



Lemniscal Corticothalamic Feedback in Auditory Scene Analysis

Natsumi Y. Homma^{1,2*} and Victoria M. Bajo³

¹ Center for Integrative Neuroscience, University of California, San Francisco, San Francisco, CA, United States, ² Coleman Memorial Laboratory, Department of Otolaryngology – Head and Neck Surgery, University of California, San Francisco, San Francisco, CA, United States, ³ Department of Physiology, Anatomy and Genetics, University of Oxford, Oxford, United Kingdom

Sound information is transmitted from the ear to central auditory stations of the brain via several nuclei. In addition to these ascending pathways there exist descending projections that can influence the information processing at each of these nuclei. A major descending pathway in the auditory system is the feedback projection from layer VI of the primary auditory cortex (A1) to the ventral division of medial geniculate body (MGBv) in the thalamus. The corticothalamic axons have small glutamatergic terminals that can modulate thalamic processing and thalamocortical information transmission. Corticothalamic neurons also provide input to GABAergic neurons of the thalamic reticular nucleus (TRN) that receives collaterals from the ascending thalamic axons. The balance of corticothalamic and TRN inputs has been shown to refine frequency tuning, firing patterns, and gating of MGBv neurons. Therefore, the thalamus is not merely a relay stage in the chain of auditory nuclei but does participate in complex aspects of sound processing that include top-down modulations. In this review, we aim (i) to examine how lemniscal corticothalamic feedback modulates responses in MGBv neurons, and (ii) to explore how the feedback contributes to auditory scene analysis, particularly on frequency and harmonic perception. Finally, we will discuss potential implications of the role of corticothalamic feedback in music and speech perception, where precise spectral and temporal processing is essential.

OPEN ACCESS

Edited by:

David K. Ryugo,
Garvan Institute of Medical Research,
Australia

Reviewed by:

Jennifer F. Linden,
University College London,
United Kingdom
Kexin Yuan,
Tsinghua University, China

*Correspondence:

Natsumi Y. Homma
Natsumi.Homma@ucsf.edu

Specialty section:

This article was submitted to
Auditory Cognitive Neuroscience,
a section of the journal
Frontiers in Neuroscience

Received: 11 June 2021

Accepted: 30 July 2021

Published: 19 August 2021

Citation:

Homma NY and Bajo VM (2021)
Lemniscal Corticothalamic Feedback
in Auditory Scene Analysis.
Front. Neurosci. 15:723893.
doi: 10.3389/fnins.2021.723893

Keywords: descending projections, layer VI cortical neurons, thalamus, medial geniculate body, tonotopy, harmonicity, speech, music

INTRODUCTION

In everyday life we are constantly analyzing our acoustic environment, which is filled with sounds from many different sources. For example, we can listen to a person next to us while others in the room are chatting. With little effort, we can treat the voice of that talker as the desired foreground signal and segregate it from the background of all other sounds in the room. We can also listen to the melody in a symphony while focusing on the parts played by different musical instruments. These listening abilities are based on the process of “auditory scene analysis” (Bregman, 1990), which is essential for grouping or segregating sound mixtures into perceptually meaningful categories while perceiving the whole auditory environment.

In auditory scene analysis, we categorize different, simultaneously occurring sounds based on several acoustic properties such as onset asynchrony (Darwin, 1984; Darwin and Ciocca, 1992),

harmonicity (Moore et al., 1985; Hartmann et al., 1990), and spatial information (Cherry, 1953; Eramudugolla et al., 2008). We can also group sound sequences, for example, based on their rhythmic structure (Miller and Heise, 1950; Bregman and Campbell, 1971; van Noorden, 1975). In addition to exploiting various acoustic attributes, scene analysis relies on attention, learning and memory and incorporates information from other sensory systems (Shamma et al., 2011; Maddox et al., 2015; Atilgan et al., 2018).

Human speech, animal vocalizations, and the sounds produced by many musical instruments are all periodic sounds comprised of a fundamental frequency (F0) plus its multiple integer harmonics (**Figure 1**). This condition is referred as “harmonicity” and is one of the basic acoustic properties for auditory scene analysis. In addition, harmonicity is considered as a basis of the perceptual attribute, “pitch,” that allow us to order sounds from low to high. Intonation, rising or falling of pitch, expresses grammatical meaning or emotion in speech. Pitch also helps in discriminating voices of different speakers, identifying different musical instruments, or conveying the melodic line in music. F0 differences between concurrent vowels contribute to indicate and segregate different talkers (Culling and Darwin, 1993; de Cheveigné, 1995; Arehart et al., 2011). The harmonic components of vowels show a unique energy distribution of frequencies, known as the spectral envelope. The peaks in the envelope (“formant frequencies”) characterize phonemes and provide critical spectral information to discriminate them (Klatt, 1982; Swanepoel et al., 2012). Similarly, the spectral envelope is characteristic for an individual musical instrument and defines its sound quality (“timbre”) (**Figure 1D**; Town and Bizley, 2013). Harmonicity and spectral regularity can serve as a strong grouping cue and play a key role for music and speech perception (McDermott and Oxenham, 2008; Micheyl and Oxenham, 2010). The harmonic structure of the sound spectrum helps the listener keep track of a speaker in competing, simultaneous speech signals (de Cheveigné et al., 1997; Popham et al., 2018). Moreover, harmonicity can contribute to longer lasting memory storage by connecting several aspects of spectral information to the behavior of a single attribute, namely F0 (McPherson and McDermott, 2020). Although spectral regularity indicates harmonic relations, it can be extended to simply regular intervals of spectral components that deviate from multiple integer harmonics of F0 but have equal spacing or that include shifting of phases on each component, while holding grouping effect and behavioral relevance (Roberts and Bailey, 1996; So et al., 2020).

Natural sounds, such as human speech, animal vocalizations, and many environmental sounds, are comprised of complex spectrotemporal modulations (Chi et al., 1999; Elliott and Theunissen, 2009; McDermott and Simoncelli, 2011). Modulation energy in speech, for example, is captured by a decaying low-pass distribution below ~ 16 Hz temporal modulation frequencies and ~ 2 cycle/octave spectral modulation frequencies (Chi et al., 1999; Elliott and Theunissen, 2009). Word recognition is severely impaired when temporal modulation frequencies below ~ 8 – 12 Hz are unavailable (Drullman et al., 1994; Elliott and Theunissen, 2009). This upper limit closely corresponds to the syllabic rate in speech (Greenberg et al.,

2003; Pellegrino et al., 2011). Identifying syllables is essential for distinguishing words. The analysis of Western music revealed that temporal modulation for music is similar to speech; however, the peak energy is shifted down from 5 Hz for speech to 3 Hz for music reflecting the typical tempo (beats or rhythms) of music (Ding et al., 2017). Both for speech and music, most of modulation energy is < 32 Hz, although rapid temporal modulations > 50 Hz are critical for the perception of aspects such as pitch, lexical meaning, formant and timbre patterns, and their detections in noisy environments (Rosen, 1992; Shamma and Lorenzi, 2013).

Sound information is transmitted from the cochlea, via the medulla, pons, midbrain, and thalamus to the cortex. In parallel to the ascending pathways, descending pathways project information back to each stage (Winer, 2006; Souffi et al., 2021). Feedback projection activity can alter neural excitability (Villa et al., 1991; He, 1997, 2003; Lohse et al., 2020) and tuning properties for frequency (Yan and Ehret, 2002; Luo et al., 2011), intensity (Yan and Ehret, 2002; Ma and Suga, 2007), sound onset (Luo et al., 2008), sound duration (Ma and Suga, 2007), or sound source location (Nakamoto et al., 2008) as has been shown for the cochlear nuclei, inferior colliculus (IC), or medial geniculate body (MGB). Behavioral experiments have demonstrated that descending feedback projections modulate perceptual abilities, including detection and discrimination of sound frequency (Guo et al., 2017), harmonicity (Homma et al., 2017), and location (Bajo et al., 2010) (for review; Lohse et al., 2019).

The projection from layer VI neurons of primary auditory cortex (A1) to the ventral division of medial geniculate body (MGBv) is one of the major feedback pathways (**Figure 2A**). A1-MGBv corticothalamic feedback projections have small excitatory terminals (Ojima, 1994), which are thought to act as “modulators” that regulate gain and firing patterns in the thalamus (Sherman and Guillery, 1998, 2011). The balance of excitation from the corticothalamic neurons and inhibition from thalamic reticular nucleus (TRN) neurons, which receives collateral projections from thalamocortical and corticothalamic neurons, modulates the tuning properties, gating and firing patterns of MGBv neurons (Zhang and Suga, 2000; Tang et al., 2012; Guo et al., 2017; Lohse et al., 2020).

In this review, we first introduce the anatomy and physiology of A1-MGBv corticothalamic feedback projections, then we summarize the feedback modulations of thalamic responses related to spectral and temporal information for speech and music processing. Finally, we explore perceptual effects of corticothalamic feedback, particularly on frequency and harmonicity analysis, and discuss how A1-MGBv corticothalamic feedback could contribute to our music and speech perception.

A1-MGBv CORTICOTHALAMIC PATHWAY

Ascending Thalamocortical Projections

The main auditory nucleus of the thalamus, i.e., the MGB, is subdivided into three distinct areas. The ventral division is part of the lemniscal pathway and shows a tonotopic organization,

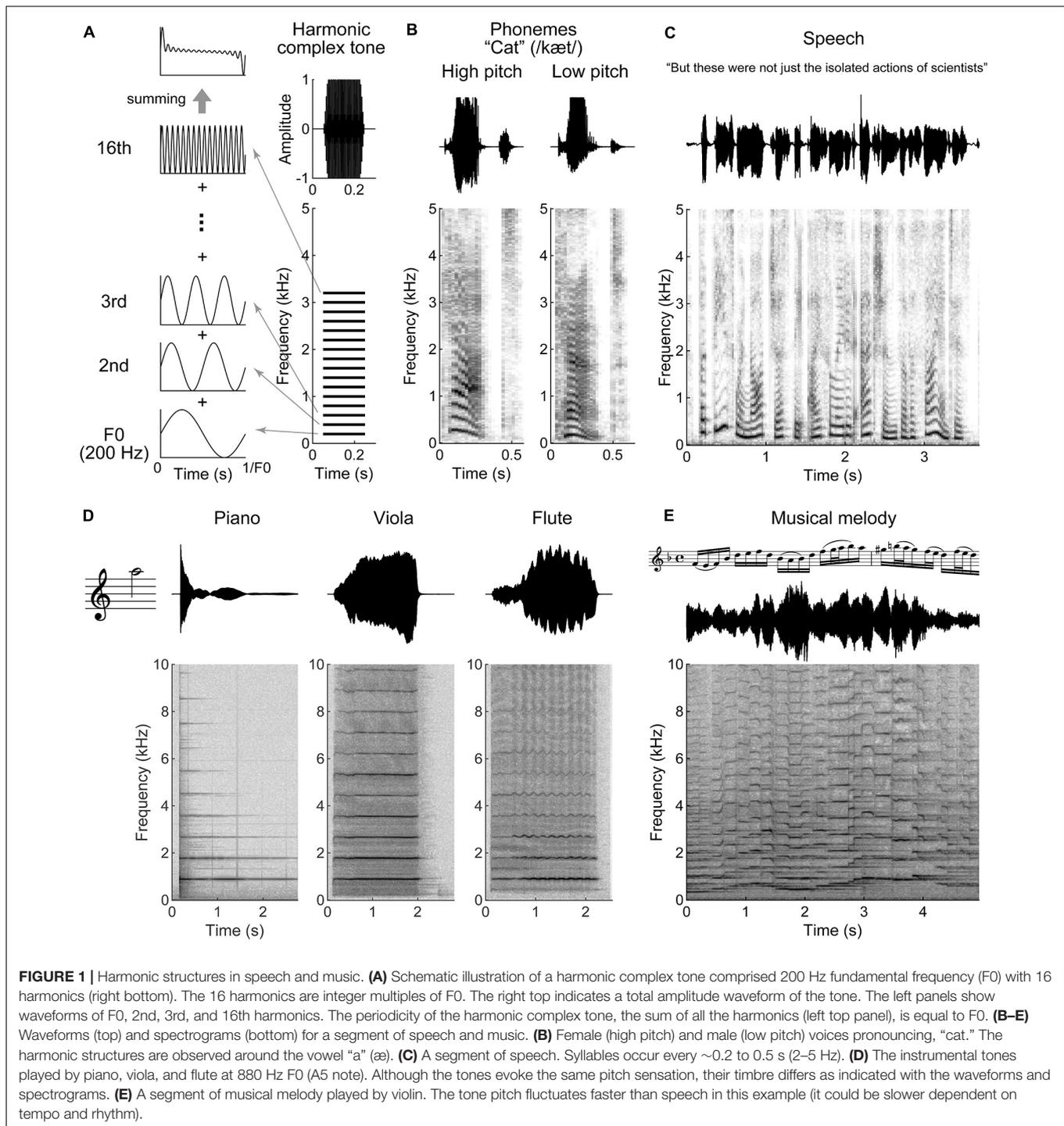
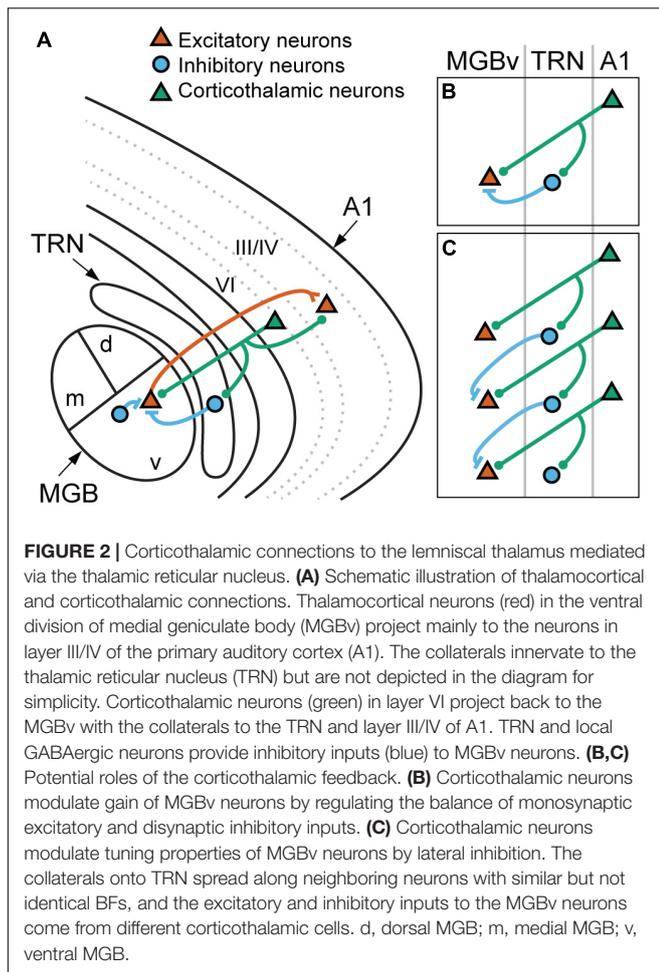


FIGURE 1 | Harmonic structures in speech and music. **(A)** Schematic illustration of a harmonic complex tone comprised 200 Hz fundamental frequency (F0) with 16 harmonics (right bottom). The 16 harmonics are integer multiples of F0. The right top indicates a total amplitude waveform of the tone. The left panels show waveforms of F0, 2nd, 3rd, and 16th harmonics. The periodicity of the harmonic complex tone, the sum of all the harmonics (left top panel), is equal to F0. **(B–E)** Waveforms (top) and spectrograms (bottom) for a segment of speech and music. **(B)** Female (high pitch) and male (low pitch) voices pronouncing, "cat." The harmonic structures are observed around the vowel "a" (æ). **(C)** A segment of speech. Syllables occur every ~0.2 to 0.5 s (2–5 Hz). **(D)** The instrumental tones played by piano, viola, and flute at 880 Hz F0 (A5 note). Although the tones evoke the same pitch sensation, their timbre differs as indicated with the waveforms and spectrograms. **(E)** A segment of musical melody played by violin. The tone pitch fluctuates faster than speech in this example (it could be slower dependent on tempo and rhythm).

in which characteristic frequencies of neurons are arranged in a dorsolateral to ventromedial topographic gradient from low to high frequencies (Aitkin and Webster, 1972; Calford, 1983; **Figure 2A**). MGBv mainly receives inputs from the ipsilateral central nucleus of the inferior colliculus (CNIC) (Calford and Aitkin, 1983; LeDoux et al., 1985; Rouiller and de Ribaupierre, 1985) with glutamatergic but also GABAergic connections (Winer et al., 1996; Peruzzi et al., 1997) and projects to A1

(Andersen et al., 1980a; Lee and Winer, 2008). The cerebral cortex consists of different types of cells that are functionally organized into a laminar structure, and MGBv neurons mainly target layer III/IV of A1 but also other cortical layers and especially layer I (Huang and Winer, 2000; Kimura et al., 2003; Smith et al., 2012; Vasquez-Lopez et al., 2017; Xie et al., 2018). MGBv also receives a small amount of input from the shell regions of IC (Kudo and Niimi, 1980; LeDoux et al., 1985) and



innervates other “core” cortical areas that usually express a tonotopic organization and, to a lesser extent, “belt” areas that receive major non-lemniscal thalamic inputs (Andersen et al., 1980a; Morel and Imig, 1987; Redies et al., 1989; Velenovsky et al., 2003; De La Mothe et al., 2006; Polley et al., 2007; Lee and Winer, 2008; Hackett et al., 2011; Saldeitis et al., 2014). Thalamic inputs sourcing different locations of MGBv in the caudal-to-rostral dimension project to the targeted locations of the core cortical areas in the ventral-to-dorsal dimension without much overlap (Storace et al., 2010, 2011; Read et al., 2011), suggesting distinct functional segregations within the lemniscal thalamocortical system.

By contrast, the medial and dorsal divisions of the auditory thalamus (MGBm and MGBd) are considered non-lemniscal areas and do not have a clear tonotopic organization compared to MGBv (Calford, 1983; Rouiller et al., 1989; Hackett et al., 2011). Comparative studies of these three divisions and non-lemniscal corticothalamic feedback have been reviewed previously (Bartlett, 2013; Lee, 2015). Briefly, MGBm mainly receives inputs from the external cortex of the IC and projects to all layers of the auditory cortex whereas MGBd is innervated by dorsal cortex of IC and projects to layers I, III/IV, VI of belt and parabelt regions and only weakly to core auditory areas (Andersen et al., 1980b;

Kudo and Niimi, 1980; LeDoux et al., 1985). Non-lemniscal MGB also receives inputs from superior colliculus (Holstege and Collewyn, 1982), and sends outputs to the striatum and amygdala (LeDoux et al., 1991), integrating multimodal sensory and emotional information (Weinberger, 2011).

Descending Corticothalamic Projections

The tonotopic organization can also be a hallmark for descending pathways and is preserved in descending lemniscal axons. In principle, lemniscal cortical regions project back to subcortical lemniscal stations, whereas non-lemniscal cortical regions target non-lemniscal stations (Winer, 2006). Thus, lemniscal corticothalamic neurons project back from layer VI to MGBv with minor inputs to MGBm and MGBd (Andersen et al., 1980a; Rouiller and de Ribaupierre, 1985; Bajo et al., 1995; Budinger et al., 2013) in parallel to the tonotopic organization of lemniscal ascending thalamocortical projections (Redies et al., 1989; Rodrigues-Dagaëff et al., 1989; Velenovsky et al., 2003; Read et al., 2008; Hackett et al., 2011; Storace et al., 2011). The information flow is organized in a layer specific manner; layer VI corticothalamic neurons send collaterals to layer III/IV (Figure 2A; Ojima, 1994; Llano and Sherman, 2008), where the axon terminals of thalamocortical projections are mainly found. The A1-MGBv corticothalamic projections are mainly ipsilateral and form the focal topographic reciprocal connections; however, a minority showed non-reciprocal inputs (Winer and Larue, 1987; Winer et al., 2001). In addition, A1 also projects to MGBd as a descending feedforward projection, which is thought to transmit sound information from layer V of A1 to non-primary higher cortices (Ojima, 1994; Llano and Sherman, 2008). This cortico-thalamo-cortical connection is essential for corticocortical communication and processing higher-order sound features (Lee, 2015; Williamson and Polley, 2019).

