

音乐知觉的神经基础：脑成像研究的元分析*

赖 寒¹ 徐 苗¹ 宋宜颖¹ 刘 嘉^{1,2}

(¹北京师范大学认知神经科学与学习国家重点实验室, 北京 100875)

(²中国科学院心理健康重点实验室, 北京 100101)

摘 要 本研究根据音乐加工的层级结构, 对现有的脑成像研究进行了元分析, 探讨了音乐知觉的神经基础。具体而言, 对特异于音乐知觉加工的两个层级, 音程分析和结构分析的神经基础进行了分析, 并在此基础上对比了参与两个层级加工的脑区。结果发现, 音程分析主要的激活分布在双侧颞上回和右侧额下回, 在中央前回、角回和脑岛等脑区也有分布。音程分析在颞上回激活最多, 可能表明颞上回为音程分析的核心区域。结构分析激活分布较广, 主要激活颞上回、颞横回和前额叶区域, 此外, 还激活了下顶叶、缘上回和舌回等顶枕区域。结构分析在前额叶激活最多, 可能表明前额叶为结构分析的核心区域。最后, 对比两层级激活的脑区发现, 二者仅在后侧颞上回存在着重合, 而在绝大部分脑区则表现出分离, 这暗示了音程分析和结构分析通过颞上回进行交流, 并负责音乐不同层面的加工。

关键词 音乐知觉; 音程分析; 结构分析; 功能磁共振; 元分析

分类号 B842

1 引言

音乐能力是人类最基本的能力之一。她具有漫长的历史——在德国费尔斯窟发现的智人骨笛表明人类欣赏和演奏音乐的历史在 40 万年以上 (Higham et al., 2012)。其次, 她具有跨文化的普遍性, 从现代社会到非洲原始部落, 是人类社会所共有的特征 (Nettle & Romaine, 2000); 同时, 她也是人类与生俱来的能力——新生儿已经具有了加工音乐的基本能力, 如提取声音特征 (Moon, Cooper, & Fifer, 1993); 而后天的学习和训练使人们能够更准确的感知音乐, 演唱、演奏或创作音乐。在认知加工层面上, 音乐能力涉及到知觉、记忆、情感、动作控制甚至社会交往等多种认知过程。其中, 音乐的知觉加工是音乐能力的基础——音乐知觉是一个复杂的认知过程, 它将在物理层面上的声音流逐步转换为在心理层面上具有一定意义的旋律感知。本研究将试图从认知神经科学的角度, 对现有

的关于音乐知觉的脑成像研究进行元分析, 通过分析参与音乐知觉过程的脑区来探讨音乐知觉的神经基础。

与语言加工类似, 音乐知觉也存在着层级加工。它包含对初级声音特征(声源、音高、音色等)的基本组成元素进行提取、组织, 以及对高级结构 (Structure) 进行加工 (Bharucha & Krumhansl, 1983; Friederici, 2002; Koelsch & Siebel, 2005; Krumhansl & Cuddy, 2010; Patel, 2003)。根据 Koelsch 等在 2011 年 (Koelsch, 2011) 提出的音乐知觉加工模型, 当物理的声音传入人耳后, 第一个加工阶段是对声音特征的提取阶段 (Auditory feature extraction)。基底神经核团(如上/下丘、膝状体、丘脑)和初级听觉皮层通过对声音信息的解码和分析来提取并分析频率、音色、强度和声源等声音特征。第二个阶段是格式塔片段形成阶段 (Gestalt formation)。此阶段根据格式塔规则, 如相似性、接近性等原则, 对频率类似、节奏类似或音色类似的声音特征进行重构, 将声音

收稿日期: 2012-11-05

* 国家自然科学基金(31230031, 91132703)、国家自然科学基金青年科学基金项目(31100808)和国家重点基础研究计划(2010CB833903)资助。

通讯作者: 徐苗, E-mail: xumiao930@gmail.com

特征组织为具有整体特性的格式塔片段,且开始进行旋律轮廓(Contour)信息的加工。第三阶段为音程分析阶段(Interval analysis),即在格式塔片段的基础上进一步分析旋律轮廓信息的变化,以及和弦内部或旋律内部音(Pitch)与音之间的音程关系,以对旋律的音程进行表征。最后是对音乐结构的加工(Music structure analysis)。在这个阶段,大脑依据音调系统和旋律和谐等原则,通过对和弦(Chord)功能及其关系进行分析,从而对音乐的旋律信息进行进一步的整合(Koelsch, 2011; Koelsch & Siebel, 2005)。在以上四个音乐的层级加工阶段中,第一阶段的声音信息分析和第二阶段的声音特征组合是声音刺激加工的基础阶段,而第三阶段的音程分析和第四阶段的结构分析则是对音乐旋律更为精细的加工。由于音乐知觉是一个更为广泛的过程,包含了从声音特征提取、格式塔形成、音程分析、结构分析等多个过程(Koelsch, 2011; Koelsch & Siebel, 2005)。大量研究表明,声音特征提取和格式塔形成可能为音乐和语言加工所共享(Friederici, 2002; Koelsch, 2011; Tervaniemi et al., 2000; Warren & Griffiths, 2003; Wong, Skoe, Russo, Dees, & Kraus, 2007),而音程分析和结构分析则更特异于音乐知觉——有研究表明如果脑损伤病人在音程分析或音乐结构分析这两级加工的受损,会表现出音乐加工的特异性损伤(Liu et al., 2012; Piccirilli, Sciarra, & Luzzi, 2000; Polk & Kertesz, 1993)。因此,本研究将专注对特异于音乐知觉的音程分析与结构分析的研究,探讨音乐特有的两个层级的神经基础,并在此基础上初步探索两个层级加工的关系,从而揭示音乐知觉的神经基础。

音程分析属于音乐知觉过程中较为高级的加工阶段。音程分析不仅需要对旋律轮廓进行加工,以了解旋律音程变化的整体趋势,而且还需要对旋律内部及和弦内部的音与音之间的音高关系进行分析,了解相对音高的变化。旋律轮廓信息与音程信息通常被整合起来(Patel & Balaban, 2000; Peretz & Zatorre, 2005),完成对音程关系的表征。以往对音程分析的研究发现,让被试利用与旋律轮廓有关的信息判断一段旋律是升序还是降序时,在颞上回(superior temporal gyrus, STG)、额下回的三角部(pars triangularis of the inferior frontal gyrus, IFGt)和脑岛(insular, Ins)等脑区表现出激活(Schwenzer & Mathiak, 2011)。并且,改变和弦音程的大小,让被试判断和弦是否发生改变时,也激活了颞上回和

额下回的三角部等区域(Doeller et al., 2003; Klein & Zatorre, 2011)。此外,在旋律轮廓基本不变的情况下,改变旋律音程的大小,让被试判断旋律是否发生改变时,颞上回、中央前回(precentral gyrus, Prec)、额下回的三角部和角回(angular gyrus, AG)等脑区均有显著的激活(Brown & Martinez, 2007; Foster & Zatorre, 2010; Wehrum et al., 2011)。音程分析的研究所采用的范式不同,所激活的加工区域也不尽相同,音程分析的加工机制及其核心区域仍值得进一步探讨。

结构分析属于音乐知觉的高级加工阶段。结构分析不仅需要和对弦的功能(和谐性、规则性、稳定性)进行分析与评价,还需要根据音调结构与和谐原则对旋律结构进行分析、建构与评估,最终完成音乐结构的加工。关于音乐结构分析的研究发现,当让被试判断一段旋律在变调前后是否不同时,颞横回(Heschl's gyrus, HG)、中央前回和缘上回(supramarginal gyrus, SMG)等区域产生激活(Foster & Zatorre, 2010)。在判断旋律是否走调的音调结构分析任务中,额下回的岛盖部(pars opercularis of the inferior frontal gyrus, IFGo)、额极(frontal pole, Fpole)、额下回的眶部(pars orbitalis of the inferior frontal cortex, IFGor)、中央前回和缘上回等脑区有激活(Schulze, Zysset, Mueller, Friederici, & Koelsch, 2011; Wehrum et al., 2011)。此外,有关和弦与旋律和谐程度的研究在颞横回、额下回、中央前回和额极等脑区发现激活(Foss, Altschuler, & James, 2007; Fujisawa & Cook, 2011; Schmithorst, 2005),暗示这些脑区可能与音乐结构的和谐加工有关。同音程分析一样,结构分析的研究由于其所采用的研究范式不同,其所激活的加工区域也不一致,结构分析的加工机制及其核心区域也仍值得进一步的研究。

通过上述研究可以发现,音程分析和结构分析既有临近的激活区域,也有各自独特的激活区域。那么这两个加工层级的神经机制之间的关系是怎样的呢?目前对二者神经机制关系的探讨主要集中在脑损伤的研究上。一些研究表明,两个加工过程表现出了双分离:当个体的音乐结构分析能力(如:音调感知)受损时,其音程分析能力表现正常(Peretz, 1993; Peretz & Zatorre, 2005; Tramo, Bharucha, & Musiek, 1990);而当音程加工能力受损时,其音乐结构分析能力表现正常(Peretz & Zatorre, 2005; Tramo, Cariani, Delgutte, & Braidà, 2003)。而另一些研究发现二者具有共同神经机制:

Schön 等人 2004 年对脑损伤病人的研究发现, 当个体的右侧颞叶受损时, 其音程分析和结构分析的能力均受到损害(Schön, Lorber, Spacal, & Semenza, 2004)。研究结果的不一致使得音程分析和结构分析这两个加工过程的关系还需进一步研究。

由此可见, 以往对音乐知觉的研究还有两个没有解决的问题。首先, 在音程分析和结构分析这两个加工层级内部, 由于不同研究采用的研究范式和分析方法不尽相同, 目前没有得到稳定的参与各层级加工的核心区域。更重要的是, 音程分析和结构分析神经基础间的关系还没有定论。神经影像是讨论两个阶段的神经机制是分离还是重叠更为直接的证据, 然而目前还缺乏相应的研究。因此, 本研究希望通过元分析的方法以找到稳定参与各层级加工的脑区; 同时, 本研究也希望在探讨音乐两层级加工机制的基础上, 初步探索音乐两层级加工的关系。

2 方法

2.1 文献搜索策略和纳入标准

本研究使用 Pubmed 数据库, 检索自 1986 年到 2011 年 10 月期间的英文文献, 检索关键词为“music”和“fMRI (Functional magnetic resonance imaging)”共检索到 483 篇英文文献。通过阅读每一篇文献的摘要和方法, 将符合以下 5 项特征的文献纳入本研究的元分析中:

(1)研究被试为正常人。以脑损伤病人、乐盲、或音乐家等大脑结构和功能已发生显著改变的特殊群体为被试的研究不纳入元分析中;

(2)被试完成与音程分析和/或结构分析有关的任务, 并进行功能磁共振扫描;

(3)选择特异于音程分析和结构分析这两个加工层级的实验对比条件(Contrast); 去除音乐知觉无关的实验对比条件, 如将不同情绪状态的旋律进行对比或将一段旋律与简单的按键反应进行对比;

(4)研究所发现的激活脑区均报告了其激活峰值(Peak)在标准大脑中的三维坐标;

