



Beyond the ears: A review exploring the interconnected brain behind the hierarchical memory of music

Yiren Ren¹ · Thackery I. Brown¹

Accepted: 22 August 2023 / Published online: 18 September 2023
© The Psychonomic Society, Inc. 2023

Abstract

Music is a ubiquitous element of daily life. Understanding how music memory is represented and expressed in the brain is key to understanding how music can influence human daily cognitive tasks. Current music-memory literature is built on data from very heterogeneous tasks for measuring memory, and the neural correlates appear to differ depending on different forms of memory function targeted. Such heterogeneity leaves many exceptions and conflicts in the data underexplained (e.g., hippocampal involvement in music memory is debated). This review provides an overview of existing neuroimaging results from music-memory related studies and concludes that although music is a special class of event in our lives, the memory systems behind it do in fact share neural mechanisms with memories from other modalities. We suggest that dividing music memory into different levels of a hierarchy (structural level and semantic level) helps understand overlap and divergence in neural networks involved. This is grounded in the fact that memorizing a piece of music recruits brain clusters that separately support functions including—but not limited to—syntax storage and retrieval, temporal processing, prediction versus reality comparison, stimulus feature integration, personal memory associations, and emotion perception. The cross-talk between frontal-parietal music structural processing centers and the subcortical emotion and context encoding areas explains why music is not only so easily memorable but can also serve as strong contextual information for encoding and retrieving nonmusic information in our lives.

Keywords Music memory · Brain · Hippocampus · Inferior frontal gyrus · Auditory cortex · Music encoding · Music recognition

Introduction

The ability to understand, to appreciate and even to create music is an inherent aptitude for many people. Thus, understanding how the human brain perceives and engages with music can provide valuable knowledge for optimizing the integration of music into humans' daily lives. In recent decades, researchers have provided substantial evidence regarding the neural mechanisms behind music perception. Music is a complex event taking the form of sequentially varying tones and rhythm, sometimes from multiple sound sources (e.g., instruments). Despite the intricate nature of music, many individuals exhibit the capacity to understand, recognize, learn and recall music efficiently, irrespective of their training backgrounds, ages, or educational levels

(Bigand & Poulin-Charronnat, 2006; Trehub, 2006; Trehub et al., 1984). However, how the memory for a piece of music is formed and developed in the brain with such efficiency is understudied. Moreover, currently there are surprisingly limited formal empirical data of the connections between music memory and other types of memory, especially concerning the neural systems supporting each. One consequence of that is the debate over whether the music memory system is distinct from those supporting other types of long-term memory.

One influential piece of evidence contributing to this debate comes from lesion cases that demonstrated damage to the medial temporal lobe, an area known for encoding and consolidating general declarative memory, did not prevent learning novel music (Esfahani-Bayerl et al., 2019). However, considering music's intricate and multifaceted structure, it is important to consider that music memory may comprise hierarchical components, potentially reliant on diverse brain centers that individually support distinct memory processes. Thus, depending on how novelty is controlled

✉ Yiren Ren
evraine7@gmail.com

¹ Georgia Institute of Technology, College of Science, School of Psychology, Atlanta, GA, USA

and how memory is assessed, different impairments may or may not be expected in response to a lesion. Indeed, in his review, Altenmüller proposed that as music representations grow in complexity, progressing from pure auditory presentation to the incorporation of multimodal elements like visual performance, the implicated brain areas “graduate” from purely lower level processing in sensory areas to higher level prefrontal areas (E. O. Altenmüller, 2001).

In order to seek support for or against such a hypothesis of hierarchical neural networks for music memory, this review synthesizes current knowledge on the brain’s contributions to music memory. This involves categorizing music memory into different levels of information processing and representation. The goal is to outline brain networks underpinning different music components and phases of music memory and to discuss the connection between music memory’s neural mechanisms and the underlying areas’ roles in other memory domains. Notably, we delve into the role of the medial temporal lobe, a well-known declarative memory hub, within memory formation or retrieval. One of our main hypotheses is that, instead of there being a “unique” music memory system independent of the medial temporal lobe, music memory is upheld by a “syntax processing network” (*syntax network*), comprising mainly the auditory cortex and the frontal-parietal regions, alongside a “contextual associates network” (*context network*): mainly situated in subcortical limbic and reward regions. Within the hierarchical

representation of a music composition, we propose that the *syntax network* preprocesses and encodes the first level of musical pattern/structure, while the *context network* supports comprehension of higher level of music meaning and emotion (“music semantics”: a music literature terminology; shown in Fig. 1). Notably, this taxonomy discussion reveals overlaps between neural networks involved in music memory and those identified by decades of systematic neuroscience studies of nonmusic memory. Given society’s widespread interest in using music as an environmental tool to help humans memorize other events (e.g., from its use in advertising to therapeutic settings), establishing connections between the neuroscience of memory and the neuroscience of music becomes imperative.

This figure shows the two major brain networks proposed to support music memory. In the left hemisphere, it shows the inferior frontal gyrus, superior temporal gyrus, and inferior parietal lobule, which together operate as the “music syntax network” that supports memory of music syntactical structure and provide first-level understanding of music sequential patterns. Separately, in the right hemisphere, the figure shows the limbic system (hippocampus and amygdala) and the striatum, inside which the ventral striatum plays an important role in reward feedback. These three areas together support the memory of associated

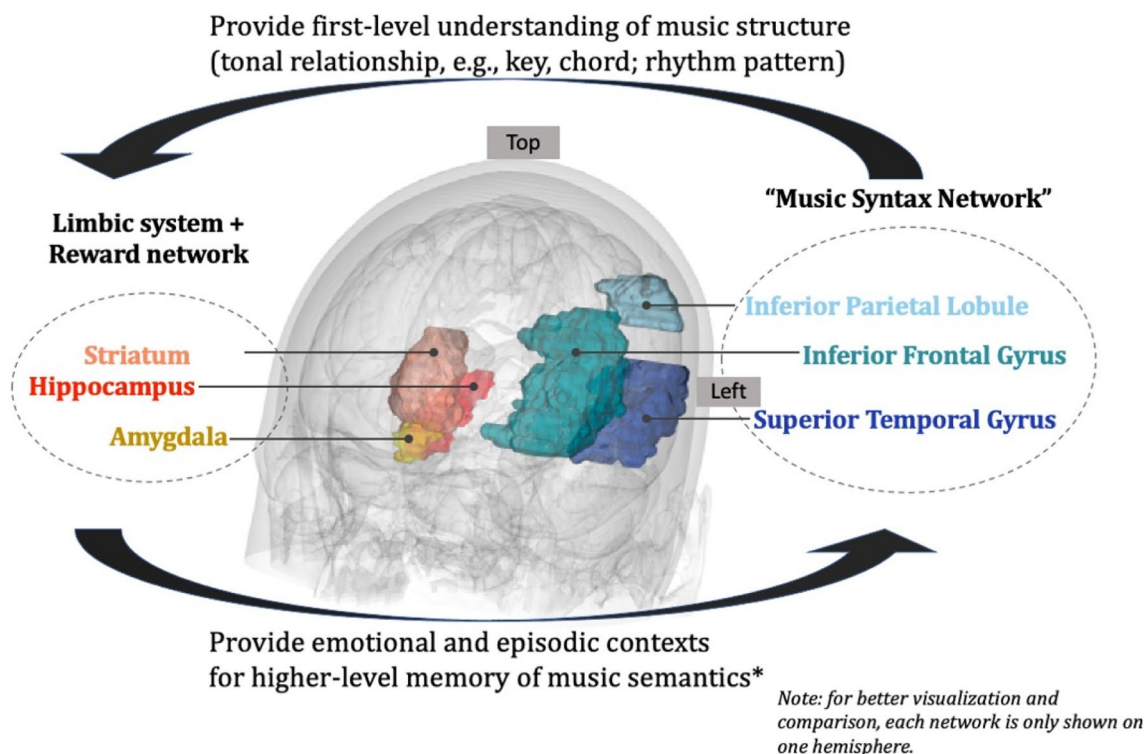


Fig. 1 Anatomical summary of the two main music-memory neural networks

emotional/episodic contexts evoked by and surrounding the music itself. As described in the text, the two types of memories supported by the two main brain networks might interact such that the “syntax network” provides first-level structural understanding of music pattern for further aesthetic/higher level analysis of music such as the emotion, processed by the limbic and reward system. Conversely, the limbic and reward system adds emotional and contextual elements to the music itself and potentially makes the music structure more memorable. We depicted the two networks separately on left and right hemispheres only for better visualization and we did not imply any lateralization of these networks.

**to be noted, the word “semantics” is used in both memory and music literature and it has slightly different meanings. In music literature, the “semantics” of music stands for the higher level information it conveys (Koelsch, 2009), in similar way to the language. Semantic memory, on the other hand, refers to the memory of a collection of general/nonepisodic knowledge. In this review, we will typically use the term “music semantics” in the manner of the music literature, and refer it as the emotional and aesthetic meaning of music. (Color figure online)*

Overview: Hierarchical music memory

Music encompasses various traditional memory categories, including semantic memory and episodic memory (Tulving, 1972), making it a multifaceted event. Music learning initiates with sound processing. Summarized in detail by a few review chapters (Ginsborg, 2004; Jäncke, 2019), the initial learning took place at a sensory level during which the brain forms separate sensory memories for streams of tones and temporal intervals (the “meter pattern”). To perceive music as a “whole” integrated event, tonal sequences are chunked into melodies until then the listeners are able to perceive the auditory input as music, not noise or speech. The ability to distinguish music from other auditory patterns is inherent (Deutsch, 1969; Trainor et al., 2004; Trehub, 1987; Winkler et al., 2009). This ability is commonly attributed to motor and sensory areas in the brain, including but not limited to the auditory cortex (BA 41/42, superior temporal gyrus), presupplementary motor area (pre-SMA), and others (Gaab et al., 2003; Groussard et al., 2010; Jäncke et al., 2003; Koelsch et al., 2009; Kunert et al., 2015). These areas, which serve as first-level processors for diverse auditory inputs, are commonly implicated in most music-listening tasks. However, this review’s primary focus is on the mechanisms

of remembering a piece of music (and its associates), and therefore it will not extensively elaborate on these important sensory systems.

Beyond sensory processing, knowledge and the memory of a specific piece of music is hierarchical (Ginsborg, 2004): from being able to recognize a familiar music composition, to being able to sing along (retrieve) parallel lyrics with the music, to knowing the associated title, composer, or the attached story and personal experiences with the music. However, many studies have defined long term memory (LTM) for music in different ways, yielding variability in the reported neural correlates. Some literature has tried to categorize music LTM using the traditional Tulving’s model and suggested that the memory of a piece of music can be implicit (e.g., procedural memory such as playing an instrument), semantic (lexicon knowledge) or episodic (e.g., “when” or “where”; Jäncke, 2019). But it has become clear that it is very difficult to test the differences between these forms of memory expressed by the brain in response to music in an experimental setting. For instance, sometimes music is so powerful in cueing associated events/emotion, when retrieving the “semantic” knowledge of a piece of music (recognizing if one piece of music is heard before), the possibility of retrieving the episodes paired with a piece of music simultaneously cannot be excluded. Similarly, it is also difficult to tease apart the implicit and semantic (explicit) components of “knowing a piece of music.” For instance, learning to perform a piece of music (such as playing the instrument) involves both explicit and implicit learning—instrument performance is a type of implicit procedural memory, but it also requires semantic understanding of the music structure and tonal lexicon. Thus, instead of using a traditional categorical memory model from psychology to explain music memory, here we propose an alternative perspective: scholars can more reliably delineate music properties and its neural correlates by evaluating music LTM at two levels according to the nature of music components being learned/memorized.

The first level is the *music syntactical structure memory*—the syntactical pattern and structure of music. A large proportion of existing neuroscience studies on music memory concentrate on this level, investigating memory of melody+rhythm patterns by posing questions such as “Can you recognize this melody?” or “Which ending tone corresponds to the previously learned music?” (Halpern & O’Connor, 2000; Sikka et al., 2015). We label this second level as *contextual associates memory*—incorporating non-music structural elements that still contribute to the extended memory representation of the music, such as the facts of the music piece (e.g., title), the associated lyrics, the paired episodic memory traces and our emotions elicited by or surrounding experiences with the composition. These associative elements provide higher level of semantic/aesthetic

meanings, enriching the understanding and memory of the music and carrying autobiographical significance.

This framework is important to a hierarchical understanding of music memory because sometimes how listeners process and encode the music structure is associated with how they understand the attached semantic meaning or autobiographical significance—for example, any story or emotion the tonal relationships of music express (Krumhansl, 2002; Meyer, 2008). We hypothesize that this hierarchical aspect is mirrored in the organization of the neural correlates—music structure/lexicon processing and encoding primarily occurs in “syntax processing” regions (NB, such processing is still inherently tied to memory, much like linguistic syntax processing), while associations between music and contexts are likely accomplished by broader regions, such as the hippocampus, a region important for binding events and building declarative relational associations (Olsen et al., 2012). Adopting this perspective of music memory, rather than traditional memory distinctions, might explain the seemingly disparate neural correlates of music memory in the literature, given that music literature has used highly variable paradigms, and focused on various stages of memory processes.

The following section of this review will categorize and summarize various task types used in the past, giving an overview on how different brain regions appear to support remembering different components of music, from initial encoding/preprocessing to long-term memory retrieval. It is worth noting that perspectives on some stages of music memory have only been minimally addressed in the literature. This is a challenge for making inferences about what the neural correlates signify, but it is also an opportunity for future research to focus on those types of designs to fill gaps in our understanding.

Music syntactical structure memory

The essence of music structure memory—A rule-governed system

Encoding music structure in memory is highly dependent on a syntax knowledge system which combines variations over multiple attributes/dimensions (e.g., pitch, meter) across time into an integrated event (Krumhansl, 1991; Levitin & Tirovolas, 2009). Music theorists term this multidimensional processing the hierarchy of music—a stable sound structure that determines when an auditory sequence qualifies as “tonal” (Krumhansl, 1991; Leonard, 1956; Narmour, 1983). This hierarchical structure differentiates music from noise, making music memorable, recognizable, and meaningful in a manner similar to language.

This syntax processing is inherently a memory-related phenomenon—like language, we understand music according to the mappings between what we hear and prior compositions and rules stored in our memory. Interestingly, most people appear to possess an innate familiarity with certain music syntax. For example, infants prefer consonant over dissonant music pieces as adults do (Trainor & Heinmiller, 1998). Similarly, the ability to recognize whether adjacent tones deviate from “tonality” was observed as early as Day 2 or 3 postpartum (Stefanics et al., 2009). Infants can also detect violations in temporal patterns. During initial exposure, infants could segment tones based on the temporal intervals using Gestalt’s Principle strategy, similar to how adults do (Hannon & Trehub, 2005; Trehub, 1987). The inherent detection of dissonance, whether tonal or temporal, is occasionally referred to as sensory/tonal dissonance and may directly lead to the stimulation of early sensory organs such as the basilar membrane (Johnson-Laird et al., 2012; Tillmann et al., 2014).