The corticothalamic feedback projections from layer VI and the cortico-thalamo-cortical signaling from layer V are distinguished by their morphological characteristics. Tracer injection studies showed that the first-order corticothalamic feedback neurons in layer VI have small distal terminals with thin axons and convergent endings in MGBv while the higher-order corticothalamic feedforward neurons in layer V have large boutons in MGBd (Rouiller and Welker, 1991; Ojima, 1994; Bajo et al., 1995; Bartlett et al., 2000; Winer and Prieto, 2001). These two thalamic terminal types are characteristic for Class1 and Class2 glutamatergic projection neurons, respectively, based on structural and physiological properties (Sherman and Guillery, 1998, 2011). Corticothalamic feedforward neurons in layer V are classified as Class1, which have been characterized as “drivers” due to their function of relaying information to the cortex. On the other hand, corticothalamic feedback neurons in layer VI are classified as Class2 or “modulators” controlling how relay neurons transmit their information. They activate type I metabotropic glutamate receptors, which are identified as a characteristic of modulator synapses, and show paired-pulse facilitation with small excitatory post-synaptic potentials (EPSPs) in MGBv and layer IV cortical neurons (Bartlett and Smith, 2002; Lee and Sherman, 2009; Lee et al., 2012). A recent study also supports the idea that corticothalamic feedback

projection neurons in layer VI are “modulators” with high selectivity of information flow while corticothalamic feedforward projections from layer V are “drivers” that integrate and transmit information (Williamson and Polley, 2019). Additionally, layer VI corticothalamic neurons can be activated by preparatory motor actions that trigger reward and auditory inputs in behaving mice, supporting the modulatory role for active listening (Clayton et al., 2021). Overall, layer VI corticothalamic neurons modulate MGBv neurons by regulating their excitability, voltage gated conductance, and synaptic potentials.

Inhibitory Inputs and Intracortical Local Circuits

In addition to MGBv, corticothalamic feedback neurons send their axonal projection terminals to layer IV and the TRN (Jones, 2007; Sherman and Guillery, 2013; **Figure 2A**). TRN consists of GABAergic cells and is located between the cortex and the thalamus wrapping the thalamic nuclei with a sheet structure. The auditory sector is identified at the posterior-ventral part of TRN and receives thalamocortical collaterals (Rouiller et al., 1985; Villa, 1990; Conley et al., 1991). MGBv neurons receive inhibitory inputs from TRN in addition to local and IC GABAergic inputs (Winer and Larue, 1996; Winer et al., 1996; Arcelli et al., 1997; Peruzzi et al., 1997; Zhang et al., 2008; Clarke and Lee, 2018). The proportion of local GABAergic neurons in MGBv present species-specific variations, being almost absent in rodents and ~25% of MGBv neurons in carnivores and primates (Winer and Larue, 1996). Corticothalamic projections to MGBv and TRN generally preserve topographic connections (Conley et al., 1991; Kimura et al., 2005; Cotillon-Williams et al., 2008); therefore, the feedback from layer VI shapes thalamic tuning by modulating the balance between converging excitation and inhibition.

In the cortex, the thalamocortical inputs are received mainly in layer III/IV. The thalamorecipient neurons then innervate the upper or supragranular cortical layers. Recent studies suggest that corticothalamic projection neurons in layer VI induce overall gain change across all cortical layers by recruiting local fast-spiking inhibitory neurons to modulate cortical oscillation (Olsen et al., 2012; Bortone et al., 2014; Kim et al., 2014; Guo et al., 2017). Taken together, this corticothalamic loop arborizing through A1, MGBv, and TRN shapes the receptive field structures and modulates information flow both in the thalamus and cortex and, thus, embodies a crucial aspect of the thalamocortical interface.

CORTICAL MANIPULATION ONTO MGBv NEURONS

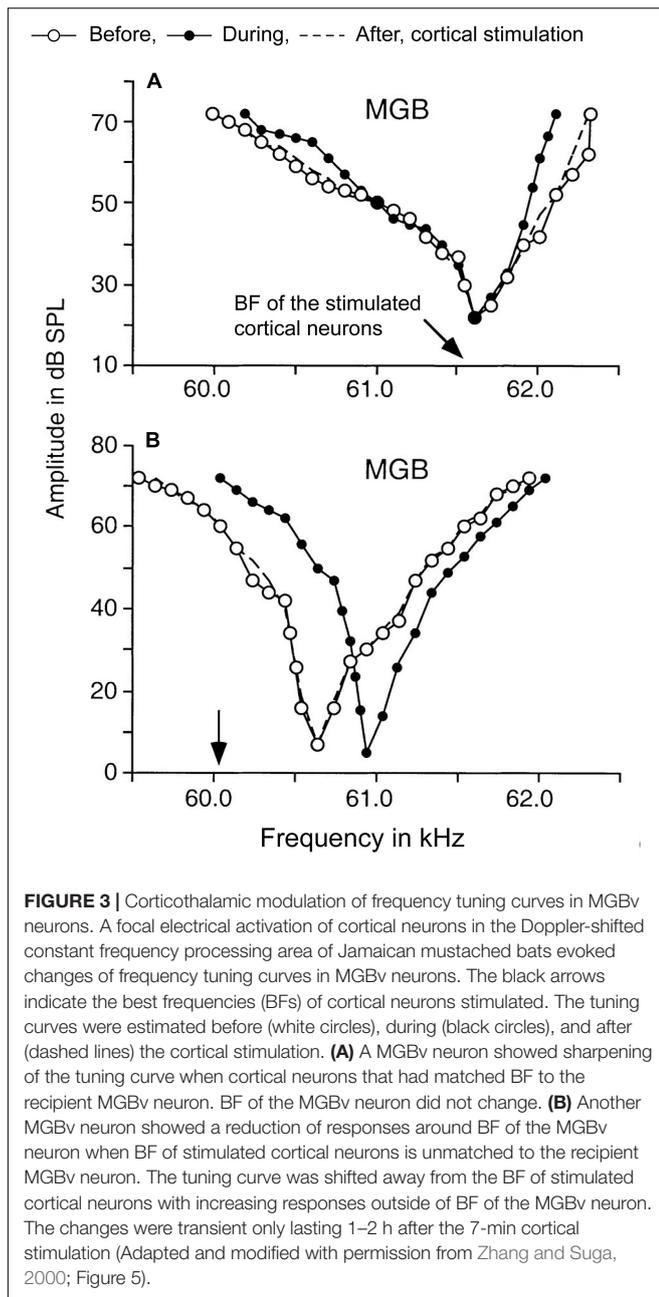
Changes in Spectral Tuning

In laboratory experiments, single frequency tones (“pure tones”) are generated as a sound with a sinusoidal waveform. A simple spectral tuning curve is typically obtained by measuring the change in response to each pure tone at different combinations of frequency and intensity. Neurons in MGBv predominantly show narrow tuning curves, indicating high frequency selectivity,

and short response latencies (cats: Calford, 1983; Morel et al., 1987; Miller et al., 2002; rats: Bordi and LeDoux, 1994; guinea pigs: Edeline et al., 1999; mice: Anderson and Linden, 2011; marmosets: Bartlett et al., 2011). Overall, MGBv neurons show sharper frequency tuning compared to A1 neurons by 0.1–0.3 octave (Miller et al., 2002; Bartlett et al., 2011). It is unknown at this time whether the thalamic frequency tuning properties in anesthetized or passively listening animals are predominantly inherited from the midbrain or shaped by local circuits with or without corticothalamic feedback.

Thalamic tuning properties can be altered by manipulating the corticothalamic activities in layer VI. Modulation of frequency tuning by corticofugal activity is dependent on the relationship between the best frequencies (BFs), i.e., the frequency that evokes the highest firing rate, of subcortical and cortical neurons (e.g., Zhang and Suga, 2000; **Figure 3**). Observed tuning changes in MGBv neurons following A1 manipulation fall into two main categories. First, when the BF of a corticofugal neuron is matched to that of the recipient thalamic neuron, the frequency tuning at the latter is sharpened by facilitation of the responses at BF and reduction of responses to frequencies away from BF (**Figure 3A**). Conversely, when the BF of corticofugal neurons differs from that of the affected MGBv neuron, responses at the BF of MGBv neuron are reduced and responses to surrounding frequencies are enhanced, shifting the tuning curve away from the BF of the stimulated corticofugal neuron (**Figure 3B**). Cortical focal electric stimulation has been shown to induce both types of changes in MGBv of bats (Zhang and Suga, 2000; Tang et al., 2012). Similar changes are also induced in MGBv by electrical stimulation of the cholinergic nucleus basalis or by behavioral conditioning, and are abolished with inactivation of the auditory cortex (guinea pigs: Edeline and Weinberger, 1991; bats: Zhang et al., 1997; mice: Zhang and Yan, 2008; Luo et al., 2011; Nelson et al., 2015). This suggests that the effects on subcortical activity are controlled by the balance of local excitation and inhibition including the influence of various neuromodulators.

These two types of response changes potentially could arise from different, converging projections. When corticothalamic and corticoreticular pathways are strictly reciprocal, the balance of excitation and inhibition on a thalamic relay cell could be modulated by monosynaptic inputs from layer VI corticothalamic cells and disynaptic inputs from TRN mediated via the collateral of the same corticothalamic cells (**Figure 2B**). When the collaterals spread to the neighboring thalamic cells that have different tuning properties (**Figure 2C**), the corticoreticular pathway could sharpen thalamic receptive fields by lateral inhibition. Although the cascades of information flow from the periphery to the first-order thalamic nuclei differ among different sensory systems, the anatomy and physiology of thalamocortical and corticothalamic neurons are generally comparable (Rouiller and Welker, 2000; Sherman and Guillery, 2013). In the somatosensory system, most corticothalamic and corticoreticular projections are organized in a reciprocal manner and contribute to gain control on the thalamic cells, and some projections diverge to neighboring cells that have different tuning properties and show lateral inhibition (Temereanca and Simons, 2004; Li and Ebner, 2007; Lam and Sherman, 2010; Crandall et al.,



2015). Orientation tuning of visual thalamic receptive fields also is shifted by focal pharmacological activation in layer VI of the primary visual cortex (Wang et al., 2016). Thus, a shift of tuning properties is a universal aspect of corticothalamic interactions and likely is mediated by disinaptic inhibition via diverged corticoreticular projections.

The ability to alter receptive field tuning may enable an attentive listener to focus on a specific speaker, other sound attributes of current interest, or adjust to changes in the sound environment by adapting to sound statistics in the environment similar to what has been demonstrated in cortical neurons (Fritz et al., 2003; Holdgraf et al., 2016; Homma et al., 2020).

A major purpose of the modulation of receptive field selectivity by corticothalamic projections could be an increased discrimination ability, e.g., for frequency and musical pitch, to enhance scene analysis in complex sound environments.

Gain Changes

In addition to the modulation of tuning properties, it was observed that lemniscal corticothalamic feedback generally facilitates excitability of MGBv neurons (cats: Ryugo and Weinberger, 1976; Villa et al., 1991; He, 1997; guinea pigs: He et al., 2002; Yu et al., 2004; mice: Guo et al., 2017; Lohse et al., 2020; marmosets: Zhang et al., 2021). Deactivation of the auditory cortex decreases the spontaneous firing rate in MGBv neurons (Ryugo and Weinberger, 1976; Villa et al., 1991; but see Zhang et al., 2021) while cortical activation enhances their responses to sounds (He, 1997; He et al., 2002; Guo et al., 2017). The effect is, however, heterogeneous, and a minority of MGBv neurons can show suppression. The proportion of excitation and inhibition differs among species due to different density of local GABAergic interneurons in MGBv (Winer and Larue, 1996; Arcelli et al., 1997). Since local inhibitory cells as well as long-range inhibitory inputs from the TRN play a key role in shaping thalamic processing, the interspecies differences in inhibitory capacity may bear on potential differences in the ability to modify processing via corticothalamic inputs.

A recent study showed that activation of corticothalamic neurons in layer VI enhanced or suppressed activity in A1, MGBv, and TRN neurons with a dependence on relative timing of optogenetic and sound stimulation (Guo et al., 2017; **Figure 4**). The authors further found that the facilitated activities in MGBv could improve performance on a behavioral frequency detection task by increasing sound-evoked responses whereas better frequency discrimination was more closely related to suppressed cortical stimulus response (Olsen et al., 2012; Bortone et al., 2014). These observations indicate that corticothalamic feedback modulations are diverse and dependent on stimulus and/or task context.

The gain control provided by corticothalamic feedback induces a change in tuning sharpness by shifting the overall excitability of MGBv neurons but generally without affecting BF. As anatomical and physiological studies have shown (see section “A1-MGBv Corticothalamic Pathway”; Guo et al., 2017), corticothalamic feedback modulates not only MGBv but also A1 and TRN. Furthermore, corticothalamic feedback via TRN inhibiting MGBv seems to determine whether A1 neurons respond to weak tones or not (Ibrahim et al., 2021). Future investigations will need to explore the joint effects of the combined thalamocortical-corticothalamic-corticoreticular-intracortical loop. In addition, forward and feedback effects must be studied in the context of natural stimuli, such as communication sounds, and with the consideration of specific task goals and motivations.

Temporal Representation and Precision

MGBv neurons are highly sensitive to temporally fluctuating sound, such as amplitude modulated tones or noise. The majority of these neurons show strongly synchronized firing patterns to temporal modulations of 20–40 Hz, but phase-locking is typically

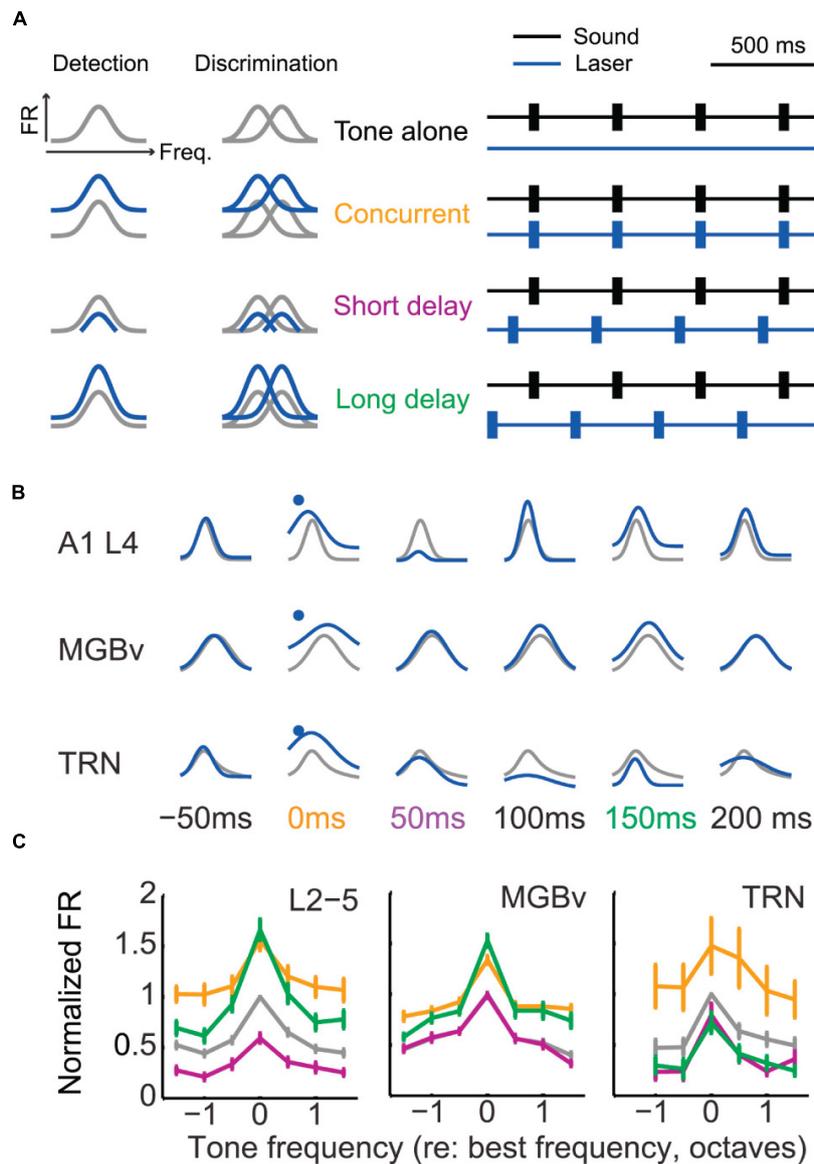


FIGURE 4 | Corticothalamic gain control on frequency tuning curves of A1, MGBv, and TRN neurons. **(A)** Schematic of corticothalamic modulation on the tuning curves of A1 cortical neurons and optogenetic laser stimulation paradigm. Lemniscal layer VI corticothalamic neurons were activated by expressing ChR2 in A1 bilaterally in Ntsr1-Cre mice and using pulses of blue laser light. Depending on the duration of the sound stimulation delay following laser activation of corticothalamic neurons in layer VI, the tuning curves were predicted to be modulated distinctively and to define tone detection and discrimination behaviors. The animals were trained to detect or discriminate sounds using an avoidance task. **(B)** Representative modulation effect on the tuning curves of A1, MGBv, and TRN neurons (gray, tone-alone; blue, tone-and-laser). Sound-evoked responses enhanced in A1 and MGBv neurons for the concurrent stimulation of tones and laser (orange). For the tone presentation with a short or long delay following the corticothalamic stimulation (purple or green), corticothalamic modulation effects differed in A1, MGBv, and TRN neurons, showing enhancement for one station while showing suppression for the others. **(C)** Tone-evoked firing rates (mean \pm SEM) were normalized and compared to the values evoked for the BF of the tone-alone condition. When tone presentation and corticothalamic stimulation were concurrent, all A1, MGBv, and TRN neurons increased firing rates (paired *t* test, $p < 0.05$). When tones were presented with a short delay following laser stimulation, the firing rates decreased in A1 and TRN neurons ($p < 0.05$) but no change was found in MGBv neurons ($p > 0.05$). For the long delayed condition, A1 and MGBv neurons increased firing rates ($p < 0.05$), whereas TRN neurons reduced it ($p < 0.05$) (Adapted and modified with permission from Guo et al., 2017; Figures 4B, 5D,E).

limited to modulation frequencies below ~ 60 to 200 Hz (guinea pigs: Creutzfeldt et al., 1980, cats: Rouiller et al., 1981; Miller et al., 2002, non-human primates: Preuss and Müller-Preuss, 1990; Bartlett and Wang, 2007, 2011). Those values are two to four times higher for MGBv neurons than for A1 neurons (Miller

et al., 2002; Bartlett and Wang, 2007) reflecting the progressively reduced ability of envelope synchronization along the ascending pathway (Joris et al., 2004). While $\sim 40\%$ of MGBv neurons with BFs mostly > 1.5 kHz show exclusively a rate code, another $\sim 40\%$ of MGBv neurons show both synchronized responses

and increased firing in a single neuron dependent on repetition rates, and the proportion was five times larger than that for A1, suggesting that MGBv is a transition stage for this computation (Bartlett and Wang, 2007; but see Yin et al., 2011). For some thalamocortical transmissions, the temporal code is transformed to a rate code due to the synaptic interactions, in which excitatory and inhibitory inputs both inherit a temporal code from MGBv but the spiking response loses synchronization if they are in-phase (Bartlett and Wang, 2007; Gao and Wehr, 2015). Furthermore, higher auditory cortical fields encode temporal fluctuations predominantly in firing rate, with the exception of very low temporal modulation frequencies (Scott et al., 2011; Hullett et al., 2016; Johnson et al., 2020).