(5)不包含落在大脑外部、白质部位、小脑和中脑纹状体部位的正性激活, 以及全脑负性激活的脑区。

最终有 16 篇文献, 共计 34 个实验对比条件符合上述标准, 纳入了本研究的元分析(表 1)。

2.2 数据提取和文献描述

所搜索的 16 篇文献共报告了 34 个实验对比条件和 280 个峰值。音程分析的研究有 8 篇文献涉及, 包含了 15 个实验对比条件, 共报告了 63 个峰值。15 个音程分析的对比条件中, 包含了旋律轮廓、旋律音程与和弦音程等三种类型的任务:(1)旋律轮廓有关的任务要求被试判断旋律的轮廓是升序还是降序(Lee, Janata, Frost, Hanke, & Granger, 2011; Schwenzer & Mathiak, 2011); (2)旋律音程有关的任务要求被试辨别固定音序列中音程信息发生变化的变异音(Doeller et al., 2003), 或在旋律轮廓基本不变的情况下, 要求被试利用相对音高信息(Relative pitch)来判断旋律音程是否发生变化(Brown & Martinez, 2007, [Melody discrimination task]; Foster & Zatorre, 2010; Hyde, Peretz, & Zatorre, 2008; Wehrum et al., 2011); (3)和弦音程有关的任务通过在不改变

表 1 16 篇文章的详细信息

音程分析			结构加工		
文献	对比条件	峰值数量	文献	对比条件	峰值数量
(Brown & Martinez, 2007)	1	20	(Foss et al., 2007)	1	7
(Doeller et al., 2003)	3	6	(Foster & Zatorre, 2010)	3	19
(Foster & Zatorre, 2010)	1	11	(Fujisawa & Cook, 2011)	6	111
(Hyde et al., 2008)	2	2	(Janata et al., 2002)	1	2
(Klein & Zatorre, 2011)	3	4	(Koelsch et al., 2005)	1	19
(Lee et al., 2011)	1	3	(Levitin & Menon, 2003)	1	3
(Schwenzer & Mathiak, 2011b)	3	6	(Schmithorst, 2005)	2	23
(Wehrum et al., 2011)	1	11	(Schulze et al., 2011)	1	1
			(Tillmann et al., 2006)	1	12
			(Wehrum et al., 2011)	2	20
总					
8	15	63	10	19	217

和弦轮廓的条件下改变和弦内部的音程关系来测试被试能否区分两个和弦(Klein & Zatorre, 2011)。

有 10 篇文献涉及结构分析, 包含 19 个实验对比条件, 共报告了 217 个峰值。结构分析包括音调结构、旋律和谐及和弦和谐等三种类型的任务: (1) 音调结构有关任务要求被试判断旋律是否走调(Schulze et al., 2011; Wehrum et al., 2011), 或判断旋律变调前后是否不同(Foster & Zatorre, 2010); (2) 旋律和谐有关任务要求被试判断旋律是否和谐(Foss et al., 2007; Janata et al., 2002; Schmithorst, 2005), 或判断旋律结构是否符合旋律和谐原则(Koelsch, Fritz, Schulze, Alsop, & Schlaug, 2005; Levitin & Menon, 2003; Schmithorst, 2005; Tillmann et al., 2006); (3) 和弦和谐有关任务要求被试判断两和弦在和谐程度或稳定程度上的异同(Foss et al., 2007; Fujisawa & Cook, 2011)。

2.3 峰值标准化

根据峰值坐标所依据的标准大脑模板, 有的研究采用的是 Talairach 坐标系, 有的采用的是 MNI (Montreal neurological institute) 坐标系。因此, 我们首先把从 16 篇文章中提取的激活脑区的峰值坐标全部统一到 MNI 坐标系, 然后使用 Caret 软件(Van Essen et al., 2001)将各激活点的坐标投射到 Caret 自带的 FLIRT 模板上, 得到音乐加工中音程分析和结构分析的峰值坐标分布图(图 1)。

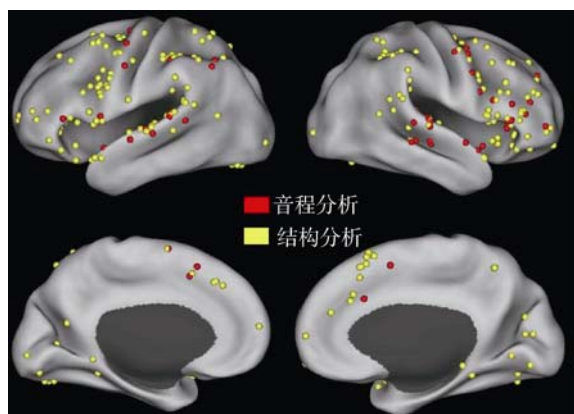


图 1 音程分析与结构分析大脑激活峰值坐标分布图

2.4 峰值聚类

首先, 我们将探讨在对音程分析的多个研究中稳定出现的脑功能区。具体而言, 本研究参照 Vigneau 等在研究中采用的元分析方法(Vigneau et al., 2006; Vigneau et al., 2011), 分别对音程分析所激活的峰值坐标进行聚类分析。具体而言, 我们首

先用 Matlab 软件采用“Ward”算法分别对位于全脑的五个脑区(额叶、颞叶、顶叶、枕叶和脑岛)上参与音程分析脑区的峰值坐标进行层级聚类分析, 使聚类得到的簇(Cluster)满足以下三个标准: (1) 簇内部各峰值间尽可能接近(在本研究中, 我们设定为簇内 X、Y、Z 方向上的标准差均小于 10 mm), (2) 簇与簇间欧式距离(Euclidean distance, ED)尽可能远(在本研究中, 各簇间间距均大于 21 mm), 和(3) 簇具有较高的信度(在本研究中, 我们剔除峰值个数小于 3 或少于 2 篇文章报告的簇)。之后, 我们用 SPSS 软件计算簇中心坐标, 即将簇内各激活点在 x、y、z 方向上的值进行平均, 并采用 FSL (FMRIB Software Library) 软件(Smith et al., 2004; Woolrich et al., 2009)查找每个簇中心坐标所在的脑区, 以此脑区给该簇命名。最后, 为了量化参与音程分析的脑区在大脑皮层的分布情况, 我们计算了参与音程分析的脑区在额叶、颞叶、顶叶、枕叶和脑岛区域的峰值分布比率。计算方法为区域内簇的峰值总数除以全脑的峰值总数。同样的流程被采用来寻找参与结构分析的稳定脑区。

2.5 音程分析与结构分析区域的重叠

为了探查音程分析和结构分析的神经基础之间的关系, 我们进一步分析了分属两个音乐层级的簇在大脑皮层上是否有重叠。具体而言, 重叠分析由以下三个步骤构成。首先, 我们将音程分析与结构分析的簇两两随机进行配对, 并计算出每对簇的中心坐标的欧式距离, 筛选出簇间间距小于 7 mm (约 3 个体素以内) 的配对。随后, 我们将筛选出的配对中的两个簇的峰值混合, 采用 K-means 聚类分析方法, 指定聚类数目为 $K=2$, 迭代次数 $T=10$ 。若重新聚类后, 两簇的峰值混合在一起, 则从描述上表明这两个簇所在的脑区上可能有重合。最后, 为了从统计上进一步验证配对中的两个簇是否存在重合, 我们对由 K-means 聚类筛选出的配对做了多元方差分析。具体而言, 我们对配对中的两个簇的峰值在 X、Y、Z 方向上分别进行了多元方差分析。若配对的簇在 X、Y、Z 方向上的组间差异均不显著, 则可认为音程分析和结构加工在这两个簇所在的脑区上有重合(Reich, Szwed, Cohen, & Amedi, 2011)。

3 结果

首先, 为考察音乐的两个加工层级所激活的区域, 我们将以前研究所发现的与音程分析和结构分

析有关的脑区峰值坐标, 标记在 Caret 自带的 FLIRT 上。结果如图 1 所示, 参与音程分析的脑区在额叶、颞叶、顶叶及脑岛区域存在激活(红色的点), 而结构分析则在额叶、颞叶、顶叶、枕叶及脑岛均存在激活, 且具有更多分布更广的激活(黄色的点)。

由于音程分析和结构分析均激活了大量脑区, 为了获得参与音乐各层级加工的脑区, 我们分别对音程分析和结构分析所激活的峰值坐标进行了层级聚类分析。

具体而言, 对于音程分析, 额叶的 29 个峰值聚为了 8 个簇, 剔除了 2 个峰值数小于 3 的簇; 颞叶的 27 个峰值聚为了 5 个簇, 剔除了 1 个峰值数为 2 的簇; 顶叶的 4 个峰值聚为了 1 个簇; 脑岛的 3 个峰值聚为了 1 个簇。最终, 音程分析共剔除了 5 个峰值, 峰值剔除率为 8%, 共得到 12 个簇。

各簇分布及坐标位置具体如图 2 所示, 各簇中心坐标值及簇内 x、y、z 方向上的标准差见表 2 和表 3。其中, 参与音程分析的 12 个簇, 左右侧各 6 个, 分布在额叶、颞叶、顶叶和脑岛。其中 6 个额叶的簇分布在双侧中央前回、双侧额下回的三角区、左侧前辅助运动皮层(formerly supplementary motor cortex, Pre-SMA)和右侧额下回的岛盖部; 4 个颞叶的簇分布在双侧颞上回后部(posterior of the STG, STGp)以及左侧颞上回前部(anterior of the STG, STGa); 此外, 还有两个簇分布在左侧顶叶角回及右侧脑岛。通过计算各区域的峰值分布比率发现, 音程分析所激活的脑区主要分布在额叶(47%)和颞叶的颞上回(43%)。

对于结构分析, 额叶的 108 个峰值聚为了 20 个簇, 剔除了 6 个峰值数小于 3 或少于 2 篇文章报告的簇; 顶叶的 44 个峰值聚为了 9 个簇, 剔除了 3 个峰值数小于 3 的簇; 颞叶的 35 个峰值聚为了 7 个簇, 剔除了 2 个峰值数小于 3 的簇; 枕叶的 21 个峰值聚为了 7 个簇, 剔除了 5 个峰值数小于 3 或少于 2 篇文章报告的簇; 脑岛的 9 个峰值则聚为了 2 个簇。最终, 结构分析共剔除了 31 个峰值, 峰值剔除率为 14%, 共得到 29 个簇。

参与结构分析的簇在左右半球上各有 15 和 14 个, 分布在额叶、顶叶、颞叶、枕叶和脑岛。其中 14 个额叶的簇分布在双侧额下回的岛盖部、左侧额下回的眶部(pars orbitalis of IFG, IFGor)、双侧额中回(middle frontal gyrus, MFG)、双侧额极、双侧中央前回、左侧旁扣带回(paracingulate gyrus, PCG)

和右侧扣带前回(anterior cingulate gyrus, ACC); 6 个顶叶的簇主要分布在双侧缘上回前部(the anterior of the supramarginal gyrus, SMGa)、双侧缘上回后部(the posterior of the supramarginal gyrus, SMGp)和双侧下顶叶(superior parietal lobule, SPL); 5 个颞叶的簇分布在双侧颞极(temporal pole, Tpole)、左侧颞平面(planum temporal, PT)、左侧颞横回(Heschl's gyrus, HG)和右侧颞上回后部; 2 个枕叶的簇分布在左侧枕叶梭状回(occipital fusiform gyrus, OFG)和右侧舌回(lingual gyrus, LG); 此外还有 2 个簇分布在双侧脑岛。计算各区域的峰值分布比率发现, 结构分析相关峰值主要分布在前额叶(prefrontal cortex, PFC, 34%), 其余的分布于顶叶(22%)、颞叶(18%)、额叶前运动区(13%)、脑岛(5%)、边缘系统(4%)和枕叶(4%)。

