

• 研究构想(Conceptual Framework) •

预期视角下音乐节拍结构的认知与神经机制*

孙丽君¹ 杨玉芳^{2,3}

(¹南京航空航天大学艺术学院, 南京 211106) (²中国科学院心理研究所行为科学重点实验室, 北京 100101)
(³中国科学院大学心理学系, 北京 100101)

摘要 节拍结构作为音乐在时间维度的组织框架, 不仅是作曲家的创作基础, 还是人们欣赏音乐美感、体验音乐情绪以及理解音乐意义的前提。在预期编码理论的框架下, 本文拟围绕节拍结构的预期与整合两方面, 通过行为实验和脑电技术, 探查音乐节拍结构的认知神经基础。具体包括以下4个方面研究: (1)考察在节奏序列展开过程中, 听者构建节拍结构心理表征从而建立预期的动态神经响应; (2)考察听者通过预期错误实现节拍结构预期更新的神经机制; (3)以乐句为结构单元, 考察听者在短时间尺度上整合多层级节拍结构的认知神经机制; (4)在乐段水平上, 考察听者如何依据远距离依存关系整合嵌套节拍结构。以上研究将有利于揭示音乐结构认知的一般机制, 为音乐认知神经模型的构建奠定基础。同时, 相关研究结论还将为音乐鉴赏活动与音乐美学研究提供客观依据, 在音乐学领域具有潜在的应用前景。

关键词 音乐认知, 文艺心理学, 神经机制, 节拍结构, 脑电图

分类号 B842

1 问题提出

音乐是人类文化的结晶, 其多样性和复杂性为人类社会所特有。与文学和绘画等艺术形式不同, 音乐既不具有语言的语义性, 也不具有绘画的具象性(蒋存梅, 2016)。作为一门听觉的艺术, 音乐实质上就是一组在时间中展开、具有结构组织的声音序列。任何一部音乐作品都是由长短各异的声音和休止构成, 这些声音事件总是按照一定的强弱规律交替组合, 并以此为基础重复发展。节拍结构即指声音序列中强拍和弱拍循环往复的组织形式, 它是音乐的时间框架(Vuust et al., 2014)。作曲家通过节拍结构, 把这些具有内在音高关系的声音元素组织成一个有机的整体, 以表达音乐的情绪和意义。

音乐的节拍结构不仅为作曲家的创作提供框

架, 还是人们欣赏音乐美感、体验音乐情绪、参与社会性活动以及音乐治疗的支持和保障。人们在聆听音乐的过程中, 需要在声音流中获取其潜在的时间组织规则并形成节拍结构, 进而理解音乐所传递的内涵(欧阳珏, 戴志强, 2010; Nave-Blodgett et al., 2021)。进一步的研究表明, 在音乐聆听的过程中, 与时间结构的同步是情绪诱发的机制之一。当音乐响起时, 人们经常会自发地产生随着音乐起舞的冲动, 或者具有节奏性的拍手、跺脚等行为。在这种与音乐节拍同步的行为中, 心跳、呼吸、皮肤电、神经递质等生理系统的指标发生改变, 同时听者体验到正性的情绪和愉悦的心理状态(Juslin, 2013; Trost & Vuilleumier, 2013)。另外, 音乐在某些特定的场合中常常作为一种社交互动的媒介。当人们需要聚集在一起庆祝或哀悼时, 处于相同音乐背景中的个体更容易体验同呼吸共命运的感觉, 音乐节拍诱发的同步使大脑前庭系统和边缘系统产生更多连接, 奖赏系统的参与促使其建立更为深刻的社交关系(Trost et al., 2007)。基于节拍同步的生理和心理机制, 节拍在一些特定的音乐治疗方法中也发挥着

收稿日期: 2024-04-07

* 国家自然科学基金青年项目(32300865)、教育部人文社会科学研究青年基金项目(22YJC760086)。

通信作者: 杨玉芳, E-mail: yangyf@psych.ac.cn

重要的作用。例如,旋律语调疗法是针对非流畅性失语症患者的一种有效且无创的治疗手段,其核心是借助与语言音韵相一致的旋律恢复言语产生能力。这种失语症治疗方法的关键在于节拍,因此,与音高相比,时间方面的训练可能更为重要(孙丽君等, 2017)。

那么,人们如何感知音乐的节拍结构?其背后的认知神经机制又是什么?音乐聆听是大脑持续预期和整合的主动认知过程(Koelsch et al., 2019)。在一个音乐事件出现之前,听者依据已经输入的声音信息构建声音结构的心理表征,从而预期即将出现的音乐事件。在这个音乐事件出现之后,听者需要将其与先前的音乐背景整合联系,进一步更新心理表征从而更好地预期之后的音乐进行。在音乐的时间维度上,对节拍结构进行自上而下的预期和整合是音乐聆听的核心过程。大量研究围绕音乐的音高维度展开研究,发现当和弦违反和声结构或和声句法时,会在大脑中诱发出早期右前负波(early right anterior negativity, ERAN)、N5 以及 P600 等脑电成分,分别反映了预期违反和结构整合(Koelsch, 2012; Koelsch et al., 2000; Zendel et al., 2015)。然而,音乐时间维度的加工较少受到研究者关注(江俊等, 2014),其中涉及音乐节拍结构的研究更是非常有限。同时,由于音乐时间组织的复杂性,节拍结构成为音乐认知研究中的难点和薄弱环节,没有得到充分的研究。

综上所述,本文拟采用脑电技术,在预期编码理论的框架下研究大脑主动加工音乐节拍结构的认知过程,探讨节拍结构预期形成和更新的动态神经响应,以及在乐句和乐段等大尺度音乐单元中整合节拍结构的认知神经机制。研究结果将深化人们对音乐时间结构加工以及音乐认知过程的认识,有助于理解人类复杂认知活动的一般机制,为音乐鉴赏活动与音乐美学研究提供客观依据,在音乐领域具有潜在的应用前景。