MGBv neurons respond strongly to species-specific and other animal vocalizations, and the response preferences are consistent with the values estimated by conventional artificial stimuli synthesized to represent acoustic properties contained in vocalizations (Creutzfeldt et al., 1980; Symmes et al., 1980; Philibert et al., 2005; Suta et al., 2007). Phase-locking to the F0 of harmonic vocalization is observed when neurons have the capacity of phase locking to pure tones in the range of the F0 (Wallace et al., 2007). The temporal spike pattern of MGBv neurons more closely matches the spectrogram of vocalizations and shows higher decoding performance than A1 neurons (Huetz et al., 2009; Souffi et al., 2020), suggesting that some information of the vocalization content is still present in a temporal code at the thalamic level.

Thalamocortical neurons have been shown to switch between two distinct firing modes: “burst” and “tonic” (Steriade et al., 1993; Sherman, 2001; Llinás and Steriade, 2006). When the principal cells are depolarized by incoming inputs and switch into “tonic mode,” voltage-gated T-type Ca^{2+} channels become inactivated. In this mode, the firing patterns show a linear relationship to the input strength and spikes occur with high temporal precision (Mease et al., 2014; Hasse and Briggs, 2017); therefore, the “tonic mode” is more suitable for sound discrimination. After a period of depolarization, cells switch into “burst mode” by re-activation of inward Ca^{2+} current. In this mode, cells are hyperpolarized and prone to produce spikes with lowered threshold and less temporal precision. Thus, the input-output relationship is highly non-linear (McCormick and Feese, 1990; Zhan et al., 1999) and is thought to be more helpful for sound detection (Hu et al., 1994; Bartlett and Smith, 1999; but see Massaux et al., 2004).

Stimulation of corticothalamic neurons in layer VI induces depolarization in MGBv neurons shortly after hyperpolarization (Bartlett and Smith, 2002; Yu et al., 2004), indicating that corticothalamic feedback can induce MGBv neurons to act in the “tonic mode” as has been seen in other sensory systems (Mease et al., 2014; but see Denman and Contreras, 2015). Furthermore, corticothalamic activation reduces adaptation to rapid repetitive stimulation as excitatory and inhibitory synaptic depression differs between the first-order visual/somatosensory thalamus and TRN (Mease et al., 2014; Crandall et al., 2015). In the auditory thalamus, corticothalamic activity seems to reduce adaptation to less-salient modulated noise since inactivation of corticothalamic neurons blocked the reduction (Kommajosyula

et al., 2021). Therefore, corticothalamic neurons may be helpful for maintaining precise temporal responses in sequential or rapidly fluctuating sounds, which is characteristic of music and speech, and especially in less-salient sounds.

The assumption that corticothalamic neurons modulate temporal precision in MGBv is mainly based on intracellular recording of responses to pure tones or broadband noise. The likelihood of tonic and burst modes in the thalamus is affected by the different brain states of waking, sleep, attentiveness, and anesthesia (Weyand et al., 2001; Massaux et al., 2004; Gent et al., 2018). In the auditory thalamocortical system, burst mode is suppressed when spectrotemporally modulated broadband noise is presented compared to spontaneous or tone-driven activity (Miller and Schreiner, 2000). Further investigations are needed that use stimulus with more naturalistic modulations and behaving animals in order to dissect how the corticothalamic feedback affects the perception of naturally modulated sound, such as speech and music.

Overall, corticothalamic neurons in layer VI have potentially three major physiological functions in their effects on MGBv neurons: (i) refining the receptive field structure, (ii) modulating response gain, and (iii) controlling temporal precision, by regulating the balance of monosynaptic excitation and disynaptic inhibition. These three aspects are not operating independently, and the causes and effects of these interactions remain to be explored in more detail. This is relevant in the context of the next question, namely how corticothalamic modulations shape sound analysis and auditory perception.

EFFECTS OF CORTICOTHALAMIC FEEDBACK ON SOUND PERCEPTION

Frequency Analysis

Although some physiological functions of lemniscal auditory corticothalamic neurons have been gradually revealed in the past decades, it remains to be determined how the feedback affects hearing abilities in common, natural tasks. It has been technically challenging to selectively target corticothalamic neurons for recording and/or manipulations while, at the same time, measuring perceptual attributes in awake, behaving animals. Although a recent study using optogenetic phototagging in awake mice showed that layer VI corticothalamic neurons have narrower frequency tuning and higher selectivity of information flow compared to layer V corticofugal neurons (Williamson and Polley, 2019), the studies for corticothalamic modulations discussed above were largely based on recordings under anesthesia. It has been postulated that corticothalamic feedback is, in particular, required for more complex sound processing in behaving animals since ablation of auditory cortex revealed performance deficits in discrimination of frequency modulated tones but not for simple frequency tones (Ohl et al., 1999; Ono et al., 2006). However, none of these studies could dissect the separate roles of cortico-cortical vs. thalamo-cortical vs. cortico-thalamic contributions. More recently, layer specific electrical microstimulation showed that modulation of signal detection and cortical frequency

processing do appear to involve recurrent cortico–thalamo–cortical interactions (Happel et al., 2014; Saldeitis et al., 2021). As mentioned above (see section “Gain Changes”), optogenetic activation of corticothalamic neurons in layer VI modulated sound detection and frequency discrimination abilities of mice (Guo et al., 2017). It demonstrated that the physiological effects of layer VI corticothalamic feedback, such as refining tuning curves and controlling gain of MGBv and A1 neurons, can contribute to modulate frequency perception. Collectively, the present understanding is that layer VI corticothalamic neurons that project to MGBv neurons can affect spectral perception and are likely critical for perceptually demanding situations.

Harmonic Structure Analysis

Mistuning Detection

We hypothesized that corticothalamic feedback from layer VI is particularly important for complex sound processing and focused our study on harmonic structure. Harmonicity perception is potentially regulated by all the three physiological effects of layer VI corticothalamic feedback proposed above. Harmonicity is a strong grouping cue in speech/music perception and scene analysis. Sounds with a harmonic structure are typically perceived as one single entity associated with a specific pitch (Roberts and Bregman, 1991; Roberts and Bailey, 1993), despite containing many different frequency components. When background noise has harmonic structure, detection of foreground sound is improved (Deroche and Culling, 2011; Steinmetzger and Rosen, 2015; Guest and Oxenham, 2019). In order to explore roles of harmonicity, inharmonic sound stimuli have been generated in several different ways by perturbing a regular frequency interval of harmonic structure. It can be achieved by simply shifting all the harmonics to lower or higher frequencies to the same degree (i.e., the frequency interval no longer matches to F0 but regular), stretching out frequency intervals of harmonic components (i.e., each interval differs in the harmonic series), or randomly shifting each harmonic in a small degree. Since the former two cases preserve some levels of spectral regularity, this moderately contributes to fused perception (Roberts and Brunstrom, 1998, 2001). Using harmonic and randomly inharmonic synthetic vowels, harmonicity is shown to improve the segregation of concurrent vowels (Culling and Darwin, 1994; de Cheveigné, 1995; de Cheveigné et al., 1997). When inharmonicity is artificially introduced by jittering each harmonic of speech, the accurate segregation of speech from competing speech or speech-like noise is impaired (Popham et al., 2018).

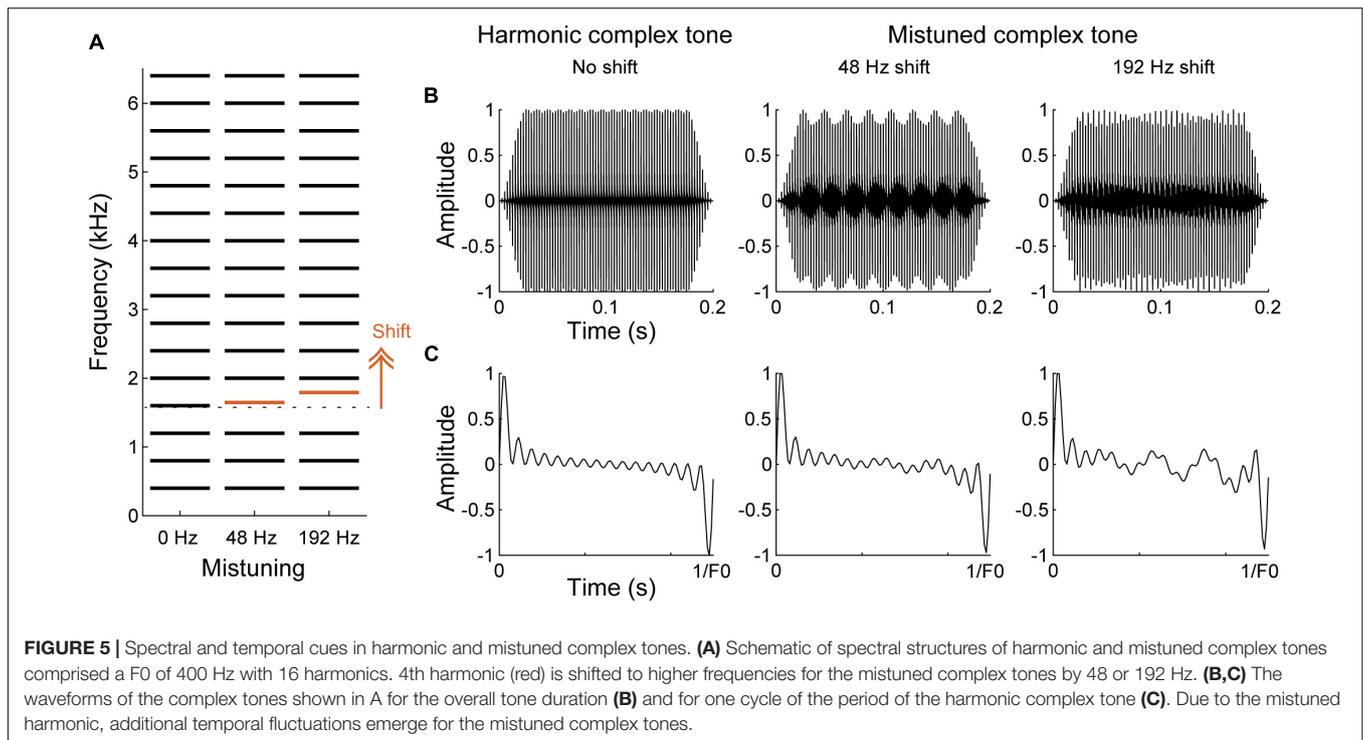
One broadly studied inharmonic tone paradigm is a mistuned complex tone, which comprises a harmonic shifted to lower or higher frequency in an otherwise harmonic complex tone (i.e., “mistuning”) (Figure 5). Specifically, the shifted component can be heard as standing out as a separate tone for low frequency harmonics or produce a sensation of roughness in the sound quality for high frequency harmonics (Moore et al., 1985, 1986; Hartmann et al., 1990). Perception of the shifted component as a separate tone is dependent on the frequencies of shifted components and degrades for higher F0s (Hartmann et al., 1990; Gockel and Carlyon, 2018). Based on the cochlear filtering model,

a harmonic is assumed resolved when it falls in a single filter bank, while it is considered unresolved when several harmonics excite the same filter (Houtsma and Smurzynski, 1990; Shackleton and Carlyon, 1994). Thus, mistuning may be detected as a deviated spectral component in a spectral template of resolved harmonics (“harmonic template”) that is expected to be a series of multiple integers of the F0 of a harmonic sound (“spectral cue”) (Goldstein, 1973; Terhardt, 1974; Lin and Hartmann, 1998). Alternatively, the sensation of roughness or “beating,” which is produced by an interaction of adjacent frequency components within the same cochlea filter, has been thought to help detect disruption of harmonicity (Assmann and Summerfield, 1994; Culling and Darwin, 1994; but see de Cheveigné, 1999). Detecting an inharmonic sound may also be assisted by the temporal excitation patterns synchronizing to the envelope fluctuations (“temporal cue”) (Licklider, 1951; Meddis and Hewitt, 1991a,b). Both, spectral and temporal cues can be used for detecting a change of harmonicity.

In terms of harmonicity perception by animals, they have been found to be sensitive to harmonic structures (Kalluri et al., 2008) and can perform behavioral tasks for F0 judgment (Tomlinson and Schwarz, 1988; Walker et al., 2009; Osmanski et al., 2013) and mistuning detection (Lohr and Dooling, 1998; Klinge and Klump, 2009; Homma et al., 2016). Smaller animals are more likely to rely on temporal cues due to their generally broader cochlea filters (Sumner et al., 2018; Walker et al., 2019).

Neural Responses to Harmonic Complex Sounds

Neural responses can encode both spectral and temporal cues to harmonic and inharmonic complex sound (for review; Micheyl and Oxenham, 2010). At the auditory nerve, neurons respond to individual frequency components and can follow the F0 of complex tones with phase-locked responses up to ~1 kHz (Horst et al., 1986; Cariani and Delgutte, 1996; Sinex et al., 2003; Cedolin and Delgutte, 2005). At higher stations of the central auditory system, temporal cues are degraded, and sound information is integrated. For harmonic sounds, neurons in IC and A1 increase firing rate when resolved harmonics are close to the neurons’ best frequencies and both stations show some phase-locking to the envelope periodicity (F0) for unresolved harmonics (Schwarz and Tomlinson, 1990; Steinschneider et al., 1998; Fishman et al., 2013; Su and Delgutte, 2019, 2020). The phase-locking limits to the F0, however, decrease tenfold from IC to A1. For inharmonic sounds, the neurons show phase-locking to the fine structure of the envelope as well as the periodicity of interactions between mistuned and neighboring harmonics (“beating”) (Sinex et al., 2002, 2005; Fishman and Steinschneider, 2010; Homma et al., 2017). While synchronized responses can be observed regardless of the distance between a mistuned harmonic and a frequency that MGBv or IC neurons are tuned to, changes of temporal patterns are weaker for A1 neurons when a mistuned harmonic is far away from a tuned frequency. In addition, the changes of firing rates occur to inharmonic sounds. The neurons increase their firing rates compared to the responses to the harmonic sound that has the same spectral components except the mistuned harmonic. Although a proportion of the neurons shows opposite decreasing trend, it results in enhanced



responses for IC, MGBv and A1 neurons on average. Similar to temporal patterns, firing rates to mistuning are specifically enhanced in A1 when a mistuned harmonic is closer to a neuron's tuned-frequency, and the frequency specific changes of temporal patterns and firing rates in A1 are thought to correlate with "standing-out" perception of a mistuned harmonic in humans (Fishman and Steinschneider, 2010). It is unknown whether thalamocortical and/or corticothalamic projections contribute to form the specificity to the mistuned harmonic frequency in A1.

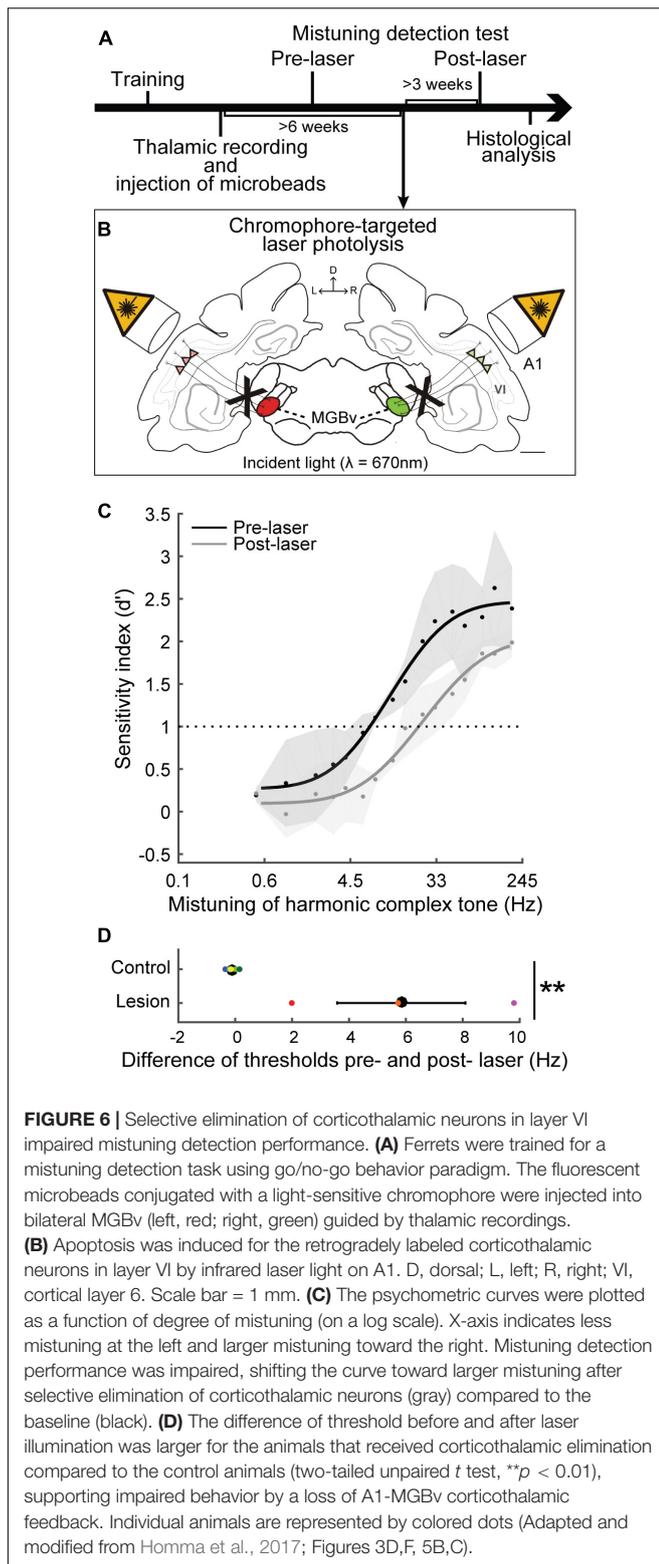
The corticothalamic modulation likely enhances spectral analysis of harmonicity in MGBv by sharpening the spectral tuning, i.e., increasing the sensitivity to small frequency shifts of mistuning. Furthermore, the corticothalamic modulation may assist detection of precise temporal excitation patterns of mistuning by switching to a "tonic mode" and improving temporal representation. Moreover, enhanced encoding of harmonic components and periodicities may refine the analysis of harmonic complex tones via spectral and temporal cues, respectively. Consequently, corticothalamic feedback may improve pitch discrimination.

In human and non-human primate auditory cortex, a population of neurons has been found to be specialized for harmonicity processing. Those neurons are excited by the F0 of a harmonic complex sound even when the actual F0 is omitted ("missing fundamental") or by periodic broadband noise stimulus that evoke pitch sensation, and were identified at the low frequency border of A1 and neighboring core region as well as at the adjacent belt regions (human: Patterson et al., 2002; Penagos et al., 2004; non-human primate: Bendor and Wang, 2005). For humans, non-primary auditory cortex is particularly critical for detecting pitch saliency and changes in

pitch (Patterson et al., 2002; Penagos et al., 2004). Neurons responding to a subset of harmonics in harmonic complex tones ("harmonic template"), independent from their responses to pure tones, are scattered throughout the core regions of non-human primate (Feng and Wang, 2017). No equivalent type of response has been unequivocally identified at subcortical stations. It remains to be explored how the F0 representation emerges in the auditory system, whether the transformation is achieved by subcortical or cortical processing, and whether corticothalamic feedback is required.