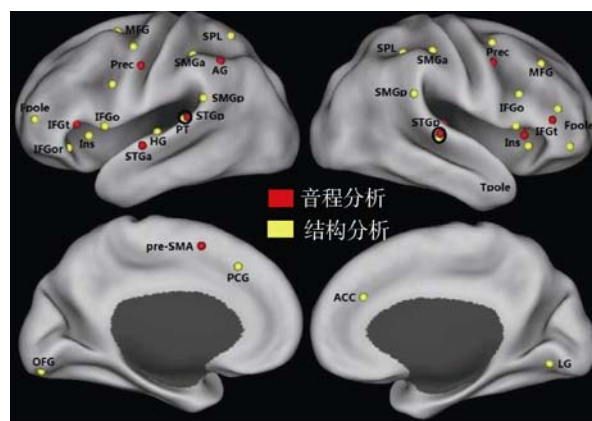


图 2 音程分析与结构分析大脑激活脑区分布图。蓝色圆圈标出的区域为音程分析与结构分析在大脑上空间重合的区域, 位于 Tpole 的簇位于皮层内侧。IFGor: pars orbitalis of the inferior frontal cortex, 额下回的眶部; IFGt: pars triangularis of the inferior frontal gyrus, 额下回的三角区; IFGo: pars opercularis of the inferior frontal gyrus, 额下回的岛盖部; Ins: insular, 脑岛; Fpole: frontal pole, 额极; Prec: precentral gyrus, 中央前回; MFG: middle frontal gyrus, 额中回; SPL: superior parietal lobule, 下顶叶; SMG: supramarginal gyrus, 缘上回; AG: angular gyrus, 角回; STG: superior temporal gyrus, 颞上回; PT: planum temporal, 颞平面; HG: Heschl's gyrus, 颞横回; Tpole: temporal pole, 颞极; pre-SMA, formerly supplementary motor cortex, 前辅助运动皮层; PCG: paracingulate gyrus, 旁扣带回; ACC: anterior cingulate gyrus, 扣带前回; LG: lingual gyrus, 舌回; a: anterior, 前; p: posterior, 后

表 2 音程分析的激活脑区坐标及峰值分布

区域	半球	簇所在位置	簇中心坐标(MNI)			峰值数量	包含研究数	分布率
			X (sd)	Y (sd)	Z (sd)			
额叶	左	Pre-SMA	-9.17(1.15)	3.94(1.13)	53.53(9.60)	4	2	47%
	左	Prec	-46.44(2.73)	-5.72(1.40)	38.20(1.64)	3	3	
	左	IFGt	-47.75(1.42)	22.75(1.16)	5.54(1.75)	3	3	
	右	IFGo	42.83(5.13)	20.44(4.29)	11.63(2.98)	6	2	
	右	IFGt	49.28(7.28)	35.02(9.43)	7.46(7.85)	6	2	
	右	Prec	49.16(5.91)	2.20(4.60)	39.55(1.10)	5	3	
颞叶	左	STGp	-60.00(5.22)	-34.25(4.52)	12.47(2.54)	4	3	43%
	左	STGa	-55.57(3.71)	-9.92(1.03)	-1.76(6.21)	7	3	
	右	STGp	51.19(4.34)	-21.14(4.01)	2.88(7.02)	8	6	
	右	STGp	64.58(1.53)	-22.88(4.86)	3.61(4.41)	6	3	
顶叶	左	AG	-48.13(4.96)	-50.73(1.14)	44.17(5.61)	3	3	5%
脑岛	右	Ins	32.76(3.10)	16.99(8.36)	1.61(7.23)	3	3	5%

注：pre-SMA：formerly supplementary motor cortex，前辅助运动皮层；Prec：precentral gyrus，中央前回；IFGt：pars triangularis of the inferior frontal gyrus，额下回的三角区；IFGo：pars opercularis of the inferior frontal gyrus，额下回的岛盖部；STG：superior temporal gyrus，颞上回；AG：angular gyrus，角回；Ins：insular，脑岛；a：anterior，前；p：posterior，后

为了探查参与音程分析与结构分析的脑区之间的关系，我们对这两个加工层级的簇两两进行了配对，并筛选出簇间间距小于 7 mm 的配对。我们发现，符合这一要求的配对共有 2 对，分别是左侧音程分析的簇，后侧颞上回，与左侧结构分析的簇，颞平面，以及右侧音程分析的簇，后侧颞上回，与右侧结构分析的簇，后侧颞上回，具体峰值坐标如表 4 所示。

进一步分别对筛选出来的配对做 K-means 聚类分析，以描述性地分别探查这两个配对的两个簇所在的脑区上是否有重合。结果发现，这两个配对的簇的峰值均混合在一起，表明这两个配对可能有重合。

再次，为了从统计上进一步验证配对中的两个簇是否存在重合，我们对由 K-means 聚类筛选出的配对做了多元方差分析(表 4)，左侧 STGp 与左侧 PT 在 X、Y、Z 方向上的 *F* 值分别为 1.129 (*p*=0.323)、0.418 (*p*=0.538)、0.392 (*p*=0.551)，右侧 STGp 与右侧 STGp 在 X、Y、Z 方向上的 *F* 值分别为 4.837 (*p*=0.052)、0 (*p*=0.995)、0.031 (*p*=0.863)。这两个配对在 X、Y、Z 方向上组间差异均不显著，表明了这两个配对的两个簇所在的脑区的重合，双侧后侧颞上回为音程分析与结构分析的重合脑区。此外，在大脑的其他区域，簇间距离均大于 7 mm，这两

个加工层级并不存在重合。

4 讨论

本研究根据音乐加工的层级结构，对已有的脑成像研究采用元分析，从而探讨了音乐知觉的神经基础。具体而言，我们对特异于音乐知觉的两个层级，音程分析和结构分析的激活峰值进行了聚类分析，得到了较为稳定的参与音乐层级加工的脑功能区。结果表明，音程分析激活的主要区域为双侧颞上回(43%)；同时，额下回和左侧角回也有稳定激活；此外，音程分析还激活了双侧的中央前回和辅助运动区(Pre-SMA)，以及右侧的脑岛。结构分析激活的区域较音程分析更广，在额叶、颞叶、顶叶、枕叶、脑岛和边缘系统均有分布。最主要的激活区域位于前额叶，包括额下回、额中回和额极。此外，额叶区域的中央前回也有稳定激活。在颞叶区域，与音程分析类似，结构分析激活了颞上回及其临近的颞横回和颞平面等听觉皮层(auditory cortex，AC)。在顶叶结构分析的激活峰值主要分布在下顶叶和缘上回。在枕叶的激活峰值分布在舌回和梭状回。在脑岛和边缘系统也存在较少的结构分析的稳定激活。基于以上结果，我们进一步探讨了参与两个层级加工的共享神经基础，发现音程分析和结构

表 3 结构分析的激活脑区坐标及峰值分布

区域	半球	簇所在位置	簇中心坐标(MNI)			峰值数目	包含研究数	分布率
			X (sd)	Y (sd)	Z (sd)			
前额叶	左	IFGo	-54.11(3.15)	10.79(0.47)	2.70(5.07)	4	2	34%
	左	IFGor	-38.45(6.57)	24.02(7.28)	-4.36(1.00)	9	5	
	左	MFG	-29.38(4.02)	3.85(7.85)	51.20(5.60)	9	2	
	左	Fpole	-38.84(6.87)	49.01(8.49)	6.49(3.34)	5	4	
	右	IFGo	49.05(4.58)	14.33(8.11)	25.65(4.62)	9	3	
	右	IFGo	51.26(4.83)	16.18(3.87)	4.11(5.60)	8	6	
	右	MFG	40.24(8.22)	28.02(1.01)	42.31(1.92)	4	2	
	右	Fpole	40.88(7.81)	42.90(1.10)	-9.65(7.50)	7	3	
	右	Fpole	46.16(9.53)	39.02(4.62)	15.36(8.16)	8	5	
边缘系统	左	PCG	-4.36(1.25)	24.60(8.53)	39.41(3.77)	4	3	4%
	右	ACC	12.36(2.17)	31.18(8.02)	24.59(5.04)	3	3	
前运动区	左	Prec	-45.33(4.21)	-4.50(8.59)	51.98(5.51)	6	4	13%
	左	Prec	-48.73(4.61)	5.52(4.18)	27.63(5.07)	12	5	
	右	Prec	40.98(9.73)	1.99(7.77)	50.40(4.64)	6	3	
颞叶	左	PT	-56.26(5.27)	-32.76(2.34)	10.54(5.68)	5	4	18%
	左	HG	-52.95(6.94)	-17.59(5.75)	4.43(3.57)	9	3	
	左	Tpole	-47.34(3.97)	7.03(7.79)	-12.48(2.12)	4	4	
	右	STGp	58.83(6.21)	-22.86(6.92)	3.19(3.84)	6	3	
	右	Tpole	45.96(5.89)	10.13(9.89)	-13.94(5.69)	8	4	
顶叶	左	SMGp	-57.45(2.68)	-40.95(5.47)	24.57(6.11)	7	3	22%
	左	SMGa	-43.93(3.21)	-35.38(7.71)	44.09(2.75)	8	2	
	左	SPL	-27.89(1.04)	-55.38(8.64)	57.35(5.05)	6	3	
	右	SMGp	60.48(5.69)	-35.28(4.28)	25.66(6.48)	6	3	
	右	SMGa	43.28(7.45)	-29.67(5.90)	40.49(2.10)	4	3	
	右	SPL	41.49(4.70)	-44.24(4.16)	50.82(6.03)	9	3	
枕叶	左	OFG	-14.42(1.08)	-85.77(5.37)	-9.27(7.40)	4	3	4%
	右	LG	18.81(1.29)	-72.53(1.16)	-3.12(4.63)	4	2	
脑岛	左	Ins	-36.46(9.67)	17.40(8.98)	3.82(7.68)	3	2	5%
	右	Ins	35.73(6.13)	21.57(4.81)	-3.56(7.79)	6	2	

注：IFGo: pars opercularis of the inferior frontal gyrus, 额下回的岛盖部; IFGor: pars orbitalis of the inferior frontal gyrus, 额下回的眶部; MFG: middle frontal gyrus, 额中回; Fpole: frontal pole, 额极; Prec: precentral gyrus, 中央前回; PCG: paracingulate gyrus, 旁扣带回; ACC: anterior cingulate gyrus, 扣带前回; PT: planum temporal, 颞平面; HG: Heschl's gyrus, 颞横回; Tpole: temporal pole, 颞极; STG: superior temporal gyrus, 颞上回; SMG: supramarginal gyrus, 缘上回; SPL: superior parietal lobule, 下顶叶; OFG: occipital fusiform gyrus, 枕叶梭状回; LG: lingual gyrus, 舌回; Ins: insular, 脑岛; a: anterior, 前; p: posterior, 后