2 研究现状与评述

2.1 音乐时间结构加工的认知神经机制

尽管音乐的时间维度比音高维度更为重要(江俊等, 2014),然而,纵观已有文献,相对于音高研究对音乐时间维度的加工较少受到研究者的关注(Overy & Turner, 2009)。Jones (1976)提出的

动态注意理论(Dynamic Attention Theory, DAT)将知觉领域的研究焦点置于时间维度。该理论强调内部注意能量与外在时间序列同步变化的关系。当事件出现的时间节点具有规则性时,人们通过注意资源分配获得内部计时知觉。在此基础上,研究者进一步提出神经共振理论(Neural Resonance Theory),解释时间追踪的脑神经基础(Large & Snyder, 2009)。该理论认为大脑中存在多种频率的神经振荡活动,它们分别与不同时间节律的外部事件形成锁相反应,即神经夹带(neural entrainment)。由此可见,无论注意能量还是神经活动,均通过内外偶联机制对具有规则性的时间信息进行编码。由于音乐时间结构具有高度的规则性(张巍, 2019),因此,以上时间知觉相关的理论模型为探究音乐时间结构提供了理论支撑。

在音乐时间知觉的实证研究中,研究者主要围绕时长分辨(Barnes & Jones, 2000; Rao et al., 2001)、节奏知觉(Geiser et al., 2009; Habibi et al., 2014)、拍子抽取(Grahn & Brett, 2007)等方面进行了探讨。人们在聆听时间序列时,不仅能够用手或脚跟随拍子同步敲击(Rankin et al., 2009; Van der Steen & Keller, 2013),还能对即将出现的拍子形成准确的预期(Geiser et al., 2012; Repp & Su, 2013)。最近,一些研究者运用脑电(Electroencephalogram, EEG)和脑磁图(Magnetoencephalography, MEG)等技术手段,发现大脑中与拍子同步的神经振荡活动是拍子感知的神经基础(Arnal et al., 2014; Fujioka et al., 2012; Fujioka et al., 2015; Nozaradan et al., 2012)。然而,拍子抽取仅是音乐时间结构感知的基础,将拍子依据强弱规律组织成小节间循环的节拍结构才是音乐时间结构加工的关键。

目前为止,直接探讨节拍结构的研究数量非常有限,且主要围绕节拍结构的分辨与识别(Hannon et al., 2004; Liégeois-Chauvel et al., 1998)。虽然已有研究初步探讨了节拍结构加工的神经基础,发现大脑右侧额下回、额中回、中央前回与壳核等关键脑区参与了节拍结构加工(Fujioka et al., 2015; Li et al., 2019; Thaut et al., 2014),但是研究所使用的实验材料大多为简单重复或等时间隔的节奏刺激,导致结果难以推广到真实的音乐聆听中,不足以揭示节拍结构加工的本质。

由于节拍结构的复杂性,对节拍结构认知机制的探讨是音乐结构加工中的薄弱环节。基于此,

本文采用更具有生态效度的音乐材料,系统考察听者对节拍结构的认知与神经机制,揭示音乐时间结构加工的本质,更全面地了解音乐聆听的认知神经过程。

2.2 音乐节拍结构的预期加工

近20年来,越来越多的研究者认为音乐聆听是大脑持续不断地预期声音、解释声音的认知迭代过程(Huron, 2008; Margulis, 2005; Pearce, 2018; Rohrmeier & Koelsch, 2012),其核心机制是依据声音信号和长时记忆中存储的音乐结构组织规则,构建合适的心理表征来预期即将出现的事件,通过将预期事件与实际发生的事件比对,更新心理表征从而更好地预期下一个事件(Koelsch et al., 2019; Vuust et al., 2022)。因此,预期是音乐认知的重要阶段,这为探究音乐节拍结构加工提供了很好的模型参照。

音乐中有限的音符和休止符能够组合出无限的节奏模式(Lerdahl & Jackendoff, 1983)。面对如此复杂的表层节奏,听者需要逐步抽取出结构骨架上重要的声音事件,并赋予其一定的强弱关系,才能构建节拍结构的心理表征,形成对未来事件的预期。已有研究表明,在节奏序列的进行中听者的注意能量峰值与强拍趋于吻合,显现出与节拍结构的同步(Jones & Boltz, 1989; Large & Jones, 1999)。进一步的研究发现,节奏序列会逐步诱发大脑神经元周期性的神经夹带活动(Harding et al., 2019; Large & Snyder, 2009),而不同节律的时间信息可能依赖不同频率的神经振荡活动(Doelling & Poeppel, 2015; Lakatos et al., 2019; Wu et al., 2024)。那么,在音乐展开中,听者如何循序渐进地形成节拍结构的心理表征?不同频率的神经夹带又是如何协同作用,以支持听者构建不同类型的节拍结构?目前,尚未有研究探讨节拍结构预期形成的动态过程及相应的神经响应。

节拍结构的预期更新与预期错误紧密相关。经典的预期编码(Predictive Coding)理论提出了以预期机制为核心的大脑功能模型。为了处理连续的信息流,大脑会持续生成不同层级的预期,将高层级的预期与较低层级的信息输入进行比较,若预期与输入之间出现不匹配,即产生了预期错误。预期错误促使大脑整合输入信息,更新原有心理表征,实现预期错误最小化。预期的可预测性越强,赋予预期错误的权重越大,大脑越倾向

于向上反馈预期错误,进而更新预期(Friston & Kiebel, 2009)。在这个理论框架下,研究者进一步提出音乐的预期理论(Koelsch et al., 2019),强调音乐聆听是人们将预期事件与真实事件反复比较并实时更新预期的过程。据此,我们推测节拍结构预期更新的主要推动力是预期错误,且受到可预测性的影响。

语言学的相关研究已对预期编码理论进行验证,成功分离出预期形成和预期更新两个加工阶段及其对应的神经基础(Kuperberg et al., 2020; Li et al., 2017)。在音乐领域,尽管已有研究通过结构违反范式对音乐预期进行了大量探索(Koelsch et al., 2000; Sun et al., 2018; Sun, Thompson et al., 2020; You et al., 2021),但是均忽略了预期的形成阶段及其对预期更新的调节作用(张晶晶 等, 2020)。在音乐聆听中,听者是否依据节拍结构表征提前建立预期从而预先激活目标刺激?预期形成与预期更新又是如何相互作用从而支持大脑对节拍结构的预期加工?以上问题仍有待进一步探究。