Selective Elimination of Corticothalamic Neurons

We have demonstrated that selective elimination of layer VI corticothalamic neurons using chromophore targeted laser photolysis impairs the ability of ferrets to detect mistuned complex tones (Homma et al., 2017). In the study, ferrets were trained in a go/no-go task to detect an inharmonic tone, which comprises the mistuned 4th harmonic in an otherwise harmonic complex tone of 16 harmonics with a F0 of 400 Hz (Figures 5, 6A). Then, fluorescent microbeads conjugated with a light-sensitive chromophore were injected bilaterally in MGBv, and >6 weeks later apoptosis was induced in the retrogradely labeled corticothalamic neurons in layer VI by focusing a infrared laser beam on A1 at the depth of the targeted layer (Figure 6B). About 60% of corticothalamic neurons were selectively eliminated. Mistuning sensitivity was measured behaviorally before and after the elimination of corticothalamic neurons, and the psychometric curve was constructed as a function of degree of mistuning (Figure 6C). Shifts of the 4th harmonic to a higher frequency ranged from 0 to 192 Hz. After the elimination of layer VI corticothalamic



neurons, the psychometric functions were displaced to larger degrees of mistuning with reduced sensitivity (d'), indicating a deficit in discrimination ability. In addition, the threshold

of detecting mistuning increased for the animals that received the corticothalamic lesion (**Figure 6D**). Although the lack of corticothalamic feedback could have reduced excitability of MGBv and decreased overall hearing sensitivity, there was no difference between lesion and control animals in the baseline performance of detecting an inharmonic tone with maximum degree of mistuning reinforced by a level difference of reference and target tones. These suggest A1-MGBv corticothalamic neurons are essential for successfully processing at least one important contributor to auditory scene analysis, namely for determining the harmonic structure of complex sounds. It remains to be elucidated how exactly the physiological changes of MGBv neurons by layer VI corticothalamic feedback contribute to mistuning detection related the temporal and spectral cues in mistuned complex tones. While reshaping of the receptive field structures is expected to improve spectral analysis of resolved harmonic components in MGBv, enhanced temporal precision in MGBv is predicted to refine temporal representation of periodicities in inharmonic complex tones. In addition, the modulation of gain titrates excitability of MGBv, which may control the focus to mistuned harmonic. All are plausible to increase the acuity of mistuning perception.

We postulate that corticothalamic feedback benefits other aspects of auditory scene analysis too. For example, enhanced signal detection by corticothalamic feedback seems to be robust for less salient inputs (Happel et al., 2014; Kommajosyula et al., 2021). Corticothalamic feedback may also contribute to signal-in-noise processing by amplifying weak foreground sounds by controlling the gain of MGBv neurons. Noise invariance also emerges from MGBv to A1 (Las et al., 2005; Rabinowitz et al., 2012; Schneider and Woolley, 2013; Souffi et al., 2020) and may be controlled by corticothalamic feedback. Finally, although speculative, the feedback may help in detecting onset synchrony, discriminating consonance and timbre, or assessing reverberation effects, by enhancing spectral and temporal processing in the MGBv.

SPEECH AND MUSIC PROCESSING IN MGBv

In this section, we will discuss potential roles of lemniscal corticothalamic feedback in speech and music processing. The ability of modulating frequency and harmonicity perception suggests layer VI corticothalamic projections can be involved in regulating speech and music recognition. Although we will mainly focus on MGBv, we briefly summarize findings in the human auditory cortex for an overall view of the auditory forebrain system (for review; Zatorre and Schönwiesner, 2011; Leonard and Chang, 2014). Growing evidence indicates that the left hemisphere is specialized for speech processing while the right hemisphere is dedicated for music processing (e.g., Zatorre et al., 1994; Griffiths et al., 1999; Tervaniemi et al., 2000; Albouy et al., 2020). This asymmetry is supported by finer temporal representation for speech in the left and superior spectral representation for music in the right (Zatorre et al., 2002; Poeppel, 2003). Furthermore, cortical regions are hierarchically

organized and functionally segregated. Spectrotemporal and phonological analyses for speech takes place in the left superior temporal gyrus (STG) and superior temporal sulcus (STS), which fall in the traditional “Wernicke’s area” (Scott, 2000; Davis and Johnsrude, 2003; Mesgarani et al., 2014). Human voice identities are encoded in the right STS (Belin et al., 2000; Schall et al., 2015). While low-level sound features are mainly represented in the core regions, high-order sound features in speech and music are encoded in non-primary regions (Norman-Haignere et al., 2015). Although frequency information is first processed in the lemniscal core regions with their tonotopic organizations, non-primary regions play key roles in pitch and melody perception (Patterson et al., 2002; Penagos et al., 2004; Albouy et al., 2020). Stronger and more selective responses to a single speaker in competing simultaneous speech are observed in non-primary areas than in A1 (Ding and Simon, 2012; Mesgarani and Chang, 2012; O’Sullivan et al., 2019). Thus, the auditory cortex, particularly high-order regions, are essential for phonological, semantic, and melodic processing of speech and music.

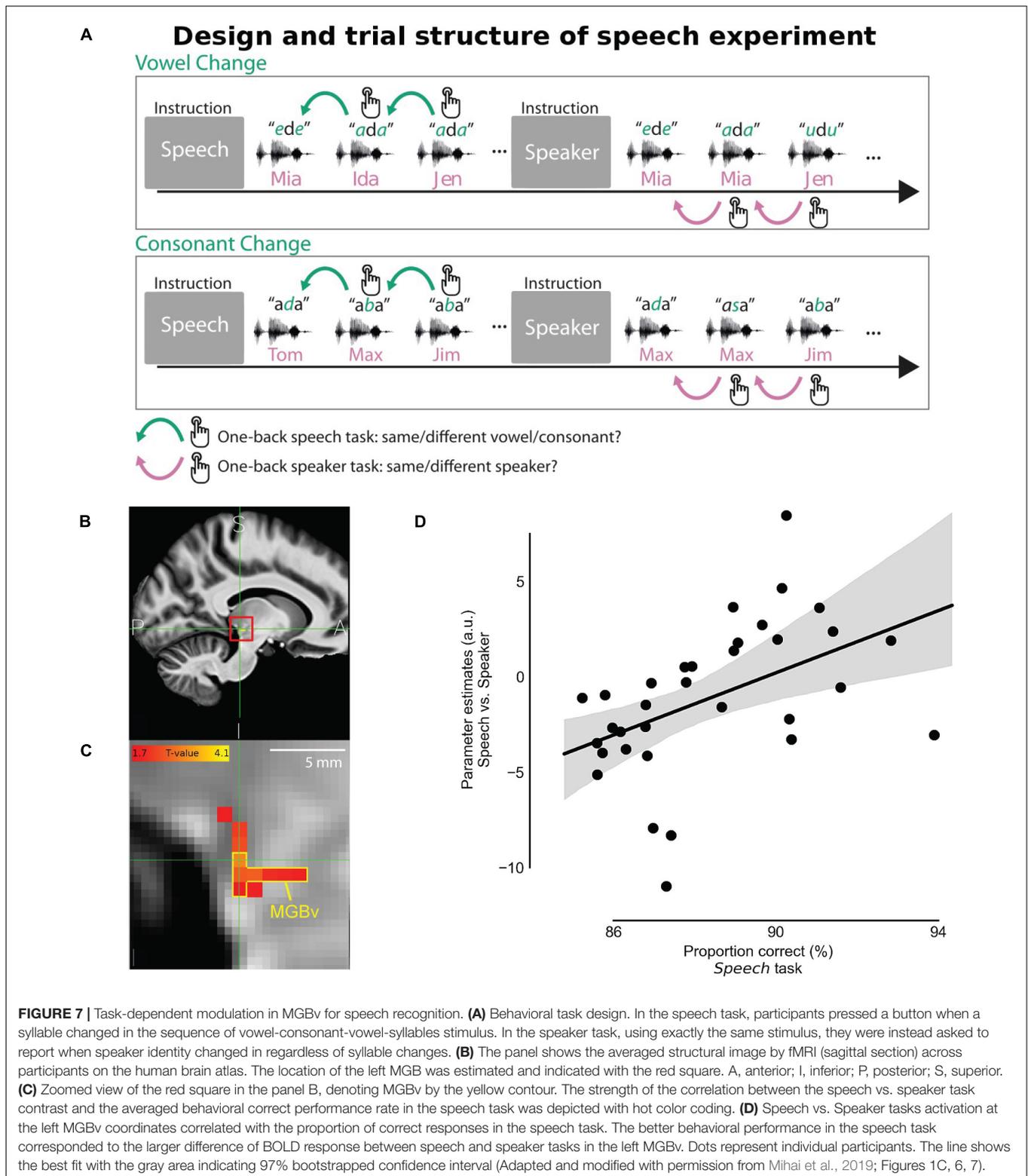
Speech Processing

The human auditory thalamus is involved in some aspects of language processing. People with developmental dyslexia have difficulty in reading and writing and often exhibit impaired auditory and visual timing processing in various cortical and subcortical regions (for review; Ozernov-Palchik and Gaab, 2016; Stein, 2019). Some of those deficits could be traced to changes in the MGB (Galaburda et al., 1994; Diaz et al., 2012). Dyslexics often experience hindrances in auditory signal processing and sensorimotor processing. Psychoacoustic testing and auditory evoked potential studies show reduced sensitivities to discriminating temporally/spectrally modulated sound, or syllables (Stein and McAnally, 1995; Kraus et al., 1996; Menell et al., 1999; Goswami et al., 2002), and neuroimaging studies indicate deficits in rhythmic perception and audio-motor integration for dyslexics, which is in line with different neural phase alignment and consistency in the delta band compared to the control group (Hämäläinen et al., 2012; Colling et al., 2017). Functional magnetic resonance imaging (fMRI) revealed decreased activity in the left MGB for dyslexics in a syllabic discrimination task but not in passive listening condition (Diaz et al., 2012), supporting the possibility that top-down modulation, potentially corticothalamic feedback, differs between dyslexics and controls. Morphological examination showed that cells are smaller in the left than in the right MGB for dyslexics while no asymmetry is observed for control subjects (Galaburda et al., 1994). Dyslexics can exhibit malformation of cortical structures, “microgyria.” An animal model expressing microgyria also exhibited abnormal anatomical changes in the MGB as well as temporal processing deficits in behavioral tasks similar to human dyslexics (Fitch et al., 1994; Herman et al., 1997; Peiffer et al., 2002; Anderson and Linden, 2016). Irregular connections between MGBv and microgyri may perturb temporal processing in dyslexics. Although further investigations are required of the anatomical and functional changes in MGB that may contribute to phonological skills, the ability to process

syllables or words with high temporal precision for spoken language in MGB appears to be closely linked to processing of written language.

Other evidence that the auditory thalamus is involved in speech processing arises from the modulation of thalamic activity by cognitive demands during speech-based tasks via top-down feedback (Alain et al., 2005; Christensen et al., 2008; von Kriegstein et al., 2008; Mihai et al., 2019). Although speech representation in the human auditory cortex has been extensively studied (for review; Zatorre and Schönwiesner, 2011; Leonard and Chang, 2014), investigations in the human thalamus with non-invasive methods have been challenging due to its relatively small volume and the deep anatomical position in the brain. Thus, the thalamic activities measured by positron emission tomography (PET) (Salvi et al., 2002) or fMRI (Alain et al., 2005; Tervaniemi et al., 2006; Christensen et al., 2008), often only reflect overall responses of the thalamic complex to speech signal. Although language processing is lateralized to the left hemisphere in the cortex (Zatorre et al., 2002), the presence of an equivalent thalamic lateralization is not fully established. Contrasting activations to a consonant-vowel-consonant-vowel pseudoword were observed for a change of duration in left thalamus and for a change of frequency in right thalamus (Tervaniemi et al., 2006). In addition, only the activity in right thalamus significantly differed between diotic and dichotic attentive listening conditions while both sides were activated for one to three syllable nouns compared to reversed speech (Christensen et al., 2008). Two other studies support left lateralization for sentence and vowel processing in thalamus (Salvi et al., 2002; Alain et al., 2005). The left thalamus was activated for trials with successful identifications of two vowels that were concurrently presented (Alain et al., 2005) further supporting that the thalamus is involved in F0 discrimination with top-down modulation. It is, thus, likely that some basic speech processing aspects are lateralized to the left thalamus.

Recent fMRI studies with finer spatial resolution successfully identified MGB and captured the tonotopic organization in its ventral division (Moerel et al., 2015; Mihai et al., 2019). The activated responses for discriminating speech signal was observed in both sides of MGB, however, activity correlated to the behavioral performance in a speech recognition task was only observed in the left MGB (von Kriegstein et al., 2008; Mihai et al., 2019). Mihai et al. (2019) assigned two different attentional tasks while listening to an identical set of sound stimuli. They asked participants to report a change in either the presented syllables or the speaker identity (**Figure 7A**). Changes between syllables were reported for the speech task while detecting change of F0s was used for the speaker task. Although there was no significant difference between speech and speaker task for the blood-oxygen-level-dependent (BOLD) responses in MGBv, the correlation between the evoked activity and the behavior performance of the speech task was only found in the left MGBv (**Figures 7B–D**). This suggests that top-down modulations enhance speech processing in the left MGBv. Then, introducing speech-shaped white noise as background, Mihai et al. (2021) showed that the enhanced top-down modulation



on speech recognition is strongly observed in left MGBv when listening condition is challenging.

Corticothalamic feedback potentially strengthens spectral and temporal processing in MGBv and modulates frequency

and harmonicity perception. Thus, syllable discrimination may have relied on enhanced frequency tuning and more precise spike representation via the feedback modulation to characterize individual syllables. In particular, attention could

have modulated corticothalamic gain, affecting the excitability of MGBv neurons. Furthermore, although F0 discrimination did not show correlations with a lateralized change in activity in the study by Mihai et al. (2019), attentional modulation may be more relevant in harder discrimination tasks, such as segregation of competing simultaneous speech. We demonstrated that corticothalamic feedback was essential for processing spectral and/or temporal cues of the harmonic complex tones and detecting mistuning. A similar mechanism could assist speaker discrimination, mainly based on F0 discrimination using spectral structure and periodicity. F0 discrimination also helps to segregate foreground sounds from background “noise.” Thus, top-down modulation via corticothalamic feedback may assist segregating simultaneously presented signals, which is an essential function of auditory scene analysis.

It would be interesting to examine what features of speech are extracted at the lemniscal thalamus and how the processing is modulated by attention or task demands. The temporal resolution of magnetoencephalography (MEG), electroencephalography (EEG), or electrocorticography (ECoG) (<10 msec) is finer than fMRI (<5 s), and recent studies showed that activity in deep subcortical structures can be detected by MEG (Müller et al., 2019; Pizzo et al., 2019). Technical advances in temporal and spatial resolutions are expected to dissect in greater detail human speech processing mechanisms along the auditory pathways from subcortical to cortical stations. In the human auditory cortex, attentional switching between low and high frequencies changes the activated locations of fMRI voxels corresponding to the attended frequencies in the primary auditory areas (Da Costa et al., 2013). In animal studies using extracellular recordings, task engagement increases or decreases responses to the behavioral target sound and reshapes the receptive field structures in A1 neurons compared to passive listening and dependent on task difficulty (Fritz et al., 2003; Atiani et al., 2009; Lee and Middlebrooks, 2011; Niwa et al., 2012; Schwartz and David, 2018). Attentive modulation was, however, larger in the belt/parabelt areas than in the core areas (Atiani et al., 2014; Niwa et al., 2015; Elgueda et al., 2019). Thus, corticothalamic feedback may be a gate of top-down modulations for complex cortical processing, supporting precise acoustic representations in A1 via the aforementioned physiological functions. Future studies are necessary to elucidate how corticothalamic projections contribute to modulate representations of sound signals in A1 and higher cortical fields as well as in MGBv.

Music Processing

Music processing studies in the human auditory system often involve a comparison between musicians and non-musicians since musical training is believed to induce plastic changes of structure and function in cortical and subcortical regions. Professional musicians tend to start receiving perceptual and motor training in their early childhood; therefore, structural and functional changes could reflect enhanced music processing. Plastic changes to musical training are indeed observed in the thalamus of musicians. Pianists have greater gray matter volume for the right thalamus (Vaquero et al., 2016), and drummers

show increased oscillatory activities between the thalamus and premotor cortex/posterior parietal cortex (Krause et al., 2010). Furthermore, the thalamus is involved in musical imagery, i.e., an evoked sensation of music without external source. The right thalamus was activated for this music imagination phenomenon (Goycoolea et al., 2007) and melody recall (Zatorre et al., 1996). In addition, the ventral thalamus was activated during melody or sentence generation (Brown et al., 2006). Pleasant feelings associated with music listening have been shown to activate the thalamus, especially the mediodorsal thalamus, which regulates emotional processing (Blood and Zatorre, 2001; Klepzig et al., 2020). Contrasting to the left-hemispheric lateralization for speech processing, these studies support right-hemispheric lateralization for music processing.

It is, however, still not well investigated how MGBv processes musical signals and what are the main functional roles of corticothalamic feedback for music processing. For example, increased activation to urban noise, including music, were observed in the MGB of schizophrenic patients, supporting a role for the thalamus in sensory gating (Tregellas et al., 2009). Subcortical auditory structures, MGB and IC, with a strong corticofugal input, showed greater synchronization and responses to pieces of music compared to a scrambled version of the music or ripple noise, suggesting top-down modulation specific to music perception as opposed to basic sound perception (Abrams et al., 2013; Jiang et al., 2013). The thalamus was more highly activated to changes of musical chords or timbre (different musical instruments), when they deviated from an expected musical flow (Koelsch et al., 2002), again pointing a role of expectation or surprise in guiding corticothalamic feedback.

Musical chords of Western music consist of multiple harmonic complex tones (>2); in other words, combinations of pitches/harmonics. If two tones are separated by 1 octave, the frequency ratio is 2:1 (“unison,” e.g., A, 220 Hz and 440 Hz), preserving harmonicity and resulting in sounding pleasant (“consonant”). A combination of two tones with the ratio of 3:2 (“perfect 5th,” e.g., C and G) maintains the regularity of spectral components; therefore, it is consonant and usually evokes positive valence emotions. The ratio of 6:5 (“minor third,” e.g., A and C) has imperfect consonance and association to sadness (a typical difference between major and minor chords). At the other extreme, if the frequency ratio of two tones is 16:15 (“minor second,” e.g., C and C#), it sounds unpleasant (“dissonant”) and can evoke all variety of negative valence emotions. Musicians showed refined representation of musical chords in the auditory brainstem response compared to non-musicians, suggesting that top-down modulation by corticofugal projections could optimize subcortical activities to efficiently process music (Lee et al., 2009). The sensation of roughness or beating has been thought to contribute to dissonance perception, and the difference of consonance and dissonance is reflected in the phase-locking in A1 to the frequency interactions of spectral components (Plomp and Levelt, 1965; Fishman et al., 2001). It seems, however, spectral regularity, i.e., harmonicity or periodicity, plays a key role while the judgment of pleasantness is also dependent on Western musical experience and cultural environment (Tramo et al., 2001; Bidelman and Krishnan, 2009; McDermott et al., 2010, 2016;

Bowling et al., 2018). A recent study suggests that harmonicity octave judgment of melodic lines, a sequence of tones, shares a similar mechanism to fused perception for octave harmonic structures (Demany et al., 2021). Given its apparent importance for perception of harmonic structure, corticothalamic feedback may help to discriminate musical chords and perceive musical melodies. Moreover, it may potentially help in the perception of rhythmic activity and its coordination between cortex and thalamus (Lee, 2013; Musacchia et al., 2014).

High-resolution human imaging and recording techniques are expected to shed light on the role of thalamocortical activity for a variety of music-based sound aspects in the near future.