分析在双侧颞上回存在着重合, 表明颞上回可能是音程分析和结构分析的层级交汇点, 暗示了二者在精细听觉加工过程上可能存在的联系。

下面, 我们将详细探讨参与音程分析和结构分析的激活区域及其构成的神经环路, 然后再在环路分析的基础上, 探讨两者的关系。

表 4 音程分析与结构分析的重合脑区坐标、簇间距离及 F 检验统计值

颞叶	音乐层级	重合簇	簇中心坐标(MNI)			ED (mm)	F (p)		
			x	y	z		x	y	z
左侧	音程分析	STGp	-60.00	-34.25	12.47	4.5	1.129	0.418	0.392
	结构分析	PT	-56.26	-32.76	10.54		(p=0.323)	(p=0.538)	(p=0.551)
右侧	音程分析	STGp	64.58	-22.88	3.61	5.8	4.837	0	0.031
	结构分析	STGp	58.83	-22.86	3.19		(p=0.052)	(p=0.995)	(p=0.863)

注：STG：superior temporal gyrus, 颞上回; PT：planum temporal, 颞平面; p：posterior, 后

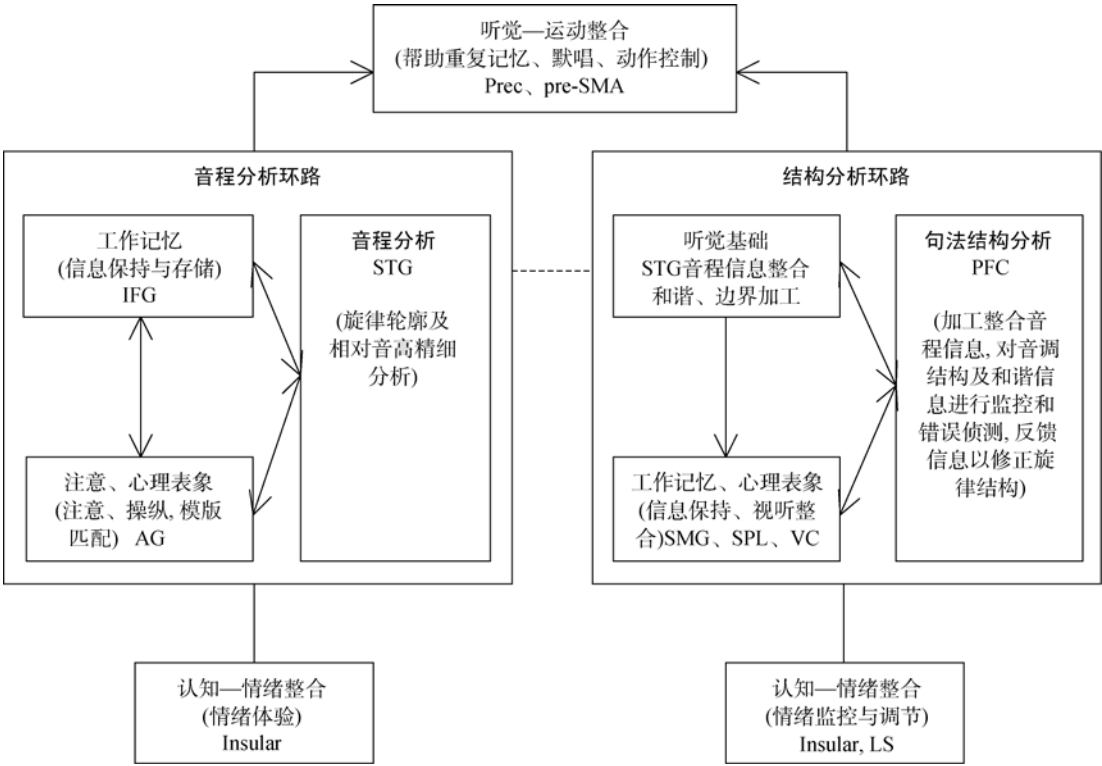


图 3 音程分析和结构分析过程示意图, 虚线表示联系。Prec: precentral gyrus, 中央前回; pre-SMA: formerly supplementary motor cortex, 前辅助运动皮层; IFG: inferior frontal gyrus, 额下回; AG: angular gyrus, 角回; STG: superior temporal gyrus, 颞上回; PFC: prefrontal cortex, 前额叶; SMG: supramarginal gyrus, 缘上回; SPL: superior parietal lobule, 下顶叶; VA: visual area, 视觉区域; Insular: 脑岛; LS: limbic system, 边缘系统

4.1 音程分析的神经基础

结合音程分析的任务及其所激活脑区在其他认知过程中的功能, 我们认为, 音程分析主要涉及 3 条加工环路: 即音程分析环路、听觉—运动环路和认知—情绪环路(图 3)。

4.1.1 音程分析环路

已有研究表明, 与听觉皮层相连的颞上回, 与听觉刺激的加工有关, 不仅参与语言的语音加工 (Pulvermüller & Fadiga, 2010; Vigneau et al., 2006), 而且在音乐知觉的音程分析中发挥着十分重要的作用(Koelsch, 2011; Koelsch & Siebel, 2005; Peretz & Zatorre, 2005)。例如, 颞上回参与音程分析阶段

旋律轮廓的加工, 知觉旋律音程变化的整体趋势, 建立对旋律的整体印象(Hyde et al., 2008; Lee et al., 2011; Schwenzer & Mathiak, 2011); 同时, 它也参与和旋的音程加工, 分析和旋内部音与音间的音高关系, 知觉和旋内部的相对音高变化, 对和旋内部的音程关系进行精细加工(Doeller et al., 2003; Klein & Zatorre, 2011)。并且, 它还参与旋律的音程加工, 分析旋律内部音与音间的音高关系, 知觉旋律内部的相对音高变化, 对旋律的音程关系进行精细分析(Brown & Martinez, 2007; Doeller et al., 2003; Foster & Zatorre, 2010; Wehrum et al., 2011)。此外, 来自脑损伤病人的研究表明, 颞上回受损将影响音

程各方面的加工,如个体对音高及旋律的知觉能力(Liégeois-Chauvel, Peretz, Babai, Laguitton, & Chauvel, 1998)。有关音程分析训练的研究也发现,音程分析训练能够提高个体颞上回的激活(Ellis et al., 2012)。因此,我们认为,颞上回可能为音程分析的核心区域,它不仅在音程分析中所占的峰值比例最多(有 43%的峰值分布于颞上回),而且其在旋律轮廓、和弦音程及旋律音程加工的任务中均有激活,主要负责旋律轮廓、和弦与旋律音程变化的知觉,以及相对音高关系的精细分析。

额下回则与和弦(Janata, Tillmann, & Bharucha, 2002)与旋律(Koelsch, 2011; Koelsch et al., 2008; Koelsch & Siebel, 2005; Nan, Knösche, Zysset, & Friederici, 2008)的工作记忆加工有关。在音程分析中,额下回在旋律轮廓(Schwenzer & Mathiak, 2011)、和弦音程(Doeller et al., 2003)和旋律音程(Brown & Martinez, 2007; Wehrum et al., 2011)的任务中均有激活,反映了音程分析中,额下回对和弦及旋律信息的保持与存储过程,是颞上回完成音程分析的基础。

最后,在音程分析中,在涉及相同/不同判断的任务(Brown & Martinez, 2007; Klein & Zatorre, 2011; Lee, Janata, Frost, Hanke, & Granger, 2011)中还激活了角回。角回不仅参与旋律的注意加工(Nan et al., 2008),而且它还与心理旋转(Halari et al., 2006)、心算(Grabner, Ansari, Koschutnig, Reishofer, & Ebner, 2011; Grabner et al., 2009; Price & Ansari, 2011; Zago et al., 2008)、语言转换(Newman, Bavelier, Corina, Jezzard, & Neville, 2002; Venkatraman, Siong, Chee, & Ansari, 2006)及动作控制(Farrer et al., 2008)等涉及操纵的心理表象加工过程有关。因此,角回的参与不仅反映了它在完成音程分析时,对和弦及旋律信息的注意及保持过程,也反映了其通过操纵抽象的声音刺激,采用模版匹配策略完成实验任务的过程。

基于上述分析,我们认为参与较为高级的听觉加工过程的颞上回,参与工作记忆加工过程的额下回和参与注意及心理表象过程的角回构成了音程分析环路。它是音程分析的核心环路,负责对旋律轮廓及相对音高关系进行精细分析,对和弦及旋律的音程信息进行注意、保持与存储,并使用模版匹配策略完成实验任务的过程。

4.1.2 听觉—运动环路

已有研究表明,听觉和运动在音乐知觉中相互

联系。对普通人的研究也发现,长期的音乐器乐训练不仅有助于运动皮层的发展,同时也促进了听觉皮层的发展(Hyde et al., 2009; Kraus & Chandrasekaran, 2010),表明听觉皮层和运动皮层可能存在关联。并且,有研究通过对比听觉—运动训练组被试和纯听觉训练组被试在音乐训练前后的变化发现,在进行音乐判别任务时,听觉—运动训练组被试比纯听觉组被试产生更强的失匹配负波(mismatch negativity, MMN),听觉运动训练比纯听觉训练更有效,表明了听觉和运动在音乐知觉中相互联系(Pantev, Lappe, Herholz, & Trainor, 2009)。此外,研究还发现,聆听音乐会促使前运动皮层的镜像神经元(Mirror neurons)的激活,他们负责对旋律有关的动作进行控制,进一步表明听觉和运动在音乐知觉中的联系(Levitin & Tirovolas, 2009)。在音程分析中,音程信息的加工激活了前运动皮层的中央前回和前辅助运动皮层,前者主要负责口舌和颞的运动(Mohammadi et al., 2009; Pulvermüller et al., 2006; Schubotz, Anwander, Knösche, Von Cramon, & Tittgemeyer, 2010; Wong, Dziedzic, Talavage, Romito, & Byrd, 2011),后者主要负责动作的控制(Nachev, Wydell, O'Neill, Husain, & Kennard, 2007; Richard Ridderinkhof, Forstmann, Wylie, Burle, & van den Wildenberg, 2010)。在音程分析中,中央前回和前辅助运动皮层均在以旋律为材料的实验任务(Brown & Martinez, 2007; Foster & Zatorre, 2010; Schwenzer & Mathiak, 2011; Wehrum et al., 2011)中存在激活,反映了在旋律加工中,中央前回和前辅助运动皮层可能担任了听觉—运动联合区的角色,整合听觉—运动的加工,以助于重复和记忆旋律信息,促进旋律的加工。语言的元分析表明,中央前回参与语言听觉—运动环路的加工,反映了其在语音加工中所存在的默读的加工过程(Vigneau et al., 2006)。与语音加工类似,我们认为,在音程分析中,中央前回的激活可能反映了旋律加工中所存在的默唱的加工过程。这在音乐的结构分析中也有体现。

基于上述分析,我们认为参与听觉—运动整合过程的中央前回和前辅助运动皮层构成了听觉—运动环路。它是音程分析的外围环路,负责听觉—运动的整合,以助于重复和记忆旋律信息,并控制默唱的加工过程。