Beyond such seemingly innate rule processing, people continue to learn and update their memory of music syntax through their experiences, which in turn can facilitate new music encoding and analysis of the hierarchical structure of a novel piece of music. People are faster in learning a piece of music from their own culture (Demorest et al., 2010), suggesting utilization of a “learned syntax” schema when learning music of familiar styles. This “learned syntax” could develop during early stages of life—a study found infants as young as 4 months old already showed a preference for rhythm style of their own culture (Soley & Hannon, 2010). An EEG study also found that children display greater prediction error signals when exposed to musical input that violates their cultural musical syntax (Jentschke et al., 2008). Behavioral and modeling studies suggest that novel music learning always involves an interaction between old–new syntax processing. That is, from memory, listeners generate predictions during music listening based on prior knowledge of music syntax, and any “surprising” or “prediction violating” tones trigger listeners to update their probability calculation of the tonal relationship, resulting in updated music syntax and more effective encoding of new music pieces (Deutsch & Feroe, 1981; Leman, 2012; Pearce & Wiggins, 2012). Due to the critical role of grounding syntax processing in long-term memory, when we investigate the brain areas that support memory of music structure, studies should not only focus on how a specific piece of music is learned. It is also important to consider how the syntax associated with a composition is learned and updated, especially in terms of how the “old syntax” from prior music experiences interacts with “novel music.”

Given the extensive and intricate range of topics addressed regarding music structure memory in this review, we have included Table 1 as a concise point of reference.

Table 1 Music structure memory and its neural correlates. Summary of common tasks used in the literature related to music structure memory, and the main brain areas reported

Functions	Brain Regions Involved	Tasks Used	Example Studies
Initial Preprocessing	inferior frontal gyrus (IFG); superior temporal gyrus (STG)	Irregular/Dissonant tone/chord listening	(Sammler et al., 2009)
Syntactical error detection (for tones)	orbitofrontal gyrus; hippocampus; amygdala; anterior cingulate cortex (ACC)		(Blood et al., 1999; Blood & Zatorre, 2001; Sammler et al., 2007)
Emotion and rewarding feedback	inferior parietal lobule (IPL); basal ganglia (BG)	Irregular rhythm sequences listening	(Lappe et al., 2013; Limb et al., 2006)
Error detection (for temporal patterns) and time perception			
Memory Encoding	IFG; STG	Cross-culture music learning	(Demorest et al., 2010; Nan et al., 2008)
Encode music with familiar syntax (prior knowledge/schemas)	angular gyrus (IPL); insula; IFG; STG		
Encode music with unfamiliar syntax + new syntax learning	right IFG; IPL	Novel music (tonal) grammar learning	(Loui et al., 2011)
Encode new musical syntax/rules	left IFG		(Jarret et al., 2019)
Encode new musical sequences based on newly learned syntax	right IFG; IPL; planum temporale	Novel rhythm pattern learning	(Notter et al., 2019)
Extract and encode novel rhythm patterns	IFG; STG; BG; hippocampus; cerebellum	Repeated listening and music clips learning	(Burunat et al., 2014)
Learn particular novel music piece			
Memory Retrieval			
Recognize music based on sense of familiarity; emotional item recognition	IFG; STG; superior frontal lobe; superior parietal lobe; insula; ACC	Famous music recognition task	(E. Altenmüller et al., 2014; Klostermann et al., 2009; Satoh et al., 2006)
Recognize the previously learned music	IFG; hippocampus; cerebellum	Novel music recognition after encoding phase	(E. Altenmüller et al., 2014; Watanabe et al., 2008)
Unattended music encoding and recognition	dorsolateral prefrontal cortex; IFG; IPL; medial temporal lobe (MTL)	Mere exposure effect	(Green et al., 2012; Samson & Peretz, 2005)

IFG = inferior frontal gyrus; STG = superior temporal gyrus; ACC = anterior cingulate cortex; OFC = orbitofrontal cortex; BG = basal ganglia; IPL = inferior parietal lobule; SFG = superior frontal gyrus; SPL = superior parietal lobule; MTL = medial temporal lobe

This table summarizes the common tasks employed in prior studies concerning each stage of music structure memory, along with the primary brain areas frequently reported in the literature. For a visual representation, please refer to the corresponding section on the left side of Fig. 2, which presents a brain map of music structure memory.

Initial syntactical analysis for memory encoding—Preprocessing

Before delving into the literature on memory encoding per se, it is important to emphasize that the initial learning of novel music involves a comparison between auditory input and the schema of music syntax—an abstracted memory and rule set for the anticipated structure of compositions. Examining brain activity during novel music listening and comparing syntax-violating compositions to syntax-consistent music ones can reveal the neural basis of encoding novel music structure and how pre-existing schemas support this process.

A common paradigm for such comparison is the chord progression paradigm. Such tasks manipulate chord structure of a piece of musical sequence in a way to violate the

specific harmonic progression syntax (e.g., Western-music tonality system) at certain points. The subjects either solely listen to the deviant sequence or select from provided chords to make the music more consonant (consonant music: the combination of tones that is musically pleasing and melodious). Using this paradigm, EEG data show error-related signal potentials when listeners encountered irregular tones, violating the schema to which the composition has been mapped through prior learning. This includes the early right anterior negativity (ERAN) signal, an ERP effect that happens around 200 ms after a syntactically irregular chord (Kalda & Minati, 2012; Koelsch et al., 2005; Koelsch et al., 2007; Pagès-Portabella & Toro, 2020), and mismatch negativity (MMN) (fronto-central negativity), a similar signal that is evoked most strongly for an “error” tone (Lappe et al., 2013; Rohrmeier & Koelsch, 2012). A recent study using multiple source analysis identified Broca’s area, along with its right homolog as the dominant source for ERAN. The primary auditory cortex, particularly the superior temporal gyrus (STG), emerged as the dominant cortical source for MMN during processing harmonically irregular chords (Villarreal et al., 2011). Moreover, an intracranial EEG study provided more direct spatial measures, pinpointing

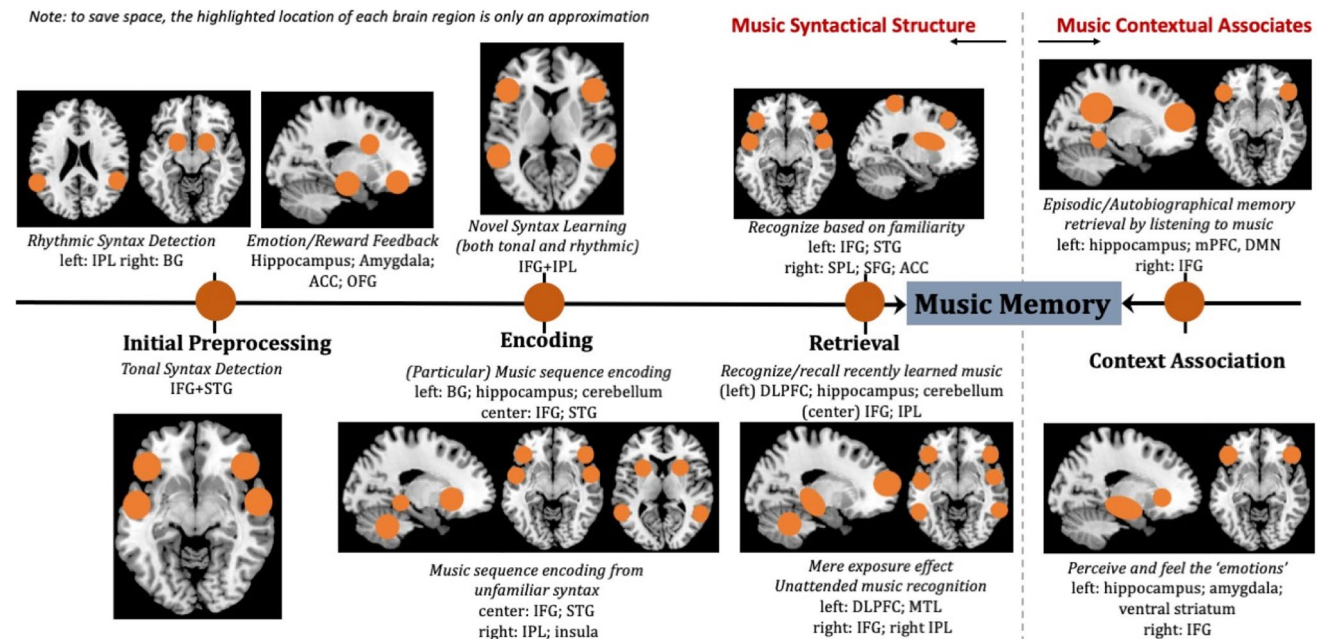


Fig. 2 Detailed visualization of brain regions supporting music memory formation. This figure shows the major brain regions that (1) support each stage of music structure memory, from initial preprocessing to retrieval and (2) the contextual associations, including episodes, semantic, and emotion association. Except for inferior frontal gyrus, which is recruited during all stages of both music memory types, the figure shows that for music structure memory, predominantly neocortical areas are involved whereas the contexts associated with the music are supported by the limbic system and subcortical reward circuitry. Note. *This is not a formal meta-analysis result, the*

highlighted brain areas those most frequently reported throughout the literature—modest variations in specific coordinates notwithstanding. IFG = inferior frontal gyrus; STG = superior temporal gyrus; ACC = anterior cingulate cortex; OFC = orbitofrontal gyrus; BG = basal ganglia; IPL = inferior parietal lobule; SFG = superior frontal gyrus; SPL = superior parietal lobule; MTL = medial temporal lobe; DMN = default mode network; mPFC = medial prefrontal cortex; DLPPFC = dorsolateral prefrontal cortex; OFG: orbitofrontal gyri. (Color figure online)

the superior temporal gyrus (STG) and inferior frontal gyrus (IFG), including Broca's area, as syntactical error detection signal sources for both music and language processing (Sammler et al., 2009). The fMRI studies during irregular chords listening revealed strong activities in these areas (STG and IFG) plus the premotor cortex, prefrontal cortex, and emotion and long-term memory related areas including orbitofrontal cortex, amygdala, hippocampus, anterior cingulate gyrus (Bravo et al., 2020; Koelsch, 2006, 2009; Koelsch et al., 2008; Koelsch et al., 2005a; Kunert et al., 2015).

An emergent theme of agreement from such work is the sensitivity of neural correlates, especially the STG and IFG, to syntax-violated musical sequence detection. Interestingly, however, some of this sensitivity might be innate. A recent study used fMRI on 1–3-day-old newborns while they passively listened to tonal music versus altered dissonant music and showed increased hemodynamic responses in right auditory cortex and left inferior frontal cortex to dissonant music excerpts (Perani et al., 2010). Thus, it is reasonable to postulate that during initial music structure processing, detecting deviation in tonal relationship or comparing the perceived musical pattern with stored syntax begins with the primary auditory cortex (mainly STG) and the IFG, which are sensitive even with minimal or no prior learning. These areas have long been found to relate to language processing including grammar and phonic processing (Binder et al., 1996; Koelsch et al., 2005b; Mesgarani et al., 2014), and their primal sensitivity to basic structure of auditory sequences may be why they also support learning new music pieces according to whether they follow already-learned syntax/styles.

Chord progression tasks mainly focus on the effects of violating music tonality. The above results suggest that the syntactical relationship between tones is processed in the similar brain system for language (Kunert et al., 2015). However, fewer neuroscience investigations have explored other musical structural components, such as the temporal relationship or metric processing. This gap is noteworthy because temporal intervals between tones are pivotal in defining a piece of music's unique identity. Indeed, keeping the order of tones the same but changing the temporal intervals between them can flip one song to be perceived as another. In this sense, remembering the structure of a piece of music also requires the encoding of the temporal pattern in addition to the tonal pattern. Although understudied, certain research teams have provided important insights into how this dimension of music structural encoding is supported during early stages by using *rhythmic progression designs* during which they played participants music with irregular temporal intervals. The fMRI results comparing listening to regular versus randomly timed rhythmic patterns for music tones revealed differential activation in the

left inferior parietal lobule (IPL) and frontal operculum, but not superior temporal areas (the auditory cortex; Limb et al., 2006). In subcortical areas, the basal ganglia was found to be robustly activated by regular beats in comparison to irregular rhythm sequences (Grahn, 2009). Another group used magnetoencephalography (MEG) and compared passive listening to both melodic and rhythmic deviations, and interestingly, they identified inferior parietal lobule (IPL) for temporal irregular music processing but not tonal processing (Lappe et al., 2013). The same group later used EEG and asked the participants to detect melodic and rhythmic errors during music listening. Results indicated that the rhythmic deviations activated inferior and superior parietal areas, along with the supplementary motor area. Conversely, melodic errors recruited the tonal syntax areas, primarily the STG and IFG (Lappe et al., 2016). Although limited, these data are important as they indicate a distinction in the neural substrate for metric and melodic rule processing and analysis for early stage of music memory encoding: while the former relies on motion-related areas such as the parietal areas and striatum, the latter may predominantly recruit the superior temporal cortex and inferior frontal cortex. A few other paradigms investigating temporal processing using same-different rhythm discriminations (Kuck et al., 2003; Thaut et al., 2014) and beat perception/reproduction tasks (Chen et al., 2008; Grahn & Brett, 2007; Konoike et al., 2012) suggest a parietal-motor-cerebellum network for rhythm processing, encoding and representation. Outside the music domain, these areas, including IPL, pre-SMA, and cerebellum, are closely associated with time perception (Assmus et al., 2003; Koch et al., 2007; Lee et al., 2007; Livesey et al., 2007). In particular, IPL has been associated with the integration of space and time. This potentially forms a core mechanism underlying rhythm learning in music processing, as memory formation for it relies on integrating tonal and temporal relationships.

In summary, the initial stage of encoding a novel music piece can engage a dynamic interplay between perception/processing and preexisting knowledge/memory of syntax. Both melodic and rhythmic structure processing studies demonstrate a cortical network for such syntax-based musical initial processing. This neocortical cluster (depicted in Figs. 1 and 2) supports the unfolding and understanding of continuous musical input over time. This, as detailed above, serves as a foundational element for the hierarchical encoding of music as an integrated memory. This system operates early in the process of music input perception and provides initial support for listeners to scan and “preprocessing” music—organizing the component music representations to be ready for subsequent encoding of the music into long-term memory. Concurrently, the process of contrasting established syntax schemas with new perception of musical sequences can foster continued updates to the syntactical

representation. Following this progression, our subsequent section will delve into evidence on the mechanisms underlying the acquisition of novel syntax.

Memory encoding—Music and syntax learning

The fidelity (precision and detail) of memory encoding in the brain—the process of storing information into long-term memory for later retrieval—highly correlates with whether the memory can be successfully recalled later (Hasselmo, 2006). While memory encoding for music structure has been explored using behavioral measures, limited neuroimaging data are available on the underlying neural correlates. Here, we consolidate findings from three types of music encoding-related task commonly used to help uncover the brain mechanisms for music structure encoding in memory.

