snyder, 2009). As mentioned above, oscillations appear to reflect cyclical variations in neuronal excitability and are used as instruments of brain operation. Recent work has built on this idea to clarify the relationship between neuronal excitability and ongoing oscillations evident in the electroencephalogram (EEG) measured at the brain surface or scalp, as well as in similar signals measured as "local field potentials" (LFP), within the brain. One of the central tenets of our hypothesis is rhythmic oscillatory modulation of driving tonal input. Neuronal oscillations reflect cyclical variations in neuronal excitability and have been shown to relate to a wide range of cognitive and behavioral operations (Chrobak and Buzsaki, 1998; Fries et al., 2007; Fan et al., 2007; Klimesch et al., 2009; Schroeder and Lakatos, 2009a). Previous work has clarified the basis of the relationship between neuronal excitability and ongoing oscillations using electroencephalogram (EEG) activity measured at the brain surface or scalp, as well as in similar signals measured as "local field potentials" (LFP), within the brain. Specifically, findings across species and brain regions show a tight coupling between the phase of ongoing EEG/LFP oscillations and neuronal firing (Steriade et al., 1993b; Lakatos et al., 2005; Buzsaki, 2006; Slezia et al., 2011) (Fig. 2). These data show that that neuronal firing is coupled to a "preferred phase" of the oscillatory EEG activity and that the final output of neuronal excitability modulation (i.e., firing) is controlled by oscillatory fluctuations. The coupling between oscillations and neuronal excitability is illustrated here in three animal species and sensory cortices (Fig. 3). In the first example (Fig. 3A), spontaneous multi-unit neuronal activity (MUA) in rat somatosensory cortex occurs at a consistent phase of the ongoing EEG (Slezia et al., 2011). In this study, a neuron's firing exhibited phase coupling, over successive cycles of LFP oscillation. This coupling shed light on earlier observations, shown in Fig. 3B, that delta band bursting of corticothalamic neurons in cat parietal cortex was grouped according to slow (0.3-0.4 Hz) rhythms (Steriade et al., 1993b). The coupling of neuronal excitability with EEG phase is preserved in higher primates (Fig. 3C), with thetaband (5-9 Hz) oscillatory activity in the supragranular layers of macaque primary auditory cortex coincident with MUA recorded from the same site (Lakatos et al., 2005). These findings, together with those discussed above (re: Fig. 3), suggest that thalamocortical rhythmic input to cortical supragranular layers would organize the excitability state around a beat, providing rhythm-related time windows of enhanced firing probability.

# 3. A caveat to the rhythm-tone processing and integration hypothesis

Despite the evidence in support of the hypothesis we advance here, it is incomplete for the following reason: If the system operated as simply as we have described it here, we would likely hear sound content that occurs "on the beat" very well, and by the same token, sound content occurring "off the beat" very poorly. This proves to be the case for simple transient stimuli (Large and Jones, 1999), but given the smooth melodic flow that we perceive in music, it is likely that the neuronal entrainment process we describe also orchestrates more subtle and complex processes. Lacking any firm information on this matter, we can nonetheless offer a speculation that is amenable to empirical examination, and may thus stimulate future studies on the topic.

We conjecture that entrainment is but one mechanism in a set of dynamic oscillatory mechanisms that operate at longer timescales, are sensitive to specific stimulus features, and form a recurrent loop with attentional mechanisms that together form a complex network of attention-controlled enhancement and suppression. Thus, different elements of the neuronal population responsive to tones may encode additional information such as tonal movements and progressions. Empirical evidence supports this notion. For example, the spiking of auditory cortical neurons in monkeys listening to extended duration naturalistic stimuli was investigated with specific attention to the precise phase of the ongoing local slow oscillation at which spiking occurred (Kayser et al., 2009). The authors found that in these circumstances, individual neurons tended to fire characteristically at different phases of the local slow (reference) oscillation. They concluded that such a nested code combining spike-train patterns with the phase of firing optimized the information carrying capacity of the cortical ensemble, and provided the best reconstruction of the input signal. This representation of frequency-based information in spike-phase coding is strongly reminiscent of the encoding of spatial (movement) information by theta rhythm phase precession in rodent hippocampal place cells (Huxter et al., 2003). Spike phase coding is an appealing mechanism for overcoming the limitations in our model, but further experimentation will be necessary to flesh-out this idea.

#### 4. Outstanding issues

A number of testable predictions based on the proposed model are outlined below. We also highlight some questions that remain to be addressed.