Finally, speech and music processing are closely related and show overlaps of their functions. For example, musical training can improve language processing for children, adults or patients with language disorders (for review; Kraus and Slater, 2015; Coffey et al., 2017). Elementary school children (~8 years old) who received musicianship classes, which included lessons of pitch and rhythm identification, and instrumental classes for 2 years showed better ability to correctly hear out speech from speech-shaped background noise than the controls that had only 1-year training (Slater et al., 2015). Musical training improves not only pitch encoding at subcortical and cortical levels for music and speech (Schön et al., 2004; Wong et al., 2007) but also speech-in-noise performance with enhanced temporal and spectral representation of speech (Parbery-Clark et al., 2011; Kraus et al., 2014; Swaminathan et al., 2015; Zendel et al., 2015). Since non-invasive techniques used in those studies did not explicitly identify which subcortical stations were involved, it remains to be examined what neural circuits are contributing for the improvement. As one of major corticofugal projections, layer VI A1-MGBv corticothalamic feedback could enhance spectral and temporal encoding of music and speech in the thalamus although other corticofugal connections from layer V to MGB and IC also may be critical for top-down modulations and plastic changes to musical training.

CLINICAL RELEVANCE

Along with musical training, corticothalamic feedback modulation may generally reinforce experience-dependent sound processing. Although speech processing ability degrades with aging, which is associated with impaired temporal coding and altered inhibitory signaling in the auditory system (Casparly et al., 2008; Gordon-Salant et al., 2011; Richardson et al., 2013; Presacco et al., 2016), temporal precision is actually improved in the MGB of aged-animals, which could indicate a compensation for the degraded hearing abilities via top-down modulation (Kommajosyula et al., 2019, 2021; Quraishie et al., 2020). Aged-musicians showed less degraded performance on the tasks that typically decline with aging (e.g., signal-in-noise, gap or mistuning detection) (Parbery-Clark et al., 2009; Zendel and Alain, 2012), which may reflect compensation via enhanced corticothalamic feedback. Thus, understanding corticofugal

modulation may guide rehabilitation and training schemes for hearing impaired patients and therefore have clinical relevance.

CONCLUDING REMARKS

It has been proposed that corticothalamic neurons in layer VI (i) refine the receptive field of MGBv neurons, (ii) control gain of sound information flow, and (iii) increase temporal precision, by regulating the balance of excitation and inhibition. Lemniscal corticothalamic feedback can modulate the perception of frequency and harmonic structures, which is a basis for complex sound processing utilizing spectral and temporal cues and assisting auditory scene analysis of segregating concurrent speech or extracting signal from background noise. Task-related modulation was observed in human MGBv for speech processing particularly in noisy listening conditions. Although music processing in MGBv largely remains to be explored, corticothalamic feedback is expected to improve pitch perception and support musical appreciation. The future investigations will need to examine what aspects of speech and music corticothalamic feedback can modulate, but also to build up our understanding of the corticothalamic circuits including corticoreticular and intracortical pathways. The cutting-edge techniques of dissecting neural microcircuits in behaving animals and neuroimaging with finer spectral temporal resolutions in humans are expected to advance it. Ultimately, the better understanding of descending modulation may help improve rehabilitation for hearing impaired patients and musical training.

AUTHOR CONTRIBUTIONS

NH and VB both conceptualized the review. NH wrote the initial draft and prepared the figures. VB supported and reviewed the manuscript. Both authors made a substantial, direct and intellectual contribution to the review, edited and revised the manuscript, and approved the submitted version for publication.

FUNDING

This work was supported by the Japan Society for the Promotion of Science (Overseas Research Fellowship) to NH, by the Wellcome Trust UK (WT108369/Z/2015/Z), and by RNID funding to VB (S52_Bajo).

ACKNOWLEDGMENTS

We thank Christoph E. Schreiner and Andrew J. King for helpful discussions and comments during the preparation of the manuscript, the University of Iowa Musical Instrument Samples (<http://theremin.music.uiowa.edu/MIS.html>) for freely available tone samples of musical instruments, and the two reviewers for their helpful comments.

REFERENCES

- Abrams, D. A., Ryali, S., Chen, T., Chordia, P., Khouzam, A., Levitin, D. J., et al. (2013). Inter-subject synchronization of brain responses during natural music listening. *Eur. J. Neurosci.* 37, 1458–1469. doi: 10.1111/ejn.12173
- Aitkin, L. M., and Webster, W. R. (1972). Medial geniculate body of the cat: organization and responses to tonal stimuli of neurons in ventral division. *J. Neurophysiol.* 35, 365–380. doi: 10.1152/jn.1972.35.3.365
- Alain, C., Reinke, K., McDonald, K. L., Chau, W., Tam, F., Pacurar, A., et al. (2005). Left thalamo-cortical network implicated in successful speech separation and identification. *Neuroimage* 26, 592–599. doi: 10.1016/j.neuroimage.2005.02.006
- Albouy, P., Benjamin, L., Morillon, B., and Zatorre, R. J. (2020). Distinct sensitivity to spectrotemporal modulation supports brain asymmetry for speech and melody. *Science* 367, 1043–1047. doi: 10.1126/science.aaz3468
- Andersen, R. A., Knight, P. L., and Merzenich, M. M. (1980a). The thalamocortical and corticothalamic connections of AI, AII, and the anterior auditory field (AFF) in the cat: evidence for two largely segregated systems of connections. *J. Comp. Neurol.* 194, 663–701. doi: 10.1002/cne.901940312
- Andersen, R. A., Roth, G. L., Aitkin, L. M., and Merzenich, M. M. (1980b). The efferent projections of the central nucleus and the pericentral nucleus of the inferior colliculus in the cat. *J. Comp. Neurol.* 194, 649–662. doi: 10.1002/cne.901940311
- Anderson, L. A., and Linden, J. F. (2011). Physiological differences between histologically defined subdivisions in the mouse auditory thalamus. *Hear. Res.* 274, 48–60. doi: 10.1016/j.heares.2010.12.016
- Anderson, L. A., and Linden, J. F. (2016). Mind the Gap: two dissociable mechanisms of temporal processing in the auditory system. *J. Neurosci.* 36, 1977–1995. doi: 10.1523/JNEUROSCI.1652-15.2016
- Arcelli, P., Frassoni, C., Regondi, M., Biasi, S. D., and Spreafico, R. (1997). GABAergic neurons in mammalian thalamus: A marker of thalamic complexity? *Brain Res. Bull.* 42, 27–37. doi: 10.1016/S0361-9230(96)00107-4
- Arehart, K. H., Souza, P. E., Muralimanohar, R. K., and Miller, C. W. (2011). Effects of age on concurrent vowel perception in acoustic and simulated electroacoustic hearing. *J. Speech Lang. Hear. Res.* 54, 190–210.
- Assmann, P. F., and Summerfield, Q. (1994). The contribution of waveform interactions to the perception of concurrent vowels. *J. Acoust. Soc. Am.* 95, 471–484. doi: 10.1121/1.408342
- Atiani, S., David, S. V., Elgueda, D., Locastro, M., Radtke-Schuller, S., Shamma, S. A., et al. (2014). Emergent selectivity for task-relevant stimuli in higher-order auditory cortex. *Neuron* 82, 486–499. doi: 10.1016/j.neuron.2014.02.029
- Atiani, S., Elhilali, M., David, S. V., Fritz, J. B., and Shamma, S. A. (2009). Task difficulty and performance induce diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields. *Neuron* 61, 467–480. doi: 10.1016/j.neuron.2008.12.027
- Atilgan, H., Town, S. M., Wood, K. C., Jones, G. P., Maddox, R. K., Lee, A. K. C., et al. (2018). Integration of visual information in auditory cortex promotes auditory scene analysis through multisensory binding. *Neuron* 97, 640–655.e4. doi: 10.1016/j.neuron.2017.12.034
- Bajo, V. M., Nodal, F. R., Moore, D. R., and King, A. J. (2010). The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nat. Neurosci.* 13, 253–260. doi: 10.1038/nn.2466
- Bajo, V. M., Rouiller, E. M., Welker, E., Clarke, S., Villa, A. E., de Ribaupierre, Y., et al. (1995). Morphology and spatial distribution of corticothalamic terminals originating from the cat auditory cortex. *Hear. Res.* 83, 161–174. doi: 10.1016/0378-5955(94)00199-Z
- Bartlett, E., and Smith, P. (2002). Effects of paired-pulse and repetitive stimulation on neurons in the rat medial geniculate body. *Neuroscience* 113, 957–974. doi: 10.1016/S0306-4522(02)00240-3
- Bartlett, E., Stark, J., Guillery, R., and Smith, P. (2000). Comparison of the fine structure of cortical and collicular terminals in the rat medial geniculate body. *Neuroscience* 100, 811–828. doi: 10.1016/S0306-4522(00)00340-7
- Bartlett, E. L. (2013). The organization and physiology of the auditory thalamus and its role in processing acoustic features important for speech perception. *Brain Lang.* 126, 29–48. doi: 10.1016/j.bandl.2013.03.003
- Bartlett, E. L., Sadagopan, S., and Wang, X. (2011). Fine frequency tuning in monkey auditory cortex and thalamus. *J. Neurophysiol.* 106, 849–859. doi: 10.1152/jn.00559.2010
- Bartlett, E. L., and Smith, P. H. (1999). Anatomic, intrinsic, and synaptic properties of dorsal and ventral division neurons in rat medial geniculate body. *J. Neurophysiol.* 81, 1999–2016. doi: 10.1152/jn.1999.81.5.1999
- Bartlett, E. L., and Wang, X. (2007). Neural representations of temporally modulated signals in the auditory thalamus of awake primates. *J. Neurophysiol.* 97, 1005–1017. doi: 10.1152/jn.00593.2006
- Bartlett, E. L., and Wang, X. (2011). Correlation of neural response properties with auditory thalamus subdivisions in the awake marmoset. *J. Neurophysiol.* 105, 2647–2667. doi: 10.1152/jn.00238.2010
- Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., and Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature* 403, 309–312. doi: 10.1038/35002078
- Bendor, D., and Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature* 436, 1161–1165. doi: 10.1038/nature03867
- Bidelman, G. M., and Krishnan, A. (2009). Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *J. Neurosci.* 29, 13165–13171. doi: 10.1523/JNEUROSCI.3900-09.2009
- Blood, A. J., and Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc. Natl. Acad. Sci. U.S.A.* 98, 11818–11823. doi: 10.1073/pnas.191355898
- Bordi, F., and LeDoux, J. E. (1994). Response properties of single units in areas of rat auditory thalamus that project to the amygdala - II. Cells receiving convergent auditory and somatosensory inputs and cells antidromically activated by amygdala stimulation. *Exp. Brain Res.* 98, 275–286. doi: 10.1007/BF00228415
- Bortone, D. S., Olsen, S. R., and Scanziani, M. (2014). Translaminar inhibitory cells recruited by layer 6 corticothalamic neurons suppress visual cortex. *Neuron* 82, 474–485. doi: 10.1016/j.neuron.2014.02.021
- Bowling, D. L., Purves, D., and Gill, K. Z. (2018). Vocal similarity predicts the relative attraction of musical chords. *Proc. Natl. Acad. Sci. U.S.A.* 115, 216–221. doi: 10.1073/pnas.1713206115
- Bregman, A. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: The MIT Press.
- Bregman, A. S., and Campbell, J. (1971). Primary auditory stream segregation and perception of order in rapid sequences of tones. *J. Exp. Psychol.* 89, 244–249. doi: 10.1037/h0031163
- Brown, S., Martinez, M. J., and Parsons, L. M. (2006). Music and language side by side in the brain: a PET study of the generation of melodies and sentences. *Eur. J. Neurosci.* 23, 2791–2803. doi: 10.1111/j.1460-9568.2006.04785.x
- Budinger, E., Brosch, M., Scheich, H., and Mylius, J. (2013). The subcortical auditory structures in the mongolian gerbil: II. Frequency-related topography of the connections with cortical field AI. *J. Comp. Neurol.* 521, 2772–2797. doi: 10.1002/cne.23314
- Calford, M. (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. doi: 10.1523/JNEUROSCI.03-11-02350.1983
- Calford, M., and Aitkin, L. (1983). Ascending projections to the medial geniculate body of the cat: evidence for multiple, parallel auditory pathways through thalamus. *J. Neurosci.* 3, 2365–2380. doi: 10.1523/JNEUROSCI.03-11-02365.1983
- Cariani, P. A., and Delgutte, B. (1996). Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *J. Neurophysiol.* 76, 1698–1716. doi: 10.1152/jn.1996.76.3.1698
- Caspary, D. M., Ling, L., Turner, J. G., and Hughes, L. F. (2008). Inhibitory neurotransmission, plasticity and aging in the mammalian central auditory system. *J. Exp. Biol.* 211, 1781–1791. doi: 10.1242/jeb.013581
- Cedolin, L., and Delgutte, B. (2005). Pitch of complex tones: rate-place and interspike interval representations in the auditory nerve. *J. Neurophysiol.* 94, 347–362. doi: 10.1152/jn.01114.2004
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *J. Acoust. Soc. Am.* 25, 975–979. doi: 10.1121/1.1907229
- Chi, T., Gao, Y., Guyton, M. C., Ru, P., and Shamma, S. (1999). Spectro-temporal modulation transfer functions and speech intelligibility. *J. Acoust. Soc. Am.* 106, 2719–2732. doi: 10.1121/1.428100
- Christensen, T. A., Antonucci, S. M., Lockwood, J. L., Kittleson, M., and Plante, E. (2008). Cortical and subcortical contributions to the attentive processing of speech. *Neuroreport* 19, 1101–1105. doi: 10.1097/WNR.0b013e3283060a9d
- Clarke, B. A., and Lee, C. C. (2018). Inhibitory projections in the mouse auditory tectothalamic system. *Brain Sci.* 8:103. doi: 10.3390/brainsci8060103

- Clayton, K. K., Williamson, R. S., Hancock, K. E., Tasaka, G., Mizrahi, A., Hackett, T. A., et al. (2021). Auditory corticothalamic neurons are recruited by motor preparatory inputs. *Curr. Biol.* 31, 310–321.e5. doi: 10.1016/j.cub.2020.10.027
- Coffey, E. B. J., Mogilever, N. B., and Zatorre, R. J. (2017). Speech-in-noise perception in musicians: a review. *Hear. Res.* 352, 49–69. doi: 10.1016/j.heares.2017.02.006
- Colling, L. J., Noble, H. L., and Goswami, U. (2017). Neural entrainment and sensorimotor synchronization to the beat in children with developmental dyslexia: an EEG study. *Front. Neurosci.* 11:360. doi: 10.3389/fnins.2017.00360
- Conley, M., Kupersmith, A. C., and Diamond, I. T. (1991). The organization of projections from subdivisions of the auditory cortex and thalamus to the auditory sector of the thalamic reticular nucleus in Galago. *Eur. J. Neurosci.* 3, 1089–1103. doi: 10.1111/j.1460-9568.1991.tb00044.x
- Cotillon-Williams, N., Huetz, C., Hennevin, E., and Edeline, J.-M. (2008). Tonotopic control of auditory thalamus frequency tuning by reticular thalamic neurons. *J. Neurophysiol.* 99, 1137–1151. doi: 10.1152/jn.01159.2007
- Crandall, S. R., Cruikshank, S. J., and Connors, B. W. (2015). A corticothalamic switch: controlling the thalamus with dynamic synapses. *Neuron* 86, 768–782. doi: 10.1016/j.neuron.2015.03.040
- Creutzfeldt, O., Hellweg, F. C., and Schreiner, C. (1980). Thalamocortical transformation of responses to complex auditory stimuli. *Exp. Brain Res.* 39, 87–104. doi: 10.1007/BF00237072
- Culling, J. F., and Darwin, C. J. (1993). Perceptual separation of simultaneous vowels: Within and across-formant grouping by F 0. *J. Acoust. Soc. Am.* 93, 3454–3467. doi: 10.1121/1.405675
- Culling, J. F., and Darwin, C. J. (1994). Perceptual and computational separation of simultaneous vowels: cues arising from low-frequency beating. *J. Acoust. Soc. Am.* 95, 1559–1569. doi: 10.1121/1.408543
- Da Costa, S., van der Zwaag, W., Miller, L. M., Clarke, S., and Saenz, M. (2013). Tuning in to sound: frequency-selective attentional filter in human primary auditory cortex. *J. Neurosci.* 33, 1858–1863. doi: 10.1523/JNEUROSCI.4405-12.2013
- Darwin, C. J. (1984). Perceiving vowels in the presence of another sound: constraints on formant perception. *J. Acoust. Soc. Am.* 76, 1636–1647. doi: 10.1121/1.391610
- Darwin, C. J., and Ciocca, V. (1992). Grouping in pitch perception: effects of onset asynchrony and ear of presentation of a mistuned component. *J. Acoust. Soc. Am.* 91, 3381–3390. doi: 10.1121/1.402828
- Davis, M. H., and Johnsrude, I. S. (2003). Hierarchical processing in spoken language comprehension. *J. Neurosci.* 23, 3423–3431. doi: 10.1523/JNEUROSCI.23-08-03423.2003
- de Cheveigné, A. (1999). Waveform interactions and the segregation of concurrent vowels. *J. Acoust. Soc. Am.* 106, 2959–2972. doi: 10.1121/1.428115
- de Cheveigné, A. (1995). Identification of concurrent harmonic and inharmonic vowels: a test of the theory of harmonic cancellation and enhancement. *J. Acoust. Soc. Am.* 97, 3736–3748. doi: 10.1121/1.412389
- de Cheveigné, A., McAdams, S., and Marin, C. M. H. (1997). Concurrent vowel identification. II. Effects of phase, harmonicity, and task. *J. Acoust. Soc. Am.* 101, 2848–2856. doi: 10.1121/1.419476
- De La Mothe, L. A., Blumell, S., Kajikawa, Y., and Hackett, T. A. (2006). Thalamic connections of the auditory cortex in marmoset monkeys: core and medial belt regions. *J. Comp. Neurol.* 496, 72–96. doi: 10.1002/cne.20924
- Demany, L., Monteiro, G., Semal, C., Shamma, S., and Carlyon, R. P. (2021). The perception of octave pitch affinity and harmonic fusion have a common origin. *Hear. Res.* 404:108213. doi: 10.1016/j.heares.2021.108213
- Denman, D. J., and Contreras, D. (2015). Complex effects on *in vivo* visual responses by specific projections from mouse cortical layer 6 to dorsal lateral Geniculate Nucleus. *J. Neurosci.* 35, 9265–9280. doi: 10.1523/JNEUROSCI.0027-15.2015
- Deroche, M. L. D., and Culling, J. F. (2011). Narrow noise band detection in a complex masker: masking level difference due to harmonicity. *Hear. Res.* 282, 225–235. doi: 10.1016/j.heares.2011.07.005
- Diaz, B., Hintz, F., Kiebel, S. J., and von Kriegstein, K. (2012). Dysfunction of the auditory thalamus in developmental dyslexia. *Proc. Natl. Acad. Sci. U.S.A.* 109, 13841–13846. doi: 10.1073/pnas.1119828109
- Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., and Poeppel, D. (2017). Temporal modulations in speech and music. *Neurosci. Biobehav. Rev.* 81, 181–187. doi: 10.1016/j.neubiorev.2017.02.011
- Ding, N., and Simon, J. Z. (2012). Emergence of neural encoding of auditory objects while listening to competing speakers. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11854–11859. doi: 10.1073/pnas.1205381109
- Drullman, R., Festen, J. M., and Plomp, R. (1994). Effect of temporal envelope smearing on speech reception. *J. Acoust. Soc. Am.* 95, 1053–1064. doi: 10.1121/1.408467
- Edeline, J. M., Manunta, Y., Nodal, F. R., and Bajo, V. M. (1999). Do auditory responses recorded from awake animals reflect the anatomical parcellation of the auditory thalamus? *Hear. Res.* 131, 135–152. doi: 10.1016/s0378-5955(99)00026-x
- Edeline, J.-M., and Weinberger, N. M. (1991). Thalamic short-term plasticity in the auditory system: associative retuning of receptive fields in the ventral medial geniculate body. *Behav. Neurosci.* 105, 618–639. doi: 10.1037/0735-7044.105.5.618
- Elgueda, D., Duque, D., Radtke-Schuller, S., Yin, P., David, S. V., Shamma, S. A., et al. (2019). State-dependent encoding of sound and behavioral meaning in a tertiary region of the ferret auditory cortex. *Nat. Neurosci.* 22, 447–459. doi: 10.1038/s41593-018-0317-8
- Elliott, T. M., and Theunissen, F. E. (2009). The modulation transfer function for speech intelligibility. *PLoS Comput. Biol.* 5:e1000302. doi: 10.1371/journal.pcbi.1000302
- Eramudugolla, R., McAnally, K. I., Martin, R. L., Irvine, D. R. F., and Mattingley, J. B. (2008). The role of spatial location in auditory search. *Hear. Res.* 238, 139–146. doi: 10.1016/j.heares.2007.10.004
- Feng, L., and Wang, X. (2017). Harmonic template neurons in primate auditory cortex underlying complex sound processing. *Proc. Natl. Acad. Sci. U.S.A.* 114, E840–E848. doi: 10.1073/pnas.1607519114
- Fishman, Y. I., Michey, C., and Steinschneider, M. (2013). Neural representation of harmonic complex tones in primary auditory cortex of the awake monkey. *J. Neurosci.* 33, 10312–10323. doi: 10.1523/JNEUROSCI.0020-13.2013
- Fishman, Y. I., and Steinschneider, M. (2010). Neural correlates of auditory scene analysis based on inharmonicity in monkey primary auditory cortex. *J. Neurosci.* 30, 12480–12494. doi: 10.1523/JNEUROSCI.1780-10.2010
- Fishman, Y. I., Volkov, I. O., Noh, M. D., Garell, P. C., Bakken, H., Arezzo, J. C., et al. (2001). Consonance and dissonance of musical chords: neural correlates in auditory cortex of monkeys and humans. *J. Neurophysiol.* 86, 2761–2788. doi: 10.1152/jn.2001.86.6.2761
- Fitch, R. H., Tallal, P., Brown, C. P., Galaburda, A. M., and Rosen, G. D. (1994). Induced microgyria and auditory temporal processing in rats: A model for language impairment? *Cereb. Cortex* 4, 260–270. doi: 10.1093/cercor/4.3.260
- Fritz, J., Shamma, S., Elhilali, M., and Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat. Neurosci.* 6, 1216–1223. doi: 10.1038/nn1141
- Galaburda, A. M., Menard, M. T., and Rosen, G. D. (1994). Evidence for aberrant auditory anatomy in developmental dyslexia. *Proc. Natl. Acad. Sci. U.S.A.* 91, 8010–8013. doi: 10.1073/pnas.91.17.8010
- Gao, X., and Wehr, M. (2015). A coding transformation for temporally structured sounds within auditory cortical neurons. *Neuron* 86, 292–303. doi: 10.1016/j.neuron.2015.03.004
- Gent, T. C., Bandarabadi, M., Herrera, C. G., and Adamantidis, A. R. (2018). Thalamic dual control of sleep and wakefulness. *Nat. Neurosci.* 21, 974–984. doi: 10.1038/s41593-018-0164-7
- Gockel, H. E., and Carlyon, R. P. (2018). Detection of mistuning in harmonic complex tones at high frequencies. *Acta Acust. United Acust.* 104, 766–769. doi: 10.3813/AAA.919219
- 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. doi: 10.1121/1.1978261
- Gordon-Salant, S., Fitzgibbons, P. J., and Yeni-Komshian, G. H. (2011). Auditory temporal processing and aging: implications for speech understanding of older people. *Audiol. Res.* 1:e4. doi: 10.4081/audiores.2011.e4
- Goswami, U., Thomson, J., Richardson, U., Stainthorpe, R., Hughes, D., Rosen, S., et al. (2002). Amplitude envelope onsets and developmental dyslexia: a new hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 99, 10911–10916. doi: 10.1073/pnas.122368599
- Goycoolea, M. V., Mena, I., Neubauer, S. G., Levy, R. G., Fernández Grez, M., and Berger, C. G. (2007). Musical brains: a study of spontaneous and evoked musical