4.1.3 认知—情绪环路

大量研究表明,同语言加工一样(Jones, Ward,

& Critchley, 2010; Mutschler et al., 2009), 音乐的加工会引发个体情绪体验的产生(Han, Rho, Jun, & Hwang, 2010; Huq, Bello, & Rowe, 2010; Kim et al., 2010; Omar et al., 2011; Yang, Lin, Su, & Chen, 2008), 并且这种情绪体验的加工大多需要脑岛的参与(Boso, Politi, Barale, & Emanuele, 2006; Jones et al., 2010; Koelsch, 2010; Koelsch & Siebel, 2005)。并且, Koelsch (2011)等人的研究发现, 音乐对情绪的唤起在很早的阶段就已发生, 在音程分析阶段也有体现。在音程分析过程中, 脑岛只在以旋律为材料的实验任务(Foster & Zatorre, 2010; Schwenzer & Mathiak, 2011; Wehrum et al., 2011)中存在激活, 反映了脑岛参与旋律音程加工的情绪唤起过程。

基于上述分析, 我们认为参与情绪唤起过程的脑岛构成了认知—情绪环路。它也是音程分析的外围环路, 负责情绪体验的唤起和加工。

4.2 结构分析的神经基础

结合结构分析的任务及其所激活脑区在其他认知过程中的功能, 我们认为, 结构分析主要涉及3条加工环路: 即结构分析环路、听觉—运动环路和认知—情绪环路(图3)。

4.2.1 结构分析环路

已有研究表明, 位于初级听觉皮层的颞横回与和旋的和谐加工有关(Fishman et al., 2001; Peretz & Zatorre, 2005; Steinschneider & Fishman, 2010)。颞横回受损会影响个体对和旋和谐信息的感知能力(Peretz, Blood, Penhune, & Zatorre, 2001; Peretz, Gagnon, & Bouchard, 1998; Peretz & Zatorre, 2005; Tramo et al., 1990)。在结构分析中, 颞横回在与和弦与旋律和谐有关的任务(Foster & Zatorre, 2010; Fujisawa & Cook, 2011; Wehrum et al., 2011)中存在激活, 可能反映了颞横回在结构分析中负责感知和旋和谐信息的功能。研究表明, 颞平面和颞上回也参与和旋的和谐加工(Peretz et al., 1998), 他们与颞横回一起, 为前额叶进一步进行旋律句法结构层面的和谐加工提供基础。此外, 位于次级听觉皮层的颞平面还与旋律片段的边界加工有关(Nan et al., 2008)。在结构分析中, 颞平面主要在与旋律加工有关的任务中(Foster & Zatorre, 2010; Koelsch et al., 2005; Tillmann et al., 2006)存在激活, 可能反映了颞平面在旋律加工过程中的边界信息加工作用, 为前额叶进一步进行旋律句法结构的分析提供基础。已有研究表明, 颞上回还与旋律句法结构的语音信

息整合有关(Koelsch, 2009; Koelsch & Siebel, 2005)。在结构分析中, 颞上回在和弦加工(Fujisawa & Cook, 2011)及旋律加工(Foster & Zatorre, 2010; Koelsch, Fritz, Schulze, Alsop, & Schlaug, 2005)中均有激活, 反映了颞上回参与经由音程分析后的音高信息的整合并形成复杂有序的声音序列的过程, 为前额叶进一步进行旋律句法结构的分析提供基础。

前额叶, 特别是额下回区域, 与句法结构的加工紧密相关(Fadiga, Craighero, & D'Ausilio, 2009), 不仅参与语言的句法结构分析(Fadiga et al., 2009; Griffiths, Marslen-Wilson, Stamatakis, & Tyler, 2012; Segaert, Menenti, Weber, Petersson, & Hagoort, 2012), 也参与旋律句法结构的加工(Fadiga et al., 2009; Janata et al., 2002; Koelsch, 2011; Koelsch & Siebel, 2005; Perani et al., 2010; Rohrmeier & Koelsch, 2012; Tillmann, 2005)。在结构分析中, 不仅位于前额叶的额下回在音调结构(Levitin & Tirovolas, 2009; Wehrum et al., 2011)、和弦和谐(Foss et al., 2007; Fujisawa & Cook, 2011)及旋律和谐(Koelsch et al., 2005; Schmithorst, 2005; Tillmann et al., 2006)等与结构分析有关的任务中均存在激活, 而且其周围的额中回和额极等前额叶区域也参与音调结构(Foster & Zatorre, 2010; Janata et al., 2002; Schulze et al., 2011; Wehrum et al., 2011)、和弦和谐(Foss et al., 2007; Fujisawa & Cook, 2011)及旋律和谐(Koelsch et al., 2005; Schmithorst, 2005; Tillmann et al., 2006)的加工。前额叶在结构分析中所占的峰值分布比例最多, 约1/3的峰值均分布于前额叶, 暗示了前额叶在结构分析中的重要作用。我们认为, 前额叶可能为音乐结构分析的核心区域。前额叶接收由颞上回整合而来的声音序列信息, 通过工作记忆对声音信息进行保持(D'Esposito, 2007; Grady, Yu, & Alain, 2008; Janata et al., 2002; Nan et al., 2008), 同时, 利用内隐的音调系统知识(Ettlinger, Margulis, & Wong, 2011; Tillmann, 2005; Tillmann, Bharucha, & Bigand, 2000), 进行旋律句法结构的分析与建构。前额叶通过其所具有的对错误信息的监控和侦测(Carrasco, 2012; Nee, Kastner, & Brown, 2011; Sokhadze et al., 2012)的执行功能, 对和旋的稳定程度与和谐程度进行侦测和评估, 参与旋律加工的和弦的功能, 并在此基础上对旋律结构的规则程度与和谐程度进行侦测和评估, 根据所反馈信息对所构建的旋律结构进行调整, 完成旋律

结构的分析。

在结构分析中, 和弦层面的加工(Foss et al., 2007; Fujisawa & Cook, 2011)与旋律层面的加工(Foster & Zatorre, 2010; Koelsch et al., 2005; Schmithorst, 2005; Tillmann et al., 2006; Wehrum et al., 2011)均激活了与听觉和言语工作记忆有关的缘上回(Kirschen, Chen, & Desmond, 2010; Koelsch et al., 2008; Nee et al., 2012; Schulze & Koelsch, 2012)及听觉—视觉整合有关的下顶叶(Molholm et al., 2006; Moran, Molholm, Reilly, & Foxe, 2008), 以及与视觉有关的枕叶梭状回(Meppelink et al., 2009; Pourtois, Schwartz, Spiridon, Martuzzi, & Vuilleumier, 2009; Stokes, Thompson, Cusack, & Duncan, 2009)和舌回(Meppelink et al., 2009; Stokes et al., 2009), 可能反映了结构分析中所存在的心理表象过程。我们认为, 在结构分析中, 缘上回可能担任了旋律句法结构工作记忆辅助加工区的角色, 与前额叶一起, 对和弦及旋律信息进行保持和存储, 同时, 它与由视觉区域(枕叶梭状回、舌回)产生的旋律句法结构表象信息一起, 在下顶叶进行交流和整合, 共同辅助前额叶完成旋律结构的分析。

基于上述分析, 我们认为参与听觉加工过程的颞上回、颞平面和颞横回, 参与旋律结构分析过程的前额叶、参与听觉工作记忆过程的缘上回、及参与心理表象过程的下顶叶、枕叶梭状回和舌回, 构成了结构分析环路。它是结构分析的核心环路, 负责感知和弦的和谐信息, 完成音程信息的整合, 并对音调及句法结构的规则程度与和谐程度进行监控和错误侦测, 根据所反馈信息对所构建的音乐结构进行调整, 并完成听觉—视觉的整合。

4.2.2 听觉—运动环路和认知—情绪环路

结构分析的听觉—运动环路主要通过中央前回实现, 与音程分析的听觉—运动环路类似。

而结构分析的认知—情绪环路主要通过脑岛和边缘系统(扣带前回、旁扣带回)实现。在结构分析中, 脑岛与边缘系统在音调结构(Foster & Zatorre, 2010; Wehrum et al., 2011)、和弦和谐(Foss, Altschuler, & James, 2007; Fujisawa & Cook, 2011)和旋律和谐(Levitin & Menon, 2003; Schmithorst, 2005; Tillmann et al., 2006)的任务中均表现出激活。

同音程分析一样, 脑岛参与旋律结构分析可能反映了伴随结构分析的情绪唤起过程。研究表明, 与边缘系统相连的扣带前回(Carballedo et al., 2011; Cox et al., 2012; Etkin, Egner, & Kalisch, 2011;

Völlm et al., 2006; Yoshimura et al., 2009)和旁扣带回(Vogt, 2005; Völlm et al., 2006)均与负性情绪的加工有关。并且, 扣带前回(Fan, Hof, Guise, Fossella, & Posner, 2008; Pessoa, 2008; Stevens, Hurley, Taber, & Hayman, 2011)还具有认知—情绪功能整合的作用, 它除了与边缘系统相连, 感知负性情绪外, 扣带前回(Dennis, 2010; Ichikawa et al., 2011)还与前额叶发生联系, 参与情绪的调节过程。同时, 扣带前回作为情绪调节的重要区域, 还具有监控和侦测(Abutalebi et al., 2012; Walsh, Buonocore, Carter, & Mangun, 2011)冲突, 及优先处理情绪冲突的功能(Kanske & Kotz, 2010)。在旋律结构分析中, 扣带前回可能担任了冲突监控辅助加工区的角色, 与前额叶一起, 对和弦及旋律的不规则与不和谐信息进行监控和侦测, 并与旁扣带回一起, 对旋律句法结构分析中所引发的情绪冲突进行调节和控制。

基于上述分析, 我们认为参与情绪唤起过程的脑岛和参与情绪调节过程的扣带前回和旁扣带回, 构成了认知—情绪环路, 负责情绪体验的加工, 以及情绪冲突信息的侦测与调节。

4.3 音程分析和结构分析在神经环路上的联系与区别

我们通过对音程分析和结构分析的元分析发现, 音程分析和结构分析均包含三条神经环路, 分别是核心环路(音程分析环路 VS. 结构分析环路)、听觉—运动环路和认知—情绪环路。在神经环路上, 二者既联系, 又区别, 共同完成音乐的知觉加工。

4.3.1 音程分析和结构分析在神经环路上的联系

音程分析和结构分析在神经环路上的联系主要体现在以下两个方面。

第一、音程分析和结构分析在双侧后侧颞上回有重合, 表明音程分析和结构分析存在听觉加工上的联系。参与高级听觉加工的颞上回可能为音程分析和结构分析的层级交汇点。在音程分析阶段, 颞上回负责对旋律轮廓、和弦与旋律音程的变化进行知觉, 并对其相对音高关系进行精细分析, 而在结构分析阶段, 颞上回与颞横回和颞平面一起, 对和弦的和谐信息进行感知, 并整合经由音程分析后的音高信息, 形成复杂有序的声音序列, 完成对音高信息的整合, 为旋律结构的分析提供基础。