- Neurophysiological predictions
  - Selective rhythmic stimulation of Matrix neurons (e.g. electrically or optogenetically) should have purely modulatory effects on neocortex, most prominent in supragranular layers, while stimulation of Core neurons should produce stereotypic driving responses primarily in Layer 4.
  - Co-stimulation of Core and Matrix neurons should produce enhancement of cortical driving responses, stronger transmission of rhythmic driving patterns to the extragranular layers and to higher cortical regions.
  - Counterphase rhythmic stimulation of Core and Matrix should suppress driving responses, especially outside of Layer 4.
  - Disruption of Matrix functioning should disrupt the emergent perception of the beat/pulse in music.
  - Matrix nuclei should have a wider interval range of accurate phase-locking to sound onset than Core neurons (e.g. can more accurately encode faster temporal intervals).
- Open questions
  - Does music-related plasticity have differential effects on Core and Matrix thalamocortical circuits? What about rhythm training vs. acoustic content training?
  - Do disorders having specific sensory deficits of tone processing (e.g. amusia) or rhythmic processing (e.g. central pattern generation) show specific Core vs. Matrix thalamocortical impact?
  - > What is the role of pitch contour and consonance in rhythmic perception and processing?
  - Can the developmental trajectory of Core and Matrix circuit development in humans help explain acquisition of tonal

(e.g. segmental) vs. rhythmic (e.g. suprasegmental) linguistic features?

## 5. Concluding remarks

We have outlined a physiological hypothesis for an integration of tone and rhythm in which thalamic activity around accent/beat patterns of a musical rhythm, conveyed by thalamocortical Matrix projections, dynamically modulates the phase of ongoing cortical oscillatory rhythms. Because oscillatory phase translates to excitability state in a neuronal ensemble, a consequence of the modulatory phase resetting that underlies synchrony induction and maintenance is that specific content associated with musical accents (e.g., notes or words) will be physiologically amplified, in a way that is similar to that believed to occur in language perception (Schroeder et al., 2008; Zion Golumbic et al., 2012). Whether or not rhythm-based enhancement of neuronal firing confers perceptual accents, however, may be a more complex issue. For example, it is known that perceptual accents of sound do not always coincide with an established meter (Repp, 2010). It is likely that entrainment mechanisms function as a base operation to modulate attentionally-mediated mechanisms that, together with intrinsic state properties, enhance perceptual sensitivity and selectivity. Thus, not only are there reasonably obvious neuronal substrates for the encoding and representation of tonal and rhythmic components of music, but there are clear means of integrating the component representations, using neuronal systems that are themselves modifiable by the experiencing of music. This model posits a concrete mechanism that may help to explain how rhythm information is represented and transmitted to organize neural activity in a wide network of areas (Alluri et al., 2012) including multiple sensory (Chen et al., 2006), motor (Konoike et al., 2012) and frontal cortices. As noted in the preceding section, the hypothesis is rudimentary, and does not yet fully account for all aspects of the neural encoding of music, and as yet, direct empirical support for the precise mechanism we propose is not in hand. However the hypothesis makes clear physiological predictions than are amenable to empirical validation (preceding section), and can help to guide additional experimentation.

Our hypothesis may provide a physiological explanation for leftright differences observed at the level of the BOLD signal and scalp measurements. For example, multiple lines of evidence have supported a spectro-temporal theory of auditory processing that posits better temporal resolution in left auditory cortical areas and better spectral resolution right hemisphere homologs (Zatorre et al., 2002). If this idea holds, the biasing of rhythm and tone within the Matrix and Core thalamic nuclei and projections may show similar hemispheric asymmetries. Specifically, we suggest that a difference in physiology of the left and right non primary, Matrix thalamic nuclei pathways may give rise to asymmetry, with primary "Core" nuclei on the left specialized for complex sound processing, and non primary "Matrix" nuclei on the right specialized for repeated stimulation and oscillatory phase reset mechanisms. It may be that the thalamic non primary system, encoding rhythm, is particularly sensitive to changes in context and that activity in the right thalamus preferentially processes rhythmic "contextual" cues. According to our hypothesis this would be explained by the oscillatory phase reset mechanisms or termination differences in the left and right Matrix systems. Evidence for this theory have been demonstrated in animal models, showing that the nonprimary thalamic pathway may be specifically tuned to code low-frequency temporal information present in acoustic signals (Abrams et al., 2011). However, there may be an additional level of complexity to consider, as leftward asymmetry for speech processing findings have been observed in both primary-and non-primary thalamic structures (King et al., 1999). Therefore, the relationship between the Core/Matrix and spectrotemporal hypotheses remains speculative, and more investigation is needed to understand the contributions of left/right specializations of the subcortical primary and non-primary pathways.