One exciting approach is a cross-culture music-learning design, which asks how humans process and learn music pieces from different cultural styles. These studies could reveal mechanisms for both single novel-music piece-encoding and new syntactical-regularity encoding by comparing how people learn new music pieces from familiar and unfamiliar cultures, particularly considering behavioral results highlighting that humans are better at encoding, recognizing, and understanding music from their own cultures (Demorest et al., 2016; Drake & El Heni, 2003; Morrison et al., 2008; Soley & Hannon, 2010). Results from this line of research imply that a proportion of the regularity present in the structure of outside-culture music is not already represented in the listener's long-term memory, and thus this provides us a window into structure learning that is more *de novo* when encoding a particular piece than is possible with the common use of within-culture music stimuli. In the brain, compared with culturally familiar music, learning unfamiliar-style music prompted increased activity in angular gyrus, middle frontal gyrus, insula, cerebellum, and paracingulate cortex (Demorest et al., 2010; Nan et al., 2008). In particular, the angular gyrus has been linked to novel rule learning functions such as language syntax processing as well as mathematical rule learning (Bemis & Pylkkänen, 2013; Pyke et al., 2015; Seghier, 2013). Encoding both culturally familiar and unfamiliar music recruited activity in STG, IFG (Nan et al., 2008, reported left hemisphere while Demorest et al., 2010, reported right hemisphere) and planum temporale, an area posterior to the auditory cortex (Demorest et al., 2010; Morrison et al., 2003; Nan et al., 2008). Interestingly, the similar engagement of STG and IFG in learning both familiar and unfamiliar cultural music suggests that these regions primarily facilitate the learning any specific music sequences based on existing syntax knowledge. By contrast, the involvement of a broader network including the angular gyrus/IPL appears when grappling with the challenge

of learning new music styles and regularities, demanding a concurrent analysis of their unfamiliar structural principles and encoding.

Because this is a collective of areas reproduced from a number of studies sharing this syntax learning and syntax using approach, to simplify the text we will refer this cluster of areas as the “music syntax network” in the following text (as shown in Fig. 1), noting that it appears to be consistently organized with inferior frontal areas and superior temporal areas, supplemented by the inferior parietal cortex (IPL; because the IPL also supports syntax processing/learning for music structure, as discussed in the following sections). Within this music syntax network, STG has been argued by some to correlate with the ability to learn new music in relation to prior knowledge. Morrison et al. (2003) compared musicians vs. untrained controls when they listened to music from their culture versus an unfamiliar culture. The results showed that, behaviorally, musicians were better at encoding the music from their own cultures than nonmusicians and this group showed higher activity in STG and midfrontal regions during encoding (Morrison et al., 2003). This implies a training-related plasticity in STG and its role in supporting the acquisition of novel music based on established regularity.

A second type of experimental design relating to music structure memory encoding uses a “new syntax learning paradigm” and focuses on the function of “statistical learning” in music. Statistical learning is an important cognitive function reflecting rule extraction from repetitive exposure to a sequence. This concept has been widely studied in language research—for example, how infants learned and integrated linguistic regularities (Romberg & Saffran, 2010). Similarly, investigating statistical learning of musical sequences might reveal how listeners learn a specific pattern of music structure with minimal prior knowledge. Such studies created artificial (novel) grammar for tonal sequences and exposed them to the listeners. During the learning task, one MEG data revealed substantial oscillatory entrainment between temporo-frontal areas that positively correlated with learning performance (Moser et al., 2021). Using diffusion tensor imaging (DTI), a study identified a positive correlation between pitch-related grammar learning performance and the tract volume between right IFG and right middle temporal gyrus, again suggesting an important role of temporo-frontal coordination in rule learning of music (Loui et al., 2011). They also reported higher white matter volume in the supramarginal gyrus, a proportion of IPL (the other being angular gyrus), correlating with individual differences in learning performance. Interestingly, one lesion study suggested that damage to left IFG would not prevent successful grammar learning of pitch sequences (Jarret et al., 2019)—when this is taken together with the Loui et al. (2011) result mentioned above (Loui et al., 2011), this implied there may

be a right lateralization of IFG in *encoding* brand new musical syntax. Meanwhile, another fMRI study, conducted during *recognition* of newly learned sequences after statistical learning, suggested left IFG activation in distinguishing learned tonal sequences from random sequences. Overall, these findings suggest IFG's pivotal role in musical syntax learning.

An earlier section of this review emphasized the temporal sequence structure being another core component of music—however, only a few studies have focused on how the brain encodes novel musical rhythm patterns. One paper suggested that right IFG, bilateral supramarginal gyrus and planum temporale contributed to extracting rhythm patterns from novel rhythmic music sequence, by extracting and summarizing the interval ratio between tones (Notter et al., 2019). It indicates that the “music syntax network” does not only process the tonal syntax for new learning but might serve as a general musical rule processing center for both the tonal and temporal relationship of sounds. Most other rhythm pattern learning tasks used nonmusic stimuli (e.g., visual stimuli present at a specific beat) and have underscored the involvement of time-processing areas such as the basal ganglia and cerebellum in rhythm encoding (Janata & Grafton, 2003; Ramnani & Passingham, 2001; Sakai et al., 2004). The basal ganglia and cerebellum have also been observed in music listening tasks, especially when the music is rhythmically regular and leads to motor reactions such as tapping or dancing (Molinari et al., 2007; Zatorre et al., 2007). However, interpretation of the contributions of these subcortical areas from nonmusic rhythm studies is challenging due to the nature of their task design, which often involves corresponding motor reactions to the stimuli. More investigation is still needed to address the question of whether these subcortical regions are critical areas supporting novel musical rhythm pattern learning independent of sometimes implicit motoric responding aligned with the rhythm.

The two types of designs discussed above, which emphasize tone and rhythm structure, typically leave open the question of how the brain encodes a specific piece of song or melody as a unit that can be *explicitly* recalled later. Unfortunately, there are relatively few studies that have utilized neuroimaging methods during the encoding stage of novel music learning with a postlearning declarative recall task. An fMRI study presented novel music with different repeated motifs embedded while the participants performed a “phrase segmentation task,” which asked the listeners to use their instinct and press buttons to indicate boundaries between music phrases during listening. Motifs with more repetition and exposure, which should lead to stronger memory, elicited increased activities in the SMA, basal ganglia, hippocampus, IFG and cerebellum (Burunat et al., 2014). This suggests potential involvement of these

areas in encoding specific melodic sequences. Another MEG study presented 68 participants a highly structured classical music prelude repeatedly and later tested their memories for the music. During encoding, the auditory cortex, insula, hippocampus and basal ganglia were engaged in the task (Bonetti et al., 2021). Notably, musical experts exhibited stronger activity in STG and insula, aligning with the syntactical learning findings mentioned above and suggesting training-related plasticity in these areas (mainly the auditory cortex) for music encoding functions. Although these types of studies are infrequent, they suggested that the “music syntax network,” including the IFG and the auditory cortex, along with the insula, hippocampus, and the basal ganglia might collectively support the encoding of specific music encoding.

In summary, the overview of various studies relating to music structure memory encoding revealed a network that spans the neocortex (centered in IFG, STG, insula, and IPL), inner cortex (the hippocampus) and the subcortex (cerebellum and basal ganglia). If we juxtapose the emphasis on primary auditory cortex from data on basic musical input processing discussed in the last section, a pattern emerges: music memory encoding draw upon downstream and higher order areas that deal with sequential learning and integration. Interestingly, the reader will note the limited mention of the hippocampus, a crucial region in memory research. While it has a well-established role in sequential learning and declarative memory encoding in nonmusic literature (Shapiro & Eichenbaum, 1999; Wallenstein et al., 1998), its involvement in music structure memory encoding is rarely reported. From the existing literature, the hippocampus only appears to come into play in specific instances, such as encoding a particular melody that was repeatedly listened to during the task. We provide further discussion of these observations and the potentially interesting conditions for hippocampal involvement in music structure memory in a later section.

Memory retrieval—Recognition memory tasks

Many of the extant “music memory” neuroscience studies have primarily focused on the retrieval stage of memory. Among these, a commonly used approach involves recognition tasks where participants listen to compositions and determine whether they have heard a particular piece before. Recognizing melody has been associated with a wide range of brain regions across the neocortex (Jäncke, 2019; Peretz & Zatorre, 2005). Notably, a meta-analysis paper examined studies that compared listening to unfamiliar versus familiar music and found that the superior frontal gyrus (SFG) exhibited the most significant involvement during listening to familiar music (Freitas et al., 2018). However, this paper also emphasized the absence of consistent activations in

any areas across different studies. There are many possible reasons, including inconsistency in music styles, task designs and participants demographics the selected studies used. Moreover, the nature of the memory retrieval being tested can vary in demands depending on the task method and difficulty. For instance, the amount of recognition cues provided or the degree of similarity between target compositions and lures might influence the retrieval process. Some debates within the field of nonmusic recognition memory might provide hints: for example, evidence suggests distinctive neural substrates for familiarity (“recognize it”) and recollection (“recall it”) (Duarte et al., 2004; Yonelinas, 2002). In a similar vein, looking into the method of each music recognition task and categorizing them by retrieval levels might offer a more nuanced understanding of the variable neural correlates during music structure memory retrieval.

One interesting breakdown of the underlying studies is that many music recognition studies used music stimuli from popular listening charts. These studies measured brain activity while participants listened to and recognized famous music selected from these databases. These music clips were typically pretested and validated for familiarity through pilot tests. In this way, these studies used music that was (presumably) frequently listened to and decisions could be made with a sense of familiarity in the recognition paradigm. This process might require less effort than recalling the music from memory without any cue (Yonelinas, 2002). Recognizing familiar (famous) music, as opposed to unfamiliar non-famous music, showed significant activation in the STG, IFG, as well as in the superior frontal gyrus (SFG), superior parietal lobule (SPL), orbitofrontal lobe (OFC), insula, anterior cingulate gyrus (ACC) and parahippocampal gyrus (E. Altenmüller et al., 2014; Freitas et al., 2018; Jacobsen et al., 2015; Klostermann et al., 2009; Satoh et al., 2006; Sikka et al., 2015). Some of these areas fall outside the syntax network (which is, IFG, STG and IPL) are commonly observed in nonmusic recognition and familiarity tasks (Aggleton et al., 2011; Haxby et al., 1996; Morita et al., 2014; O’Connor et al., 2010). For example, in item recognition, superior parietal regions seem to support a sense of familiarity while more ventral lateral parietal and temporal regions, which were not found in these music familiar–unfamiliar recognition contrasts, are more involved in memory recollection (Yonelinas, 2002). Areas such as insula, anterior cingulate and superior frontal regions are not as consistently associated in episodic recognition tasks, though, do seem to relate to emotion-related recognition task such as facial recognition (Campbell et al., 2015; Haxby et al., 1996; Morita et al., 2014) and music recognition, as implied here. Given most music carries some degree of affective valence, one study focusing on how familiarity of music modulated induced-emotion showed higher activity in emotion-memory-related areas such as the cingulate cortex,

thalamus, amygdala, and hippocampus (Pereira et al., 2011). This suggests an association between “knowing a piece of music” and “liking the music,” and indicated a reward-emotion circuits behind music long-term memory. This connection between music, emotion, and memory is a topic of further discussion in later sections on music contextual memory).

Overall, compared with episodic memory familiarity, which many studies argue to rely more on medial temporal lobe cortex (the perirhinal cortex in particular; Yonelinas, 2002), most probes of music structural familiarity reveal a frontal-parietal cortical network of activity. The “music syntax network” consistently emerges as being associated with familiar music recognition, as it had been involved in encoding (Demorest et al., 2010; Ford et al., 2016; Gagnepain et al., 2017; Groussard et al., 2010; Jacobsen et al., 2015; Pereira et al., 2011; Plailly et al., 2007; Sikka et al., 2015). Considering the need to process music syntactical structure in order to successfully recognize music as “old,” it remains unclear and a matter for continued research whether this network is primarily involved in the fundamental levels of musical stimuli processing across retrieval and encoding, or whether it serves a distinctive functional role for music retrieval compared with its role in encoding.

A few studies of the neural systems supporting music recognition have used less known music, and implemented an encoding or familiarizing phase before the memory recognition task. While famous music might easily cue a sense of familiarity, in these tasks using novel music, recognition might require more in-depth processing and effortful comparison between to-be-recognized stimuli versus old memory. Consequently, such tasks might engage brain areas that support old memory recollection. By contrasting correct recognition(hits) and incorrect rejections(misses), one study revealed increased activity in the right IFG and left cerebellum, indicating a correlation with successful music structure LTM retrieval (E. Altenmüller et al., 2014). In another study contrasting hits versus correct rejections (old–new), revealed more activity in left IFG and the hippocampus for hits, implying their role in representing and retrieving old memory (Watanabe et al., 2008). Comparing these two studies raises the question: is there any functional difference between left and right IFG contributions to music retrieval? One study used a cross-culture design in which participants learned music of their own culture as well as that from an unfamiliar culture. During the postlearning recognition task, they found more right IFG activation during recognition of music from familiar culture and more left IFG when recognizing music from unfamiliar culture (Demorest et al., 2010). This might imply that the right IFG is more involved in remembering and recognizing patterns from existing/well-learned music syntax while left IFG contributes to general regular sequence processing and retrieval. Studies

comparing lyrical incongruity (e.g., the adjacent words in lyrics do not make sense together) and melodic incongruity also revealed left only IFG activity in both conditions, and they suggested this area serves as a general sequence rule processor for both music and language (Koelsch et al., 2000; Peretz, 2002).

Besides the retrieval/recognition of music structure, there is another behavioral phenomenon tied to music structure memory recognition: the mere exposure effect. The mere exposure effect shows that prior exposure to music can enhance an individual's preference for the music's composition (Fang et al., 2007; Peretz et al., 1998). Various studies of this effect used repeated passive music listening paradigms; tracking how the preference for music compositions increased is associated with recognition memory of music (Green et al., 2012; Samson & Peretz, 2005; Wang & Chang, 2004). Neuroimaging data from these paradigms showed an increase in dorsolateral prefrontal cortex (DLPFC), IFG, and IPL activity (Green et al., 2012) as listeners increased their preference to the music as a function of repeated exposure. This again implicates a frontal-parietal cortical network involved in music recognition, even it is unattended development of preferences associated with that familiarity. In addition to the IFG and IPL, which supports the music syntax, the DLPFC is known to support working memory (Barbey et al., 2013; Levy & Goldman-Rakic, 2000)—a general but critical memory function during continuous music encoding. Other than the areas mentioned above, the medial temporal lobe might also be critical for certain types of music learning and recognition. Evidence from medial temporal lobe (MTL) lesions showed that right MTL damage would result in failure of the mere exposure effect for music (Samson & Peretz, 2005). These patients also displayed impaired explicit recognition abilities for the music itself, as measured after the listening paradigm. This suggests that the MTL might be necessary for successful music encoding and future explicit recognition success, even though this area is less commonly identified in neuroimaging studies of music recognition function.