- sensations without external auditory stimuli. *Acta Otolaryngol.* 127, 711–721. doi: 10.1080/00016480601053057
- Greenberg, S., Carvey, H., Hitchcock, L., and Chang, S. (2003). Temporal properties of spontaneous speech—a syllable-centric perspective. *J. Phon.* 31, 465–485. doi: 10.1016/j.wocn.2003.09.005
- Griffiths, T. D., Johnsrude, I., Dean, J. L., and Green, G. G. R. (1999). A common neural substrate for the analysis of pitch and duration pattern in segmented sound? *Neuroreport* 10, 3825–3830. doi: 10.1097/00001756-199912160-00019
- Guest, D. R., and Oxenham, A. J. (2019). The role of pitch and harmonic cancellation when listening to speech in harmonic background sounds. *J. Acoust. Soc. Am.* 145, 3011–3023. doi: 10.1121/1.5102169
- Guo, W., Clause, A. R., Barth-Marion, A., and Polley, D. B. (2017). A corticothalamic circuit for dynamic switching between feature detection and discrimination. *Neuron* 95, 180–194.e5. doi: 10.1016/j.neuron.2017.05.019
- Hackett, T. A., Barkat, T. R., O'Brien, B. M. J., Hensch, T. K., and Polley, D. B. (2011). Linking topography to tonotopy in the mouse auditory thalamocortical circuit. *J. Neurosci.* 31, 2983–2995. doi: 10.1523/JNEUROSCI.5333-10.2011
- Hämäläinen, J. A., Rupp, A., Soltész, F., Szűcs, D., and Goswami, U. (2012). Reduced phase locking to slow amplitude modulation in adults with dyslexia: an MEG study. *Neuroimage* 59, 2952–2961. doi: 10.1016/j.neuroimage.2011.09.075
- Happel, M. F. K., Deliano, M., Handschuh, J., and Ohl, F. W. (2014). Dopamine-modulated recurrent corticothalamic feedback in primary sensory cortex promotes detection of behaviorally relevant stimuli. *J. Neurosci.* 34, 1234–1247. doi: 10.1523/JNEUROSCI.1990-13.2014
- Hartmann, W. M., McAdams, S., and Smith, B. K. (1990). Hearing a mistuned harmonic in an otherwise periodic complex tone. *J. Acoust. Soc. Am.* 88, 1712–1724. doi: 10.1121/1.400246
- Hasse, J. M., and Briggs, F. (2017). Corticogeniculate feedback sharpens the temporal precision and spatial resolution of visual signals in the ferret. *Proc. Natl. Acad. Sci. U.S.A.* 114, E6222–E6230. doi: 10.1073/pnas.1704524114
- He, J. (1997). Modulatory effects of regional cortical activation on the onset responses of the cat medial geniculate neurons. *J. Neurophysiol.* 77, 896–908. doi: 10.1152/jn.1997.77.2.896
- He, J. (2003). Corticofugal modulation on both ON and OFF responses in the nonlemniscal auditory thalamus of the guinea pig. *J. Neurophysiol.* 89, 367–381. doi: 10.1152/jn.00593.2002
- He, J., Yu, Y.-Q., Xiong, Y., Hashikawa, T., and Chan, Y.-S. (2002). Modulatory effect of cortical activation on the lemniscal auditory thalamus of the Guinea pig. *J. Neurophysiol.* 88, 1040–1050. doi: 10.1152/jn.2002.88.2.1040
- Herman, A. E., Galaburda, A. M., Fitch, R. H., Carter, A. R., and Rosen, G. D. (1997). Cerebral microgyria, thalamic cell size and auditory temporal processing in male and female rats. *Cereb. Cortex* 7, 453–464. doi: 10.1093/cercor/7.5.453
- Holdgraf, C. R., de Heer, W., Pasley, B., Rieger, J., Crone, N., Lin, J. J., et al. (2016). Rapid tuning shifts in human auditory cortex enhance speech intelligibility. *Nat. Commun.* 7:13654. doi: 10.1038/ncomms13654
- Holstege, G., and Collewyn, H. (1982). The efferent connections of the nucleus of the optic tract and the superior colliculus in the rabbit. *J. Comp. Neurol.* 209, 139–175. doi: 10.1002/cne.902090204
- Homma, N. Y., Bajo, V. M., Happel, M. F. K., Nodal, F. R., and King, A. J. (2016). Mistuning detection performance of ferrets in a go/no-go task. *J. Acoust. Soc. Am.* 139, EL246–EL251. doi: 10.1121/1.4954378
- Homma, N. Y., Happel, M. F. K., Nodal, F. R., Ohl, F. W., King, A. J., and Bajo, V. M. (2017). A role for auditory corticothalamic feedback in the perception of complex sounds. *J. Neurosci.* 37, 6149–6161. doi: 10.1523/JNEUROSCI.0397-17.2017
- Homma, N. Y., Hullett, P. W., Atencio, C. A., and Schreiner, C. E. (2020). Auditory cortical plasticity dependent on environmental noise statistics. *Cell Rep.* 30, 4445–4458.e5. doi: 10.1016/j.celrep.2020.03.014
- Horst, J. W., Javel, E., and Farley, G. R. (1986). Coding of spectral fine structure in the auditory nerve. I. Fourier analysis of period and interspike interval histograms. *J. Acoust. Soc. Am.* 79, 398–416. doi: 10.1121/1.393528
- Houtsma, A. J. M., and Smurzynski, J. (1990). Pitch identification and discrimination for complex tones with many harmonics. *J. Acoust. Soc. Am.* 87, 304–310. doi: 10.1121/1.399297
- Hu, B., Senatorov, V., and Mooney, D. (1994). Lemniscal and non-lemniscal synaptic transmission in rat auditory thalamus. *J. Physiol.* 479, 217–231. doi: 10.1113/jphysiol.1994.sp020290
- Huang, C. L., and Winer, J. A. (2000). Auditory thalamocortical projections in the cat: laminar and areal patterns of input. *J. Comp. Neurol.* 427, 302–331.
- Huetz, C., Philibert, B., and Edeline, J.-M. (2009). A spike-timing code for discriminating conspecific vocalizations in the thalamocortical system of anesthetized and awake guinea pigs. *J. Neurosci.* 29, 334–350. doi: 10.1523/JNEUROSCI.3269-08.2009
- Hullett, P. W., Hamilton, L. S., Mesgarani, N., Schreiner, C. E., and Chang, E. F. (2016). Human superior temporal Gyrus Organization of spectrotemporal modulation tuning derived from speech stimuli. *J. Neurosci.* 36, 2014–2026. doi: 10.1523/JNEUROSCI.1779-15.2016
- Ibrahim, B. A., Murphy, C. A., Yuditseva, G., Shinagawa, Y., Banks, M. I., and Llano, D. A. (2021). Corticothalamic gating of population auditory thalamocortical transmission in mouse. *eLife* 10:e56645. doi: 10.7554/eLife.56645
- Jiang, F., Stecker, G. C., and Fine, I. (2013). Functional localization of the auditory thalamus in individual human subjects. *Neuroimage* 78, 295–304. doi: 10.1016/j.neuroimage.2013.04.035
- Johnson, J. S., Niwa, M., O'Connor, K. N., and Sutter, M. L. (2020). Amplitude modulation encoding in the auditory cortex: comparisons between the primary and middle lateral belt regions. *J. Neurophysiol.* 124, 1706–1726. doi: 10.1152/jn.00171.2020
- Jones, E. G. (2007). *The Thalamus*, 2nd Edn. Cambridge: Cambridge University Press.
- Joris, P. X., Schreiner, C. E., and Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiol. Rev.* 84, 541–577. doi: 10.1152/physrev.00029.2003
- Kalluri, S., Depireux, D. A., and Shamma, S. A. (2008). Perception and cortical neural coding of harmonic fusion in ferrets. *J. Acoust. Soc. Am.* 123, 2701–2716. doi: 10.1121/1.2902178
- Kim, J., Matney, C. J., Blankenship, A., Hestrin, S., and Brown, S. P. (2014). Layer 6 Corticothalamic Neurons Activate a Cortical Output Layer, Layer 5a. *J. Neurosci.* 34, 9656–9664. doi: 10.1523/JNEUROSCI.1325-14.2014
- Kimura, A., Donishi, T., Okamoto, K., and Tamai, Y. (2005). Topography of projections from the primary and non-primary auditory cortical areas to the medial geniculate body and thalamic reticular nucleus in the rat. *Neuroscience* 135, 1325–1342. doi: 10.1016/j.neuroscience.2005.06.089
- Kimura, A., Donishi, T., Sakoda, T., Hazama, M., and Tamai, Y. (2003). Auditory thalamic nuclei projections to the temporal cortex in the rat. *Neuroscience* 117, 1003–1016. doi: 10.1016/s0306-4522(02)00949-1
- Klatt, D. H. (1982). “Prediction of perceived phonetic distance from critical-band spectra: a first step,” in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, Paris, 1278–1281. doi: 10.1109/ICASSP.1982.1171512
- Klepzig, K., Horn, U., König, J., Holtz, K., Wendt, J., Hamm, A. O., et al. (2020). Brain imaging of chill reactions to pleasant and unpleasant sounds. *Behav. Brain Res.* 380:112417. doi: 10.1016/j.bbr.2019.112417
- Klinge, A., and Klump, G. M. (2009). Frequency difference limens of pure tones and harmonics within complex stimuli in Mongolian gerbils and humans. *J. Acoust. Soc. Am.* 125, 304–314. doi: 10.1121/1.3021315
- Koelsch, S., Gunter, T. C., Cramon, D. Y., Zysset, S., Lohmann, G., and Friederici, A. D. (2002). Bach speaks: a cortical “language-network” serves the processing of music. *Neuroimage* 17, 956–966. doi: 10.1006/nimg.2002.1154
- Kommajosyula, S. P., Bartlett, E. L., Cai, R., Ling, L., and Caspary, D. (2021). Corticothalamic projections deliver enhanced-responses to medial geniculate body as a function of the temporal reliability of the stimulus. *bioRxiv* [Preprint]. doi: 10.1101/2021.05.07.443156
- Kommajosyula, S. P., Cai, R., Bartlett, E., and Caspary, D. M. (2019). Top-down or bottom up: decreased stimulus salience increases responses to predictable stimuli of auditory thalamic neurons. *J. Physiol.* 597, 2767–2784. doi: 10.1113/JP277450
- Kraus, N., McGee, T. J., Carrell, T. D., Zecker, S. G., Nicol, T. G., and Koch, D. B. (1996). Auditory neurophysiologic responses and discrimination deficits in children with learning problems. *Science* 273, 971–973. doi: 10.1126/science.273.5277.971

