Piper E 生物化学与生物物理进展 Progress in Biochemistry and Biophysics 2015, 42(7): 647~661

www.pibb.ac.cn

听觉选择性注意的可塑性及其神经机制研究*

朱莉1)郑奋1)邓娟2)闫铮3)徐晓玲1)**

(¹⁾南昌大学信息工程学院,南昌 330031;²⁾中国医学科学院、北京协和医科大学生物医学工程研究所,天津 300192; ³⁾华侨大学信息科学与工程学院,厦门 361021)

摘要本文旨在研究音乐训练是否能增强基于音高和空间位置的听觉选择性注意力,以及音乐训练对听觉可塑性的神经机制.听觉感知实验中,受试者根据音高差异或空间位置差异,选择两个同时播放的数字之一.听觉认知实验在安静和噪声环境中播放频率分辨率不同的复合音,记录受试者听觉脑干频率跟随响应(frequency-following responses, FFRs).本文提出分析 FFR 的四种方法,即包络相关频率跟随响应(envelope-related frequency-following response, FFR_{EN})的短时锁相值、瞬时相位差 极性图、相位差均值矢量以及时间细节结构相关频率跟随响应(temporal-fine-structure-related frequency-following response, FFR_{ITS})的幅度谱信噪比.实验结果表明,在完成基于音高的任务时,受过音乐训练的受试者准确率更高、反应时间更短.外 界噪声不影响两组人群在基频(fundamental frequency, *F*0)的神经元锁相能力,但是显著降低了谐波处的神经元锁相能力.受 过音乐训练的受试者的神经元在基频处的锁相能力和谐波处抗噪能力均增强,且其 FFR_{ITS}幅度谱信噪比与基于音高的行为学 准确率呈正相关.因此,受过音乐训练的受试者其音高选择性注意感知能力的提高取决于认知神经能力的增强,经过音乐训 练后,*F*0 处 FFR_{ITS}的锁相能力、谐波处 FFR_{ITS}的抗噪和持续锁相能力以及谐波处 FFR_{ITS}幅度谱信噪比均明显增强.音乐训 练对听觉选择性注意具有显著的可塑性.

关键词 听觉可塑性,频率跟随响应,听觉选择性注意,神经机制,音高,音乐训练
 学科分类号 Q42,R338
 DOI: 10.16476/j.pibb.2015.0125

日常环境中常常有多个声源并存,这对人脑是 一个挑战.在大多数情况下,人的听觉神经系统采 取的策略是会选择性地注意所感兴趣的声源,同时 忽略其他声源^[1-2].听觉选择性注意力是听觉研究 的难点和热点.选择声源的主要依据是不同声音的 音高和空间位置这两个特征存在差异.通过音高和 空间位置实现选择性注意的神经机制有所不同.尽 管音高可较好地区分声源,但是,在大多数声场环 境中,音高的差异不如空间位置差异显著^[3].选择 性注意可分为时间选择性注意和空间选择性注意^[4]. 人脑主要通过音高、包络、时间细节结构等实现时 间选择性注意,通过双耳时间差和强度差实现空间 选择性注意.

近期研究表明,音乐训练的经验对人音高感知 能力具有可塑性,即受过音乐训练的受试者对音高 的辨识能力优于没有经过系统音乐训练的受试者^[9] 主要原因是受过音乐训练的受试者的音乐训练内容 一般都以音高为基础,长期的训练使得其对音高的 行为学感知能力增强.另一原因是,受过音乐训练 的人经过不同音色、音律或声部等区分训练,因此 在复杂的声场中,能更有效地区分出所需声音和干 扰声音⁶⁰.与此同时,研究表明,音乐训练能改变 脑皮层下神经系统的结构和功能的适应性. Parbery-Clark等¹⁷通过实验证明了当存在噪声干扰 时,受过音乐训练的受试者对语音刺激声的脑干响 应比未受过音乐训练的受试者的脑干响应更稳健, 且潜伏期更短.此外,脑皮层下的神经系统可塑性

^{*} 国家自然科学基金(61463035), 江西省自然科学基金 (20142BAB217022, 20142BAB207004)和江西省教育厅自然科学基 金(GJJ14193)资助项目.

^{**} 通讯联系人.