In summary, music structure memory retrieval is associated with multiple brain networks. The main network supporting music recognition includes the “syntax network” as well as neighboring frontal-parietal regions. Table 1 highlights major papers contributing to this understanding, noting the varied task types used in those references. Figure 2 synthesizes the correlates of music memory formation from Table 1 into a visualization of the networks in brain space, which helps the reader to see similarities and changes in recruited networks across different phases of processing music memory. These areas may support recognizing the regular sequential pattern of the music stimuli by comparing it with prior knowledge of the syntax. Recognizing a piece of previously heard music also activate emotion-related

areas such as the orbitofrontal areas and the basal ganglia potentially due to a positive correlation between music rewarding and the familiarity (Salimpoor et al., 2015). One remaining debate is whether the hippocampus and surrounding medial temporal lobe is essential for music structure memory retrieval. Although larger proportion of studies did not report hippocampal involvement during famous or newly learned music recognition tasks, some MTL lesion data show a failure to explicitly recognize music (Samson & Peretz, 2005), and a few other memory recognition tasks using more complicated designs (such as music boundary detection tasks) actually did report a hippocampus correlation with successful recognition (Burunat et al., 2014, 2018; Knösche et al., 2005). When associating hippocampal activity with overall memory performance (overall recognition rate), one study found a positive correlation, suggesting hippocampus might still relate to the overall memory ability for music structure (Watanabe et al., 2008). These observations lead to a question of whether hippocampus and its neighboring MTL cortical areas contribute to a higher level of memory retrieval in music. Given the MTL's essential role in nonmusic declarative memory functions, it is of considerable interest to delve into the finer details of whether and how the hippocampus and surrounding MTL contribute to music memory.

Medial temporal lobe and music structure memory

The medial temporal lobe plays an essential role in successfully learning new declarative information. As the reader will also see in later sections of this review, the MTL is intimately tied to contextual and emotional dimensions of music memory. However, as is evident in the sections above, it is still under debate whether MTL is necessary for music *structure learning* per se. Although several studies reported individuals with bilateral medial temporal lobe lesions being able to learn music (Cavaco et al., 2012; Finke et al., 2012; Valtonen et al., 2014), there are potential alternative explanations. Among these studies, two cases measured music memory via instrument learning and claimed intact music memory in MTL lesions patients because they could successfully learn to play and read new music. Such an ability to play instruments is associated strongly with procedural or skill memory that is supported by brain areas outside medial temporal lobe (Squire, 1992) and is echoed by a long literature showing new motor skill learning in MTL lesion patients (Corkin, 1968; Squire & Zola-Morgan, 1991). Another study aimed at more explicit component of music structure memory using a recognition task and showed intact recognition ability for newly learned music in an MTL-lesioned cellist. They thus concluded that new music learning was independent of MTL (Finke et al., 2012). However, this study came with an interesting caveat: the patient was a

life-long cellist, and his music training history could provide an elevated music memory ability compared with an average person, bolstered by mechanisms for greater neural plasticity in cortex for information that can be related to prior knowledge/old memory (van Kesteren et al., 2012, 2014)—in his case, new music compositions—and that potentially resists the functional damage by the lesions in MTL. We speculate these “schema”-grounded cortical learning enhancements may also explain why there are more cases of intact music memory in AD patients when the patients had music training history (Baird & Samson, 2009). In order to eliminate this possibility, recently, the same group that studied the cellist patient published another case of intact newly learned music recognition ability in a bilateral hippocampal lesion patient. Although they identified this patient as a musical layperson because he only played instruments in late childhood and adolescence, in their musical processing and perception task, this patient actually outperformed the control group, suggesting a higher-than-average music ability. Along this vein, aging studies showed that people with nominal amounts of music training early in life, even if they had not played instruments for years, showed training-driven neural plasticity relating to sound processing (White-Schwoch et al., 2013). In healthy population studies, evidence of hippocampal morphological plasticity was also found in music trained group, reflected in higher grey matter volumes (Herdener et al., 2010; Teki et al., 2012). Such training-induced neural changes might correlate with better music processing and encoding ability in musicians compared with nonmusicians (Burunat et al., 2018). Thus, the lesion cases of patients with previous music training, although suggestive that music memory ability is persistent despite MTL damage (compared with other forms of memory such as visual episodic memory), may be insufficient to conclude that music structure learning functions are truly independent of MTL involvement.

It should also be stressed that there are other lesion studies that support MTL’s functional importance for music learning. Multiple case studies of bilateral or unilateral MTL lesion patients or Alzheimer’s disease patients present some extent of disability in recognizing newly learned music (Bartlett et al., 1995; Cowles et al., 2003; Samson & Peretz, 2005; Samson & Zatorre, 1992), already familiar music (Bartlett et al., 1995; Huijgen et al., 2015; Peretz, 1993, 1996), or single sounds that were recently presented (Squire et al., 2001), and even in implicit preference development via the mere exposure effect (Samson & Peretz, 2005). Many emphasized that these individuals retained normal online music structure perception/processing ability while losing long-term memory for the music they had processed (Peretz, 1993, 1996). Some patients can learn to perform the music right after perception but failed to consolidate it (Cowles et al., 2003), the others failed to hold

the music in short-term declarative memory, such as in a music timbre comparison test (Samson & Zatorre, 1994). All such data suggest memory encoding and consolidating for music structure may in fact be impoverished with MTL damage, aligning with findings of the same functions of MTL from nonmusic domains (LaBar & Phelps, 1998; Shapiro & Eichenbaum, 1999).

When looking into healthy populations, an important consideration is that music is highly regular and predictable based on prior knowledge, even for “novel” pieces and thus music learning can be adequately supported by an “old syntax” network which we review above, centered on frontal areas which were also shown to store and retrieve nonmusic old memory (Preston & Eichenbaum, 2013; van Kesteren et al., 2010). It also bears noting that MRI considerations such as signal-to-noise ratio, statistical thresholds and activity cluster size cutoffs, can contribute to false negatives in studies which aren’t designed specifically to test for MTL involvement, and given the correlational nature of neuroimaging data such null reports cannot establish a lack of causality for the MTL in music memory. As a result, the scarcity of reports indicating MTL involvement in declarative music memory does not necessarily imply that MTL does not play an role in supporting music structure memory at all.

A major reason for expecting that it should is that both human and animal studies have shown that the MTL makes core contribution to sequential learning particularly due to its functions of time and temporal context processing (Eichenbaum, 2014; MacDonald et al., 2011) and its ability to bind multidimensional components of memories together (Gheysen & Fias, 2012; Olsen et al., 2012; Wallenstein et al., 1998). In fact, damage in hippocampus bilaterally can lead to failures in the statistical learning of sequential regularities (Covington et al., 2018; Schapiro et al., 2014). On complementary side, there is a rich and growing literature on non-music memory which also delineates mechanisms for frontal-cortical memory systems to circumvent the hippocampus and form new semantic associations when the new experience maps more strongly to prior knowledge (van Kesteren et al., 2012, 2014). Inspired by the examples where MTL *is* implicated in music learning, and because music is typically a multidimensional sequence of information, requiring the ability to bind multiple elements together and to learn statistical regularities, we draw on such perspectives from the neuroscience of memory to speculate the following: that the hippocampus might collaborate with the “syntax network” during initial learning of musical structure regularity to support high-order sequence segmentation and relational binding of temporally distinct and distant elements (as it does in nonmusic memory literature). The experiences an individual has with the music structures (through their lifetime listening and exposure, or formally studying an instrument, or through familiarity with the individual song or structurally

similar compositions) lowers the burden of new and continued learning on the hippocampus. This perspective can help rectify the disparate findings mentioned earlier. It also emphasizes the importance of studies that control for various sources of direct familiarity and more schematic/thematic familiarity with music stimuli, especially when investigating hippocampal involvement.

In the interest of promoting future research in this direction, here we further unpack this idea. During initial learning of truly novel music, one important step is to extract a hierarchical structure of a continuous auditory input. Analogous to verbal story, words within a sentence are more correlated with each other than outside the sentence. Similarly, sentences adjacent to each other are more semantically related than distant sentences. To learn long and continuous music sequences, the listener needs to construct a hierarchical map by segmenting ongoing music into bars (a small segment of music that holds a few beats), phrases, periods, and sections. This segmentation is based on the statistical relationship between tones as defined by musical syntax. During this process, tones are grouped into regular patterns and patterns are segmented/grouped and ordered—this is the initial encoding of the music as an integrated piece that holds structural regularity. In nonmusic memory literature, hippocampal function shows strong associations with hierarchical regularity learning, supporting cluster segmenting and temporal boundary detecting during on-going sequence encoding (Gupta et al., 2012; Hsieh et al., 2014; Schapiro et al., 2016). Some evidence suggests this same function does exist in music sequence encoding. One study asked participants to listen to 8-minute-long music with repeated motifs embedded. The participants performed a segmentation task using motifs as cues and showed significant activation in both the hippocampus and IFG during music boundary identification (Burunat et al., 2014). They also observed strong functional connectivity between hippocampus with cerebellum during motif repetition, when the most stand-out regular pattern was detected and encoded, suggesting these two regions have a role in sequence regularity detection. In a follow-up study from the same group, they conducted a similar music boundary segmentation task and again found strong functional connectivity between hippocampus and cerebellum that positively correlated with the section segmentation performance (Burunat et al., 2018). This is aligned with evidence from visual serial sequence learning tasks showing hippocampal-cerebellum network contributes to spatio-temporal sequence encoding and prediction ability (Onuki et al., 2015). Similarly, another group designed a unique acoustic serial sequence learning paradigm during which the listeners needed to detect and learn regular acoustic patterns hidden within distracting irregular sequences (Jablonowski et al., 2018).

They observed a positive correlation between the amount of learned acoustic sequence and the left hippocampus activation, implying this area's primary involvement in sequential regularity detection and encoding during a serial of musical input. Besides detecting the sequence, the ability to segment ongoing sequences also enables better extracting and understanding of the regular patterns, especially when initially encoding a novel music sequence. One MEG task focusing on music phrase perception recorded brain changes during listening to the continuous sequences of music phrases and using source-localization reported major activity from limbic system and posterior MTL at the phrase boundaries (Knösche et al., 2005), suggesting this brain cluster is important in dividing ongoing sequences into different groups.

In addition to regularity detection and segmentation, the hippocampus might also provide support for associations between temporally “distant” elements during initial music structure learning. In nonmusic domains, the hippocampus supports both dissociation and binding in memory between temporally distant experiences—for example, building association between two events that happened at different times (Wallenstein et al., 1998). In music learning, an overall composition involves higher order associations between multiple musical subevents over a protracted time scale, and such hippocampal binding functions may be particularly necessary when the music composition is long. For example, orchestral classical music piece often contains more than one chapter and chapters can differ a lot in tempo, key and emotion. In this case, music syntactical schemas might be insufficient for relating and binding two very different chapters together (because they don't follow the same structure). One fMRI task tested a similar idea, probing learning of distinct musical sequences as same groups, and showed that lesions in left MTL led to failure in this higher order music pair learning, supporting such a role in binding distant music sequences into memory (Wilson & Saling, 2008). Overall, although the data are limited, extant literature implies that the hippocampus might be more involved during initial music encoding, when tasks require truly novel and explicit sequence structure detection, as well as binding segments of music into higher order relational memories; these results align with nonmusic memory literature and together suggest hippocampus may have an essential role in sequence temporal structure encoding across domains, particularly through its function of sequential element segmentation and association. What many studies failing to show hippocampal involvement may lack is control over a sufficient level of novelty, sequential structural complexity, and/or explicit memory demands—but more studies are needed to put this explanation to the test and truly rectify the mixed findings in the literature.

Music contextual associates memory

When asked to recall a piece of music, we often find ourselves retrieving more than just its musical structure. Music-associated information such as the title, lyrics, emotional theme, or even the autobiographical episodes connected to it can come to mind, sometimes even involuntarily. Humans love to listen to music and play it in many important life events, such as weddings. This can be attributed partially to the deeper semantic meaning conveyed by music. The story the music conveys and the emotion it induces connect the listeners to the music, leading to the encoding of self-related emotional episodic memories. Unfortunately, these contextual associate elements of music memory have rarely been systematically discussed from either a behavioral or neuronal perspective. This review has thus saved this aspect of music memory for the last part, as it can be a particularly fruitful future direction for music memory research—especially for those who are interested in using music to aid life-related behaviors such as modulating memory and emotion. Here we will summarize the limited current knowledge of the neural correlates supporting formation and retrieval of music contextual associates memory. The right side of Fig. 2 provides a visual aid for the main brain areas discussed in this section.

Music-lyrics binding

Music that has lyrics plays the predominant role in the modern music production market, and for most of us much of the everyday music we hear and sing has lyrics. Research has shown that texts are more memorable when they are embedded within a piece of music (Palisson et al., 2015; Simmons-Stern et al., 2010). Advertisers have leveraged this phenomenon by pairing their slogans with melodies to capitalize on the memorability of this combination (Yalch, 1991). One cognitive hypothesis stemming from this phenomenon suggests that when texts become an additional dimension of the music sequence, the processing mechanism of the texts switches from the language system to the music system. One evidence is that memory deficits patients who struggled to retrieve verbal memory when they were encoded or cued using speech alone showed significant improvement when the texts were learned and retrieved using music. This finding suggested that memorizing lyrics (texts) involves dissociable processors when presented alone versus as a part of music (Simmons-Stern et al., 2010). A possible cognitive mechanism might relate to the extremely regular temporal structure of music that might improve words chunking and modulate attention to

verbal structure by providing temporal scaffold (Conway et al., 2009; Wallace, 1994). Alonso and colleagues conducted multiple neuroimaging tasks to explore neural differences between learning music that was sung with lyrics versus learning music and lyrics presented separately at the same time. They found left IFG, left motor cortex and bilateral medial temporal gyrus (MTG) activation contributing to unified encoding conditions, whereas other areas including the right hippocampus, left caudate and cerebellum circuits and right IFG. during the separate presentation condition (Alonso et al., 2014, 2016). It suggests that when lyrics are learned as a part of the music, less areas' involvement is needed for encoding it, whereas when presented as a separate sequence, it recruits additional brain areas for higher order relational binding of the sequences, such as the hippocampus.

When lyrics are encoded as a part of music, do the same areas processing melodic sequences support encoding the lyrics? Behavioral and neuroscience evidence showed varied insights into this matter. From a behavioral standpoint, although tunes and texts show a bidirectional effect—such that one can cue another during retrieval (Peretz et al., 2004; Peynircioğlu et al., 2008)—some studies suggest that learning the two might compete with each other (Racette & Peretz, 2007), implying encoding the lyrical dimension and melodic dimension of music might share (and compete for) the same neural mechanism. On the other hand, one ERP data recorded while people listened to a song with manipulations of semantic congruency (lyrics congruency) or/and music syntactical congruency. The results showed dissociable neural responses to lyrical semantic violations versus music semantic violations—even though they were played at the same time as an integrated event. This supports a view of independence of lyric and tune processing when listening to sung music (Besson et al., 1998). An earlier lesion study also supported dissociation, highlighting that text recognition might be associated with left temporal lobe structures whereas melodic coding or recognition depended on both temporal lobe hemispheres (Samson & Zatorre, 1991).

Overall, the available evidence suggests that text encoding recruits fewer brain areas when it becomes a contextual associate of music. As a semantic/contextual component of the music itself, it is processed mainly by the music syntactic network, but the MTL may be important for helping bridge linguistic associates of music depending on how they are presented.