Further relationships between the current hypothesis and theories of learning can also be speculated upon. Integration of tone and rhythm through a thalamocortical oscillatory control is likely built on innate mechanisms that guide on a listener's "sensitivity" and is modifiable by learning. Considerable evidence supports this notion, showing that oscillations are innate brain mechanisms and exhibit developmental and music training-related changes. In animal models, changes in the frequency and synchronization of neural oscillations are markers of certain stages of development. Maturation of neural synchrony in each stage is compatible with changes in the myelination of cortico-cortical connections and with development of GABAergic neurotransmission (Singer, 1995; Uhlhaas et al., 2010). In babies 16-36 months of age, gamma oscillation power is associated with language, cognitive and attention measures in children and can predict cognitive and linguistic outcomes through five years of age (Benasich et al., 2008; Gou et al., 2011). In addition, induced gamma band activity becomes prominent around 4.5 years of age in children who have had 1 year of music training, but not in children of this age who have not been trained (Trainor et al., 2009). Second, musicians also show evidence of training-related changes in oscillatory activity. Extensive musical training is associated with higher gamma band evoked synchrony when listening to music (Bhattacharva et al., 2001) or notes of instruments that they themselves play (Shahin et al., 2008). A tantalizing piece of evidence for our prediction also shows that musicians have stronger interactions between the thalamus and premotor cortex (PMC) at the beta frequency during a motor synchronization task, and drummers in particular exhibit stronger interactions between thalamus and posterior parietal cortex (PPC) at alpha and beta frequencies (Krause et al., 2010; Fujioka et al., 2012).

The relationship between music training and language ability may also be considered in terms of our oscillatory hypothesis. A growing body of evidence in language-disordered populations suggest that deficits in specific bands of oscillations may be associated with speech and language deficits. For example, recent data in this special issue shows that that rhythmic entrainment at slow  $(\sim 5 \text{ Hz})$  rates is atypical in dyslexia [in this issue, (Leong and Goswami, 2013)]. Because the average syllable rate in spoken language hovers around this frequency, the authors suggest a link between the generalized mechanism of low frequency oscillations and the representation of perceived and produced syllables. This study extends previous work in school-aged children, showing that cortical encoding of slower features of speech representation (e.g. the speech envelope) was impaired in poor readers and predicted over 40% of standardized reading and phonological processing scores (Abrams et al., 2009). Because the phase of low frequency oscillations controls the amplitude of high frequency oscillatory activity and ultimately, neuronal firing, the idea that deficits in low frequency encoding could be related to the perception and representation of fast acoustic change is worth careful consideration. Our current model suggests separate, but integrated, anatomical inputs for high- and low-frequency representations, such that the relationship between high- and low-frequency encoding deficits may be dissociable according to the anatomical correlates of particular deficits and disorders. One way in which this may be observed is deficits in patterns of functional connectivity. Our hypothesis predicts strong inter-areal functional connectivity for low-frequency processing and strong intra-areal functional connectivity for

## **ARTICLE IN PRESS**

high-frequency processing in a typically developed system. This dichotomy may provide the basis for music-related increases in functional connectivity across cortices, as has been demonstrated in the audio-motor network in cross-sectional and longitudinal measures of music training [in this issue, (François and Schön, 2013)]. Our hypothesis, together with the converging speech and language results above posit a role for low-frequency oscillations as a governing mechanism for encoding acoustic features of both musical rhythm and the speech envelope (e.g. syllable rate, prosody).

There are a few key theoretical implications of our hypothesis as it relates to language development and disorders. First, our hypothesis implies that the separation and integration of rhythm and tone mechanisms would follow the developmental trajectory of neuronal maturation in the in the Matrix and Core systems. While the physical maturation of these systems is not well known in the human infant, behavioral studies suggest a sensitivity to lowfrequency acoustic features very early in life. For example, newborns are able to discriminate the prosody of their native language relative to foreign contrasts (Nazzi et al., 1998; Ramus et al., 2000). This low-frequency specialization is not restricted to perception, as the melodic contour of newborn cries has been shown to mimic the prosodic contours present in their native language (Mampe et al., 2009). Presumably, it is the low-pass filter characteristic of the womb environment that makes the low-frequency content (e.g. prosody, syllable rate, rhythm, etc.) of a mother's voice particularly salient for the developing fetus. Furthermore, a compelling line of research has shown that music experience at ages as early as 6months-of-age can impact infant rhythm preference (Hannon and Trainor, 2007). In utilizing the Core and Matrix dichotomy, we can consider the overlap of rate-related theories of language, music and brain oscillations. In our model, we posit that low-frequency oscillations (e.g. delta, theta) in the Matrix encode the rhythmic aspects of music processing, which in turn are integrated with the more rapid changes of tone and pitch, represented in the Core system. Because this system serves auditory processing in general, and not music exclusively, an analogy to language development can be easily drawn. Specifically, we consider that low-frequency brain oscillations encoding slower-rate speech acoustics (e.g. syllables) in the Matrix can be integrated with rapid changes in speech acoustics (e.g. phonemes, formants), depending on the developmental stage and fitness of each part of the respective systems. Our anatomical model and unifying theory of music, language and brain rhythms is supported by data showing overlap between music training and improvement in other faculties, such as recent findings that beat synchronization training can improve multiple auditory and cognitive functions (Tierney and Kraus, 2013). Thus, we believe our hypothesis may provide a physiological basis for further exploration of music-related language remediation, particularly in developing populations.