Tel: 0791-83969680, E-mail: xuxiaoling98@163.com 收稿日期: 2015-06-02, 接受日期: 2015-06-25

与特定的行为能力存在一定的相关性. 特定音乐训 练圈或者针对特定的听觉任务进行的高强度训练网 能够有选择性地增强人的某些神经功能. 音高及空 间位置的感知能力与人的神经对声音信号时间信息 的编码密切相关. 复合音的音高感知由听觉神经对 低频声音的时间细节结构(temporal fine structure, TFS)锁相和对高频声音的包络锁相共同决定[10-11]. 低频声音时间细节结构中的双耳时间差(interaural time difference, ITD)信息又可以有效地提供左右空 间位置信息^[12]. 听觉认知过程非常复杂,因此需要 找到一个合适的能充分反映听觉认知过程的信号. 听觉脑干响应(auditory brainstem response, ABR)就 是公认的能合理体现人脑对声音认知过程的脑电 (electroencephalography, EEG)信号. 当语音信号 达到脑皮层之前,脑皮层下的听觉神经系统能够非 常精确地复现声音刺激的某些特性(如基频、共振 峰等),这类复现声音的脑电响应可以通过头皮电 位记录下来,即听觉脑干响应 (auditory brainstem) response, ABR). 一个辅音 - 元音音节的 ABR 可 分为暂态——起始响应和稳态——频率跟随响应 (frequency following response, FFR)^[13].

锁相是系统的响应与周期性刺激的特定相位保持同步的一种物理现象.FFR 能够准确地表达刺激声每个周期的时域波形,反映听觉神经的锁相能力^[14].Krishnan^[15]研究 FFR 在 /u/、/i/ 和 /a/ 3 个元音共振峰谐波点的幅度谱峰值,发现神经元的锁相活动信息确实可以保留在 FFR 里.FFR 对声源基频的跟随能力有利于识别说话人和感受韵律^[16].除此之外,FFR 能很好地保留用于语音识别和认知的声音特征,可分析人的听觉神经系统功能,有利于比较不同环境或不同人群的听觉神经系统差异.关于 FFR 的起源至今仍有争论.虽然 FFR 能体现语音信号的基频、起始点、结束点及共振峰等特

征,但是听觉"what"(如声音内容、声音特征等) 和"where"(如声源位置)认知通路很可能起源于脑干^[17].

本文首先实现了受过和未受过音乐训练的受试 者听觉行为学感知实验.研究两组人群以音高为特 征和以空间位置为特征的听觉选择性注意的差异. 实验过程中,两个音高不同且空间位置不同的数字 同时播放,要求受试者根据特征完成数字选择.为 了进一步研究音乐训练是否对听觉神经具有生理学 的可塑性,本文对同一人群进行了脑电认知实验, 记录了受试者的 FFR、分析受试者在复合音刺激下 的包络相关频率跟随响应(envelope-related frequency-following response, FFR_{ENV})和时间细节结构相关频率跟随响应(temporal-fine-structure-related frequency-following response, FFR_{Trs}). 文献[18]和[19]提出了分析 FFR_{ENV} 的长时锁相值(phase locking value, PLV)和 FFR_{ENV} 幅度谱信噪比(signal-to-noise ratio, SNR).本文在此基础上主要利用拔靴抽样、傅里叶变换及短时傅里叶变换等统计理论和信号处理理论,提出了四种新的 FFR 分析方法:FFR_{ENV} 的短时锁相值、瞬时相位差极性图、相位差均值矢量以及 FFR_{Trs} 的幅度谱信噪比.分析了两组人群的 FFR_{ENV} 和 FFR_{Trs} 的特点及其与行为学结果的相关性.并讨论了音乐训练如何影响人的听觉神经元机制,进而提高了听觉选择性注意的行为学准确率.本文最后讨论了 FFR 可能来源于什么脑区.

1 听觉感知与认知实验

本文的听觉感知与认知实验均在美国 Boston University 的 Auditory Neuroscience Laboratory 完成.

1.1 听觉感知行为学实验

行为学实验的受试者包括8名受过音乐训练的 受试者(Musician,其中1名为女性)和8名未经过 训练的受试者(Non-musician, 其中2名为女性). 受试者年龄范围是 21~38 岁, 右利手, 母语是英 文, 听力正常(250 Hz~8 kHz 倍频程的气导听力阈 值小于 10 dB). 受试者均自述没有认知障碍, 且不 熟悉任何音调语言. 受过音乐训练的受试者选择标 准是9岁(5.7±1.8)前开始接受音乐训练,连续弹 奏乐器4年以上(12.1±5.5), 弹奏的乐器包括钢 琴、小提琴、吉他、黑管和大提琴等. 未受过音乐 训练的受试者选择标准是 14 岁前(21.0 ± 9.9)没有 接受任何音乐训练,不会弹奏任何乐器或连续弹奏 乐器不超过1年(0.49±0.29). 其中一位未受过音 乐训练的受试者未完成随后的听觉脑电 FFR 实验, 为了人数的平衡,两组人群对比时,使用7名 Musician 和7名 Non-musician 的实验数据. 全部 受试者均签署了同意由 Boston University Charles River Campus Institutional Review Board 认可的实验 程序协议,并支付了相应实验费用.