Music as context: Episodic memory binding

Music is powerful in a way that it can become a part of our personal memories and sometimes only by hearing it again can effortlessly trigger recollections of long-past personal

experiences (H. Baumgartner, 1992). Music can perform as a strong memory retrieval cue, and multiple prior studies have succeeded applying music in helping both normal people and memory dementia groups such as mild Alzheimer's patients to recall autobiographical memories (Baird & Samson, 2014; Cady et al., 2008; El Haj et al., 2012; Irish et al., 2006). Regarding the cognitive mechanisms behind it, some argued that music can add strong emotional response onto personal experience and people tend to encode and retrieve emotional memories in better efficiency (Belfi et al., 2016; Eschrich et al., 2008; Jäncke, 2008). Others suggested that during retrieval, music listening activated a top-down access to the personal memories by providing a predominant social context (Cady et al., 2008). We currently lacked neural evidence for either hypothesis and limited works investigated the brain mechanisms of music-associated personal episodic memory (autobiographical memory). In particular, no neuroimaging study has been found focusing the encoding phase of autobiographical memory with music played in the background. Future researchers interested in this question can consider comparing episodic memory being encoded with music versus control (e.g., nonmusic sound played in background) and by asking how the neural differences correlate with memory encoding performances in different conditions, people might gain a better understanding of how music-associated episodic memory formation benefit from the “music system” neuronally.

The medial prefrontal cortex (mPFC) seems to be a core area supporting personal memories *retrieval* during music listening—individuals with damages in this area showed selective impairments in recalling music-evoked autobiographical memories but not for face-evoked memories (noted, both cues are emotion-related; Belfi et al., 2016). In general population group, studies have shown increased mPFC activity during autobiographical memories recall with familiar music played (Ford et al., 2011; Janata, 2009). Additionally, during music listening, mPFC has been found to support tracking of melodic dynamics—for example, a topography of tonal space, such as different keys in Western music, was found in rostromedial prefrontal cortex (Janata et al., 2002). Beyond its role in music listening experience, mPFC has a broader function in supporting retrieval of general declarative memory and facilitating the formation of associations between memories (Euston et al., 2012). In this context, the mPFC is positioned to potentially connect the music structure memory with the episodic memory. Taken together, mPFC might specifically support music-evoked autobiographical memories by not only gathering the personal memories but also assembling the structural elements of music with the episodic content paired with it order to form and re-experience a set of episodic/affective collective memory (Janata, 2009).

MTL/hippocampus is a core area for general episodic memories retrieval (Eldridge et al., 2000). Although Janata (2009) did not observe significant MTL engagement during music-evoked autobiographical memories (MEAMs) retrieval (Janata, 2009), some evidence has indeed hinted its involvement in this process—the hippocampus might especially contribute to the episode details and specificity retrieval. Ford et al. (2011) found a positive correlation between the MTL activation and the levels of specificity of the autobiographical memory been retrieved by listeners (Ford et al., 2011), suggesting its role in recruiting details of the associated episodes paired with the music. Kubit and Janata (2018) compared when participants attended to the music familiarity versus when the attention was directed to the associated personal memories and observed a dual network that separately supports these two retrieval processes (Kubit & Janata, 2018). They identified MTL plus the default mode network (DMN), which showed a coupling contribution to episodic memory retrieval (Huijbers et al., 2011), being most engaged during personal memory attending condition whereas a frontal-parietal network with especially the IFG and pre-SMA activation for music attending recall (Kubit & Janata, 2018). The latter finding is aligned with previous discussion that music structural memory retrieval was frontal-parietal focused. In a MEAMs study comparing young and old, the authors found that young people tended to retrieve more details of the personal memories than the old and in the brain the young showed more MTL plus DMN activation whereas the old group mainly recruited the dorso-medial PFC activation (Ford et al., 2016). Behavioral studies suggested that not only young people tended to retrieve more specific events corresponding to music (Ford et al., 2014; Schulkind et al., 1999), music training might also strengthen the specific event–music binding. It was shown that during music listening, musicians tended to recall more personal memories and to relate the music to themselves (Fauvel et al., 2013; Groussard et al., 2010). MRI data revealed more hippocampal and amygdala recruitment in musicians and it was postulated by the authors that the MTL's involvement was related to its function in memory imagery and emotional contextualization during music listening (Alluri et al., 2015; Frühholz et al., 2014). Comparing musicians versus nonmusicians during familiar music listening, Groussard et al. (2010) found more anterior hippocampus as well as entorhinal cortex and parahippocampal involvement in musicians and suggested that hippocampus and other MTL areas engaged in episodic recollection processing during familiar music listening (Groussard et al., 2010). Other relevant studies compared listening to self-selected music with unfamiliar or recent learned music and exhibited hippocampal, parahippocampal and cerebellar activity only in self-selected familiar music listening (Thaut et al., 2020; Wu et al., 2019),

indicating their role in recollecting self-related emotions and memories triggered by self-selected music.

Lastly, hippocampus might also interact with the frontal cortex to retrieve the autobiographical details as a part of the music (context). The connectivity between the two areas has been highlighted in other episodic memory retrieval tasks such as spatial retrieval (Preston & Eichenbaum, 2013). In a familiar music listening task, the researchers used dynamic causal modeling to test the direction and interaction between brain areas modulated by different music familiarity levels. They observed (1) left hippocampal activation during listening to familiar music only in musicians, suggesting more episodic recruitment in this group; (2) a top-down connectivity projected from left IFG to left hippocampus modulated by music familiarity (Gagnepain et al., 2017), suggesting that during music-associated memory retrieval, hippocampal recruitment is modulated online by the retrieval of episodic memory traces whereas IFG might support music structure identification as well as project this familiarity information into the hippocampus and drive it to retrieve episodic memory (shown in Fig. 1). It provided evidence for the connectivity and communication between the two networks separately supporting “music syntactical structure memory” and “music contextual associated episodic memory” that together develop a vivid and self-related experience of music memory retrieval.

Emotion and reward

One remarkable aspect of music is its ability to express and evoke emotion. Emotion is usually an impartible component of memory not only because it colors and defines the subjective perception of the memory but also due to its power to enhance memorability and recall of the memory (Kensinger & Schacter, 2006; Tyng et al., 2017). Extensive research has explored music emotion in terms of how the association between different types of tonal structure and emotions was developed both behaviorally and neuronally. Here, we aim to discuss how the memory of emotion is encoded as well as how the evoked emotion interacts with the memory of music itself, both the structural memory and other contextual associative memory.

The emotions induced by music are highly relevant to listeners' ability to anticipate music's structure elements. The musical syntax knowledge allows listeners to predict the ongoing music and by comparing the reality and the prediction, the listeners can generate reward-related emotions (Tillmann et al., 2014). For example, the feeling of pleasantness is induced by whether the music follows the syntactical rules and in many studies, the researchers created unpleasant music by violating the music regularity (Blood et al., 1999; Koelsch et al., 2006; Tillmann et al., 2014). When comparing manipulated music versus original regular music

using fMRI, several studies observed significant changes in limbic and reward systems including but not limited to hippocampus, amygdala, parahippocampus, ventral striatum mainly the nucleus accumbens (NAc), insula and orbitofrontal cortex (Blood et al., 1999; Brown et al., 2004; Koelsch et al., 2006; Mueller et al., 2015). Among these areas, ventral striatum represented reward system seems to be pivotal in comparing expectation with the reality, which may lead to dopamine release and result in providing the rewarding and emotional value to the music stimuli (Salimpoor et al., 2011, 2013). Another study observed hippocampus exhibited similar activity to the ventral striatum participants listening to manipulated music—both were correlated with the music pleasantness (Mueller et al., 2015). This hippocampal involvement might contribute to a stronger encoding of pleasant music, regulated by the hippocampal-striatum pathway (Lisman & Grace, 2005) via dopaminergic (Salimpoor et al., 2013) and serotonergic mechanisms (Evers & Suhr, 2000). Both neurotransmitters play important roles in general learning and memory processes (Meneses & Liy-Salmeron, 2012; Wise, 2004). This music-emotion effect on memory strength might even expand to nonmusic memory. For example, it has shown that musical pleasure and hedonia can improve encoding performance for verbal episodic memory with music in the background, potentially supported by the interaction between ventral striatum/NAc and the hippocampus in a way that the dopamine release could strengthen long-term potentiation in the hippocampus, suggested by the authors (Cardona et al., 2020).

In addition to the reward system, the amygdala and hippocampus together also help distinguish and respond to the pleasant versus unpleasant music, manipulated through syntax violations (Koelsch et al., 2006)—both areas show significantly increased responses to negative music. One lesion study found that temporal lobe epilepsy patients who had amygdala and hippocampal damages could not recognize sad music and moreover, they did not show better memory for emotional music than neutral music, suggesting the role of these areas in encoding contextual emotional aspects of music memory (Samson et al., 2009). Amygdala and hippocampus also interact with the “music syntax network” in perceiving and measuring the uncertainty and surprise during musical sequential processing, suggested by a modeling study (Cheung et al., 2019). They pointed out that the amygdala–hippocampal interaction was essential in deciding the valence of the music, whereas NAc modulated the motivation of subsequent information perceiving and prediction. While amygdala and hippocampus particularly respond to the unpleasant and surprising music, music that follows prediction activated the areas relating to syntax processing including the Heschl's gyrus (part of auditory cortex) and left IFG (Koelsch et al., 2006). Music that makes people extremely satisfied and intensely pleasurable may also

activate frontal areas such as the orbitofrontal lobe (Blood & Zatorre, 2001). Overall, the rewarding system together with the limbic system might support the “music syntax network” to better encode music sequences by providing rewarding feedback.

In addition to the emotional responses modulated by music structure or syntax, the emotion conveyed by the music can sometimes be more complicated, extending beyond a simple dichotomy of pleasantness and unpleasantness (Leonard, 1956). The perception of music emotion relates to the memory in a way that the emotion adds semantic content to the music—it become a part of the story a piece of music conveys. Firstly, the emotional semantic component strongly correlates with the memorability of the music itself. For example, music excerpts with stronger emotional valence and arousal dimensions are better remembered (Ferreri & Rodriguez-Fornells, 2017). The MTL/limbic system plays the major role in perceiving and identifying emotion expressed by the music, especially complex emotions, such as pride and guilt (Frühholz et al., 2014). The hippocampus, in particular, can support temporal integration of the complex emotional semantic information and provides a contextual memory association (Fig. 1), which can further strengthen the memory of the music itself (Frühholz et al., 2014). Relating to its role in episodic memory encoding, hippocampus support self-related memory association such as scenario imagery during the music-evoked emotional experiences (Alluri et al., 2015). Evidence includes musicians tending to imagine themselves in a personal experience during music listening, and showing more hippocampus activation (Groussard et al., 2010). Amygdala and anterior hippocampus activation were also observed when emotional music was paired with neutral film clips but not during solely music listening (Eldar et al., 2007), suggesting that the existence of visual context implements the semantic/contextual association and these areas might support encoding this association during music memory encoding. Another study observed similar neural activity when people listened the music with closed eyes but not opened eyes, possibly linking to a semantic imaginary processing during music encoding (Frühholz et al., 2014; Lerner et al., 2009). In the meta-analysis, Koelsch (2020) discussed the role of the hippocampus in music emotion perception and experience. He suggested that the right hippocampus supports the perception of attachment-related emotions associated with the social bonding such as empathy function whereas left hippocampus reacts to only negative emotion of the music relating to the familiarity and expectation of the music (Koelsch, 2020). It added evidence to a frequent debate in music literature about the differences between emotion perception and (felt) emotional responses to music (Gabrielsson, 2001). While the evoked pleasantness or unpleasantness correlates to whether the music structural pattern follows our syntactical knowledge,

the perceived emotion relates more to the understanding and mirroring of the semantic meaning embedded within the tonal variation of the music. Other than the hippocampus, IFG, the music syntax area might also support higher level emotional/semantic meaning understanding. From nonmusic tasks, IFG was found to involve in empathy and social perception such as facial emotion perception, via the mirror neurons mechanism (Jabbi & Keysers, 2008; Kaplan & Iacoboni, 2006; Press et al., 2012). The mirror neurons enabled individual to understand the social meaning (such as intention and emotion) of the communication signal perceived from others (Kilner et al., 2009). Several music literatures identified this mechanism revealed by left IFG activity during music listening and music performance watching and they concluded that IFG was also engaged in understanding the semantic intention of the music to better perceive and experience the emotion induced from the music (D’Ausilio, 2009; Molnar-Szakacs & Overy, 2006).

The emotion of music not only adds semantic meaning to the music, it also provides contextual association to non-music information encoded with music. For example, emotional music played in the background implemented visual encoding including pictures and faces (T. Baumgartner et al., 2006; Jolij & Meurs, 2011; Proverbio et al., 2015). Other than visual memory, pleasant music also benefited verbal episodic memory encoding and recollection (Cardona et al., 2020). Some mechanism behind might relate to phase synchronization, a mechanism found in both human and rats relating to memory formation (Fell & Axmacher, 2011). Increased frontal theta power was observed during consonant music listening compared with dissonant music (Sammler et al., 2007). An increased synchronization between right temporal and right frontal cortex was also found to correlate with the degree of pleasantness the listeners reported to the background music (Ara & Marco-Pallarés, 2020). The memory retrieval and working memory are tightly related to such frontal oscillation (Schack et al., 2005) and thus pleasant music listening might promote memory function for information that learned in a musical context via such a mechanism. Future studies aiming at utilizing music as memory aids should consider further investigating such emotional power. For example, some suggest the famous but debated “Mozart’s effect” (that listening to certain music can improve cognitive function including attention and memory) might relate to the arousal and emotion evoked by background music (Schellenberg & Weiss, 2013).

In summary, identifying and experiencing emotion during music listening can enhance both music structure memory and associated contextual memory. Likely via modulating communication between ‘music syntax network’, the reward system, and the MTL (Fig. 1), emotion might enhance deeper encoding of episodic-like contextual associations during music memory formation. These effects, to be studied

further, could provide a powerful means for music memory to modulate other memory in applied settings.