第二、结构分析环路是对音程分析环路的拓展和提升。结构分析包含更多的加工过程, 也需要更多自上而下的调控过程的参与。例如, 与音程分析

相比, 结构分析还需要颞横回和颞平面参与和弦和和谐的加工, 也需要扣带前回和旁扣带回进行情绪冲突的监控、侦测及调节。

在音乐知觉中, 音程分析和结构分析经由后侧颞上回联系在一起, 反映了音乐知觉中从音程分析到结构分析的一个连续信息加工过程, 而结构分析相对于音程分析更为复杂的加工过程反映了结构分析在音乐知觉中更为高级的加工。

4.3.2 音程分析和结构分析在神经回路上的区别

然而, 虽然音程分析和结构分析在音乐知觉中相互联系, 但是二者的加工又表现出各自独特的特点, 在神经回路上存在以下三个方面的区别。

第一、音程分析和结构分析的核心环路的核心区域不同。音程分析的核心区域为参与高级听觉加工过程的颞上回, 它负责对旋律轮廓、和弦与旋律音程的变化进行知觉, 以及对相对音高关系进行精细分析。而结构加工的核心区域则为参与旋律结构分析过程的前额叶, 它负责对错误信息进行监控和侦测, 对和弦的稳定程度与和谐程度进行侦测和评估, 在此基础上对旋律结构的规则程度与和谐程度进行侦测和评估, 并根据所反馈信息对所构建的旋律结构进行调整。

第二、虽然音乐分析和结构分析的核心环路均包含了工作记忆和心理表象过程, 但是二者所涉及的脑区并不相同, 即使有的脑区名字相同, 但是他们却不是重合在一起的, 可能反映了他们在功能上的侧重不一样。例如, 在音程分析中参与工作记忆加工的额下回主要与旋律信息的保持与存储有关, 而它在结构分析中除参与旋律工作记忆的加工外, 还主要与旋律句法结构的分析有关。又如在音程分析中, 心理表象过程通过角回完成, 反映了采用模版匹配等策略完成实验任务的过程, 而在结构分析中, 其心理表象过程则通过下顶叶、枕叶梭状回和舌回完成, 反映了旋律信息与旋律结构表象信息的听觉—视觉整合过程。

第三, 虽然音程分析和结构分析均包含听觉—运动环路和认知—情绪环路, 但是除听觉—运动环路外, 二者在认知—情绪回路上所涉及的脑区也不相同。除脑岛外, 结构分析的认知—情绪环路还激活了与边缘系统相连的扣带前回和旁扣带回, 反映了结构分析中还存在着对情绪冲突的监控、侦测及调节的过程。

总之, 在音乐知觉中, 音程分析与结构分析既联系, 又区别, 二者各司其职, 共同完成对旋律的

加工。

5 展望

本研究通过元分析的方法, 首次对音乐知觉特有的两个加工层级, 音程分析和结构分析的神经基础以及二者的关系进行了全面的分析, 从而探讨了音乐知觉的神经基础。虽然本研究所基于的音乐知觉模型主要来自于事件相关电位(event-related potential, ERP)的研究, 但是我们基于fMRI数据所得到的元分析结果与其类似, 从而为建立基于ERP和fMRI数据的更为完善的音乐知觉模型提供了一个新的尝试。

本研究认为, 音程分析和结构分析均包含三条神经环路, 分别是核心环路(音程分析环路 VS. 结构分析环路)、听觉—运动环路和认知—情绪环路。然而, 这六条环路只是基于音程分析和结构分析的元分析结果的一个猜想, 还需要更为科学和严格的方法加以验证。未来研究有必要通过基于静息或者结构等模态的数据来进行网络分析, 探讨音程分析和结构分析的神经加工机制, 验证本文提出的六条神经环路是否正确。并且, 还需要在同一批被试中直接比较音程分析和结构加工的关系, 看看哪些脑区是特异于其中一个层级, 哪些脑区是两个层级共享, 验证本文的研究结果。同时, 由于音程分析是一个相对粗略的加工过程, 包含了旋律轮廓和音程这两类音间关系信息的加工。脑损伤病人的研究表明, 旋律轮廓和音程信息的加工也可能存在分离(Liégeois-Chauvel et al., 1998; Patterson et al., 2002; Peretz & Zatorre, 2005)。因此, 随着相关fMRI研究量的增加, 在未来的元分析中, 有必要对二者的神经基础进行比较, 进一步探查二者的关系。另外, 在本研究中, 我们只关注了音乐旋律层面的两个特有的加工过程, 而并未涉及在音乐知觉中与旋律加工同等重要的节奏层面的加工。脑损伤病人的研究表明, 音乐的旋律加工和节奏加工可能存在分离(Di Pietro et al., 2004; Piccirilli et al., 2000)。因此, 未来研究还有必要将旋律加工和节奏加工的神经基础进行分析和比较, 进一步探讨二者加工分离的区域。

此外, 对比音乐和语言的元分析结果发现(Vigneau et al., 2006; Vigneau et al., 2011), 音乐和语言均包含听觉—运动环路, 以及工作记忆等加工过程的参与, 并且, 无论是音程分析, 还是结构分析, 在各神经环路中, 均激活了参与语言加工的节

点, 如额下回和颞上回等。已有研究表明, 额上回对于语言的加工, 特别是语音和语义的加工十分重要 (Kim, Karunanayaka, Privitera, Holland, & Szaflarski, 2011; Liakakis, Nickel, & Seitz, 2011; Newman & Joanisse, 2011), 它在语言句法结构的加工中也发挥着重要作用 (Fadiga et al., 2009; Griffiths et al., 2012; Meltzer, McArdle, Schafer, & Braun, 2010; Segaert et al., 2012)。颞上回也与语言的语音和语义加工有关 (Kim et al., 2011; Pulvermüller & Fadiga, 2010; Vigneau et al., 2006; Wilson, Tregellas, Slason, Pasko, & Rojas, 2011; Zhang et al., 2011)。此外, 中央前回 (Pulvermüller & Fadiga, 2010; Vigneau et al., 2006)、角回 (Newman et al., 2002; Venkatraman et al., 2006) 和缘上回 (Kirschen et al., 2010; Koelsch et al., 2008; Nee et al., 2012; Schulze & Koelsch, 2012) 和脑岛 (Houdé, Rossi, Lubin, & Joliot, 2010; Mutschler et al., 2009) 等诸多区域也参与语言的加工。音乐与语言加工在加工区域上的相似性暗示了音乐和语言可能存在加工过程上的联系。然而, 由于目前关于音乐与语言加工的关系尚存争议, 有的研究者认为音乐和语言可能存在共有的加工机制 (Fedorenko, Patel, Casasanto, Winawer, & Gibson, 2009; Koelsch, 2011; Koelsch & Siebel, 2005; Schmithorst, 2005), 有的研究者认为音乐和语言分属不同的模块, 各有一套独特的加工机制 (Fedorenko, Behr, & Kanwisher, 2011; Liu et al., 2012; Rogalsky, Rong, Saberi, & Hickok, 2011)。

因此, 为了系统地探讨音乐和语言之间的关系, 未来研究有必要将音乐与语言的元分析结果进行比较, 从音乐与语言的各个加工层级出发, 系统的比较音乐与语言在加工区域上可能存在的重合与分离, 探查音乐和语言是否存在共有的加工机制以及这种共有机制所发生的脑区和重合模式, 探究音乐和语言是否同属同一个加工模块。

参 考 文 献

- 带*的文献进入元分析程序。
- Abutalebi, J., Della Rosa, P. A., Green, D. W., Hernandez, M., Scifo, P., Keim, R., ... Costa, A. (2012). Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cerebral Cortex*, 22(9), 2076–2086.
- Bharucha, J., & Krumhansl, C. L. (1983). The representation of harmonic structure in music: Hierarchies of stability as a function of context. *Cognition*, 13(1), 63–102.
- Boso, M., Politi, P., Barale, F., & Enzo, E. (2006). Neurophysiology and neurobiology of the musical experience. *Functional Neurology*, 21(4), 187–191.
- *Brown, S., & Martinez, M. J. (2007). Activation of premotor vocal areas during musical discrimination. *Brain and Cognition*, 63(1), 59–69.
- Carballedo, A., Scheuerecker, J., Meisenzahl, E., Schoepf, V., Bokde, A., Möller, H. J., ... Frodl, T. (2011). Functional connectivity of emotional processing in depression. *Journal of Affective Disorders*, 134(1), 272–279.
- Carrasco, M. (2012). *Cognitive role of medial PFC in error processing: Lessons learned from healthy children and pediatric OCD, anxiety, and ASD*. Unpublished doctoral dissertation, The University of Michigan.
- Cox, C. L., Uddin, L. Q., Di Martino, A., Castellanos, F. X., Milham, M. P., & Kelly, C. (2012). The balance between feeling and knowing: Affective and cognitive empathy are reflected in the brain's intrinsic functional dynamics. *Social Cognitive and Affective Neuroscience*, 7(6), 727–737.
- D'Esposito, M. (2007). From cognitive to neural models of working memory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1481), 761–772.
- Dennis, T. A. (2010). Neurophysiological markers for child emotion regulation from the perspective of emotion-cognition integration: Current directions and future challenges. *Developmental Neuropsychology*, 35(2), 212–230.
- Di Pietro, M., Laganaro, M., Leemann, B., & Schnider, A. (2004). Receptive amusia: Temporal auditory processing deficit in a professional musician following a left temporo-parietal lesion. *Neuropsychologia*, 42(7), 868–877.
- *Doeller, C. F., Opitz, B., Mecklinger, A., Krick, C., Reith, W., & Schröger, E. (2003). Prefrontal cortex involvement in preattentive auditory deviance detection: Neuroimaging and electrophysiological evidence. *Neuroimage*, 20(2), 1270–1282.
- Ellis, R. J., Norton, A. C., Overy, K., Winner, E., Alsop, D. C., & Schlaug, G. (2012). Differentiating maturational and training influences on fMRI activation during music processing. *Neuroimage*, 60(3), 1902–1912.
- Etkin, A., Egner, T., & Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends in Cognitive Sciences*, 15(2), 85–93.
- Ettlinger, M., Margulis, E. H., & Wong, P. C. M. (2011). Implicit memory in music and language. *Frontiers in Psychology*, 2, 211.
- Fadiga, L., Craighero, L., & D'Ausilio, A. (2009). Broca's area in language, action, and music. *Annals of the New York Academy of Sciences*, 1169(1), 448–458.
- Fan, J., Hof, P. R., Guise, K. G., Fossella, J. A., & Posner, M. I. (2008). The functional integration of the anterior cingulate cortex during conflict processing. *Cerebral Cortex*, 18(4), 796–805.
- Farrer, C., Frey, S. H., Van Horn, J. D., Tunik, E., Turk, D., Inati, S., & Grafton, S. T. (2008). The angular gyrus computes action awareness representations. *Cerebral Cortex*, 18(2), 254–261.
- Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences*, 108(39), 16428–16433.
- Fedorenko, E., Patel, A., Casasanto, D., Winawer, J., & Gibson, E. (2009). Structural integration in language and music: Evidence for a shared system. *Memory & Cognition*, 37(1), 1–9.