综上所述,本文拟从预期形成和预期更新两个方面,分别考察听者构建节拍结构心理表征的动态神经响应,以及如何通过预期错误实现预期更新,深入探究节拍结构的预期加工机制,为构建音乐结构加工的认知神经模型提供实证依据。

2.3 音乐节拍结构的整合加工

音乐与自然环境声音的重要区别在于其具有复杂的组织结构。音乐事件在时间的展开中以层级的方式形成具有逻辑关系的“树形结构”(Fitch, 2013; Lerdahl & Jackendoff, 1983)。对音乐层级性结构的知觉不仅制约着人们对节拍结构的加工,也是音乐情绪体验和意义理解的关键(You et al., 2023)。

就音乐的节拍结构来说,声音元素既需要依赖近距离依存关系(short-distance dependency)进行局部整合,同时还要考虑远距离依存关系(long-distance dependency)完成整体的整合加工。前者指在较小时间尺度上(如,乐句之内),相邻元素之间的依赖关系;后者指在较大的时间尺度上(如,多乐句或乐段),需要跨越层级(嵌套)的局部结构才能架构的逻辑关系(Phillips et al., 2005)。最近的研究发现,小时间尺度的信息加工主要依赖视觉或听觉皮层等单模块脑区,而大时间尺度则更多地依赖默认网络等跨模块区域(Wolff et al.,

2022)。基于人类神经系统对不同时间尺度信息的差异化加工,我们认为节拍结构在乐句与乐段的单元组织中具有不同的加工模式。

已有部分研究考察了乐句内部节拍层级结构的认知加工机制。结果发现,节拍结构违反会诱发更大的 ERAN、N5 和 P600 等脑电成分,反映了早期的自动化探测和晚期的整合加工(Sun et al., 2018; Sun, Thompson et al., 2020)。在实际的音乐作曲和演奏中,节拍结构的组织复杂且丰富,在小节、拍子、八分音符甚至更多的时值层面形成多水平的层级结构(Lerdahl & Jackendoff, 1983)。但是以上研究尚未考虑小节水平之外的其他层级对节拍结构加工的影响。依据联合层级控制理论(Coordinated Hierarchical Control, CHC),大脑需要调动不同的神经网络来处理各种层级水平的认知加工(Asano et al., 2021)。那么,低层级、中层级与高层级的节拍结构是否会诱发不同的大脑响应与加工模式?听者能否以及如何整合多种层级水平的音乐节拍结构,尚无研究涉及。

人们日常的音乐欣赏活动通常是大时间尺度的音乐单元,包含更多远距离依存关系。乐段中的嵌套结构以远距离依存关系为核心,既是音乐逻辑性的体现,更是高水平音乐理解所必需的。然而,目前尚没有研究探讨听者如何整合嵌套节拍结构。在音高维度,研究者通过操纵乐段开始或中间部分的和声功能,对嵌套和声结构进行了充分的研究,为听者进行远距离音乐事件整合的心理现实性提供了实证依据(Koelsch et al., 2013; Ma et al., 2018; Sun, Feng et al., 2020; Zhou et al., 2019)。嵌套和声结构的研究方法和结论,对探究长时程单元中节拍结构的整合具有重要的参考和借鉴价值。因此,探讨听者如何依据远距离依存关系加工嵌套节拍结构既是必要的,也是有坚实基础的。

本文拟在乐句和乐段两个时间尺度上,分别考察节拍结构的多层级整合加工机制,以及如何依据远距离依存关系在更大的时间单元中整合嵌套节拍结构。

3 研究构想

本研究将预期编码理论与音乐节拍加工结合,在预期误差框架下探究音乐结构信息如何在时间维度上整合更新。依据已有信息形成具有结构的

内部表征是感知音乐节拍结构的基础和先决条件,一旦形成结构表征,听者即处于对即将发生的音乐事件进行预期的状态。音乐事件出现之后,听者判断是否符合预期,将其整合到已有的音乐背景中,并依据预期符合或违反的情况调整结构表征,从而更好地预期。因此,预期和整合是大脑对节拍结构加工的两个必不可少的重要阶段,二者缺一不可,并互相影响,反映了大脑对节拍结构自主加工的迭代过程。

本文围绕预期和整合两个重要的认知阶段,探讨音乐节拍结构加工的认知与神经机制。重点解决两个问题:(1)节拍结构的预期加工;(2)节拍结构的整合加工。研究框架如图 1 所示,由两个模块(4 个研究)组成。模块一聚焦节拍结构的预期加工,从预期产生和更新两个方面,分别考察听者依据持续性信息输入构建节拍结构心理表征的动态过程(研究 1),以及如何通过预期错误实现预期更新的神经机制(研究 2);模块二聚焦节拍结构的整合加工,在乐句和乐段两个水平上,分别考察听者对节拍结构多层级水平的整合加工(研究 3),以及如何依据远距离依存关系,在更大的时间单元中整合嵌套的节拍结构(研究 4)。

3.1 节拍结构心理表征构建的动态过程

研究 1 在动态注意理论的框架下,探讨听者在节奏序列展开过程中实时的神经响应,揭示听者构建节拍结构心理表征的动态过程。实验刺激来源于真实音乐作品。首先,选取 2/4 拍、3/4 拍、4/4 拍、3/8 拍、6/8 拍五种常见节拍结构的音乐材料;然后,抽取其中 4 个小节,剔除所有音高信息,形成音高固定的节奏序列;最后,通过 Sibelius 软件打谱并导出音频文件。依据调性音乐生成理论(Lerdahl & Jackendoff, 1983),每种节拍类型具有特定的节拍结构和强弱关系。例如,2/4 拍结构较简单,每小节包含强、弱两个拍子;而 6/8 拍结构较复杂,每小节包含强、弱、弱、次强、弱、弱 6 个拍子。