Conclusions

Decades of music cognition studies have provided comprehensive evidence for what types of mechanisms are involved in music processing. However, there is still a lack of consistent definitions of what constitutes music memory, and a lack of consistent data on how it is formed in the brain. Importantly, few prior studies have investigated how the association between music memory and general long-term memory functions is mediated. This Review proposes a restructuring of how music memory is conceptualized for the purpose of future neuroscience research. We argue that music memory is a multidimensional and hierarchical event, within which different components might rely on different neural networks. By virtue of music's complexity, these do not fall cleanly along traditional boundaries between different forms of memory from other literature. The review presents neuroscience results organized according to (1) memory formation and (2) memory retrieval for (a) music syntactical memory (music structure memory) and (b) memory for associated nonmusic elements (music contextual associates memory). We believe it is clear that while auditory processing cortices play a central role in music-listening memory tasks, there is an expansion of this network to other areas in different tasks, with the network components brought in being modulated by levels of memory complexity. As much music literature has pointed out, music is a powerful event that requires less subcortical involvement during memory encoding and retrieval. Its unique strictly rule-followed features may tax a more restricted network during structure preprocessing, which is mainly supported by the inferior frontal gyrus and auditory cortex, the same brain areas associated with language comprehension. For music structure memory encoding, we suggest there is sufficient evidence for a shared neural circuit for it with other forms of memory—other than the “syntactic network” (IFG, auditory cortex, mainly the STG, IPL), this level of music learning also involves a neural circuit that includes the medial temporal lobe and subcortical structures such as the basal ganglia and cerebellum that generally support statistical learning, time perception/encoding and sequential learning for nonmusic events. Importantly, although some lesion literature implies that music structure memory is independent of the hippocampus, this review found numerous datapoints suggesting that music syntactical/structural memory can rely on the same medial temporal lobe mechanisms supporting declarative sequence memories of other forms. Corresponding to its role as a hierarchical sequence

associator from nonmusic memory paradigms (McKenzie et al., 2014), the hippocampus may help bind components from different dimensions and across time to help form integrated representation of the music. We argue this may be especially true during the initial stage of novel sequential pattern detection and encoding, while reviewing why it is very uncommon for a music composition to be truly novel to the brain's memory systems. In particular, one reason for evidence of “hippocampal independence” suggested by many music memory studies could be that music learning typically benefits from learned syntax and our “prior knowledge” schematic memory system, which different memory literatures hold to be supported by frontal areas and their interactions with cortex outside of MTL.

On the other hand, it is surprisingly rare for neuroimaging studies to have looked into music memory on the level of semantic declarative memory—the story it conveys—or in terms of how it becomes associated with information from other domains (emotions, episodic experience). We provide an overview of the current knowledge of neural correlates for memory of music paired with lyrics, music associated episodic events, and memory for the paired emotion. Overall, evidence suggests there is a dispersed network for this level of memory beyond the core “syntax network” that enables us to encode and retrieve music contextual associates. Areas including the limbic system and reward circuitry became involved when the memory for music demands higher levels of semantic, emotional, and autobiographical processing. Importantly, the semantic/contextual associations formed by these subcortical areas can conversely enhance the encoding and retrieval of music structural memory itself. As a result, one important take-home from the literature is that music memory is special, not only because it is associated with a schematic syntactical system that supports highly efficient and often-effortless learning but also due to the interactions between its syntactical and semantic representations: the higher-level of semantic and emotional meaning of music provides a scaffold for personal and contextual associations for the listeners. This interaction is reflected in the brain putatively through a complex cortical–subcortical collaboration (Fig. 1). Admittedly, this conclusion is derived from a somewhat limited body of neuroscience results. We hope that future research will focus on understanding this interaction, in part because it might be further applied to aid people in better encoding and retrieving of memories of other aspects of their lives (a potential already being explored in critical cases such as Alzheimer's). By bridging the gap between mechanisms for music processing and nonmusic long-term memory formation, we might be able to design “music listening experience” to better benefit our daily life.

Open practices statement None data or experiments were reported here, and none of the materials was preregistered.

References

- Aggleton, J. P., Dumont, J. R., & Warburton, E. C. (2011). Unraveling the contributions of the diencephalon to recognition memory: A review. *Learning & Memory, 18*(6), 384–400.
- Alluri, V., Brattico, E., Toiviainen, P., Burunat, I., Bogert, B., Numminen, J., & Kliuchko, M. (2015). Musical expertise modulates functional connectivity of limbic regions during continuous music listening. *Psychomusicology: Music, Mind, and Brain, 25*(4), Article 443.
- Alonso, I., Davachi, L., Valabrègue, R., Lambrecq, V., Dupont, S., & Samson, S. (2016). Neural correlates of binding lyrics and melodies for the encoding of new songs. *NeuroImage, 127*, 333–345.
- Alonso, I., Sammler, D., Valabrègue, R., Dinkelacker, V., Dupont, S., Belin, P., & Samson, S. (2014). Hippocampal sclerosis affects fMR-adaptation of lyrics and melodies in songs. *Frontiers in Human Neuroscience, 8*, 111.
- Altenmüller, E. O. (2001). How many music centers are in the brain? *Annals of the New York Academy of Sciences, 930*(1), 273–280.
- Altenmüller, E., Siggel, S., Mohammadi, B., Samii, A., & Münte, T. F. (2014). Play it again, Sam: Brain correlates of emotional music recognition. *Frontiers in Psychology, 5*, 114.
- Ara, A., & Marco-Pallarés, J. (2020). Fronto-temporal theta phase-synchronization underlies music-evoked pleasantness. *NeuroImage, 212*, 116665.
- Assmus, A., Marshall, J. C., Ritzl, A., Noth, J., Zilles, K., & Fink, G. R. (2003). Left inferior parietal cortex integrates time and space during collision judgments. *NeuroImage, 20*, S82–S88.
- Baird, A., & Samson, S. (2009). Memory for music in Alzheimer's disease: Unforgettable? *Neuropsychology Review, 19*(1), 85–101.
- Baird, A., & Samson, S. (2014). Music evoked autobiographical memory after severe acquired brain injury: Preliminary findings from a case series. *Neuropsychological Rehabilitation, 24*(1), 125–143.
- Barbey, A. K., Koenigs, M., & Grafman, J. (2013). Dorsolateral prefrontal contributions to human working memory. *Cortex, 49*(5), 1195–1205.
- Bartlett, J. C., Halpern, A. R., & Dowling, W. J. (1995). Recognition of familiar and unfamiliar melodies in normal aging and Alzheimer's disease. *Memory & Cognition, 23*(5), 531–546.
- Baumgartner, H. (1992). Remembrance of things past: Music, autobiographical memory, and emotion. *ACR North American Advances*.
- Baumgartner, T., Lutz, K., Schmidt, C. F., & Jäncke, L. (2006). The emotional power of music: How music enhances the feeling of affective pictures. *Brain Research, 1075*(1), 151–164.
- Belfi, A. M., Karlan, B., & Tranel, D. (2016). Music evokes vivid autobiographical memories. *Memory, 24*(7), 979–989.
- Bemis, D. K., & Pylkkänen, L. (2013). Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. *Cerebral Cortex, 23*(8), 1859–1873.
- Besson, M., Faita, F., Peretz, I., Bonnel, A.-M., & Requin, J. (1998). Singing in the brain: Independence of lyrics and tunes. *Psychological Science, 9*(6), 494–498.
- Bigand, E., & Poulin-Charronnat, B. (2006). Are we “experienced listeners”? A review of the musical capacities that do not depend on formal musical training. *Cognition, 100*(1), 100–130.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Rao, S. M., & Cox, R. W. (1996). Function of the left planum temporale in auditory and linguistic processing. *Brain, 119*(4), 1239–1247.
- Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences, 98*(20), 11818–11823.
- Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience, 2*(4), 382–387.
- Bonetti, L., Brattico, E., Carlomagno, F., Donati, G., Cabral, J., Haumann, N. T., Deco, G., Vuust, P., & Kringelbach, M. L. (2021). Rapid encoding of musical tones discovered in whole-brain connectivity. *NeuroImage, 245*, 118735.
- Bravo, F., Cross, I., Hopkins, C., Gonzalez, N., Docampo, J., Bruno, C., & Stamatakis, E. A. (2020). Anterior cingulate and medial prefrontal cortex response to systematically controlled tonal dissonance during passive music listening. *Human Brain Mapping, 41*(1), 46–66.
- Brown, S., Martinez, M. J., & Parsons, L. M. (2004). Passive music listening spontaneously engages limbic and paralimbic systems. *NeuroReport, 15*(13), 2033–2037.
- Burunat, I., Alluri, V., Toiviainen, P., Numminen, J., & Brattico, E. (2014). Dynamics of brain activity underlying working memory for music in a naturalistic condition. *Cortex, 57*, 254–269.
- Burunat, I., Brattico, E., Hartmann, M., Vuust, P., Särkämö, T., & Toiviainen, P. (2018). Musical training predicts cerebello-hippocampal coupling during music listening. *Psychomusicology: Music, Mind, and Brain, 28*(3), 152–163.
- Cady, E. T., Harris, R. J., & Knappenberger, J. B. (2008). Using music to cue autobiographical memories of different lifetime periods. *Psychology of Music, 36*(2), 157–177.
- Campbell, D. W., Wallace, M. G., Modirrousta, M., Polimeni, J. O., McKeen, N. A., & Reiss, J. P. (2015). The neural basis of humour comprehension and humour appreciation: The roles of the temporoparietal junction and superior frontal gyrus. *Neuropsychologia, 79*, 10–20.
- Cardona, G., Rodriguez-Fornells, A., Nye, H., Rifà-Ros, X., & Ferreri, L. (2020). The impact of musical pleasure and musical hedonia on verbal episodic memory. *Scientific Reports, 10*(1), 1–13.
- Cavaco, S., Feinstein, J. S., van Twillert, H., & Tranel, D. (2012). Musical memory in a patient with severe anterograde amnesia. *Journal of Clinical and Experimental Neuropsychology, 34*(10), 1089–1100.
- Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex, 18*(12), 2844–2854.
- Cheung, V. K. M., Harrison, P. M. C., Meyer, L., Pearce, M. T., Haynes, J.-D., & Koelsch, S. (2019). Uncertainty and surprise jointly predict musical pleasure and amygdala, hippocampus, and auditory cortex activity. *Current Biology, 29*(23), 4084–4092.
- Conway, C. M., Pisoni, D. B., & Kronenberger, W. G. (2009). The importance of sound for cognitive sequencing abilities. *Current Directions in Psychological Science, 18*(5), 275–279.
- Corkin, S. (1968). Acquisition of motor skill after bilateral medial temporal-lobe excision. *Neuropsychologia, 6*(3), 255–265.
- Covington, N. V., Brown-Schmidt, S., & Duff, M. C. (2018). The necessity of the hippocampus for statistical learning. *Journal of Cognitive Neuroscience, 30*(5), 680–697.
- Cowles, A., Beatty, W. W., Nixon, S. J., Lutz, L. J., Paulk, J., Paulk, K., & Ross, E. D. (2003). Musical skill in dementia: A violinist presumed to have Alzheimer's disease learns to play a new song. *Neurocase, 9*(6), 493–503.
- D'Ausilio, A. (2009). Mirror-like mechanisms and music. *The Scientific World Journal, 9*, 1415–1422.
- Demorest, S. M., Morrison, S. J., Nguyen, V. Q., & Bodnar, E. N. (2016). The influence of contextual cues on cultural bias in music

- memory. *Music Perception: An Interdisciplinary Journal*, 33(5), 590–600.
- Demorest, S. M., Morrison, S. J., Stambaugh, L. A., Beken, M., Richards, T. L., & Johnson, C. (2010). An fMRI investigation of the cultural specificity of music memory. *Social Cognitive and Affective Neuroscience*, 5(2/3), 282–291.
- Deutsch, D. (1969). Music recognition. *Psychological Review*, 76(3), 300.
- Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review*, 88(6), 503.
- Drake, C., & El Heni, J. B. (2003). Synchronizing with music: Intercultural differences. *Annals of the New York Academy of Sciences*, 999(1), 429–437.
- Duarte, A., Ranganath, C., Winward, L., Hayward, D., & Knight, R. T. (2004). Dissociable neural correlates for familiarity and recollection during the encoding and retrieval of pictures. *Cognitive Brain Research*, 18(3), 255–272.
- Eichenbaum, H. (2014). Time cells in the hippocampus: A new dimension for mapping memories. *Nature Reviews Neuroscience*, 15(11), 732–744.
- El Haj, M., Postal, V., & Allain, P. (2012). Music enhances autobiographical memory in mild Alzheimer's disease. *Educational Gerontology*, 38(1), 30–41.
- Eldar, E., Ganor, O., Admon, R., Bleich, A., & Hendler, T. (2007). Feeling the real world: Limbic response to music depends on related content. *Cerebral Cortex*, 17(12), 2828–2840.
- Eldridge, L. L., Knowlton, B. J., Furmanski, C. S., Bookheimer, S. Y., & Engel, S. A. (2000). Remembering episodes: A selective role for the hippocampus during retrieval. *Nature Neuroscience*, 3(11), 1149–1152.
- Eschrich, S., Münte, T. F., & Altenmüller, E. O. (2008). Unforgettable film music: The role of emotion in episodic long-term memory for music. *BMC Neuroscience*, 9(1), 1–7.
- Esfahani-Bayerl, N., Finke, C., Kopp, U., Moon, D.-U., & Ploner, C. J. (2019). Musical memory and hippocampus revisited: Evidence from a musical layperson with highly selective hippocampal damage. *Cortex*, 119, 519–527.
- Euston, D. R., Gruber, A. J., & McNaughton, B. L. (2012). The role of medial prefrontal cortex in memory and decision making. *Neuron*, 76(6), 1057–1070.
- Evers, S., & Suhr, B. (2000). Changes of the neurotransmitter serotonin but not of hormones during short time music perception. *European Archives of Psychiatry and Clinical Neuroscience*, 250(3), 144–147.
- Fang, X., Singh, S., & Ahluwalia, R. (2007). An examination of different explanations for the mere exposure effect. *Journal of Consumer Research*, 34(1), 97–103.
- Fauvel, B., Groussard, M., Eustache, F., Desgranges, B., & Platel, H. (2013). Neural implementation of musical expertise and cognitive transfers: Could they be promising in the framework of normal cognitive aging? *Frontiers in Human Neuroscience*, 7, 693.
- Fell, J., & Axmacher, N. (2011). The role of phase synchronization in memory processes. *Nature Reviews Neuroscience*, 12(2), 105–118.
- Ferreri, L., & Rodriguez-Fornells, A. (2017). Music-related reward responses predict episodic memory performance. *Experimental Brain Research*, 235(12), 3721–3731.
- Finke, C., Esfahani, N. E., & Ploner, C. J. (2012). Preservation of musical memory in an amnesic professional cellist. *Current Biology*, 22(15), R591–R592.
- Ford, J. H., Rubin, D. C., & Giovanello, K. S. (2016). The effects of song familiarity and age on phenomenological characteristics and neural recruitment during autobiographical memory retrieval. *Psychomusicology: Music, Mind, and Brain*, 26(3), 199.
- Ford, J. H., Addis, D. R., & Giovanello, K. S. (2011). Differential neural activity during search of specific and general autobiographical memories elicited by musical cues. *Neuropsychologia*, 49(9), 2514–2526.
- Ford, J. H., Rubin, D. C., & Giovanello, K. S. (2014). Effects of task instruction on autobiographical memory specificity in young and older adults. *Memory*, 22(6), 722–736.
- Freitas, C., Manzano, E., Burini, A., Taylor, M. J., Lerch, J. P., & Anagnostou, E. (2018). Neural correlates of familiarity in music listening: A systematic review and a neuroimaging meta-analysis. *Frontiers in Neuroscience*, 12, 686.
- Frühholz, S., Trost, W., & Grandjean, D. (2014). The role of the medial temporal limbic system in processing emotions in voice and music. *Progress in Neurobiology*, 123, 1–17.
- Gaab, N., Gaser, C., Zaehle, T., Jancke, L., & Schlaug, G. (2003). Functional anatomy of pitch memory—An fMRI study with sparse temporal sampling. *NeuroImage*, 19(4), 1417–1426.
- Gabrielsson, A. (2001). Emotion perceived and emotion felt: Same or different? *Musicae Scientiae*, 5(1_Suppl), 123–147.
- Gagnepain, P., Fauvel, B., Desgranges, B., Gaubert, M., Viader, F., Eustache, F., Groussard, M., & Platel, H. (2017). Musical expertise increases top-down modulation over hippocampal activation during familiarity decisions. *Frontiers in Human Neuroscience*, 11, 472.
- Gheysen, F., & Fias, W. (2012). Dissociable neural systems of sequence learning. *Advances in Cognitive Psychology*, 8(2), 73.
- Ginsborg, J. (2004). Strategies for memorizing music. In A. Williamson (Ed.), *Musical excellence: Strategies and techniques to enhance performance* (pp. 123–141). Oxford Academic Press.
- Grahn, J. A. (2009). The role of the basal ganglia in beat perception: Neuroimaging and neuropsychological investigations. *Annals of the New York Academy of Sciences*, 1169(1), 35–45.
- Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, 19(5), 893–906.
- Green, A. C., Bærentsen, K. B., Stødkilde-Jørgensen, H., Roepstorff, A., & Vuust, P. (2012). Listen, learn, like! Dorsolateral prefrontal cortex involved in the mere exposure effect in music. *Neurology Research International*, 2012.
- Groussard, M., Rauchs, G., Landeau, B., Viader, F., Desgranges, B., Eustache, F., & Platel, H. (2010). The neural substrates of musical memory revealed by fMRI and two semantic tasks. *NeuroImage*, 53(4), 1301–1309.
- Gupta, A. S., Van Der Meer, M. A. A., Touretzky, D. S., & Redish, A. D. (2012). Segmentation of spatial experience by hippocampal theta sequences. *Nature Neuroscience*, 15(7), 1032–1039.
- Halpern, A. R., & O'Connor, M. G. (2000). Implicit memory for music in Alzheimer's disease. *Neuropsychology*, 14(3), 391.
- Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science*, 16(1), 48–55.
- Hasselmo, M. E. (2006). The role of acetylcholine in learning and memory. *Current Opinion in Neurobiology*, 16(6), 710–715.
- Haxby, J. V., Ungerleider, L. G., Horwitz, B., Maisog, J. M., Rapoport, S. I., & Grady, C. L. (1996). Face encoding and recognition in the human brain. *Proceedings of the National Academy of Sciences*, 93(2), 922–927.
- Herdener, M., Esposito, F., di Salle, F., Boller, C., Hilti, C. C., Habermeyer, B., Scheffler, K., Wetzel, S., Seifritz, E., & Cattapan-Ludewig, K. (2010). Musical training induces functional plasticity in human hippocampus. *Journal of Neuroscience*, 30(4), 1377–1384.
- Hsieh, L.-T., Gruber, M. J., Jenkins, L. J., & Ranganath, C. (2014). Hippocampal activity patterns carry information about objects in temporal context. *Neuron*, 81(5), 1165–1178.
- Huijbers, W., Pennartz, C. M. A., Cabeza, R., & Daselaar, S. M. (2011). The hippocampus is coupled with the default network during