A mechanism for tone and rhythm integration also has strong implications for clinical populations who lose the ability to generate or continue rhythmic motion. While this paper focused on externally driven entrainment, a similar mechanism is likely to operate in accordance with internally generated periodicity. In the auditory domain, the perception of the regular pulse in a sequence of temporal intervals is associated with activity in the basal ganglia (Grahn and Brett, 2007, 2009). A further study suggests that the striatum is a locus of internal beat generation, with the generation and prediction of a perceived beat associated with larger striatal activity than in initial beat perception or adjustment (Grahn and Rowe, 2009). The striatum receive inputs from all cortical areas and, throughout the thalamus, project principally to frontal lobe areas (prefrontal, premotor and supplementary motor areas) which are concerned with motor planning (Herrero et al., 2002). According to our hypothesis, anatomical or physiological disturbances that result in stereotypies or a decrease in motor pattern generation could be associated not only with basal ganglia deficits, but also with dysfunction of the thalamocortical Matrix mechanism that we believe transmits this information.

#### References

- Abrams, D.A., Nicol, T., Zecker, S., Kraus, N., 2009. Abnormal cortical processing of the syllable rate of speech in poor readers. J. Neurosci. 29 (24), 7686–7693.
- Abrams, D.A., Nicol, T., Zecker, S., Kraus, N., 2011. A possible role for a paralemniscal auditory pathway in the coding of slow temporal information. Hear. Res. 272, 125–134.
- Alluri, V., Toiviainen, P., Jaaskelainen, I.P., Glerean, E., Sams, M., Brattico, E., 2012. Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. Neuroimage. 59, 3677–3689.
- Anderson, L.A., Wallace, M.N., Palmer, A.R., 2007. Identification of subdivisions in the medial geniculate body of the guinea pig. Hear. Res. 228, 156–167.
- Bailey, J.A., Penhune, V.B., 2010. Rhythm synchronization performance and auditory working memory in early- and late-trained musicians. Exp. Brain Res. 204, 91– 101.
- Benasich, A.A., Gou, Z., Choudhury, N., Harris, K.D., 2008. Early cognitive and language skills are linked to resting frontal gamma power across the first 3 years. Behav. Brain Res. 195, 215–222.
- Berger, E.D., 1929. Uber das Elektrenkephalogramm des Menschen. Arch. Psychiatr. Nervenkr., 527–570.
- Besle, J., Schevon, C.A., Mehta, A.D., Lakatos, P., Goodman, R.R., McKhann, G.M., Emerson, R.G., Schroeder, C.E., 2011. Tuning of the human neocortex to the temporal dynamics of attended events. J. Neurosci 31 (9), 3176–3185.
- Besson, M., Schon, D., Moreno, S., Santos, A., Magne, C., 2007. Influence of musical expertise and musical training on pitch processing in music and language. Restor. Neurol. Neurosci. 25, 399–410.
- Bhatara, A., Tirovolas, A.K., Duan, L.M., Levy, B., Levitin, D.J., 2011. Perception of emotional expression in musical performance. J. Exp. Psychol. Hum. Percept. Perform. 37, 921–934.