- Kraus, N., and Slater, J. (2015). Music and language: relations and disconnections. *Handb. Clin. Neurol.* 129, 207–222. doi: 10.1016/B978-0-444-62630-1.00012-3
- Kraus, N., Slater, J., Thompson, E. C., Hornickel, J., Strait, D. L., Nicol, T., et al. (2014). Music enrichment programs improve the neural encoding of speech in at-risk children. *J. Neurosci.* 34, 11913–11918. doi: 10.1523/JNEUROSCI.1881-14.2014
- Krause, V., Schnitzler, A., and Pollok, B. (2010). Functional network interactions during sensorimotor synchronization in musicians and non-musicians. *Neuroimage* 52, 245–251. doi: 10.1016/j.neuroimage.2010.03.081
- Kudo, M., and Niimi, K. (1980). Ascending projections of the inferior colliculus in the cat: an autoradiographic study. *J. Comp. Neurol.* 191, 545–556. doi: 10.1002/cne.901910403
- Lam, Y.-W., and Sherman, S. M. (2010). Functional organization of the somatosensory cortical layer 6 feedback to the thalamus. *Cereb. Cortex* 20, 13–24. doi: 10.1093/cercor/bhp077
- Las, L., Stern, E. A., and Nelken, I. (2005). Representation of tone in fluctuating maskers in the ascending auditory system. *J. Neurosci.* 25, 1503–1513. doi: 10.1523/JNEUROSCI.4007-04.2005
- LeDoux, J. E., Farb, C. R., and Romanski, L. M. (1991). Overlapping projections to the amygdala and striatum from auditory processing areas of the thalamus and cortex. *Neurosci. Lett.* 134, 139–144. doi: 10.1016/0304-3940(91)90526-Y
- LeDoux, J. E., Ruggiero, D. A., and Reis, D. J. (1985). Projections to the subcortical forebrain from anatomically defined regions of the medial geniculate body in the rat. *J. Comp. Neurol.* 242, 182–213. doi: 10.1002/cne.902420204
- Lee, C. C. (2013). Thalamic and cortical pathways supporting auditory processing. *Brain Lang.* 126, 22–28. doi: 10.1016/j.bandl.2012.05.004
- Lee, C. C. (2015). Exploring functions for the non-lemniscal auditory thalamus. *Front. Neural Circuits* 9:69. doi: 10.3389/fncir.2015.00069
- Lee, C. C., Lam, Y.-W., and Sherman, S. M. (2012). Intracortical convergence of layer 6 neurons. *Neuroreport* 23, 736–740. doi: 10.1097/WNR.0b013e328356c1aa
- Lee, C.-C., and Middlebrooks, J. C. (2011). Auditory cortex spatial sensitivity sharpens during task performance. *Nat. Neurosci.* 14, 108–114. doi: 10.1038/nn.2713
- Lee, C. C., and Sherman, S. M. (2009). Modulator property of the intrinsic cortical projection from layer 6 to layer 4. *Front. Syst. Neurosci.* 3:3. doi: 10.3389/neuro.06.003.2009
- Lee, C. C., and Winer, J. A. (2008). Connections of cat auditory cortex: I. Thalamocortical system. *J. Comp. Neurol.* 507, 1879–1900. doi: 10.1002/cne.21611
- Lee, K. M., Skoe, E., Kraus, N., and Ashley, R. (2009). Selective Subcortical Enhancement of Musical Intervals in Musicians. *J. Neurosci.* 29, 5832–5840. doi: 10.1523/JNEUROSCI.6133-08.2009
- Leonard, M. K., and Chang, E. F. (2014). Dynamic speech representations in the human temporal lobe. *Trends Cogn. Sci.* 18, 472–479. doi: 10.1016/j.tics.2014.05.001
- Li, L., and Ebner, F. F. (2007). Cortical Modulation of Spatial and Angular Tuning Maps in the Rat Thalamus. *J. Neurosci.* 27, 167–179. doi: 10.1523/JNEUROSCI.4165-06.2007
- Licklider, J. C. R. (1951). A duplex theory of pitch perception. *Experientia* 7, 128–134. doi: 10.1007/BF02156143
- Lin, J. Y., and Hartmann, W. M. (1998). The pitch of a mistuned harmonic: evidence for a template model. *J. Acoust. Soc. Am.* 103, 2608–2617. doi: 10.1121/1.422781
- Llano, D. A., and Sherman, S. M. (2008). Evidence for nonreciprocal organization of the mouse auditory thalamocortical-corticothalamic projection systems. *J. Comp. Neurol.* 507, 1209–1227. doi: 10.1002/cne.21602
- Llinás, R. R., and Steriade, M. (2006). Bursting of Thalamic Neurons and States of Vigilance. *J. Neurophysiol.* 95, 3297–3308. doi: 10.1152/jn.00166.2006
- Lohr, B., and Dooling, R. J. (1998). Detection of changes in timbre and harmonicity in complex sounds by zebra finches (*Taeniopygia guttata*) and budgerigars (*Melopsittacus undulatus*). *J. Comp. Psychol.* 112, 36–47. doi: 10.1037/0735-7036.112.1.36
- Lohse, M., Bajo, V. M., and King, A. J. (2019). Development, organization and plasticity of auditory circuits: lessons from a cherished colleague. *Eur. J. Neurosci.* 49, 990–1004. doi: 10.1111/ejn.13979
- Lohse, M., Bajo, V. M., King, A. J., and Willmore, B. D. B. (2020). Neural circuits underlying auditory contrast gain control and their perceptual implications. *Nat. Commun.* 11:324. doi: 10.1038/s41467-019-14163-5
- Luo, F., Liu, X., Wang, C., and Yan, J. (2011). The pedunclopontine tegmental nucleus: a second cholinergic source for frequency-specific auditory plasticity. *J. Neurophysiol.* 105, 107–116. doi: 10.1152/jn.00546.2010
- Luo, F., Wang, Q., Kashani, A., and Yan, J. (2008). Corticofugal modulation of initial sound processing in the brain. *J. Neurosci.* 28, 11615–11621. doi: 10.1523/JNEUROSCI.3972-08.2008
- Ma, X., and Suga, N. (2007). Multiparametric corticofugal modulation of collicular duration-tuned neurons: modulation in the amplitude domain. *J. Neurophysiol.* 97, 3722–3730. doi: 10.1152/jn.01268.2006
- Maddox, R. K., Atilgan, H., Bizley, J. K., and Lee, A. K. C. (2015). Auditory selective attention is enhanced by a task-irrelevant temporally coherent visual stimulus in human listeners. *eLife* 4:e04995. doi: 10.7554/eLife.04995
- Massaux, A., Dutrieux, G., Cotillon-Williams, N., Manunta, Y., and Edeline, J.-M. (2004). Auditory thalamus bursts in anesthetized and non-anesthetized states: contribution to functional properties. *J. Neurophysiol.* 91, 2117–2134. doi: 10.1152/jn.00970.2003
- McCormick, D. A., and Feese, H. R. (1990). Functional implications of burst firing and single spike activity in lateral geniculate relay neurons. *Neuroscience* 39, 103–113. doi: 10.1016/0306-4522(90)90225-S
- McDermott, J. H., Lehr, A. J., and Oxenham, A. J. (2010). Individual differences reveal the basis of consonance. *Curr. Biol.* 20, 1035–1041. doi: 10.1016/j.cub.2010.04.019
- McDermott, J. H., and Oxenham, A. J. (2008). Music perception, pitch, and the auditory system. *Curr. Opin. Neurobiol.* 18, 452–463. doi: 10.1016/j.conb.2008.09.005
- McDermott, J. H., Schultz, A. F., Undurraga, E. A., and Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature* 535, 547–550. doi: 10.1038/nature18635
- McDermott, J. H., and Simoncelli, E. P. (2011). Sound texture perception via statistics of the auditory periphery: evidence from sound synthesis. *Neuron* 71, 926–940. doi: 10.1016/j.neuron.2011.06.032
- McPherson, M. J., and McDermott, J. H. (2020). Time-dependent discrimination advantages for harmonic sounds suggest efficient coding for memory. *Proc. Natl. Acad. Sci. U.S.A.* 117, 32169–32180. doi: 10.1073/pnas.2008956117
- Mease, R. A., Krieger, P., and Groh, A. (2014). Cortical control of adaptation and sensory relay mode in the thalamus. *Proc. Natl. Acad. Sci. U.S.A.* 111, 6798–6803. doi: 10.1073/pnas.1318665111
- 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. doi: 10.1121/1.400725
- 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.* 89:2883. doi: 10.1121/1.400726
- Menell, P., McAnally, K. I., and Stein, J. F. (1999). Psychophysical Sensitivity and Physiological Response to Amplitude Modulation in Adult Dyslexic Listeners. *J. Speech Lang. Hear. Res.* 42, 797–803. doi: 10.1044/jslhr.4204.797
- Mesgarani, N., and Chang, E. F. (2012). Selective cortical representation of attended speaker in multi-talker speech perception. *Nature* 485, 233–236. doi: 10.1038/nature11020
- Mesgarani, N., Cheung, C., Johnson, K., and Chang, E. F. (2014). Phonetic Feature Encoding in Human Superior Temporal Gyrus. *Science* 343, 1006–1010. doi: 10.1126/science.1245994
- Micheyl, C., and Oxenham, A. J. (2010). Pitch, harmonicity and concurrent sound segregation: psychoacoustical and neurophysiological findings. *Hear. Res.* 266, 36–51. doi: 10.1016/j.heares.2009.09.012
- Mihai, P. G., Moerel, M., de Martino, F., Trampel, R., Kiebel, S., and von Kriegstein, K. (2019). Modulation of tonotopic ventral medial geniculate body is behaviorally relevant for speech recognition. *Elife* 8:e44837. doi: 10.7554/eLife.44837
- Mihai, P. G., Tschentscher, N., and von Kriegstein, K. (2021). Modulation of the primary auditory thalamus when recognising speech with background noise. *J. Neurosci.* doi: 10.1523/JNEUROSCI.2902-20.2021 [Epub ahead of print].
- Miller, G. A., and Heise, G. A. (1950). The Trill Threshold. *J. Acoust. Soc. Am.* 22:637. doi: 10.1121/1.1906663

- Miller, L. M., Escabi, M. A., Read, H. L., and Schreiner, C. E. (2002). Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex. *J. Neurophysiol.* 87, 516–527. doi: 10.1152/jn.00395.2001
- Miller, L. M., and Schreiner, C. E. (2000). Stimulus-based state control in the thalamocortical system. *J. Neurosci.* 20, 7011–7016. doi: 10.1523/JNEUROSCI.20-18-07011.2000
- Moerel, M., De Martino, F., Uğurbil, K., Yacoub, E., and Formisano, E. (2015). Processing of frequency and location in human subcortical auditory structures. *Sci. Rep.* 5:17048. doi: 10.1038/srep17048
- Moore, B. C., Glasberg, B. R., and Peters, R. W. (1986). Thresholds for hearing mistuned partials as separate tones in harmonic complexes. *J. Acoust. Soc. Am.* 80, 479–483. doi: 10.1121/1.394043
- Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (1985). Thresholds for the detection of inharmonicity in complex tones. *J. Acoust. Soc. Am.* 77, 1861–1867. doi: 10.1121/1.391937
- Morel, A., and Imig, T. J. (1987). Thalamic projections to fields A, AI, P, and VP in the cat auditory cortex. *J. Comp. Neurol.* 265, 119–144. doi: 10.1002/cne.902650109
- Morel, A., Rouiller, E., de Ribaupierre, Y., and de Ribaupierre, F. (1987). Tonotopic organization in the medial geniculate body (MGB) of lightly anesthetized cats. *Exp. Brain Res.* 69, 24–42. doi: 10.1007/BF00247026
- Müller, F., Niso, G., Samiee, S., Pfitz, M., Baillet, S., and Kupers, R. (2019). A thalamocortical pathway for fast rerouting of tactile information to occipital cortex in congenital blindness. *Nat. Commun.* 10:5154. doi: 10.1038/s41467-019-13173-7
- Musacchia, G., Large, E. W., and Schroeder, C. E. (2014). Thalamocortical mechanisms for integrating musical tone and rhythm. *Hear. Res.* 308, 50–59. doi: 10.1016/j.heares.2013.09.017
- Nakamoto, K. T., Jones, S. J., and Palmer, A. R. (2008). Descending projections from auditory cortex modulate sensitivity in the midbrain to cues for spatial position. *J. Neurophysiol.* 99, 2347–2356. doi: 10.1152/jn.01326.2007
- Nelson, S. L., Kong, L., Liu, X., and Yan, J. (2015). Auditory cortex directs the input-specific remodeling of thalamus. *Hear. Res.* 328, 1–7. doi: 10.1016/j.heares.2015.06.016
- Niwa, M., Johnson, J. S., O'Connor, K. N., and Sutter, M. L. (2012). Active engagement improves primary auditory cortical neurons' ability to discriminate temporal modulation. *J. Neurosci.* 32, 9323–9334. doi: 10.1523/JNEUROSCI.5832-11.2012
- Niwa, M., O'Connor, K. N., Engall, E., Johnson, J. S., and Sutter, M. L. (2015). Hierarchical effects of task engagement on amplitude modulation encoding in auditory cortex. *J. Neurophysiol.* 113, 307–327. doi: 10.1152/jn.00458.2013
- Norman-Haignere, S., Kanwisher, N. G., and McDermott, J. H. (2015). Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. *Neuron* 88, 1281–1296. doi: 10.1016/j.neuron.2015.11.035
- Ohl, F. W., Wetzel, W., Wagner, T., Rech, A., and Scheich, H. (1999). Bilateral ablation of auditory cortex in Mongolian gerbil affects discrimination of frequency modulated tones but not of pure tones. *Learn. Mem.* 6, 347–362. doi: 10.1101/lm.6.4.347
- Ojima, H. (1994). Terminal morphology and distribution of corticothalamic fibers originating from layers 5 and 6 of cat primary auditory cortex. *Cereb. Cortex* 4, 646–663. doi: 10.1093/cercor/4.6.646
- Olsen, S. R., Bortone, D. S., Adesnik, H., and Scanziani, M. (2012). Gain control by layer six in cortical circuits of vision. *Nature* 483, 47–52. doi: 10.1038/nature10835
- Ono, K., Kudoh, M., and Shibuki, K. (2006). Roles of the auditory cortex in discrimination learning by rats. *Eur. J. Neurosci.* 23, 1623–1632. doi: 10.1111/j.1460-9568.2006.04695.x
- Osmanski, M. S., Song, X., and Wang, X. (2013). The role of harmonic resolvability in pitch perception in a vocal nonhuman primate, the common marmoset (*Callithrix jacchus*). *J. Neurosci.* 33, 9161–9168. doi: 10.1523/JNEUROSCI.0066-13.2013
- O'Sullivan, J., Herrero, J., Smith, E., Schevon, C., McKhann, G. M., Sheth, S. A., et al. (2019). Hierarchical encoding of attended auditory objects in multi-talker speech perception. *Neuron* 104, 1195–1209.e3. doi: 10.1016/j.neuron.2019.09.007
- Ozernov-Palchik, O., and Gaab, N. (2016). Tackling the “dyslexia paradox”: reading brain and behavior for early markers of developmental dyslexia. *Wiley Interdiscip. Rev. Cogn. Sci.* 7, 156–176. doi: 10.1002/wcs.1383
- Parbery-Clark, A., Skoe, E., and Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *J. Neurosci.* 29, 14100–14107. doi: 10.1523/JNEUROSCI.3256-09.2009
- Parbery-Clark, A., Strait, D. L., and Kraus, N. (2011). Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia* 49, 3338–3345. doi: 10.1016/j.neuropsychologia.2011.08.007
- 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. doi: 10.1016/S0896-6273(02)01060-7
- Peiffer, A. M., Rosen, G. D., and Fitch, R. H. (2002). Rapid auditory processing and MGN morphology in microgyric rats reared in varied acoustic environments. *Dev. Brain Res.* 138, 187–193. doi: 10.1016/S0165-3806(02)00472-8
- Pellegrino, F., Coupe, C., and Marsico, E. (2011). A cross-language perspective on speech information rate. *Language (Baltim)* 87, 539–558. doi: 10.2307/23011654
- 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. doi: 10.1523/JNEUROSCI.0383-04.2004
- Peruzzi, D., Bartlett, E., Smith, P. H., and Oliver, D. L. (1997). A monosynaptic GABAergic input from the inferior colliculus to the medial geniculate body in rat. *J. Neurosci.* 17, 3766–3777. doi: 10.1523/JNEUROSCI.17-10-03766.1997
- Philibert, B., Laudanski, J., and Edeline, J. M. (2005). Auditory thalamus responses to guinea-pig vocalizations: a comparison between rat and guinea-pig. *Hear. Res.* 209, 97–103. doi: 10.1016/j.heares.2005.07.004
- Pizzo, F., Roehri, N., Medina Villalon, S., Trébuchon, A., Chen, S., Lagarde, S., et al. (2019). Deep brain activities can be detected with magnetoencephalography. *Nat. Commun.* 10:971. doi: 10.1038/s41467-019-08665-5
- Plomp, R., and Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *J. Acoust. Soc. Am.* 38, 548–560. doi: 10.1121/1.1909741
- Poeppl, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as ‘asymmetric sampling in time.’ *Speech Commun.* 41, 245–255. doi: 10.1016/S0167-6393(02)00107-3
- Polley, D. B., Read, H. L., Storace, D. A., and Merzenich, M. M. (2007). Multiparametric auditory receptive field organization across five cortical fields in the albino rat. *J. Neurophysiol.* 97, 3621–3638. doi: 10.1152/jn.01298.2006
- Popham, S., Boebinger, D., Ellis, D. P. W., Kawahara, H., and McDermott, J. H. (2018). Inharmonic speech reveals the role of harmonicity in the cocktail party problem. *Nat. Commun.* 9:2122. doi: 10.1038/s41467-018-04551-8
- Presacco, A., Simon, J. Z., and Anderson, S. (2016). Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. *J. Neurophysiol.* 116, 2346–2355. doi: 10.1152/jn.00372.2016
- Preuss, A., and Müller-Preuss, P. (1990). Processing of amplitude modulated sounds in the medial geniculate body of squirrel monkeys. *Exp. Brain Res.* 79, 207–211. doi: 10.1007/BF00228890
- Quraishe, S., Newman, T., and Anderson, L. (2020). Auditory temporal acuity improves with age in the male mouse auditory thalamus: A role for perineuronal nets? *J. Neurosci. Res.* 98, 1780–1799. doi: 10.1002/jnr.24537
- Rabinowitz, N. C., Willmore, B. D. B., Schnupp, J. W. H., and King, A. J. (2012). Spectrotemporal contrast kernels for neurons in primary auditory cortex. *J. Neurosci.* 32, 11271–11284. doi: 10.1523/JNEUROSCI.1715-12.2012
- Read, H. L., Miller, L. M., Schreiner, C. E., and Winer, J. A. (2008). Two thalamic pathways to primary auditory cortex. *Neuroscience* 152, 151–159. doi: 10.1016/j.neuroscience.2007.11.026
- Read, H. L., Nauen, D. W., Escabi, M. A., Miller, L. M., Schreiner, C. E., and Winer, J. A. (2011). Distinct core thalamocortical pathways to central and dorsal primary auditory cortex. *Hear. Res.* 274, 95–104. doi: 10.1016/j.heares.2010.11.010
- Redies, H., Brandner, S., and Creutzfeldt, O. D. (1989). Anatomy of the auditory thalamocortical system of the guinea pig. *J. Comp. Neurol.* 282, 489–511. doi: 10.1002/cne.902820403
- Richardson, B. D., Ling, L. L., Uteshev, V. V., and Caspary, D. M. (2013). Reduced GABA receptor-mediated tonic inhibition in aged rat auditory thalamus. *J. Neurosci.* 33, 1218–1227. doi: 10.1523/JNEUROSCI.3277-12.2013
- Roberts, B., and Bailey, P. J. (1993). Spectral pattern and the perceptual fusion of harmonics. II. A special status for added components? *J. Acoust. Soc. Am.* 94, 3165–3177. doi: 10.1121/1.407222