- memory retrieval but not during memory encoding. *PLOS ONE*, 6(4), e17463.
- Huijgen, J., Dellacherie, D., Tillmann, B., Clément, S., Bigand, E., Dupont, S., & Samson, S. (2015). The feeling of familiarity for music in patients with a unilateral temporal lobe lesion: A gating study. *Neuropsychologia*, 77, 313–320.
- Irish, M., Cunningham, C. J., Walsh, J. B., Coakley, D., Lawlor, B. A., Robertson, I. H., & Coen, R. F. (2006). Investigating the enhancing effect of music on autobiographical memory in mild Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, 22(1), 108–120.
- Jabbi, M., & Keysers, C. (2008). Inferior frontal gyrus activity triggers anterior insula response to emotional facial expressions. *Emotion*, 8(6), 775.
- Jablonowski, J., Taesler, P., Fu, Q., & Rose, M. (2018). Implicit acoustic sequence learning recruits the hippocampus. *PLOS ONE*, 13(12), e0209590.
- Jacobsen, J.-H., Stelzer, J., Fritz, T. H., Chételat, G., La Joie, R., & Turner, R. (2015). Why musical memory can be preserved in advanced Alzheimer's disease. *Brain*, 138(8), 2438–2450.
- Janata, P. (2009). The neural architecture of music-evoked autobiographical memories. *Cerebral Cortex*, 19(11), 2579–2594.
- Janata, P., Birk, J. L., Van Horn, J. D., Leman, M., Tillmann, B., & Bharucha, J. J. (2002). The cortical topography of tonal structures underlying Western music. *Science*, 298(5601), 2167–2170.
- Janata, P., & Grafton, S. T. (2003). Swinging in the brain: Shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience*, 6(7), 682–687.
- Jäncke, L. (2008). Music, memory and emotion. *Journal of Biology*, 7(6), 1–5.
- Jäncke, L. (2019). Music and Memory. *The Oxford Handbook of Music and the Brain*, 237.
- Jäncke, L., Specht, K., Shah, J. N., & Hugdahl, K. (2003). Focused attention in a simple dichotic listening task: An fMRI experiment. *Cognitive Brain Research*, 16(2), 257–266.
- Jarret, T., Stockert, A., Kotz, S. A., & Tillmann, B. (2019). Implicit learning of artificial grammatical structures after inferior frontal cortex lesions. *PLOS ONE*, 14(9), e0222385.
- Jentschke, S., Koelsch, S., Sallat, S., & Friederici, A. D. (2008). Children with specific language impairment also show impairment of music-syntactic processing. *Journal of Cognitive Neuroscience*, 20(11), 1940–1951.
- Johnson-Laird, P. N., Kang, O. E., & Leong, Y. C. (2012). On musical dissonance. *Music Perception: An Interdisciplinary Journal*, 30(1), 19–35.
- Jolij, J., & Meurs, M. (2011). Music alters visual perception. *PLOS ONE*, 6(4), e18861.
- Kalda, T., & Minati, L. (2012). Detecting scale violations in absence of mismatch requires music-syntactic analysis: A further look at the early right anterior negativity (ERAN). *Brain Topography*, 25(3), 285–292.
- Kaplan, J. T., & Iacoboni, M. (2006). Getting a grip on other minds: Mirror neurons, intention understanding, and cognitive empathy. *Social Neuroscience*, 1(3/4), 175–183.
- Kensinger, E. A., & Schacter, D. L. (2006). Amygdala activity is associated with the successful encoding of item, but not source, information for positive and negative stimuli. *Journal of Neuroscience*, 26(9), 2564–2570.
- Kilner, J. M., Neal, A., Weiskopf, N., Friston, K. J., & Frith, C. D. (2009). Evidence of mirror neurons in human inferior frontal gyrus. *Journal of Neuroscience*, 29(32), 10153–10159.
- Klostermann, E. C., Loui, P., & Shimamura, A. P. (2009). Activation of right parietal cortex during memory retrieval of nonlinguistic auditory stimuli. *Cognitive, Affective, & Behavioral Neuroscience*, 9(3), 242–248.
- Knösche, T. R., Neuhaus, C., Haueisen, J., Alter, K., Maess, B., Witte, O. W., & Friederici, A. D. (2005). Perception of phrase structure in music. *Human Brain Mapping*, 24(4), 259–273.
- Koch, G., Oliveri, M., Torriero, S., Salerno, S., Gerfo, E. L., & Caltagirone, C. (2007). Repetitive TMS of cerebellum interferes with millisecond time processing. *Experimental Brain Research*, 179(2), 291–299.
- Koelsch, S. (2006). Significance of Broca's area and ventral premotor cortex for music-syntactic processing. *Cortex*, 42(4), 518–520.
- Koelsch, S. (2009). Neural substrates of processing syntax and semantics in music. In R. Haas & V. Brandes (Eds.), *Music that works* (pp. 143–153). Springer.
- Koelsch, S. (2020). A coordinate-based meta-analysis of music-evoked emotions. *NeuroImage*, 223, 117350.
- Koelsch, S., Fritz, T., & Schlaug, G. (2008). Amygdala activity can be modulated by unexpected chord functions during music listening. *NeuroReport*, 19(18), 1815–1819.
- Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005a). Adults and children processing music: An fMRI study. *NeuroImage*, 25(4), 1068–1076.
- Koelsch, S., Gunter, T. C., Wittfoth, M., & Sammler, D. (2005b). Interaction between syntax processing in language and in music: An ERP study. *Journal of Cognitive Neuroscience*, 17(10), 1565–1577.
- Koelsch, S., Fritz, T., & v. Cramon, D. Y., Müller, K., & Friederici, A. D. (2006). Investigating emotion with music: An fMRI study. *Human Brain Mapping*, 27(3), 239–250.
- Koelsch, S., Jentschke, S., Sammler, D., & Mietschen, D. (2007). Untangling syntactic and sensory processing: An ERP study of music perception. *Psychophysiology*, 44(3), 476–490.
- Koelsch, S., Maess, B., & Friederici, A. D. (2000). Musical syntax is processed in the area of Broca: An MEG study. *NeuroImage*, 11(5), S56.
- Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Müller, K., & Gruber, O. (2009). Functional architecture of verbal and tonal working memory: An FMRI study. *Human Brain Mapping*, 30(3), 859–873.
- Konoike, N., Kotozaki, Y., Miyachi, S., Miyauchi, C. M., Yomogida, Y., Akimoto, Y., Kuraoka, K., Sugiura, M., Kawashima, R., & Nakamura, K. (2012). Rhythm information represented in the fronto-parieto-cerebellar motor system. *NeuroImage*, 63(1), 328–338.
- Krumhansl, C. L. (1991). Music psychology: Tonal structures in perception and memory. *Annual Review of Psychology*, 42(1), 277–303.
- Krumhansl, C. L. (2002). Music: A link between cognition and emotion. *Current Directions in Psychological Science*, 11(2), 45–50.
- Kubit, B., & Janata, P. (2018). Listening for memories: Attentional focus dissociates functional brain networks engaged by memory-evoking music. *Psychomusicology: Music, Mind, and Brain*, 28(2), 82.
- Kuck, H., Grossbach, M., Bangert, M., & Altenmüller, E. (2003). Brain processing of meter and rhythm in music: Electrophysiological evidence of a common network. *Annals of the New York Academy of Sciences*, 999(1), 244–253.
- Kunert, R., Willems, R. M., Casasanto, D., Patel, A. D., & Hagoort, P. (2015). Music and language syntax interact in Broca's area: An fMRI study. *PLOS ONE*, 10(11), e0141069.
- LaBar, K. S., & Phelps, E. A. (1998). Arousal-mediated memory consolidation: Role of the medial temporal lobe in humans. *Psychological Science*, 9(6), 490–493.
- Lappe, C., Lappe, M., & Pantev, C. (2016). Differential processing of melodic, rhythmic and simple tone deviations in musicians-an MEG study. *NeuroImage*, 124, 898–905.

- Lappe, C., Steinsträter, O., & Pantev, C. (2013). Rhythmic and melodic deviations in musical sequences recruit different cortical areas for mismatch detection. *Frontiers in Human Neuroscience*, 7, 260.
- Lee, K.-H., Egleston, P. N., Brown, W. H., Gregory, A. N., Barker, A. T., & Woodruff, P. W. R. (2007). The role of the cerebellum in subsecond time perception: Evidence from repetitive transcranial magnetic stimulation. *Journal of Cognitive Neuroscience*, 19(1), 147–157.
- Leman, M. (2012). *Music and schema theory: Cognitive foundations of systematic musicology* (Vol. 31). Springer Science & Business Media.
- Leonard, M. (1956). *Emotion and meaning in music*. University of Chicago Press.
- Lerner, Y., Papo, D., Zhdanov, A., Belozersky, L., & Hendler, T. (2009). Eyes wide shut: Amygdala mediates eyes-closed effect on emotional experience with music. *PLOS ONE*, 4(7), e6230.
- Levitin, D. J., & Tirovolas, A. K. (2009). Current advances in the cognitive neuroscience of music. *Annals of the New York Academy of Sciences*, 1156(1), 211–231.
- Levy, R., & Goldman-Rakic, P. S. (2000). Segregation of working memory functions within the dorsolateral prefrontal cortex. In W. X. Schneider, A. M. Owen, & J. Duncan (Eds.), *Executive control and the frontal lobe: Current issues* (pp. 23–32). Springer.
- Limb, C. J., Kemeny, S., Ortigoza, E. B., Rouhani, S., & Braun, A. R. (2006). Left hemispheric lateralization of brain activity during passive rhythm perception in musicians. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology: An Official Publication of the American Association of Anatomists*, 288(4), 382–389.
- Lisman, J. E., & Grace, A. A. (2005). The hippocampal-VTA loop: Controlling the entry of information into long-term memory. *Neuron*, 46(5), 703–713.
- Livesey, A. C., Wall, M. B., & Smith, A. T. (2007). Time perception: Manipulation of task difficulty dissociates clock functions from other cognitive demands. *Neuropsychologia*, 45(2), 321–331.
- Loui, P., Li, H. C., & Schlaug, G. (2011). White matter integrity in right hemisphere predicts pitch-related grammar learning. *NeuroImage*, 55(2), 500–507.
- MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). Hippocampal “time cells” bridge the gap in memory for discontinuous events. *Neuron*, 71(4), 737–749.
- McKenzie, S., Frank, A. J., Kinsky, N. R., Porter, B., Rivière, P. D., & Eichenbaum, H. (2014). Hippocampal representation of related and opposing memories develop within distinct, hierarchically organized neural schemas. *Neuron*, 83(1), 202–215.
- Meneses, A., & Liy-Salmeron, G. (2012). Serotonin and emotion, learning and memory. *Reviews in the Neurosciences*, 23(5/6), 543–553.
- Mesgarani, N., Cheung, C., Johnson, K., & Chang, E. F. (2014). Phonetic feature encoding in human superior temporal gyrus. *Science*, 343(6174), 1006–1010.
- Meyer, L. B. (2008). *Emotion and meaning in music*. University of Chicago Press.
- Molinari, M., Leggio, M. G., & Thaut, M. H. (2007). The cerebellum and neural networks for rhythmic sensorimotor synchronization in the human brain. *The Cerebellum*, 6(1), 18–23.
- Molnar-Szakacs, I., & Overy, K. (2006). Music and mirror neurons: From motion to ‘e’ motion. *Social Cognitive and Affective Neuroscience*, 1(3), 235–241.
- Morita, T., Tanabe, H. C., Sasaki, A. T., Shimada, K., Kakigi, R., & Sadato, N. (2014). The anterior insular and anterior cingulate cortices in emotional processing for self-face recognition. *Social Cognitive and Affective Neuroscience*, 9(5), 570–579.
- Morrison, S. J., Demorest, S. M., Aylward, E. H., Cramer, S. C., & Maravilla, K. R. (2003). fMRI investigation of cross-cultural music comprehension. *NeuroImage*, 20(1), 378–384.
- Morrison, S. J., Demorest, S. M., & Stambaugh, L. A. (2008). Enculturation effects in music cognition: The role of age and music complexity. *Journal of Research in Music Education*, 56(2), 118–129.
- Moser, J., Batterink, L., Hegner, Y. L., Schleger, F., Braun, C., Paller, K. A., & Preissl, H. (2021). Dynamics of nonlinguistic statistical learning: From neural entrainment to the emergence of explicit knowledge. *NeuroImage*, 240, 118378.
- Mueller, K., Fritz, T., Mildner, T., Richter, M., Schulze, K., Lepsien, J., Schroeter, M. L., & Möller, H. E. (2015). Investigating the dynamics of the brain response to music: A central role of the ventral striatum/nucleus accumbens. *NeuroImage*, 116, 68–79.
- Nan, Y., Knösche, T. R., Zysset, S., & Friederici, A. D. (2008). Cross-cultural music phrase processing: An fMRI study. *Human Brain Mapping*, 29(3), 312–328.
- Narmour, E. (1983). Some major theoretical problems concerning the concept of hierarchy in the analysis of tonal music. *Music Perception*, 1(2), 129–199.
- Notter, M. P., Hanke, M., Murray, M. M., & Geiser, E. (2019). Encoding of auditory temporal gestalt in the human brain. *Cerebral Cortex*, 29(2), 475–484.
- O’Connor, A. R., Han, S., & Dobbins, I. G. (2010). The inferior parietal lobule and recognition memory: Expectancy violation or successful retrieval? *Journal of Neuroscience*, 30(8), 2924–2934.
- Olsen, R. K., Moses, S. N., Riggs, L., & Ryan, J. D. (2012). The hippocampus supports multiple cognitive processes through relational binding and comparison. *Frontiers in Human Neuroscience*, 6, 146.
- Onuki, Y., Van Someren, E. J. W., De Zeeuw, C. I., & Van der Werf, Y. D. (2015). Hippocampal–cerebellar interaction during spatio-temporal prediction. *Cerebral Cortex*, 25(2), 313–321.
- Pagès-Portabella, C., & Toro, J. M. (2020). Dissonant endings of chord progressions elicit a larger ERAN than ambiguous endings in musicians. *Psychophysiology*, 57(2), e13476.
- Palisson, J., Roussel-Baclet, C., Mailliet, D., Belin, C., Ankri, J., & Narme, P. (2015). Music enhances verbal episodic memory in Alzheimer’s disease. *Journal of Clinical and Experimental Neuropsychology*, 37(5), 503–517.
- Pearce, M. T., & Wiggins, G. A. (2012). Auditory expectation: The information dynamics of music perception and cognition. *Topics in Cognitive Science*, 4(4), 625–652.
- Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., Baldoli, C., & Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proceedings of the National Academy of Sciences*, 107(10), 4758–4763.
- Pereira, C. S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L., & Brattico, E. (2011). Music and emotions in the brain: Familiarity matters. *PLOS ONE*, 6(11), e27241.
- Peretz, I. (1993). Auditory atonalia for melodies. *Cognitive Neuropsychology*, 10(1), 21–56.
- Peretz, I. (1996). Can we lose memory for music? A case of music agnosia in a nonmusician. *Journal of Cognitive Neuroscience*, 8(6), 481–496.
- Peretz, I. (2002). Brain specialization for music. *Neuroscientist*, 8(4), 372.
- Peretz, I., Gaudreau, D., & Bonnel, A.-M. (1998). Exposure effects on music preference and recognition. *Memory & Cognition*, 26(5), 884–902.
- Peretz, I., Radeau, M., & Arguin, M. (2004). Two-way interactions between music and language: Evidence from priming recognition of tune and lyrics in familiar songs. *Memory & Cognition*, 32(1), 142–152.
- Peretz, I., & Zatorre, R. J. (2005). Brain organization for music processing. *Annual Review of Psychology*, 56, 89–114.