- Bhattacharya, J., Petsche, H., Pereda, E., 2001. Long-range synchrony in the gamma band: role in music perception. J. Neurosci. 21, 6329–6337.
- Buzsaki, G., 2006. Rhythms of the Brain. Oxford University Press, New York.
- Buzsaki, G., Draguhn, A., 2004. Neuronal oscillations in cortical networks. Science 304, 1926–1929.
- Calford, M.B., 1983. The parcellation of the medial geniculate body of the cat defined by the auditory response properties of single units. J. Neurosci. 3, 2350–2364.
- Chapin, H., Jantzen, K., Kelso, J.A., Steinberg, F., Large, E., 2010. Dynamic emotional and neural responses to music depend on performance expression and listener experience. PLoS.One. 5, e13812.
- Chen, J.L., Penhune, V.B., Zatorre, R.J., 2008. Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. J. Cogn. Neurosci. 20, 226–239.
  Chen, J.L., Zatorre, R.J., Penhune, V.B., 2006. Interactions between auditory and
- Chen, J.L., Zatorre, R.J., Penhune, V.B., 2006. Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. Neuroimage 32, 1771–1781.
- Chrobak, J.J., Buzsaki, G., 1998. Gamma oscillations in the entorhinal cortex of the freely behaving rat. J. Neurosci. 18, 388–398.
- Dawe, LA., Platt, J.R., Racine, R.J., 1995. Rhythm perception and differences in accent weights for musicians and nonmusicians. Percept. Psychophys. 57, 905–914.
- de Venecia, R.K., Smelser, C.B., Lossman, S.D., McMullen, N.T., 1995. Complementary expression of parvalbumin and calbindin D-28k delineates subdivisions of the rabbit medial geniculate body. J. Comp. Neurol. 359, 595–612.
- Devergie, A., Grimault, N., Tillmann, B., Berthommier, F., 2010. Effect of rhythmic attention on the segregation of interleaved melodies. J. Acoust. Soc. Am. 128, EL1–EL7.
- Dilley, L.C., Pitt, M.A., 2010. Altering context speech rate can cause words to appear or disappear. Psychol. Sci. 21, 1664–1670.Fan, J., Byrne, J., Worden, M.S., Guise, K.G., McCandliss, B.D., Fossella, J., et al., 2007.
- Fan, J., Byrne, J., Worden, M.S., Guise, K.G., McCandliss, B.D., Fossella, J., et al., 2007. The relation of brain oscillations to attentional networks. J. Neurosci. 27, 6197– 6206.
- François, C., Schön, D., 2013. Neural sensitivity to statistical regularities as a fundamental biological process that underlies auditory learning: the role of musical practice. Hear. Res. (Epub ahead of print).
- Freeman, W.J., Rogers, L.J., Holmes, M.D., Silbergeld, D.L., 2000. Spatial spectral analysis of human electrocorticograms including the alpha and gamma bands. J. Neurosci. Methods 95, 111–121.
- Fries, P., Nikolic, D., Singer, W., 2007. The gamma cycle. Trends Neurosci. 30, 309– 316.
- Fujioka, T., Trainor, L.J., Large, E.W., Ross, B., 2012. Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. J. Neurosci. 32, 1791– 1802.
- Fujioka, T., Trainor, L.J., Ross, B., Kakigi, R., Pantev, C., 2004. Musical training enhances automatic encoding of melodic contour and interval structure. J. Cogn. Neurosci. 16, 1010–1021.