实验的刺激声源是 4 个单音节英文数字(1, 2, 3, 4), 声源选自 TIDIGITS Corpus 数据集. 每个声源起始和结束部分加 10 ms 的 cos2 窗函数, 然后用 Pratt 声音处理软件将英文数字的基频调制为 100 Hz. 基频等价于音高,前者是信号领域的定

义,后者是听觉感知的定义.每一个实验中试验的 听觉刺激都同时包括2个数字.2个数字的音高对 称地升高或降低3个半音(如,音高变为84.1 Hz、 118.9 Hz).用 KEMAR 头模型生成的头相关传递 函数(head-related transfer function, HRTF)将2个数 字对称地放置于水平30°和-30°上,距头模型水平 1 m. 混合的刺激声源采样率为24.4 kHz,声压水 平 SPL 60 dB.

行为学实验包括空间位置注意力任务(locationattention task)和音高注意力任务(pitch-attention task),实验过程如图1所示.受试者需要根据视觉 提示,从2个同时播放的数字中选择目标数字.每 一次试验由等待期、视觉提示期、声源刺激期和受 试者响应期构成.视觉提示期间,屏幕上出现一个 提示,指示受试者依据空间位置或音高选择目标声 源.空间位置注意力任务的视觉提示中的左箭头表 示要求受试者注意左侧 30°的数字,右箭头表示注

意右侧-30°的数字. 音高注意力任务的视觉提示中 的上箭头表示要求受试者注意高音的数字,下箭头 表示注意低音的数字.视觉提示阶段,声音出现的 前 10 ms,屏幕左侧会出现一个固定圆点,提示受 试者需要集中注意力. 空间位置和音高注意力任务 要求受试者按标有"1""2""3""4"的4个 键,对应听到的4个数字.两次试验之间的等待间 隔 200 ms,每次试验视觉提示持续 300 ms,声源 时长 200 ms,包括按键在内的响应期间 1500 ms, 注意力集中时长为 200 ms+1500 ms = 1700 ms. 若 在 1500 ms 内没有按键则判为"识别错误". 每位 受试者完成四组行为学实验,每组大约持续4min, 空间位置和音高任务随机出现,但每组中各自出现 的总次数相等.实验界面用 Visual C++编写,声卡 由 System 3 Real-time Signal Processing Systems 及 编写的程序控制,使用 ER-1 入耳式耳塞,实验在 静音室中完成.



Fig. 1 Schematic description of the behavioral experiment

1.2 音乐训练对听觉认知可塑性的脑电实验

1.2.1 受试者与刺激声源

三个实验的受试者在实验前都进行了双耳纯音 听力阈值测试,测试频率为 250~8000 Hz,频率 间隔为倍频程.在各个频率下,所有受试者的单侧 耳的听力阈值都低于 15 dB,且没有中枢神经或外 周听觉损伤病史,所有受试者听力正常.受试者签 署了同意由 Boston University Charles River Campus Institutional Review Board 认可的实验程序协议,并 支付了相应实验费用.实验所用的刺激声音由 MATLAB (Natick, MA)离线生成并存储,在实验 中以 25 kHz 的采样率回放.

受试者包括行为学实验中的8名受过音乐训练 的受试者和7名未受过音乐训练的受试者(有一名 未完成实验). 由8个等强度的纯音的 sine 相位纯 音(100, 200, …, 800 Hz)组成, 基频为 100 Hz. 实验包括安静和噪声两种环境.噪声环境中,加入 了一个持续的高斯噪声(经过 3000 Hz 低通),信噪 比 SNR 等于+10 dB. 复合音长度均为 170 ms,包 括 10 ms cos2 窗函数的升降时间.声源刺激保持 80 dB 声强水平(sound pressure level, SPL),同 时向双耳播放.重复声源刺激的平均间隔为(770 ± 100) ms,单个试验之间的时间扰动目的是为了保 证与刺激无关的信号是随机相位.声音刺激以交错 极性的方式播放,即一半的试验是正极性,而另一 半的试验是负极性.