- Fishman, Y. I., Volkov, I. O., Noh, M. D., Garell, P. C., Bakken, H., Arezzo, J. C., & Steinschneider, M. (2001). Consonance and dissonance of musical chords: Neural correlates in auditory cortex of monkeys and humans. *Journal of Neurophysiology*, 86(6), 2761–2788.
- *Foss, A. H., Altschuler, E. L., & James, K. H. (2007). Neural correlates of the pythagorean ratio rules. *Neuroreport*, 18(15), 1521–1525.
- *Foster, N. E. V., & Zatorre, R. J. (2010). A role for the intraparietal sulcus in transforming musical pitch information. *Cerebral Cortex*, 20(6), 1350–1359.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6(2), 78–84.
- *Fujisawa, T. X., & Cook, N. D. (2011). The perception of harmonic triads: An fMRI study. *Brain Imaging and Behavior*, 5(2), 109–125.
- Grabner, R. H., Ansari, D., Koschutnig, K., Reishofer, G., & Ebner, F. (2011). The function of the left angular gyrus in mental arithmetic: Evidence from the associative confusion effect. *Human Brain Mapping*, doi:10.1002/hbm.21489.
- Grabner, R. H., Ischebeck, A., Reishofer, G., Koschutnig, K., Delazer, M., Ebner, F., & Neuper, C. (2009). Fact learning in complex arithmetic and figural-spatial tasks: The role of the angular gyrus and its relation to mathematical competence. *Human Brain Mapping*, 30(9), 2936–2952.
- Grady, C. L., Yu, H., & Alain, C. (2008). Age-related differences in brain activity underlying working memory for spatial and nonspatial auditory information. *Cerebral Cortex*, 18(1), 189–199.
- Griffiths, J. D., Marslen-Wilson, W. D., Stamatakis, E. A., & Tyler, L. K. (2012). Functional organization of the neural language system: Dorsal and ventral pathways are critical for syntax. *Cerebral Cortex*, 23(1), 139–147.
- Halari, R., Sharma, T., Hines, M., Andrew, C., Simmons, A., & Kumari, V. (2006). Comparable fMRI activity with differential behavioural performance on mental rotation and overt verbal fluency tasks in healthy men and women. *Experimental Brain Research*, 169(1), 1–14.
- Han, B. J., Rho, S., Jun, S., & Hwang, E. (2010). Music emotion classification and context-based music recommendation. *Multimedia Tools and Applications*, 47(3), 433–460.
- Higham, T., Basell, L., Jacobi, R., Wood, R., Ramsey, C. B., & Conard, N. J. (2012). Testing models for the beginnings of the Aurignacian and the advent of figurative art and music: The radiocarbon chronology of Geißenklösterle. *Journal of Human Evolution*, 62(6), 664–676.
- Houdé, O., Rossi, S., Lubin, A., & Joliot, M. (2010). Mapping numerical processing, reading, and executive functions in the developing brain: An fMRI meta-analysis of 52 studies including 842 children. *Developmental Science*, 13(6), 876–885.
- Huq, A., Bello, J. P., & Rowe, R. (2010). Automated music emotion recognition: A systematic evaluation. *Journal of New Music Research*, 39(3), 227–244.
- Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). Musical training shapes structural brain development. *The Journal of Neuroscience*, 29(10), 3019–3025.
- *Hyde, K. L., Peretz, I., & Zatorre, R. J. (2008). Evidence for the role of the right auditory cortex in fine pitch resolution. *Neuropsychologia*, 46(2), 632–639.
- Ichikawa, N., Siegle, G. J., Jones, N. P., Kamishima, K., Thompson, W. K., Gross, J. J., & Ohira, H. (2011). Feeling bad about screwing up: Emotion regulation and action monitoring in the anterior cingulate cortex. *Cognitive, Affective, & Behavioral Neuroscience*, 11(3), 354–371.
- *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., Tillmann, B., & Bharucha, J. J. (2002). Listening to polyphonic music recruits domain-general attention and working memory circuits. *Cognitive, Affective Behavioral Neuroscience*, 2(2), 121–140.
- Jones, C. L., Ward, J., & Critchley, H. D. (2010). The neuropsychological impact of insular cortex lesions. *Journal of Neurology, Neurosurgery & Psychiatry*, 81(6), 611–618.
- Kanske, P., & Kotz, S. A. (2010). Emotion triggers executive attention: Anterior cingulate cortex and amygdala responses to emotional words in a conflict task. *Human Brain Mapping*, 32(2), 198–208.
- Kim, K. K., Karunanayaka, P., Privitera, M. D., Holland, S. K., & Szaflarski, J. P. (2011). Semantic association investigated with functional MRI and independent component analysis. *Epilepsy Behavior*, 20(4), 613–622.
- Kim, Y. E., Schmidt, E. M., Migneco, R., Morton, B. G., Richardson, P., Scott, J., & Turnbull, D. (2010). Music emotion recognition: A state of the art review. *ISMIR, Utrecht, Netherlands*, 255–266.
- Kirschen, M. P., Chen, S. H. A., & Desmond, J. E. (2010). Modality specific cerebro-cerebellar activations in verbal working memory: An fMRI study. *Behavioural Neurology*, 23(1-2), 51–63.
- *Klein, M. E., & Zatorre, R. J. (2011). A role for the right superior temporal sulcus in categorical perception of musical chords. *Neuropsychologia*, 49(5), 878–887.
- Koelsch, S. (2009). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology*, 15(2), 207–212.
- Koelsch, S. (2010). Towards a neural basis of music-evoked emotions. *Trends in Cognitive Sciences*, 14(3), 131–137.
- Koelsch, S. (2011). Toward a neural basis of music perception—a review and updated model. *Frontiers in Psychology*, 2, 110.
- *Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: An fMRI study. *Neuroimage*, 25(4), 1068–1076.
- Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Müller, K., & Gruber, O. (2008). Functional architecture of verbal and tonal working memory: An FMRI study. *Human Brain Mapping*, 30(3), 859–873.
- Koelsch, S., & Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences*, 9(12), 578–584.
- Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, 11(8), 599–605.
- Krumhansl, C. L., & Cuddy, L. L. (2010). A theory of tonal hierarchies in music. *Music Perception*, 36, 51–87.
- *Lee, Y. S., Janata, P., Frost, C., Hanke, M., & Granger, R. (2011). Investigation of melodic contour processing in the brain using multivariate pattern-based fMRI. *Neuroimage*, 57(1), 293–300.
- *Levitin, D. J., & Menon, V. (2003). Musical structure is processed in "language" areas of the brain - a possible role

- for Brodmann Area 47 in temporal coherence. *Neuroimage*, 20(4), 2142–2152.
- 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.
- 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.
- Liakakis, G., Nickel, J., & Seitz, R. J. (2011). Diversity of the inferior frontal gyrus-a meta-analysis of neuroimaging studies. *Behavioural Brain Research*, 225(1), 341–347.
- Liu, F., Jiang, C., Thompson, W. F., Xu, Y., Yang, Y., & Stewart, L. (2012). The mechanism of speech processing in congenital amusia: Evidence from Mandarin speakers. *PLoS One*, 7(2), e30374.
- Meltzer, J. A., McArdle, J. J., Schafer, R. J., & Braun, A. R. (2010). Neural aspects of sentence comprehension: Syntactic complexity, reversibility, and reanalysis. *Cerebral Cortex*, 20(8), 1853–1864.
- Meppelink, A. M., De Jong, B. M., Renken, R., Leenders, K. L., Cornelissen, F. W., & Van Laar, T. (2009). Impaired visual processing preceding image recognition in Parkinson's disease patients with visual hallucinations. *Brain*, 132(11), 2980–2993.
- Mohammadi, B., Kollewe, K., Samii, A., Krampfl, K., Dengler, R., & Münte, T. F. (2009). Decreased brain activation to tongue movements in amyotrophic lateral sclerosis with bulbar involvement but not Kennedy syndrome. *Journal of Neurology*, 256(8), 1263–1269.
- Molholm, S., Sehatpour, P., Mehta, A. D., Shpaner, M., Gomez-Ramirez, M., Ortigue, S., & Foxe, J. J. (2006). Audio-visual multisensory integration in superior parietal lobule revealed by human intracranial recordings. *Journal of Neurophysiology*, 96(2), 721–729.
- Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, 16(4), 495–500.
- Moran, R., Molholm, S., Reilly, R., & Foxe, J. (2008). Changes in effective connectivity of human superior parietal lobule under multisensory and unisensory stimulation. *European Journal of Neuroscience*, 27(9), 2303–2312.
- Mutschler, I., Wieckhorst, B., Kowalewski, S., Derix, J., Wentlandt, J., Schulze-Bonhage, A., & Ball, T. (2009). Functional organization of the human anterior insular cortex. *Neuroscience Letters*, 457(2), 66–70.
- Nachev, P., Wydell, H., O'Neill, K., Husain, M., & Kennard, C. (2007). The role of the pre-supplementary motor area in the control of action. *Neuroimage*, 36(S2), T155–T163.
- 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.
- Nee, D. E., Brown, J. W., Askren, M. K., Berman, M. G., Demiralp, E., Krawitz, A., & Jonides, J. (2012). A meta-analysis of executive components of working memory. *Cerebral Cortex*, 23, 264–282.
- Nee, D. E., Kastner, S., & Brown, J. W. (2011). Functional heterogeneity of conflict, error, task-switching, and unexpectedness effects within medial prefrontal cortex. *Neuroimage*, 54(1), 528–540.
- Nettle, D., & Romaine, S. (2000). *The extinction of the world's languages*. Oxford: Oxford University Press.
- Newman, A. J., Bavelier, D., Corina, D., Jezard, P., & Neville, H. J. (2002). A critical period for right hemisphere recruitment in American sign language processing. *Nature Neuroscience*, 5(1), 76–80.
- Newman, R. L., & Joanisse, M. F. (2011). Modulation of brain regions involved in word recognition by homophonous stimuli: An fMRI study. *Brain Research*, 1367, 250–264.
- Omar, R., Henley, S., Bartlett, J. W., Hailstone, J. C., Gordon, E., Sauter, D. A., ... Warren, J. D. (2011). The structural neuroanatomy of music emotion recognition: Evidence from frontotemporal lobar degeneration. *Neuroimage*, 56(3), 1814–1821.
- Pantev, C., Lappe, C., Herholz, S. C., & Trainor, L. (2009). Auditory-somatosensory integration and cortical plasticity in musical training. *Annals of the New York Academy of Sciences*, 1169(1), 143–150.
- Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, 6(7), 674–681.
- Patel, A. D., & Balaban, E. (2000). Temporal patterns of human cortical activity reflect tone sequence structure. *Nature*, 404(6773), 80–84.
- Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., & Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron*, 36(4), 767–776.
- Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., ... Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proceedings of the National Academy of Sciences*, 107(10), 4758–4763.
- Peretz, I. (1993). Auditory atonalia for melodies. *Cognitive Neuropsychology*, 10(1), 21–56.
- Peretz, I., Blood, A. J., Penhune, V., & Zatorre, R. (2001). Cortical deafness to dissonance. *Brain*, 124(5), 928–940.
- Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, 68(2), 111–141.
- Peretz, I., & Zatorre, R. J. (2005). Brain organization for music processing. *Annual Review of Psychology*, 56, 89–114.
- Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, 9(2), 148–158.
- Piccirilli, M., Sciarna, T., & Luzzi, S. (2000). Modularity of music: Evidence from a case of pure amusia. *Journal of Neurology Neurosurgery Psychiatry*, 69(4), 541–545.
- Polk, M., & Kertesz, A. (1993). Music and language in degenerative disease of the brain. *Brain and Cognition*, 22(1), 98–117.
- Pourtois, G., Schwartz, S., Spiridon, M., Martuzzi, R., & Vuilleumier, P. (2009). Object representations for multiple visual categories overlap in lateral occipital and medial fusiform cortex. *Cerebral Cortex*, 19(8), 1806–1819.
- Price, G. R., & Ansari, D. (2011). Symbol processing in the left angular gyrus: Evidence from passive perception of digits. *Neuroimage*, 57(3), 1205–1211.
- Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, 11(5), 351–360.
- Pulvermüller, F., Huss, M., Kherif, F., del Prado Martin, F. M., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences*, 103(20), 7865–7870.
- Reich, L., Szwed, M., Cohen, L., & Amedi, A. (2011). A ventral visual stream reading center independent of visual experience. *Current Biology*, 21(5), 363–368.
- Richard Ridderinkhof, K., Forstmann, B. U., Wylie, S. A., Burle, B., & van den Wildenberg, W. P. M. (2010).