# **ARTICLE IN PRESS**

#### G. Musacchia et al. / Hearing Research xxx (2013) 1-10

- Gao, L., Meng, X., Ye, C., Zhang, H., Liu, C., Dan, Y., et al., 2009. Entrainment of slow oscillations of auditory thalamic neurons by repetitive sound stimuli. J. Neurosci. 29, 6013–6021.
- Gou, Z., Choudhury, N., Benasich, A.A., 2011. Resting frontal gamma power at 16, 24 and 36 months predicts individual differences in language and cognition at 4 and 5 years. Behav. Brain Res. 220, 263–270.
- Grahn, J.A., Brett, M., 2007. Rhythm and beat perception in motor areas of the brain. J. Cogn. Neurosci. 19, 893–906.
- Grahn, J.A., Brett, M., 2009. Impairment of beat-based rhythm discrimination in Parkinson's disease. Cortex 45, 54–61.
- Grahn, J.A.J., Rowe, J.B., 2009. Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. J. Neurosci. 29, 7540–7548.
- Hannon, E.E., Trainor, L.J., 2007. Music acquisition: effects of enculturation and formal training on development. Trends Cogn. Sci. 11, 466–472.
   Hashikawa, T., Molinari, M., Rausell, E., Jones, E.G., 1995. Patchy and laminar ter-
- Hashikawa, T., Molinari, M., Rausell, E., Jones, E.G., 1995. Patchy and laminar terminations of medial geniculate axons in monkey auditory cortex. J. Comp. Neurol. 362, 195–208.
- Hennig, H., Fleischmann, R., Fredebohm, A., Hagmayer, Y., Nagler, J., Witt, A., et al., 2011. The nature and perception of fluctuations in human musical rhythms. PLoS.One 6, e26457.
- Herrero, M.T., Barcia, C., Navarro, J.M., 2002. Functional anatomy of thalamus and basal ganglia. Childs Nerv.Syst. 18, 386–404.
- Huxter, J., Burgess, N., O'Keefe, J., 2003. Independent rate and temporal coding in hippocampal pyramidal cells. Nature 425, 828–832. Jacobs, J., Kahana, M.J., Ekstrom, A.D., Fried, I., 2007. Brain oscillations control timing
- Jacobs, J., Kahana, M.J., Ekstrom, A.D., Fried, I., 2007. Brain oscillations control timing of single-neuron activity in humans. J. Neurosci. 27, 3839–3844.
- Jones, E.G., 1998. A new view of specific and nonspecific thalamocortical connections. Adv. Neurol. 77, 49–71.
- Jones, E.G., 2001. The thalamic matrix and thalamocortical synchrony. Trends Neurosci. 24, 595–601.
- Jones, E.G., 2003. Chemically defined parallel pathways in the monkey auditory system. Ann. N. Y. Acad. Sci. 999, 218–233.
- Kaas, J.H., Hackett, T.A., 1999. 'What' and 'where' processing in auditory cortex. Nat. Neurosci. 2, 1045–1047.
- Kayser, C., Montemurro, M.A., Logothetis, N.K., Panzeri, S., 2009. Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. Neuron 61, 597–608.
- Kelso, J.A.S., deGuzman, G.C., Holroyd, T., 1990. The self-organized phase attractive dynamics of coordination. In: .Babloyantz, A. (Ed.), Self-organization, Emerging Properties, and Learning, pp. 41–62.
- Kimura, A., Donishi, T., Sakoda, T., Hazama, M., Tamai, Y., 2003. Auditory thalamic nuclei projections to the temporal cortex in the rat. Neuroscience 117, 1003– 1016.
- King, C., Nicol, T., McGee, T., Kraus, N., 1999. Thalamic asymmetry is related to acoustic signal complexity. Neurosci. Lett. 267, 89–92.
- Klimesch, W., Freunberger, R., Sauseng, P., 2009. Oscillatory mechanisms of process binding in memory. Neurosci. Biobehav. Rev 34 (7), 1002–1014.
- Konoike, N., Kotozaki, Y., Miyachi, S., Miyauchi, C.M., Yomogida, Y., Akimoto, Y., et al., 2012. Rhythm information represented in the fronto-parieto-cerebellar motor system. Neuroimage. 63, 328–338.
- Kraus, N., Nicol, T., 2005. Brainstem origins for cortical 'what' and 'where' pathways in the auditory system. Trends Neurosci. 28, 176–181.
- Krause, V., Schnitzler, A., Pollok, B., 2010. Functional network interactions during sensorimotor synchronization in musicians and non-musicians. Neuroimage 52, 245–251.
- Lakatos, P., Chen, C.M., O'Connell, M.N., Mills, A., Schroeder, C.E., 2007. Neuronal oscillations and multisensory interaction in primary auditory cortex. Neuron 53, 279–292.
- Lakatos, P., Karmos, G., Mehta, A.D., Ulbert, I., Schroeder, C.E., 2008. Entrainment of neuronal oscillations as a mechanism of attentional selection. Science 320, 110– 113.
- Lakatos, P., Shah, A.S., Knuth, K.H., Ulbert, I., Karmos, G., Schroeder, C.E., 2005. An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. J. Neurophysiol. 94, 1904–1911.
- Large, E.W., 2008. Resonating to musical rhythm: theory and experiment. In: .Grondin, S. (Ed.), The Psychology of Time. Emerald, West Yorkshire, pp. 189– 213.
- Large, E.W., Jones, M.R., 1999. The dynamics of attending: how we track time varying events. Psychol. Rev. 106, 119–159.
- Large, E.W., Snyder, J.S., 2009. Pulse and meter as neural resonance. Ann. N. Y. Acad. Sci. 1169, 46–57.
- Leong, V., Goswami, U., 2013. Assessment of rhythmic entrainment at multiple timescales in dyslexia: evidence for disruption to syllable timing. Hear. Res (in this issue).
- Lomber, S.G., Malhotra, S., 2008. Double dissociation of 'what' and 'where' processing in auditory cortex. Nat. Neurosci. 11, 609–616.
- London, J.M., 2004. Hearing in Time: Psychological Aspects of Musical Meter. Oxford University Press, New York, NY.
- Luo, H., Poeppel, D., 2012. Cortical oscillations in auditory perception and speech: evidence for two temporal windows in human auditory cortex. Front. Psychol. 3, 170.
- Magne, C., Schon, D., Besson, M., 2006. Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. J. Cogn. Neurosci. 18, 199–211.