本研究的分析思路包括:(1)全局能量谱(global field power, GFP)计算每一个时间点上所有头皮电极的标准差,它是不依赖于参考电极的分析方法(Jia, 2019),能够揭示脑电响应的动态变化过程(Sun, Hu, et al., 2020)。为了探究节拍结构心理表征构建的动态过程,通过 GFP 对脑电信号进行整体分析,明确节拍结构心理表征构建的动

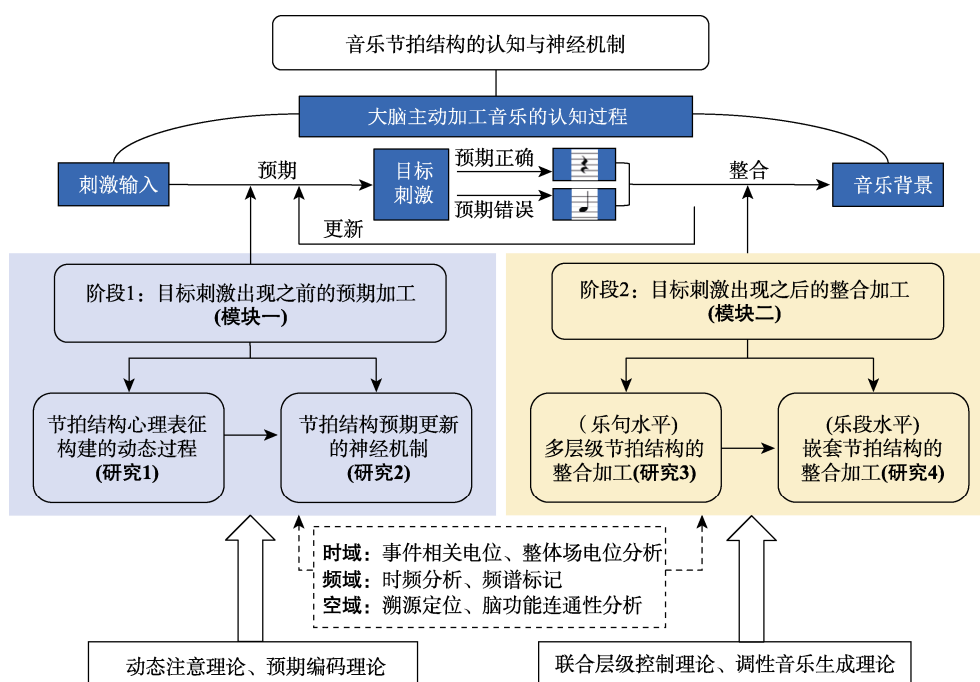


图 1 研究框架图。

态性和变化趋势。(2) 试次相干性 (intertrial coherence, ITC) 通过计算各个频率上试次之间的相位一致性来量化神经夹带 (Delorme & Makeig 2004)。基于神经夹带与节拍感知的关联性, 本研究通过频谱标记方法, 用 ITC 来量化神经夹带响应, 挖掘神经夹带和不同频率神经夹带的协同响应, 揭示客观音响刺激与神经反应的对应关系, 阐明节拍结构心理表征构建的神经特征。(3) 比较 5 种常见节拍结构的神经差异化响应, 揭示节拍类型对节拍结构心理表征构建的影响及其神经机制。

对本研究结果的预期如下:

假设 1: 依据动态注意理论 (Jones, 1976), 节拍结构心理表征的构建过程是动态的, 随着节奏序列的展开, 听者大脑皮层逐步诱发趋于规律性的整体场电位与频谱能量分布。

假设 2: 结合已有研究结果 (Doelling & Poeppel, 2015; Nozaradan et al., 2018), 节拍结构表征的神经机制在于神经夹带和不同频率夹带之间的协同作用。

假设 3: 不同节拍类型诱发的神经夹带响应之间具有差异, 分别与各自节拍频率一致。

3.2 节拍结构预期更新的神经机制

研究 2 采用结构违反范式, 考察听者对目标音乐事件预期错误所诱发的神经响应, 揭示节拍结构预期更新的认知神经机制。实验刺激为自创的节奏序列, 所有音符采用固定的音高频率, 长度为 6 个小节。实验采用 2 (可预测性: 高 vs. 低) × 2 (预期准确性: 正确 vs. 错误) 的两因素重复测量实验设计, 同时操纵关键音乐事件的可预测性和预期准确性, 形成 4 种实验条件。当节拍结构非常明确时, 目标事件的可预测性高, 当音乐背景中出现节拍结构强弱规律的局部改变时, 目标事件的可预测性降低; 当关键音乐事件处于强拍位置时, 预期正确, 当通过延音线使之处于弱拍位置时, 预期错误。

本研究的分析思路包括: (1) 本研究中包含关键和弦, 因此, 首先采用事件相关电位 (event-related potential, ERP) 技术手段分离预期形成阶段和预期更新阶段, 以目标刺激作为关键时间点, 分别截取出现前和出现后的时间段分析, 通过事件相关的时频域脑电响应探究两个阶段的相互作用/制约关系, 阐释前者对后者的调节作用, 深入揭示节拍结构预期更新背后复杂的认知神经机

制。(2)相-幅耦合(phase-amplitude coupling, PAC)是探究神经振荡通过与外界节律同步化参与知觉加工的重要神经指标之一(张雪等, 2016), 特指低频神经振荡的相位与高频成分的幅度之间存在特定的共变关系, 反映了大脑神经活动在不同时间尺度上的组织(Jensen & Colgin, 2007; Lakatos et al., 2005)。考虑到节拍结构的抽取和表征过程需要多尺度神经元振荡的协调和功能整合, 因此, 在目标音乐事件出现前的阶段, 以 PAC 量化 delta-beta 的跨频率耦合, 探讨听者在不同音乐背景中形成预期的大脑响应, 揭示可预测性对节拍结构预期的影响。