生物化学与生物物理进展

1.2.2 实验设备与电极放置

硬件设备包括: (a) 一台 PC 机用以控制实验, 包括触发声音传输、存储数据等; (b)特殊用途声 卡控制硬件(System 3 Real-time Signal Processing Systems,包括 D/A 转换与放大、Tucker Davis Technologies、Gainesville、FL),用以将声音传输 至耳塞; (c)入耳式耳塞(ER-1、Etymotic、Elk Grove Village、IL).

记录 FFR 的软件是 BioSemi Active Two System (BioSemi, Amsterdam, Netherlands),采样 率为 16.384 kHz. 传导 Ag-AgCl 头皮电极按照 International 10-20 System of Electrode Placement 标 准放置. FFR 从激活的 Cz 电极记录下来,参考信 号来自于两耳垂间电极. 此外,用两个额外的电极 监测垂直眼动信号(electrooculogram, EOG). 记录 中所有电极的偏置电压小于 20 mv.

1.2.3 实验过程

当头皮电极放置完成之后,受试者进入静音室 (single-walled Eckel C-14 booth, Cambridge, MA), 隔绝声音和电信号. 整个数据采集过程中, 受试者 静坐在椅子上观看一部有字幕的无声电影,且不注 意声源刺激. 实验1的每名受试者完成1组实验, 实验准备和数据采集时间约为2h. 实验1的声源 刺激在安静和噪声环境下分别重复播放4000次. 数据分为8个模块,每个模块1000个试验,从安 静环境开始,安静与噪声环境交替进行.每个模块 内,正负刺激极性在试验之间交替进行,从而产生 2000个重复的刺激(2种极性×2种环境). 实验2 和3的每位受试者在一天之内完成4组实验,每个 实验的实验准备和数据采集时间约为1h,4h.4 组实验分别对应4个不同的声源刺激,顺每位受 试者4组实验顺序相同.每个声源刺激在安静环 境下重复2000次,正负刺激极性在试验之间交替 进行.

2 实验数据分析方法

本文作者在文献[18]和[19]中提出了计算 FFR_{ENV}长时锁相值和 FFR_{ENV}幅度谱 SNR 的方法, 我们在此基础上提出另 4 种分析 FFR 信号的方法, 即 FFR_{ENV}短时锁相值、瞬时相位差极性图、相位 差均值矢量以及 FFR_{TIS}幅度谱 SNR.

2.1 FFR_{ENV} 短时锁相值

为了分析 FFR 是否随时间变化,我们利用短时傅里叶变换(short-term Fourier transform, STFT) 计算短时锁相值,它是频率和时间的函数,具体步骤如下.

a. 针对每位受试者 s,每种环境 x(安静/噪声),每种极性(正极性/负极性),用拔靴法,每次随机抽取 n = 400 个试验结果作为样本,形成样本集.

b. 计算第 *n* 个样本的时域响应 g(t)的短时傅 里叶变换,记为 $S_{n,x}^{s}(t, f)$,相位记为 $\phi_{n,x}^{s}(t, f)$.

$$S_{n,x}^{s}(t,f) = \left| g(\tau)w(\tau-t)e^{-j2\pi/\tau}d\tau \right|$$
(1)

$$\phi_{n,x}^{s}(t,f) = \measuredangle S_{n,x}^{s}(t,f)$$
(2)

其中,f表示频率;t和 τ 表示时间点;w(t)为 短时窗函数,本文取 Hamming 窗,窗长 25 ms, 相邻窗之间 50%重叠,频率分辨率为 4 Hz.

c. 受试者 *s*, 第*i*次抽样,环境 *x*,计算 FFR_{EW}的PLV,记为*P*^{*}_{EW ≠ i}(*t*, *f*).

$$P_{ENV, x, i}^{s}(t, f) = \frac{1}{400} \left| \sum_{n \in POS_{i}} e^{j\phi_{n,x}^{i}(t, f)} + \sum_{m \in NEG_{i}} e^{j\phi_{m,x}^{i}(t, f)} \right|$$
(3)

其中, $n \in POS_i$ 表示第 n 个样本属于第 i 次抽样中的正极性样本集 POS_i , $m \in NEG_i$ 表示第 m 个样本属于第 i 次抽样中的负极性样本集 NEG_i .

d. 每位受试者重复步骤 a~c 共 100 次, 生成 FFR_{ITS} 的 PLV 分布.

e. 计算每位受试者 s,每种环境 x 下 FFR_{ENV}
 的长时锁相值,记为Ŷ_{ENV}
 (t, f).

$$\hat{P}_{ENV, x}^{s}(t, f) = \frac{1}{100} \sum_{i=1}^{100} P_{ENV, x, i}^{s}(t, f)$$
(4)

STFT 方法的计算短时 PLV 能更精确地反映 FFR 相位随时间和频率的变化.计算量非常大,通 常 30 min 的脑电数据需要计算约 48 h. 短时 PLV 能够体现试验之间的瞬时相位变化.如果给定频 率,不同的试验之间的瞬时相位变化小,则 PLV 的数值接近1,反之则接近0.