- Makeig, S., Delorme, A., Westerfield, M., Jung, T.P., Townsend, J., Courchesne, E., et al., 2004. Electroencephalographic brain dynamics following manually responded visual targets. PLoS. Biol. 2, e176.
- Mampe, B., Friederici, A.D., Christophe, A., Wermke, K., 2009. Newborns' cry melody is shaped by their native language. Curr. Biol. 19, 1994–1997.
- Marques, C., Moreno, S., Luis, C.S., Besson, M., 2007. Musicians detect pitch violation in a foreign language better than nonmusicians: behavioral and electrophysiological evidence. J. Cogn. Neurosci. 19, 1453–1463.
- McMullen, N.T., de Venecia, R.K., 1993. Thalamocortical patches in auditory neocortex. Brain Res. 620, 317–322.
- Meinz, E.J., Hambrick, D.Z., 2010. Deliberate practice is necessary but not sufficient to explain individual differences in piano sight-reading skill: the role of working memory capacity. Psychol. Sci. 21, 914–919.Munte, T.F., Nager, W., Beiss, T., Schroeder, C., Altenmuller, E., 2003. Specialization of
- Munte, T.F., Nager, W., Beiss, T., Schroeder, C., Altenmuller, E., 2003. Specialization of the specialized: electrophysiological investigations in professional musicians. Ann. N. Y. Acad. Sci. 999, 131–139.
- Musacchia, G., Sams, M., Skoe, E., Kraus, N., 2007. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. Proc. Natl. Acad. Sci. U. S. A. 104, 15894–15898.
- Musacchia, G., Strait, D., Kraus, N., 2008. Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. Hear. Res. 241, 34–42.
- Nazzi, T., Bertoncini, J., Mehler, J., 1998. Language discrimination by newborns: toward an understanding of the role of rhythm. J. Exp. Psychol. Hum. Percept. Perform. 24, 756–766.
- Oliver, D.L., Huerta, M.F., 1992. Inferior and superior colliculi. In: Webster, D.B., Popper, A.N., Fay, R.R. (Eds.), The Mammalian Auditory Pathway. Springer-Verlag, New York, pp. 168–221.
- Pandya, D.N., Rosene, D.L., Doolittle, A.M., 1994. Corticothalamic connections of auditory-related areas of the temporal lobe in the rhesus monkey. J. Comp. Neurol. 345, 447–471.
- Pantev, C., 1995. Evoked and induced gamma-band activity of the human cortex. Brain Topogr. 7, 321–330.
- Pantev, C., Roberts, L.E., Schulz, M., Engelien, A., Ross, B., 2001. Timbre-specific enhancement of auditory cortical representations in musicians. Neuroreport 12, 169–174.
- Pantev, C., Ross, B., Fujioka, T., Trainor, L.J., Schulte, M., Schulz, M., 2003. Music and learning-induced cortical plasticity. Ann. N. Y. Acad. Sci. 999, 438–450.
- Parbery-Clark, A., Strait, D.L., Kraus, N., 2011. Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. Neuropsychologia 49, 3338–3345.
- Patel, A.D., Daniele, J.R., 2003. An empirical comparison of rhythm in language and music. Cognition 87, B35–B45.
- Quene, H., Port, R.F., 2005. Effects of timing regularity and metrical expectancy on spoken-word perception. Phonetica 62, 1–13.
- Rankin, S.K., Large, E.W., Fink, P.W., 2009. Fractal tempo fluctuation and pulse prediction. Music Percept. 26, 401–403.
- Ramus, F., Hauser, M.D., Miller, C., Morris, D., Mehler, J., 2000. Language discrimination by human newborns and by cotton-top tamarin monkeys. Science 288, 349–351.
- Repp, B.H., 2005. Sensorimotor synchronization: a review of the tapping literature. Psychon. Bull. Rev. 12, 969–992.
- Repp, B.H., 2010. Do metrical accents create illusory phenomenal accents? Atten. Percept. Psychophys. 72 (5), 1390–1403. http://dx.doi.org/10.3758/app.72.5. 1390.
- Saleh, M., Reimer, J., Penn, R., Ojakangas, C.L., Hatsopoulos, N.G., 2010. Fast and slow oscillations in human primary motor cortex predict oncoming behaviorally relevant cues. Neuron 65, 461–471.
- Schroeder, C.E., Lakatos, P., 2009a. Low-frequency neuronal oscillations as instruments of sensory selection. Trends Neurosci. 32, 9–18.
- Schroeder, C.E., Lakatos, P., 2009b. The gamma oscillation: master or slave? Brain Topogr. 22, 24–26.
- Schroeder, C.E., Lakatos, P., Kajikawa, Y., Partan, S., Puce, A., 2008. Neuronal oscillations and visual amplification of speech. Trends Cogn. Sci. 12, 106–113.
- Schroeder, C.E., Wilson, D.A., Radman, T., Scharfman, H., Lakatos, P., 2010. Dynamics of active sensing and perceptual selection. Curr. Opin. Neurobiol. 20, 172–176.
- Shah, A.S., Bressler, S.L., Knuth, K.H., Ding, M., Mehta, A.D., Ulbert, I., et al., 2004. Neural dynamics and the fundamental mechanisms of event-related brain potentials. Cereb. Cortex 14, 476–483.
- Shahin, A.J., Roberts, L.E., Chau, W., Trainor, L.J., Miller, L.M., 2008. Music training leads to the development of timbre-specific gamma band activity. Neuroimage 41, 113–122.
- Shepard, W., Ascher, L., 1973. Effects of linguistic rule conformity on free recall in children and adults. Dev. Psychol. 8, 139.
- Sherman, S.M., Guillery, R.W., 2002. The role of the thalamus in the flow of information to the cortex. Philos. Trans. R. Soc. Lond. B Biol. Sci. 357, 1695–1708.
- Shields, J.L., McHugh, A., Martin, J.G., 1974. Reaction time to phoneme target as a function of rhythmic cues in continuous speech. J. Exp. Psychol. 10, 250–255.
- Singer, W., 1995. Development and plasticity of cortical processing architectures. Science 270, 758–764.
- Slezia, A., Hangya, B., Ulbert, I., Acsady, L., 2011. Phase advancement and nucleusspecific timing of thalamocortical activity during slow cortical oscillation. J. Neurosci. 31, 607–617.