对本研究结果的预期如下:

假设 1: 依据预期编码理论(Friston & Kiebel, 2009), 在预期更新阶段, 预期准确性的主效应显著, 预期错误会诱发更大的 ERAN 效应、P600 效应和 beta 频段频谱功率的增加, 分别反映听者对节拍结构预期错误的探测、整合和更新。

假设 2: 以往研究发现, 在听觉目标刺激出现之前, delta 与 beta 频段的耦合关系有助于提升预测的准确性(Arnal et al., 2014)。依此, 在本研究中的预期形成阶段, 可预测性的主效应显著, 在高可预测性条件下诱发具有更高耦合关系的 delta 与 beta 神经振荡活动。

假设 3: 可预测性与预期准确性存在交互作用, 主要体现在预期更新阶段 P600 波幅以及 beta 频段的能量变化(Li et al., 2017), 支持可预测性对节拍结构预期更新具有调节作用的研究假设。

3.3 乐句中多层级节拍结构的整合加工

研究 3 以乐句为结构单元, 操纵乐句结束事件的节拍位置, 探讨人们如何在较小的时间尺度上整合节拍结构, 揭示听者依据相邻元素之间的依赖关系加工多层级节拍结构的机制。音乐材料为包含 6 个小节的 2/4 拍的乐句, 所有音符采用固定的音高频率。每个节奏序列均由二分音符、四分音符和八分音符等多种时值的音符构成, 具有多层级的组织方式。本研究采用单因素 3 水平(层级水平: 高层级 vs. 中层级 vs. 低层级)的重复测量实验设计。实验操纵乐句结束事件的节拍位置, 使之分别处于小节水平、四分音符水平以及八分音符水平的不同层级上, 形成高层级、中层级与低层级三种条件。高层级的结束音最稳定, 能够给乐句带来最好的完成感, 乐句整合难度最

低, 中层级次之, 低层级结束音则最不稳定, 乐句整合最为困难。为了避免中层级和低层级条件中结束事件前休止符的新异性, 同时为了提示乐句即将结束, 所有节奏序列的倒数第二小节均为休止符。实验任务要求被试在听完每个节奏序列之后, 在 5 点量表上判断整个乐句的完成感。

本研究分析思路包括: (1)以末尾事件作为关键时间点进行 ERP 分析, 探究听者在小节、拍子、八分音符三个层级水平上整合节拍结构所诱发的电生理反应, 揭示节拍结构加工的多层级性。(2)将 ERP 成分与行为结果做相关分析, 探究乐句完成感的神经指标。(3)为了验证联合层级控制理论提出的层级性加工与活动脑区之间的关联性, 借助溯源定位和脑功能连通性分析, 考察听者整合多层级节拍结构所激活的脑区, 揭示大脑在处理不同层级的节拍结构时调用的神经网络。

对本研究结果的预期如下:

假设 1: 不同层级水平的乐句结束音所诱发的大脑响应具有差异, 高层级比中层级、中层级比低层级诱发更高的完成感、更小的 P600 或 N5 效应, 支持节拍结构多层级整合加工的研究假设。如果不支持研究假设, 条件间不会出现梯度变化。

假设 2: 基于联合层级控制理论(Asano et al., 2021), 高层级音乐事件的激活脑区比低层级更靠前, 前额叶脑区与顶叶脑区之间具有更高的连接强度。

假设 3: 与和声的整合加工过程类似, 听者对不同层级水平节拍结构的整合均存在早期探测与晚期整合的双阶段加工特征, 分别体现在早期 ERAN 和晚期 N5 或 P600(Sun et al., 2018)。

3.4 乐段中嵌套节拍结构的整合加工

研究 4 采用来自真实音乐作品中的乐段材料, 考察在自然聆听状态下由嵌套节拍结构所诱发的神经响应, 揭示在大时间尺度中节拍结构整合的认知过程和神经机制。实验刺激来源于 2/4 拍的真实音乐作品, 抽取长度为 8 个小节且包含两个乐句的音乐片段, 剔除所有音高信息, 形成音高频率固定的节奏序列。本研究采用 2 (层级结构: 非嵌套 vs. 嵌套) × 2 (依存距离: 长 vs. 短) 的两因素重复测量实验设计。借鉴和声结构远距离整合的研究范式(Koelsch et al., 2013; Zhou et al., 2019), 实验通过操纵第三和第四小节的节拍结构类型, 改变两个乐句之间的结构层级关系, 形成

非嵌套节拍结构与嵌套节拍结构两种条件。同时,通过扩展第三小节和第四小节的节奏序列改变乐句长度,探讨依存距离对节拍结构加工的影响效应。为了尽可能获得自然聆听状态,要求被试做偏差音色判断的无关任务,包含偏差音色的试次不计入统计分析。

本研究的分析思路包括:(1)以末尾乐句的开始作为关键时间点,对比节拍结构整合中长距离与短距离加工,揭示工作记忆在长时程音乐单元加工中的作用。(2)探究节拍结构中结构层级和依存距离加工的共享/分离的认知机制,揭示听者如何在复杂的音乐组织中加工节拍结构、理解音乐意义。(3)对比嵌套节拍结构与嵌套和声结构加工的差异,揭示神经系统加工大时间尺度信息的一般机制。

对本研究结果的预期如下:

假设1:嵌套结构比非嵌套结构诱发更大的N5成分和beta频段的神经振荡能量(Koelsch et al., 2013; Ma et al., 2018),表明听者能够对节拍结构进行远距离整合加工,具有加工嵌套节拍结构的心理现实性。

假设2:依存距离对节拍结构加工的影响主要体现在工作记忆负荷,与嵌套结构的加工机制分离。

假设3:嵌套节拍结构与嵌套和声结构的加工具有相似的神经反应模式,反映了远距离依存关系对结构抽取的高水平需求。

4 理论建构

音乐在时间结构中流淌和发展,节拍结构的认知加工是理解音乐意义的一个核心过程。本文聚焦音乐的时间维度,从认知神经科学的视角,围绕节拍结构的预期加工与整合加工两方面,研究节拍结构预期形成和更新的动态神经机制,研究具有复杂层级和远距离依存关系的节拍结构整合的认知神经机制。本文拓展了预期编码理论在音乐认知领域的研究范畴和内容,为今后构建音乐认知神经模型奠定基础,具体体现在以下几点:

第一,音乐聆听是大脑对声音进行主动认知的过程,提前预期即将出现的音乐事件是音乐理解的重要环节。如何实现预期?首先需要根据已有的音乐信息,在大脑中构建一个具有结构的心理表征,然后才能根据结构特征在其框架内,预期下一个音乐事件。在构建节拍结构的心理表征

方面,动态注意理论虽然指明了时间结构加工的动态性特征,但不能识别节拍结构建构的过程,也不能区分音乐中各种节拍结构构建的差异(Jones, 1976)。本文运用具有高时间分辨率的EEG技术手段,考察大脑追踪节拍结构的实时响应,揭示人们如何在音乐的展开中逐步构建节拍结构心理表征及其关键神经指标。同时,厘清不同类型节拍结构诱发的神经反应,阐明人们对音乐中常见节拍结构形成预期的动态过程及其相应的神经机制。在节拍结构的预期方面,目前相关研究主要通过操纵某一个音乐事件与音乐背景的吻合性,分析在该音乐事件出现之后所诱发的与预期相关的神经响应,从而间接证实音乐预期的存在(Geiser et al., 2010; Vuust et al., 2005, 2009)。本文基于预期编码理论(Friston & Kiebel, 2009),通过聚焦音乐事件出现之前的预期阶段,分析音乐事件出现之前的脑响应,为预期形成的预激活阶段提供直接的神经证据。同时,分离预期形成与预期更新两个阶段,考察预期错误对节拍结构预期更新的作用,探究前者对后者的调节作用,揭示节拍结构预期更新背后的神经机制。

第二,依据调性音乐生成理论,层级性组织是音乐逻辑性的集中体现,是制约音乐意义理解的关键(Lerdahl & Jackendoff, 1983)。节拍结构本身就具有复杂的层级性,并直接影响音符的稳定性,推动音乐发展,给音乐带来动力性。音乐中分割音符时值最常见的法则是二分法,即按照二分之一关系分裂出更短的时值长度,每一个时值的音符具有相同的层级水平,不同层级水平的音符组合在一起,交织成复杂的层级网络。尤其是在复拍子、混合拍子等复杂的节拍类型中,节拍结构层级的多样性和复杂性尤为凸显。本文探究听者如何在复杂层级关系中整合节拍结构,对联合层级控制理论(Asano et al., 2021)进行拓展,揭示大脑如何调动不同的神经网络来加工节拍结构中的各种层级水平。另外,节拍结构处于具有层级结构的音乐组织中,因此,探究节拍结构在不同时间尺度单元中的整合加工是必要的。以往研究主要探讨小节或者乐句水平上简单的节拍结构加工(Sun et al., 2018; Sun, Thompson, et al., 2020),研究结果不能推广到多乐句或乐段的长时程音乐结构单元。在长时程音乐结构单元中,听者需要同时依据近距离和远距离的依存关系整合具有嵌

套关系的音乐事件,是对工作记忆和音乐结构知识的双重考量。在真实的音乐聆听体验中,远距离整合加工的重要性不言而喻,但当前研究仅关注了音高维度中和声结构(Koelsch et al., 2013; Ma et al., 2018)。本文探究在乐段水平上,听者如何依据远距离依存关系整合长时程音乐单元中的嵌套节拍结构,揭示复杂音乐组织中节拍结构的加工机制。

第三,本文通过音乐材料的设计,不仅能够控制心理声学因素的干扰,而且更具有生态效度。以往研究通常截取现成的音乐片段(Doelling & Poeppel, 2015),或者采用等时间间隔的简单节奏序列(Fujioka et al., 2012, 2015; Li et al., 2019)。前者混淆了节拍结构之外的音高、力度等影响因素,后者虽然能够排除声学因素的干扰,但是其枯燥和不真实性,不符合真实的音乐聆听体验,不适用于探究真实音乐加工。本文的音乐材料改编自音乐作品,根据实验设计进行严格的操纵和控制,既包含丰富的时间信息,接近真实音乐的组织形式,又能够排除无关变量的干扰,因而提高了研究结果的可靠性和推广性。

综上所述,在预期编码理论的框架下,本文围绕节拍结构预期和整合两个加工过程展开,4个研究构成一个有机整体、共同回答音乐节拍结构认知机制的科学问题。成熟的理论基础确保了研究的可行性;实验材料改编自真实音乐作品,但是经过了严格的操纵和控制,这保证了研究结果的内部效度和生态效度;EEG能够精细地捕捉大脑对音乐展开中每个音符变化的实时响应,在探究音乐节拍结构加工脑机制及其动态进程中采用该方法可行而必要的。本文不仅有助于揭示音乐结构认知的本质,为音乐认知神经模型的构建奠定基础,同时,还将为音乐鉴赏活动与音乐美学研究提供客观依据,在音乐领域具有潜在的应用前景。