- Neurocognitive mechanisms of action control: Resisting the call of the Sirens. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(2), 174–192.
- Rogalsky, C., Rong, F., Saberi, K., & Hickok, G. (2011). Functional anatomy of language and music perception: Temporal and structural factors investigated using functional magnetic resonance imaging. *Journal of Neuroscience*, 31(10), 3843–3852.
- Rohrmeier, M. A., & Koelsch, S. (2012). Predictive information processing in music cognition: A critical review. *International Journal of Psychophysiology*, 83(2), 164–175.
- Schön, D., Lörber, B., Späc, M., & Semenza, C. (2004). A selective deficit in the production of exact musical intervals following right-hemisphere damage. *Cognitive Neuropsychology*, 21(7), 773–784.
- *Schmithorst, V. J. (2005). Separate cortical networks involved in music perception: Preliminary functional MRI evidence for modularity of music processing. *Neuroimage*, 25(2), 444–451.
- Schubotz, R. I., Anwander, A., Knösche, T. R., von Cramon, D. Y., & Tittgemeyer, M. (2010). Anatomical and functional parcellation of the human lateral premotor cortex. *Neuroimage*, 50(2), 396–408.
- Schulze, K., & Koelsch, S. (2012). Working memory for speech and music. *Annals of the New York Academy of Sciences*, 1252(1), 229–236.
- *Schulze, K., Zysset, S., Mueller, K., Friederici, A. D., & Koelsch, S. (2011). Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. *Human Brain Mapping*, 32(5), 771–783.
- *Schwenzer, M., & Mathiak, K. (2011). Numeric aspects in pitch identification: An fMRI study. *BMC Neuroscience*, 12, 26.
- Segaert, K., Menenti, L., Weber, K., Petersson, K. M., & Hagoort, P. (2012). Shared syntax in language production and language comprehension—an fMRI study. *Cerebral Cortex*, 22(7), 1662–1670.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., ... Flitney, D. E. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, 23, S208–S219.
- Sokhadze, E. M., Baruth, J. M., Sears, L., Sokhadze, G. E., El-Baz, A. S., & Casanova, M. F. (2012). Prefrontal neuromodulation using rTMS improves error monitoring and correction function in autism. *Applied Psychophysiology and Biofeedback*, 37, 91–102.
- Steinschneider, M., & Fishman, Y. I. (2010). Neural correlates of sensory consonance and dissonance in primate primary auditory cortex. *The Journal of the Acoustical Society of America*, 127, 1951.
- Stevens, F. L., Hurley, R. A., Taber, K. H., & Hayman, L. A. (2011). Anterior cingulate cortex: Unique role in cognition and emotion. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 23(2), 121–125.
- Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *The Journal of Neuroscience*, 29(5), 1565–1572.
- Tervaniemi, M., Medvedev, S., Alho, K., Pakhomov, S. V., Roudas, M. S., Van Zuijen, T. L., & Näätänen, R. (2000). Lateralized automatic auditory processing of phonetic versus musical information: A PET study. *Human Brain Mapping*, 10(2), 74–79.
- Tillmann, B. (2005). Implicit investigations of tonal knowledge in nonmusician listeners. *Annals of the New York Academy of Sciences*, 1060(1), 100–110.
- Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, 107(4), 885–913.
- *Tillmann, B., Koelsch, S., Escoffier, N., Bigand, E., Lalitte, P., Friederici, A. D., & von Cramon, D. Y. (2006). Cognitive priming in sung and instrumental music: Activation of inferior frontal cortex. *Neuroimage*, 31(4), 1771–1782.
- Tramo, M. J., Bharucha, J. J., & Musiek, F. E. (1990). Music perception and cognition following bilateral lesions of auditory cortex. *Journal of Cognitive Neuroscience*, 2(3), 195–212.
- Tramo, M. J., Cariani, P., Delgutte, B., & Braid, L. (2003). Neurobiology of harmony perception. *The Cognitive Neuroscience of Music*, 127–151.
- Van Essen, D. C., Drury, H. A., Dickson, J., Harwell, J., Hanlon, D., & Anderson, C. H. (2001). An integrated software suite for surface-based analyses of cerebral cortex. *Journal of the American Medical Association*, 8(5), 443–459.
- Venkatraman, V., Siong, S. C., Chee, M. W. L., & Ansari, D. (2006). Effect of language switching on arithmetic: A bilingual fMRI study. *Journal of Cognitive Neuroscience*, 18(1), 64–74.
- Vigneau, M., Beaucoeur, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., ... Tzourio-Mazoyer, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *Neuroimage*, 30(4), 1414–1432.
- Vigneau, M., Beaucoeur, V., Hervé, P. Y., Jobard, G., Petit, L., Crivello, F., ... Tzourio-Mazoyer, N. (2011). What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing? Insights from a meta-analysis. *Neuroimage*, 54(1), 577–593.
- Vogt, B. A. (2005). Pain and emotion interactions in subregions of the cingulate gyrus. *Nature Reviews Neuroscience*, 6(7), 533–544.
- Völlm, B. A., Taylor, A. N. W., Richardson, P., Corcoran, R., Stirling, J., McKie, S., ... Elliott, R. (2006). Neuronal correlates of theory of mind and empathy: A functional magnetic resonance imaging study in a nonverbal task. *Neuroimage*, 29(1), 90–98.
- Walsh, B. J., Buonocone, M. H., Carter, C. S., & Mangun, G. R. (2011). Integrating conflict detection and attentional control mechanisms. *Journal of Cognitive Neuroscience*, 23(9), 2211–2221.
- Warren, J. D., & Griffiths, T. D. (2003). Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. *The Journal of Neuroscience*, 23(13), 5799–5804.
- *Wehrum, S., Dege, F., Ott, U., Walter, B., Stippokohl, B., Kagerer, S., ... Stark, R. (2011). Can you hear a difference? Neuronal correlates of melodic deviance processing in children. *Brain Research*, 1402, 80–92.
- Wilson, L. B., Tregellas, J. R., Slason, E., Pasko, B. E., & Rojas, D. C. (2011). Implicit phonological priming during visual word recognition. *Neuroimage*, 55(2), 724–731.
- Wong, D., Dziedzic, M., Talavage, T. M., Romito, L. M., & Byrd, K. E. (2011). Motor control of jaw movements: An fMRI study of parafunctional clench and grind behavior. *Brain Research*, 1383, 206–217.
- Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., & Kraus, N.

- (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, 10(4), 420–422.
- Woolrich, M. W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., ... Smith, S. M. (2009). Bayesian analysis of neuroimaging data in FSL. *Neuroimage*, 45(1), S173–S186.
- Yang, Y. H., Lin, Y. C., Su, Y. F., & Chen, H. H. (2008). A regression approach to music emotion recognition. *IEEE Transactions on Audio, Speech, and Language Processing*, 16(2), 448–457.
- Yoshimura, S., Ueda, K., Suzuki, S., Onoda, K., Okamoto, Y., & Yamawaki, S. (2009). Self-referential processing of negative stimuli within the ventral anterior cingulate gyrus and right amygdala. *Brain Cognition*, 69(1), 218–225.
- Zago, L., Petit, L., Turbelin, M. R., Andersson, F., Vigneau, M., & Tzourio-Mazoyer, N. (2008). How verbal and spatial manipulation networks contribute to calculation: An fMRI study. *Neuropsychologia*, 46(9), 2403–2414.
- Zhang, L., Xi, J., Xu, G., Shu, H., Wang, X., & Li, P. (2011). Cortical dynamics of acoustic and phonological processing in speech perception. *PLoS One*, 6(6), e20963.

The Neural Mechanism underlying Music Perception: A Meta-analysis of fMRI Studies

LAI Han¹; XU Miao¹; SONY Yiyi¹; LIU Jia^{1,2}

(¹ State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China)

(² Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China)

Abstract

Music is part of the human nature. Music perception involves a series of hierarchical processing levels, including auditory feature extraction, gestalt formation, interval analysis, and structure analysis. Previous functional magnetic resonance imaging (fMRI) studies have identified multiple cortical regions involved in different levels of music perception, but these studies yielded inconsistent findings possibly because different research paradigms and data analysis methods were used. Therefore, the neural mechanism underlying music perception is unclear. In this study we performed a meta-analysis to identify brain regions stably recruited by music perception. Specially, we focused on two processing levels specific to music perception, which were interval analysis and structure analysis. In addition, to explore the hierarchical structure of music perception, we examined the possible overlapping and dissociation of the regions involved in interval analysis and structure analysis.

We used meta-analysis approach to re-analyze results from fMRI studies on interval analysis and structure analysis. There were eight studies on interval analysis which included 15 contrasts and 63 peaks, and ten studies on structure analysis which included 19 contrasts and 217 peaks. The coordinates of peak voxels reported in these studies were projected onto a brain template to visualize the distributions of activations recruited by interval analysis and structure analysis respectively. To identify brain regions stably activated by music perception, peaks of each analysis level were segregated into a number of spatially distinct clusters, using a hierarchical classification algorithm that minimized the spatial extent within each cluster while maximizing the peak-to-peak distance between clusters. Then, we calculated mean coordinates for each cluster in the MNI space and used FSL View procedure to identify the anatomical label of each cluster. Finally, to explore whether clusters belonging to the two analysis levels were spatially overlapped or dissociated, we examined the pairs of clusters with distances less than 7mm (about 3 voxels) by K-means clustering.

The meta-analysis identified 12 clusters for interval analysis and 29 clusters for structure analysis. Clusters for interval analysis primarily localized in the superior temporal gyrus (STG) with the peaks distribution rate (PDR) being 43%, while clusters for structure analysis mainly localized in the prefrontal cortex with a PDR of 34%. Although both interval analysis and structure analysis involved frontal, temporal, parietal and insular areas, their activations overlapped only in the posterior portion of the superior temporal gyrus (STGp).

In sum, the meta-analysis suggests that interval analysis and structure analysis are two separate processing levels in the hierarchical structure of music perception, with largely dissociated neural activations. In addition, the two analysis levels intersected only in STGp, which might play a role in information exchanges between interval analysis and structure analysis. Thus, our study provides clues for future researches on neural basis underlying hierarchical structure of music perception.

Key words music perception; interval analysis; structure analysis; fMRI; meta-analysis