- Peynircioğlu, Z. F., Rabinovitz, B. E., & Thompson, J. L. W. (2008). Memory and metamemory for songs: The relative effectiveness of titles, lyrics, and melodies as cues for each other. *Psychology of Music*, 36(1), 47–61.
- Plailly, J., Tillmann, B., & Royet, J.-P. (2007). The feeling of familiarity of music and odors: The same neural signature? *Cerebral Cortex*, 17(11), 2650–2658.
- Press, C., Weiskopf, N., & Kilner, J. M. (2012). Dissociable roles of human inferior frontal gyrus during action execution and observation. *NeuroImage*, 60(3), 1671–1677.
- Preston, A. R., & Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Current Biology*, 23(17), R764–R773.
- Proverbio, A. M., Nasi, V. L., Arcari, L. A., De Benedetto, F., Guardamagna, M., Gazzola, M., & Zani, A. (2015). The effect of background music on episodic memory and autonomic responses: Listening to emotionally touching music enhances facial memory capacity. *Scientific Reports*, 5, 15219.
- Pyke, A., Betts, S., Fincham, J. M., & Anderson, J. R. (2015). Visuospatial referents facilitate the learning and transfer of mathematical operations: Extending the role of the angular gyrus. *Cognitive, Affective, & Behavioral Neuroscience*, 15(1), 229–250.
- Racette, A., & Peretz, I. (2007). Learning lyrics: To sing or not to sing? *Memory & Cognition*, 35(2), 242–253.
- Ramnani, N., & Passingham, R. E. (2001). Changes in the human brain during rhythm learning. *Journal of Cognitive Neuroscience*, 13(7), 952–966.
- Rohrmeier, M. A., & Koelsch, S. (2012). Predictive information processing in music cognition: A critical review. *International Journal of Psychophysiology*, 83(2), 164–175.
- Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(6), 906–914.
- Sakai, K., Hikosaka, O., & Nakamura, K. (2004). Emergence of rhythm during motor learning. *Trends in Cognitive Sciences*, 8(12), 547–553.
- Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience*, 14(2), 257–262.
- Salimpoor, V. N., van den Bosch, I., Kovacevic, N., McIntosh, A. R., Dagher, A., & Zatorre, R. J. (2013). Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science*, 340(6129), 216–219.
- Salimpoor, V. N., Zald, D. H., Zatorre, R. J., Dagher, A., & McIntosh, A. R. (2015). Predictions and the brain: How musical sounds become rewarding. *Trends in Cognitive Sciences*, 19(2), 86–91.
- Sammler, D., Grigutsch, M., Fritz, T., & Koelsch, S. (2007). Music and emotion: Electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology*, 44(2), 293–304.
- Sammler, D., Koelsch, S., Ball, T., Brandt, A., Elger, C. E., Friederici, A. D., Grigutsch, M., Huppertz, H., Knösche, T. R., & Wellmer, J. (2009). Overlap of musical and linguistic syntax processing: Intracranial ERP evidence. *Annals of the New York Academy of Sciences*, 1169(1), 494–498.
- Samson, S., Dellacherie, D., & Platel, H. (2009). Emotional power of music in patients with memory disorders: Clinical implications of cognitive neuroscience. *Annals of the New York Academy of Sciences*, 1169(1), 245–255.
- Samson, S., & Peretz, I. (2005). Effects of prior exposure on music liking and recognition in patients with temporal lobe lesions. *Annals of the New York Academy of Sciences*, 1060(1), 419–428.
- Samson, S., & Zatorre, R. J. (1994). Contribution of the right temporal lobe to musical timbre discrimination. *Neuropsychologia*, 32(2), 231–240.
- Samson, S., & Zatorre, R. J. (1991). Recognition memory for text and melody of songs after unilateral temporal lobe lesion: Evidence for dual encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(4), 793.
- Samson, S., & Zatorre, R. J. (1992). Learning and retention of melodic and verbal information after unilateral temporal lobectomy. *Neuropsychologia*, 30(9), 815–826.
- Satoh, M., Takeda, K., Nagata, K., Shimosegawa, E., & Kuzuhara, S. (2006). Positron-emission tomography of brain regions activated by recognition of familiar music. *American Journal of Neuroradiology*, 27(5), 1101–1106.
- Schack, B., Klimesch, W., & Sauseng, P. (2005). Phase synchronization between theta and upper alpha oscillations in a working memory task. *International Journal of Psychophysiology*, 57(2), 105–114.
- Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M., & Turk-Browne, N. B. (2014). The necessity of the medial temporal lobe for statistical learning. *Journal of Cognitive Neuroscience*, 26(8), 1736–1747.
- Schapiro, A. C., Turk-Browne, N. B., Norman, K. A., & Botvinick, M. M. (2016). Statistical learning of temporal community structure in the hippocampus. *Hippocampus*, 26(1), 3–8.
- Schellenberg, E. G., & W. Weiss, M. (2013). Music and cognitive abilities. *The Psychology of Music*, 499–550.
- Schulkind, M. D., Hennis, L. K., & Rubin, D. C. (1999). Music, emotion, and autobiographical memory: They're playing your song. *Memory & Cognition*, 27(6), 948–955.
- Seghier, M. L. (2013). The angular gyrus: Multiple functions and multiple subdivisions. *The Neuroscientist*, 19(1), 43–61.
- Shapiro, M. L., & Eichenbaum, H. (1999). Hippocampus as a memory map: Synaptic plasticity and memory encoding by hippocampal neurons. *Hippocampus*, 9(4), 365–384.
- Sikka, R., Cuddy, L. L., Johnsrude, I. S., & Vanstone, A. D. (2015). An fMRI comparison of neural activity associated with recognition of familiar melodies in younger and older adults. *Frontiers in Neuroscience*, 9, 356.
- Simmons-Stern, N. R., Budson, A. E., & Ally, B. A. (2010). Music as a memory enhancer in patients with Alzheimer's disease. *Neuropsychologia*, 48(10), 3164–3167.
- Soley, G., & Hannon, E. E. (2010). Infants prefer the musical meter of their own culture: A cross-cultural comparison. *Developmental Psychology*, 46(1), 286.
- Squire, L. R. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience*, 4(3), 232–243.
- Squire, L. R., Schmolck, H., & Stark, S. M. (2001). Impaired auditory recognition memory in amnesic patients with medial temporal lobe lesions. *Learning & Memory*, 8(5), 252–256.
- Squire, L. R., & Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science*, 253(5026), 1380–1386.
- Stefanics, G., Háden, G. P., Sziller, I., Balázs, L., Beke, A., & Winkler, I. (2009). Newborn infants process pitch intervals. *Clinical Neurophysiology*, 120(2), 304–308.
- Teki, S., Kumar, S., von Kriegstein, K., Stewart, L., Lyness, C. R., Moore, B. C. J., Capleton, B., & Griffiths, T. D. (2012). Navigating the auditory scene: An expert role for the hippocampus. *Journal of Neuroscience*, 32(35), 12251–12257.
- Thaut, M. H., Fischer, C. E., Leggieri, M., Vuong, V., Churchill, N. W., Fornazzari, L. R., & Schweizer, T. A. (2020). Neural basis of long-term musical memory in cognitively impaired older persons. *Alzheimer Disease & Associated Disorders*, 34(3), 267–271.
- Thaut, M. H., Trimarchi, P. D., & Parsons, L. M. (2014). Human brain basis of musical rhythm perception: Common and distinct neural substrates for meter, tempo, and pattern. *Brain Sciences*, 4(2), 428–452.

- Tillmann, B., Poulin-Charronnat, B., & Bigand, E. (2014). The role of expectation in music: From the score to emotions and the brain. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(1), 105–113.
- Trainor, L. J., & Heinmiller, B. M. (1998). The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. *Infant Behavior and Development*, 21(1), 77–88.
- Trainor, L. J., Wu, L., & Tsang, C. D. (2004). Long-term memory for music: Infants remember tempo and timbre. *Developmental Science*, 7(3), 289–296.
- Trehub, S. E. (1987). Infants' perception of musical patterns. *Perception & Psychophysics*, 41(6), 635–641.
- Trehub, S. E. (2006). *Infants as musical connoisseurs*. Oxford University Press.
- Trehub, S. E., Bull, D., & Thorpe, L. A. (1984). Infants' perception of melodies: The role of melodic contour. *Child Development*, 821–830.
- Tulving, E. (1972). Episodic and semantic memory. Organization of memory, 1(381–403), 1.
- Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in Psychology*, 8, 1454.
- Valtonen, J., Gregory, E., Landau, B., & McCloskey, M. (2014). New learning of music after bilateral medial temporal lobe damage: Evidence from an amnesic patient. *Frontiers in Human Neuroscience*, 8, 694.
- van Kesteren, M. T. R., Fernández, G., Norris, D. G., & Hermans, E. J. (2010). Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proceedings of the National Academy of Sciences*, 107(16), 7550–7555.
- van Kesteren, M. T. R., Rijpkema, M., Ruiter, D. J., Morris, R. G. M., & Fernández, G. (2014). Building on Prior Knowledge: Schema-dependent Encoding Processes Relate to Academic Performance. *Journal of Cognitive Neuroscience*, 26(10), 2250–2261.
- van Kesteren, M. T. R., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, 35(4), 211–219.
- Villarrreal, E. A. G., Brattico, E., Leino, S., Østergaard, L., & Vuust, P. (2011). Distinct neural responses to chord violations: A multiple source analysis study. *Brain Research*, 1389, 103–114.
- Wallace, W. T. (1994). Memory for music: Effect of melody on recall of text. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(6), 1471.
- Wallenstein, G. V., Hasselmo, M. E., & Eichenbaum, H. (1998). The hippocampus as an associator of discontinuous events. *Trends in Neurosciences*, 21(8), 317–323.
- Wang, M.-Y., & Chang, H.-C. (2004). The mere exposure effect and recognition memory. *Cognition & Emotion*, 18(8), 1055–1078.
- Watanabe, T., Yagishita, S., & Kikyo, H. (2008). Memory of music: Roles of right hippocampus and left inferior frontal gyrus. *NeuroImage*, 39(1), 483–491.
- White-Schwach, T., Carr, K. W., Anderson, S., Strait, D. L., & Kraus, N. (2013). Older adults benefit from music training early in life: Biological evidence for long-term training-driven plasticity. *Journal of Neuroscience*, 33(45), 17667–17674.
- Wilson, S. J., & Saling, M. M. (2008). Contributions of the right and left mesial temporal lobes to music memory: Evidence from melodic learning difficulties. *Music Perception*, 25(4), 303–314.
- Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences*, 106(7), 2468–2471.
- Wise, R. A. (2004). Dopamine, learning and motivation. *Nature Reviews Neuroscience*, 5(6), 483–494.
- Wu, K., Anderson, J., Townsend, J., Frazier, T., Brandt, A., & Karmonik, C. (2019). Characterization of functional brain connectivity towards optimization of music selection for therapy: A fMRI study. *International Journal of Neuroscience*, 129(9), 882–889.
- Yalch, R. F. (1991). Memory in a jingle jungle: Music as a mnemonic device in communicating advertising slogans. *Journal of Applied Psychology*, 76(2), 268.
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, 46(3), 441–517.
- Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, 8(7), 547–558.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.