参考文献

- 蒋存梅. (2016). *音乐心理学*. 上海: 华东师范大学出版社.
- 江俊, 王梓梦, 万璇, 蒋存梅. (2014). 音乐时间加工的影响因素. *心理科学进展*, 22(4), 650–658.
- 欧阳珣, 戴志强. (2010). 音乐节拍认知的研究评述. *心理科学进展*, 18(11), 1692–1699.
- 孙丽君, 周临舒, 阎芮平, 蒋存梅. (2017). 旋律语调疗法及其对失语症的临床应用. *心理科学*, 40(1), 231–237.
- 张晶晶, 梁啸岳, 陈伊笛, 陈庆荣. (2020). 音乐句法加工的认知机制与音乐结构的影响模式. *心理科学进展*, 28(6), 883–892.
- 张巍. (2019). 20世纪音乐节奏的研究——若干问题及分析. *音乐艺术(上海音乐学院学报)*, (1), 74–89.
- 张雪, 袁佩君, 王莹, 蒋毅. (2016). 知觉相关的神经振荡-外界节律同步化现象. *生物化学与生物物理进展*, 43(4), 308–315.
- Arnal, L. H., Doelling, K. B., & Poeppel, D. (2014). Delta-Beta coupled oscillations underlie temporal prediction accuracy. *Cerebral Cortex*, 25(9), 3077–3085.
- Asano, R., Boeckx, C., & Seifert, U. (2021). Hierarchical control as a shared neurocognitive mechanism for language and music. *Cognition*, 216, 104847.
- Barnes, R., & Jones, M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology*, 41(3), 254–311.
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21.
- Doelling, K. B., & Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences of the United States of America*, 112(45), E6233–E6242.
- Fitch, W. T. (2013). Rhythmic cognition in humans and animals: Distinguishing meter and pulse perception. *Frontiers in Systems Neuroscience*, 7, 68.
- Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of The Royal Society B Biological Sciences*, 364(1521), 1211–1221.
- Fujioka, T., Ross, B., & Trainor, L. J. (2015). Beta-band oscillations represent auditory beat and its metrical hierarchy in perception and imagery. *Journal of Neuroscience*, 35(45), 15187–15198.
- Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. *Journal of Neuroscience*, 32(5), 1791–1802.
- Geiser, E., Notter, M., & Gabrieli, J. D. (2012). A corticostriatal neural system enhances auditory perception through temporal context processing. *Journal of Neuroscience*, 32(18), 6177–6182.
- Geiser, E., Sandmann, P., Jäncke, L., & Meyer, M. (2010). Refinement of metre perception-training increases hierarchical metre processing. *European Journal of Neuroscience*, 32(11), 1979–1985.
- Geiser, E., Ziegler, E., Jäncke, L., & Meyer, M. (2009). Early electrophysiological correlates of meter and rhythm processing in music perception. *Cortex*, 45(1), 93–102.
- Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, 19(5), 893–906.
- Habibi, A., Wirantana, V., & Starr, A. (2014). Cortical

- activity during perception of musical rhythm: Comparing musicians and non-musicians. *Psychomusicology: Music, Mind, and Brain*, 24(2), 125–135.
- Hannon, E. E., Snyder, J. S., Eerola, T., & Krumhansl, C. L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, 30(5), 956–974.
- Harding, E. E., Sammler, D., Henry, M. J., Large, E. W., & Kotz, S. A. (2019). Cortical tracking of rhythm in music and speech. *NeuroImage*, 185, 96–101.
- Huron, D. B. (2008). *Sweet anticipation: Music and the psychology of expectation*. MIT Press.
- Jensen, O., & Colgin, L. L. (2007). Cross-frequency coupling between neuronal oscillations. *Trends in Cognitive Sciences*, 11(7), 267–269.
- Jia, H. (2019). Microstate Analysis. In L. Hu, & Z. Zhang (Eds.), *EEG signal processing and feature extraction*. (pp.141–158). Springer Nature Singapore Pte Ltd.
- Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83(5), 323–355.
- Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, 96(3), 459–491.
- Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, 10(3), 235–266.
- Koelsch, S. (2012). *Brain and music*. Wiley-Blackwell.
- Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: “Nonmusicians” are musical. *Journal of Cognitive Neuroscience*, 12(3), 520–541.
- Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences of the United States of America*, 110(38), 15443–15448.
- Koelsch, S., Vuust, P., & Friston, K. (2019). Predictive processes and the peculiar case of music. *Trends in Cognitive Sciences*, 23(1), 63–77.
- Kuperberg, G. R., Brothers, T., & Wlotko, E. W. (2020). A tale of two positivities and the N400: Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience*, 32(1), 12–35.
- Lakatos, P., Gross, J., & Thut, G. (2019). A new unifying account of the roles of neuronal entrainment. *Current Biology*, 29(18), R890–R905.
- Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of Neurophysiology*, 94(3), 1904–1911.
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106(1), 119–159.
- Large, E. W., & Snyder, J. S. (2009). Pulse and meter as neural resonance. *Annals of the New York Academy of Sciences*, 1169(1), 46–57.
- Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. MIT Press.
- Li, Q., Liu, G., Wei, D., Liu, Y., Yuan, G., & Wang, G. (2019). Distinct neuronal entrainment to beat and meter: Revealed by simultaneous EEG-fMRI. *NeuroImage*, 194, 128–135.
- Li, X., Zhang, Y., Xia, J., & Swaab, T. Y. (2017). Internal mechanisms underlying anticipatory language processing: Evidence from event-related-potentials and neural oscillations. *Neuropsychologia*, 102, 70–81.
- Liégeois-Chauvel, C., Peretz, I., Babai, M., Laguitton, V., & Chauvel, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain*, 121(10), 1853–1867.
- Ma, X., Ding, N., Tao, Y., & Yang, Y. F. (2018). Differences in neurocognitive mechanisms underlying the processing of center-embedded and non-embedded musical structures. *Frontiers in Human Neuroscience*, 12, 425.
- Margulis, E. H. (2005). A model of melodic expectation. *Music Perception: An Interdisciplinary Journal*, 22(4), 663–714.
- Nave-Blodgett, J. E., Snyder, J. S., & Hannon, E. E. (2021). Hierarchical beat perception develops throughout childhood and adolescence and is enhanced in those with musical training. *Journal of Experimental Psychology: General*, 150(2), 314–339.
- Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *Journal of Neuroscience*, 32(49), 17572–17581.
- Nozaradan, S., Schönwiesner, M., Keller, P. E., Lenc, T., & Lehmann, A. (2018). Neural bases of rhythmic entrainment in humans: Critical transformation between cortical and lower-level representations of auditory rhythm. *European Journal of Neuroscience*, 47(4), 321–332.
- Overly, K., & Turner, R. (2009). The rhythmic brain. *Cortex*, 45(1), 1–3.
- Pearce, M. T. (2018). Statistical learning and probabilistic prediction in music cognition: Mechanisms of stylistic enculturation. *Annals of the New York Academy of Sciences*, 1423(1), 378–395.
- Phillips, C., Kazanina, N., & Abada, S. H. (2005). ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research*, 22(3), 407–428.
- Rankin, S. K., Large, E. W., & Fink, P. W. (2009). Fractal tempo fluctuation and pulse prediction. *Music Perception: An Interdisciplinary Journal*, 26(5), 401–413.
- Rao, S. M., Mayer, A. R., & Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nature Neuroscience*, 4(3), 317–323.
- Repp, B. H., & Su, Y. H. (2013). Sensorimotor

- synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin & Review*, 20(3), 403-452.
- Rohrmeier, M. A., & Koelsch, S. (2012). Predictive information processing in music cognition: A critical review. *International Journal of Psychophysiology*, 83(2), 164-175.
- Sun, L., Feng, C., & Yang, Y. (2020). Tension experience induced by nested structures in music. *Frontiers in Human Neuroscience*, 14, 210.
- Sun, L., Hu, L., Ren, G., & Yang, Y. (2020). Musical tension associated with violations of hierarchical structure. *Frontiers in Human Neuroscience*, 14, 578112.
- Sun, L., Liu, F., Zhou, L., & Jiang, C. (2018). Musical training modulates the early but not the late stage of rhythmic syntactic processing. *Psychophysiology*, 55(2), e12983.
- Sun, L., Thompson, W. F., Liu, F., Zhou, L., & Jiang, C. (2020). The human brain processes hierarchical structures of meter and harmony differently: Evidence from musicians and nonmusicians. *Psychophysiology*, 57(9), e13598.
- Thaut, M. H., Trimarchi, P., & Parsons, L. (2014). Human brain basis of musical rhythm perception: Common and distinct neural substrates for meter, tempo, and pattern. *Brain Sciences*, 4(2), 428-452.
- Trost, W. J., Labbé, C., & Grandjean, D. (2017). Rhythmic entrainment as a musical affect induction mechanism. *Neuropsychologia*, 96, 96-110.
- Trost, W., & Vuilleumier, P. (2013). 'Rhythmic entrainment' as a mechanism for emotion induction by music: A neurophysiological perspective. In T. Cochrane, B. Fantini & K. R. Scherer (Eds.), *The emotional power of music* (pp. 213-225). Oxford University Press.
- Van der Steen, M. C., & Keller, P. E. (2013). The Adaptation and Anticipation Model (ADAM) of sensorimotor synchronization. *Frontiers in Human Neuroscience*, 7, 253.
- Vuust, P., Gebauer, L. K., & Witek, M. A. (2014). Neural underpinnings of music: The polyrhythmic brain. *Advances in Experimental Medicine and Biology*, 829, 339-356.
- Vuust, P., Heggli, O. A., Friston, K. J., & Kringelbach, M. L. (2022). Music in the brain. *Nature Reviews Neuroscience*, 23(5), 287-305.
- Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C., & Roepstorff, A. (2009). Predictive coding of music-brain responses to rhythmic incongruity. *Cortex*, 45(1), 80-92.
- Vuust, P., Pallesen, K. J., Bailey, C., van Zuijen, T. L., Gjedde, A., Roepstorff, A., & Ostergaard, L. (2005). To musicians, the message is in the meter: Pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *NeuroImage*, 24(2), 560-564.
- Wolff, A., Berberian, N., Golesorkhi, M., Gomez-Pilar, J., Zilio, F., & Northoff, G. (2022). Intrinsic neural timescales: Temporal integration and segregation. *Trends in Cognitive Sciences*, 26(2), 159-173.
- Wu, Q., Sun, L., Ding, N., & Yang, Y. (2024). Musical tension is affected by metrical structure dynamically and hierarchically. *Cognitive Neurodynamics*, 1-22.
- You, S., Sun, L., Li, X., & Yang, Y. (2021). Contextual prediction modulates musical tension: Evidence from behavioral and neural responses. *Brain and Cognition*, 152, 105771.
- You, S., Sun, L., & Yang, Y. (2023). The effects of contextual certainty on tension induction and resolution. *Cognitive Neurodynamics*, 17(1), 191-201.
- Zendel, B. R., Lagrois, M. - É., Robitaille, N., & Peretz, I. (2015). Attending to pitch information inhibits processing of pitch information: The curious case of amusia. *Journal of Neuroscience*, 35(9), 3815-3824.
- Zhou, L., Liu, F., Jiang, J., & Jiang, C. (2019). Impaired emotional processing of chords in congenital amusia: Electrophysiological and behavioral evidence. *Brain and Cognition*, 135, 103577.

The cognitive and neural mechanisms of metric structure in music: A predictive perspective

SUN Lijun¹, YANG Yufang^{2,3}

(¹ College of Arts, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

(² CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China)

(³ Department of Psychology, University of Chinese Academy of Sciences, Beijing 100101, China)

Abstract: Metric structure is the temporal framework of music. It is not only the basis for composers to create music, but also the prerequisite for people to process musical aesthetics, musical emotion and musical meaning. Hence, within the predictive coding framework, the proposed project explores the cognitive and neural mechanisms underlying the prediction and integration of metric structure, using behavioral experiments and electroencephalogram techniques. Specifically, it includes the following four studies: (1)

track the dynamic neural activity during the representation construction of metric structure and the prediction establishment as rhythmic sequences unfold, (2) explore how listeners use prediction errors to update the prediction of metric structure, (3) examine the neural mechanisms underlying the integration of multiple hierarchical metric structure at the phrase level and (4) investigate how listeners integrate nested metric structure according to long-distance dependency at the period level. The above research will reveal the general cognitive mechanism of the processing of musical structure and lay the foundation for the construction of cognitive neural model of music. Simultaneously, the results will provide empirical evidence for music appreciation activities and music aesthetics research, which has potential application prospects in the field of music.

Keywords: music cognition, literary and artistic psychology, neural mechanisms, metric structure, electroencephalogram