- Roberts, B., and Bailey, P. J. (1996). Spectral regularity as a factor distinct from harmonic relations in auditory grouping. *J. Exp. Psychol. Hum. Percept. Perform.* 22, 604–614. doi: 10.1037//0096-1523.22.3.604
- Roberts, B., and Bregman, A. S. (1991). Effects of the pattern of spectral spacing on the perceptual fusion of harmonics. *J. Acoust. Soc. Am.* 90:3050. doi: 10.1121/1.401779
- Roberts, B., and Brunstrom, J. M. (1998). Perceptual segregation and pitch shifts of mistuned components in harmonic complexes and in regular inharmonic complexes. *J. Acoust. Soc. Am.* 104, 2326–2338. doi: 10.1121/1.423771
- Roberts, B., and Brunstrom, J. M. (2001). Perceptual fusion and fragmentation of complex tones made inharmonic by applying different degrees of frequency shift and spectral stretch. *J. Acoust. Soc. Am.* 110, 2479–2490. doi: 10.1121/1.1410965
- Rodrigues-Dagaëff, C., Simm, G., De Ribaupierre, Y., Villa, A., De Ribaupierre, F., and Rouiller, E. M. (1989). Functional organization of the ventral division of the medial geniculate body of the cat: evidence for a rostro-caudal gradient of response properties and cortical projections. *Hear. Res.* 39, 103–125. doi: 10.1016/0378-5955(89)90085-3
- Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 336, 367–373. doi: 10.1098/rstb.1992.0070
- Rouiller, E., de Ribaupierre, Y., Toros-Morel, A., and de Ribaupierre, F. (1981). Neural coding of repetitive clicks in the medial geniculate body of cat. *Hear. Res.* 5, 81–100. doi: 10.1016/0378-5955(81)90028-9
- Rouiller, E. M., Colomb, E., Capt, M., and de Ribaupierre, F. (1985). Projections of the reticular complex of the thalamus onto physiologically characterized regions of the medial geniculate body. *Neurosci. Lett.* 53, 227–232. doi: 10.1016/0304-3940(85)90190-9
- Rouiller, E. M., and de Ribaupierre, F. (1985). Origin of afferents to physiologically defined regions of the medial geniculate body of the cat: ventral and dorsal divisions. *Hear. Res.* 19, 97–114. doi: 10.1016/0378-5955(85)90114-5
- Rouiller, E. M., Rodrigues-Dagaëff, C., Simm, G., De Ribaupierre, Y., Villa, A., and De Ribaupierre, F. (1989). Functional organization of the medial division of the medial geniculate body of the cat: tonotopic organization, spatial distribution of response properties and cortical connections. *Hear. Res.* 39, 127–142. doi: 10.1016/0378-5955(89)90086-5
- Rouiller, E. M., and Welker, E. (1991). Morphology of corticothalamic terminals arising from the auditory cortex of the rat: a Phaseolus vulgaris-leucoagglutinin (PHA-L) tracing study. *Hear. Res.* 56, 179–190. doi: 10.1016/0378-5955(91)90168-9
- Rouiller, E. M., and Welker, E. (2000). A comparative analysis of the morphology of corticothalamic projections in mammals. *Brain Res. Bull.* 53, 727–741. doi: 10.1016/S0361-9230(00)00364-6
- Ryugo, D. K., and Weinberger, N. M. (1976). Corticofugal modulation of the medial geniculate body. *Exp. Neurol.* 39, 377–391.
- Saldeitis, K., Happel, M. F. K., Ohl, F. W., Scheich, H., and Budinger, E. (2014). Anatomy of the auditory thalamocortical system in the mongolian gerbil: nuclear origins and cortical field-, layer-, and frequency-specificities. *J. Comp. Neurol.* 522, 2397–2430. doi: 10.1002/cne.23540
- Saldeitis, K., Jeschke, M., Budinger, E., Ohl, F. W., and Happel, M. F. K. (2021). Laser-induced apoptosis of corticothalamic neurons in layer VI of auditory cortex impact on cortical frequency processing. *Front. Neural Circuits* 15:659280. doi: 10.3389/fncir.2021.659280
- Salvi, R. J., Lockwood, A. H., Frisina, R. D., Coad, M. L., Wack, D. S., and Frisina, D. R. (2002). PET imaging of the normal human auditory system: responses to speech in quiet and in background noise. *Hear. Res.* 170, 96–106. doi: 10.1016/S0378-5955(02)00386-6
- Schall, S., Kiebel, S. J., Maess, B., and von Kriegstein, K. (2015). Voice identity recognition: functional division of the right STS and its behavioral relevance. *J. Cogn. Neurosci.* 27, 280–291. doi: 10.1162/jocn_a_00707
- Schneider, D. M., and Woolley, S. M. N. (2013). Sparse and background-invariant coding of vocalizations in auditory scenes. *Neuron* 79, 141–152. doi: 10.1016/j.neuron.2013.04.038
- Schön, D., Magne, C., and Besson, M. (2004). The music of speech: music training facilitates pitch processing in both music and language. *Psychophysiology* 41, 341–349. doi: 10.1111/1469-8986.00172.x
- Schwartz, Z. P., and David, S. V. (2018). Focal suppression of distractor sounds by selective attention in auditory cortex. *Cereb. Cortex* 28, 323–339. doi: 10.1093/cercor/bhx288
- Schwarz, D. W. F., and Tomlinson, R. W. W. (1990). Spectral response patterns of auditory cortex neurons to harmonic complex tones in alert monkey (*Macaca mulatta*). *J. Neurophysiol.* 64, 282–298. doi: 10.1152/jn.1990.64.1.282
- Scott, B. H., Malone, B. J., and Semple, M. N. (2011). Transformation of temporal processing across auditory cortex of awake macaques. *J. Neurophysiol.* 105, 712–730. doi: 10.1152/jn.01120.2009
- Scott, S. K. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123, 2400–2406. doi: 10.1093/brain/123.12.2400
- Shackleton, T. M., and Carlyon, R. P. (1994). The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination. *J. Acoust. Soc. Am.* 95, 3529–3540. doi: 10.1121/1.409970
- Shamma, S. A., Elhilali, M., and Micheyl, C. (2011). Temporal coherence and attention in auditory scene analysis. *Trends Neurosci.* 34, 114–123. doi: 10.1016/j.tins.2010.11.002
- Shamma, S., and Lorenzi, C. (2013). On the balance of envelope and temporal fine structure in the encoding of speech in the early auditory system. *J. Acoust. Soc. Am.* 133, 2818–2833. doi: 10.1121/1.4795783
- Sherman, S. M. (2001). Tonic and burst firing: dual modes of thalamocortical relay. *Trends Neurosci.* 24, 122–126. doi: 10.1016/S0166-2236(00)01714-8
- Sherman, S. M., and Guillery, R. W. (1998). On the actions that one nerve cell can have on another: distinguishing “drivers” from “modulators”. *Proc. Natl. Acad. Sci. U.S.A.* 95, 7121–7126. doi: 10.1073/pnas.95.12.7121
- Sherman, S. M., and Guillery, R. W. (2011). Distinct functions for direct and transthalamic corticocortical connections. *J. Neurophysiol.* 106, 1068–1077. doi: 10.1152/jn.00429.2011
- Sherman, S. M., and Guillery, R. W. (2013). *Functional Connections of Cortical Areas: A New View from the Thalamus*, 1st Edn. Cambridge, MA: The MIT Press.
- Sinex, D. G., Guzik, H., Li, H., Henderson Sabes, J., and Sabes, J. H. (2003). Responses of auditory nerve fibers to harmonic and mistuned complex tones. *Hear. Res.* 182, 130–139. doi: 10.1016/S0378-5955(03)00189-8
- Sinex, D. G., Henderson Sabes, J., and Li, H. (2002). Responses of inferior colliculus neurons to harmonic and mistuned complex tones. *Hear. Res.* 168, 150–162. doi: 10.1016/S0378-5955(02)00366-0
- Sinex, D. G., Li, H., and Velenovsky, D. S. (2005). Prevalence of stereotypical responses to mistuned complex tones in the inferior colliculus. *J. Neurophysiol.* 94, 3523–3537. doi: 10.1152/jn.01194.2004
- Slater, J., Skoe, E., Strait, D. L., O’Connell, S., Thompson, E., and Kraus, N. (2015). Music training improves speech-in-noise perception: longitudinal evidence from a community-based music program. *Behav. Brain Res.* 291, 244–252. doi: 10.1016/j.bbr.2015.05.026
- Smith, P. H., Uhlrich, D. J., Manning, K. A., and Banks, M. I. (2012). Thalamocortical projections to rat auditory cortex from the ventral and dorsal divisions of the medial geniculate nucleus. *J. Comp. Neurol.* 520, 34–51. doi: 10.1002/cne.22682
- So, N. L. T., Edwards, J. A., and Woolley, S. M. N. (2020). Auditory selectivity for spectral contrast in cortical neurons and behavior. *J. Neurosci.* 40, 1015–1027. doi: 10.1523/JNEUROSCI.1200-19.2019
- Souffi, S., Lorenzi, C., Varnet, L., Huetz, C., and Edeline, J.-M. (2020). Noise-sensitive but more precise subcortical representations coexist with robust cortical encoding of natural vocalizations. *J. Neurosci.* 40, 5228–5246. doi: 10.1523/JNEUROSCI.2731-19.2020
- Souffi, S., Nodal, F. R., Bajo, V. M., and Edeline, J.-M. (2021). When and how does the auditory cortex influence subcortical auditory structures? New insights about the roles of descending cortical projections. *Front. Neurosci.* 15. doi: 10.3389/fnins.2021.690223
- Stein, J. (2019). The current status of the magnocellular theory of developmental dyslexia. *Neuropsychologia* 130, 66–77. doi: 10.1016/j.neuropsychologia.2018.03.022
- Stein, J. F., and McAnally, K. (1995). Auditory temporal processing in developmental dyslexics. *Irish J. Psychol.* 16, 220–228. doi: 10.1080/03033910.1995.10558058
- Steinmetzger, K., and Rosen, S. (2015). The role of periodicity in perceiving speech in quiet and in background noise. *J. Acoust. Soc. Am.* 138, 3586–3599. doi: 10.1121/1.4936945

- Steinschneider, M., Reser, D. H., Fishman, Y. I., Schroeder, C. E., and Arezzo, J. C. (1998). Click train encoding in primary auditory cortex of the awake monkey: evidence for two mechanisms subserving pitch perception. *J. Acoust. Soc. Am.* 104, 2935–2955. doi: 10.1121/1.423877
- Steriade, M., McCormick, D., and Sejnowski, T. (1993). Thalamocortical oscillations in the sleeping and aroused brain. *Science* 262, 679–685. doi: 10.1126/science.8235588
- Storace, D. A., Higgins, N. C., and Read, H. L. (2010). Thalamic label patterns suggest primary and ventral auditory fields are distinct core regions. *J. Comp. Neurol.* 518, 1630–1646. doi: 10.1002/cne.22345
- Storace, D. A., Higgins, N. C., and Read, H. L. (2011). Thalamocortical pathway specialization for sound frequency resolution. *J. Comp. Neurol.* 519, 177–193. doi: 10.1002/cne.22501
- Su, Y., and Delgutte, B. (2019). Pitch of harmonic complex tones: rate and temporal coding of envelope repetition rate in inferior colliculus of unanesthetized rabbits. *J. Neurophysiol.* 122, 2468–2485. doi: 10.1152/jn.00512.2019
- Su, Y., and Delgutte, B. (2020). Robust rate-place coding of resolved components in harmonic and inharmonic complex tones in auditory midbrain. *J. Neurosci.* 40, 2080–2093. doi: 10.1523/JNEUROSCI.2337-19.2020
- Sumner, C. J., Wells, T. T., Bergevin, C., Sollini, J., Kreft, H. A., Palmer, A. R., et al. (2018). Mammalian behavior and physiology converge to confirm sharper cochlear tuning in humans. *Proc. Natl. Acad. Sci. U.S.A.* 115, 11322–11326. doi: 10.1073/pnas.1810766115
- Suta, D., Popelár, J., Kvasnák, E., and Syka, J. (2007). Representation of species-specific vocalizations in the medial geniculate body of the guinea pig. *Exp. Brain Res.* 183, 377–388. doi: 10.1007/s00221-007-1056-3
- Swaminathan, J., Mason, C. R., Streeter, T. M., Best, V., Kidd, G. Jr., and Patel, A. D. (2015). Musical training, individual differences and the cocktail party problem. *Sci. Rep.* 5:11628. doi: 10.1038/srep11628
- Swanepoel, R., Oosthuizen, D. J. J., and Hanekom, J. J. (2012). The relative importance of spectral cues for vowel recognition in severe noise. *J. Acoust. Soc. Am.* 132, 2652–2662. doi: 10.1121/1.4751543
- Symmes, D., Alexander, G. E., and Newman, J. D. (1980). Neural processing of vocalizations and artificial stimuli in the medial geniculate body of squirrel monkey. *Hear. Res.* 3, 133–146. doi: 10.1016/0378-5955(80)90041-6
- Tang, J., Yang, W., and Suga, N. (2012). Modulation of thalamic auditory neurons by the primary auditory cortex. *J. Neurophysiol.* 108, 935–942. doi: 10.1152/jn.00251.2012
- Temereanca, S., and Simons, D. J. (2004). Functional topography of corticothalamic feedback enhances thalamic spatial response tuning in the somatosensory whisker/barrel system. *Neuron* 41, 639–651. doi: 10.1016/S0896-6273(04)00046-7
- Terhardt, E. (1974). Pitch, consonance, and harmony. *J. Acoust. Soc. Am.* 55, 1061–1069. doi: 10.1121/1.1914648
- Tervaniemi, M., Medvedev, S. V., Alho, K., Pakhomov, S. V., Roudas, M. S., van Zuijlen, T. L., et al. (2000). Lateralized automatic auditory processing of phonetic versus musical information: a PET study. *Hum. Brain Mapp.* 10, 74–79.
- Tervaniemi, M., Szameitat, A. J., Kruck, S., Schröger, E., Alter, K., De Baene, W., et al. (2006). From air oscillations to music and speech: functional magnetic resonance imaging evidence for fine-tuned neural networks in audition. *J. Neurosci.* 26, 8647–8652. doi: 10.1523/JNEUROSCI.0995-06.2006
- Tomlinson, R. W. W., and Schwarz, D. W. F. (1988). Perception of the missing fundamental in nonhuman primates. *J. Acoust. Soc. Am.* 84, 560–565. doi: 10.1121/1.396833
- Town, S. M., and Bizley, J. K. (2013). Neural and behavioral investigations into timbre perception. *Front. Syst. Neurosci.* 7:88. doi: 10.3389/fnsys.2013.00088
- Tramo, M. J., De Cariani, P. A., Delgutte, B., and Braid, L. D. (2001). Neurobiological foundations for the theory of harmony in western tonal music. *Ann. N.Y. Acad. Sci.* 930, 92–116. doi: 10.1111/j.1749-6632.2001.tb05727.x
- Tregellas, J. R., Ellis, J., Shatti, S., Du, Y. P., and Rojas, D. C. (2009). Increased hippocampal, thalamic, and prefrontal hemodynamic response to an urban noise stimulus in Schizophrenia. *Am. J. Psychiatry* 166, 354–360. doi: 10.1176/appi.ajp.2008.08030411
- van Noorden, L. P. A. S. (1975). *Temporal Coherence in the Perception of Tone Sequences*. Ph.D. thesis. Eindhoven: Eindhoven University of Technology.
- Vaquero, L., Hartmann, K., Ripollés, P., Rojo, N., Sierpowska, J., François, C., et al. (2016). Structural neuroplasticity in expert pianists depends on the age of musical training onset. *Neuroimage* 126, 106–119. doi: 10.1016/j.neuroimage.2015.11.008
- Vasquez-Lopez, S. A., Weissenberger, Y., Lohse, M., Keating, P., King, A. J., and Dahmen, J. C. (2017). Thalamic input to auditory cortex is locally heterogeneous but globally tonotopic. *eLife* 6:e25141. doi: 10.7554/eLife.25141
- Velenovsky, D. S., Cetas, J. S., Price, R. O., Sinex, D. G., and McMullen, N. T. (2003). Functional subregions in primary auditory cortex defined by thalamocortical terminal arbors: an electrophysiological and anterograde labeling study. *J. Neurosci.* 23, 308–316.
- Villa, A. E. P. (1990). Physiological differentiation within the auditory part of the thalamic reticular nucleus of the cat. *Brain Res. Rev.* 15, 25–40. doi: 10.1016/0165-0173(90)90010-L
- Villa, A. E. P., Rouiller, E. M., Simm, G. M., Zurita, P., de Ribaupierre, Y., and de Ribaupierre, F. (1991). Corticofugal modulation of the information processing in the auditory thalamus of the cat. *Exp. Brain Res.* 86, 506–517. doi: 10.1007/BF00230524
- von Kriegstein, K., Patterson, R. D., and Griffiths, T. D. (2008). Task-dependent modulation of medial geniculate body is behaviorally relevant for speech recognition. *Curr. Biol.* 18, 1855–1859. doi: 10.1016/j.cub.2008.10.052
- Walker, K. M. M., Gonzalez, R., Kang, J. Z., McDermott, J. H., and King, A. J. (2019). Across-species differences in pitch perception are consistent with differences in cochlear filtering. *eLife* 8:e41626. doi: 10.7554/eLife.41626
- Walker, K. M. M., Schnupp, J. W. H., Hart-Schnupp, S. M. B., King, A. J., and Bizley, J. K. (2009). Pitch discrimination by ferrets for simple and complex sounds. *J. Acoust. Soc. Am.* 126, 1321–1335. doi: 10.1121/1.3179676
- Wallace, M. N., Anderson, L. A., and Palmer, A. R. (2007). Phase-locked responses to pure tones in the auditory thalamus. *J. Neurophysiol.* 98, 1941–1952. doi: 10.1152/jn.00697.2007
- Wang, W., Andolina, I. M., Lu, Y., Jones, H. E., and Sillito, A. M. (2016). Focal gain control of thalamic visual receptive fields by layer 6 corticothalamic feedback. *Cereb. Cortex* 28, 267–280. doi: 10.1093/cercor/bhw376
- Weinberger, N. M. (2011). The medial geniculate, not the amygdala, as the root of auditory fear conditioning. *Hear. Res.* 274, 61–74. doi: 10.1016/j.heares.2010.03.093
- Weyand, T. G., Boudreaux, M., and Guido, W. (2001). Burst and tonic response modes in thalamic neurons during sleep and wakefulness. *J. Neurophysiol.* 85, 1107–1118. doi: 10.1152/jn.2001.85.3.1107
- Williamson, R. S., and Polley, D. B. (2019). Parallel pathways for sound processing and functional connectivity among layer 5 and 6 auditory corticofugal neurons. *eLife* 8:e42974. doi: 10.7554/eLife.42974
- Winer, J. A. (2006). Decoding the auditory corticofugal systems. *Hear. Res.* 212, 1–8. doi: 10.1016/j.heares.2005.06.014
- Winer, J. A., Diehl, J. J., and Larue, D. T. (2001). Projections of auditory cortex to the medial geniculate body of the cat. *J. Comp. Neurol.* 430, 27–55.
- Winer, J. A., and Larue, D. T. (1987). Patterns of reciprocity in auditory thalamocortical and corticothalamic connections: study with horseradish peroxidase and autoradiographic methods in the rat medial geniculate body. *J. Comp. Neurol.* 257, 282–315. doi: 10.1002/cne.902570212
- Winer, J. A., and Larue, D. T. (1996). Evolution of GABAergic circuitry in the mammalian medial geniculate body. *Proc. Natl. Acad. Sci. U.S.A.* 93, 3083–3087. doi: 10.1073/pnas.93.7.3083
- Winer, J. A., and Prieto, J. J. (2001). Layer V in cat primary auditory cortex (AI): cellular architecture and identification of projection neurons. *J. Comp. Neurol.* 434, 379–412. doi: 10.1002/cne.1183
- Winer, J. A., Saint Marie, R. L., Larue, D. T., and Oliver, D. L. (1996). GABAergic feedforward projections from the inferior colliculus to the medial geniculate body. *Proc. Natl. Acad. Sci. U.S.A.* 93, 8005–8010. doi: 10.1073/pnas.93.15.8005
- Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., and Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat. Neurosci.* 10, 420–422. doi: 10.1038/nn1872
- Xie, F., You, L., Cai, D., Liu, M., Yue, Y., Wang, Y., et al. (2018). Fast inhibitory decay facilitates adult-like temporal processing in layer 5 of developing primary auditory cortex. *Cereb. Cortex* 28, 4319–4335. doi: 10.1093/cercor/bhx284
- Yan, J., and Ehret, G. (2002). Corticofugal modulation of midbrain sound processing in the house mouse. *Eur. J. Neurosci.* 16, 119–128. doi: 10.1046/j.1460-9568.2002.02046.x

- Yin, P., Johnson, J. S., O'Connor, K. N., and Sutter, M. L. (2011). Coding of amplitude modulation in primary auditory cortex. *J. Neurophysiol.* 105, 582–600. doi: 10.1152/jn.00621.2010
- Yu, Y.-Q., Xiong, Y., Chan, Y.-S., and He, J. (2004). Corticofugal gating of auditory information in the thalamus: an *in vivo* intracellular recording study. *J. Neurosci.* 24, 3060–3069. doi: 10.1523/JNEUROSCI.4897-03.2004
- Zatorre, R., Evans, A., and Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* 14, 1908–1919. doi: 10.1523/JNEUROSCI.14-04-01908.1994
- 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. doi: 10.1016/S1364-6613(00)01816-7
- Zatorre, R. J., Halpern, A. R., Perry, D. W., Meyer, E., and Evans, A. C. (1996). Hearing in the Mind's Ear: a PET investigation of musical imagery and perception. *J. Cogn. Neurosci.* 8, 29–46. doi: 10.1162/jocn.1996.8.1.29
- Zatorre, R. J., and Schönwiesner, M. (2011). “Cortical speech and music processes revealed by functional neuroimaging,” in *The Auditory Cortex*, eds J. A. Winer and C.E. Schreiner (Berlin: Springer), 657–677. doi: 10.1007/978-1-4419-0074-6_31
- Zendel, B. R., and Alain, C. (2012). Musicians experience less age-related decline in central auditory processing. *Psychol. Aging* 27, 410–417. doi: 10.1037/a0024816
- Zendel, B. R., Tremblay, C.-D., Belleville, S., and Peretz, I. (2015). The impact of musicianship on the cortical mechanisms related to separating speech from background noise. *J. Cogn. Neurosci.* 27, 1044–1059. doi: 10.1162/jocn_a_00758
- Zhan, X. J., Cox, C. L., Rinzel, J., and Sherman, S. M. (1999). Current clamp and modeling studies of low-threshold calcium spikes in cells of the Cat's Lateral Geniculate Nucleus. *J. Neurophysiol.* 81, 2360–2373. doi: 10.1152/jn.1999.81.5.2360
- Zhang, Y., and Suga, N. (2000). Modulation of responses and frequency tuning of thalamic and collicular neurons by cortical activation in mustached bats. *J. Neurophysiol.* 84, 325–333. doi: 10.1152/jn.2000.84.1.325
- Zhang, Y., Suga, N., and Yan, J. (1997). Corticofugal modulation of frequency processing in bat auditory system. *Nature* 387, 900–903. doi: 10.1038/43180
- Zhang, Y., Wang, X., Zhu, L., Bai, S., Li, R., Sun, H., et al. (2021). Selective corticofugal modulation on sound processing in auditory thalamus of awake marmosets. *bioRxiv* [Preprint]. doi: 10.1101/2021.03.13.435231
- Zhang, Y., and Yan, J. (2008). Corticothalamic feedback for sound-specific plasticity of auditory thalamic neurons elicited by tones paired with basal forebrain stimulation. *Cereb. Cortex* 18, 1521–1528. doi: 10.1093/cercor/bhm188
- Zhang, Z., Liu, C.-H., Yu, Y.-Q., Fujimoto, K., Chan, Y.-S., and He, J. (2008). Corticofugal projection inhibits the auditory thalamus through the thalamic reticular nucleus. *J. Neurophysiol.* 99, 2938–2945. doi: 10.1152/jn.00002.2008

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Homma and Bajo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.