# **ARTICLE IN PRESS**

G. Musacchia et al. / Hearing Research xxx (2013) 1-10

- Snyder, J.S., Large, E.W., 2005. Gamma-band activity reflects the metric structure of rhythmic tone sequences. Brain Res. Cogn. Brain Res. 24, 117–126.
- Staples, S., 1968. A paired-associates learning task utilizing music as the mediator: an exploratory study. J. Music Ther. 5, 53–57.
- Stefanics, G., Hangya, B., Hernadi, I., Winkler, I., Lakatos, P., Ulbert, I., 2010. Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. J. Neurosci. 30, 13578–13585.
   Steriade, M., McCormick, D.A., Sejnowski, T.J., 1993a. Thalamocortical oscillations in
- Steriade, M., McCormick, D.A., Sejnowski, T.J., 1993a. Thalamocortical oscillations in the sleeping and aroused brain. Science 262, 679–685.
   Steriade, M., Nunez, A., Amzica, F., 1993b. A novel slow (<1 Hz) oscillation of</li>
- Steriade, M., Nunez, A., Amzica, F., 1993b. A novel slow (<1 Hz) oscillation of neocortical neurons in vivo: depolarizing and hyperpolarizing components. J. Neurosci. 13, 3252–3265.
- Strait, D.L., Hornickel, J., Kraus, N., 2011. Subcortical processing of speech regularities underlies reading and music aptitude in children. Behav. Brain Funct. 7, 44.
- Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., Schroger, E., 2005. Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. Exp. Brain Res. 161, 1–10.

- Tierney, A.T., Kraus, N., 2013. The ability to tap to a beat relates to cognitive, linguistic, and perceptual skills. Brain Lang. 124, 225–231.
- Trainor, LJ., Shahin, A.J., Roberts, LE., 2009. Understanding the benefits of musical training: effects on oscillatory brain activity. Ann. N. Y. Acad. Sci. 1169, 133–142.
- Uhlhaas, P.J., Roux, F., Rodriguez, E., Rotarska-Jagiela, A., Singer, W., 2010. Neural synchrony and the development of cortical networks. Trends Cogn. Sci. 14, 72–80.
   Weener, P., 1971. Language structure and free recall of verbal messages by children. Dev. Psychol. 5, 237–243.
- Will, U., Berg, E., 2007. Brain wave synchronization and entrainment to periodic acoustic stimuli. Neurosci. Lett. 424, 55–60.
- Winer, J.A., Schreiner, C.E., 2005. Projections from the cochlear nuclear complex to the inferior colliculus. In: The Inferior Colliculus. Springer Science+Business Media, New York, pp. 115–131.
- Zatorre, R.J., Belin, P., Penhune, V.B., 2002. Structure and function of auditory cortex: music and speech. Trends Cogn. Sci. 6, 37–46.
   Zion Golumbic, E.M., Poeppel, D., Schroeder, C.E., 2012. Temporal context in speech
- Zion Golumbic, E.M., Poeppel, D., Schroeder, C.E., 2012. Temporal context in speech processing and attentional stream selection: a behavioral and neural perspective. Brain Lang. 77 (5), 980–991.