2.2 瞬时相位差极性图及相位差均值矢量

a. 针对每位受试者 s, 每种环境 x(安静/噪声), 每种极性(正极性/负极性), 用拔靴法, 每次随机抽 n=400 取个试验结果作为样本, 形成样本集.

b. 第*n*个样本的正 / 负极性响应 $g_{n, POS}(t)/g_{n, NEG}(t)$ 的短时傅里叶变换分别记为 $S_{n, POS, x}(t, f)/S_{n, NEG, x}(t, f)$, 对应的相位分别记为 $\theta_{n, POS, x}(t, f)/\theta_{n, NEG, x}(t, f)$.

$$S_{n, POS, x}^{s}(t, f) = \int g_{n, POS}(\tau)w(\tau-t)e^{-j2\pi j\tau}d\tau$$
(5)
$$S_{n, NEG, x}^{s}(t, f) = \int g_{n, NEG}(\tau)w(\tau-t)e^{-j2\pi j\tau}d\tau$$

$$\theta_{n, POS, x}^{s}(t, f) = \checkmark S_{n, POS, x}^{s}(t, f)$$
(6)
$$\theta_{n, NEG, x}^{s}(t, f) = \checkmark S_{n, NEG, x}^{s}(t, f)$$
(7)

c. 计算受试者 s, 第 i 次抽样,环境 x 情况下的相位差均值,记为 θ_{x} (t, f).

$$\theta_{i,x}^{s}(t,f) = \frac{1}{400} \sum_{n=1}^{400} \theta_{n,POS,x}^{s}(t,f) - \theta_{n,NEG,x}^{s}(t,f)$$
(7)

d. 重复步骤 a~c一共 100 次,得到相位差.

$$\theta_{x}^{s}(t,f) = \frac{1}{100} \sum_{i=1}^{100} \theta_{i,x}^{s}(t,f)$$
(8)

e. 将每个频率的相位 θ_x(t, f)在角度极性坐标 轴上,以响应时间和角度为变量作图,即为瞬时相 位差极性图. 它可以直观反映特定频率下的听觉锁 相能力随时间变化的情况.

f. 由相位差构成的单位矢量,在从潜伏期 t_i 开始至响应时长 T的结束的时间轴上计算该单位矢 量的积分,再相对于时间做归一化,即得到相位差 均值矢量 $\overline{\Theta}(n, f), (0 \leq \overline{\Theta}(f) \leq 1)$. 再对 $\overline{\Theta}(n, f)$ 的 n个样本求均值得到第 i 次抽样的 $\overline{\Theta}(i, f)$ 相位差矢 量.最后重复抽样 100 次,对 $\overline{\Theta}(i, f)$ 求均值,得到 相位差均值矢量 $\overline{\Theta}(f)$.

$$\overline{\Theta}(n,f) = \frac{1}{T - t_l} \int_{t_l}^{T} e^{j(\theta_{n,POS,s}(f,t) - \theta_{n,NEG,s}(f,t))} dt$$
(9)

$$\overline{\Theta}(i,f) = \frac{1}{400} \sum_{n=1}^{400} \overline{\Theta}(n,f)$$
(10)

$$\overline{\Theta}(f) = \frac{1}{100} \sum_{i=1}^{100} \overline{\Theta}(i, f)$$
(11)

2.3 FFR_{TFS}幅度谱信噪比

FFR_{TFS}幅度谱 SNR 的计算过程如下.

a. 对每位受试者 s, 每种环境 x(安静 / 噪声), 每种极性(正 / 负极性), 拔靴法随机抽取 n=400 个 试验作为样本, 形成样本集.

b. 计算每个正极性试验的时域 FFR 的 DFT, 记为 $D_{n,x}^{s}(f)$; 每个负极性试验的时域 FFR 的离散 时间傅里叶变换,记为 $D_{n,x}^{s}(f)$.

c. 受试者 s, 第 i 次抽样,环境 x, 计算 n=
 400 个 FFR_{TFS} 的平均对数幅度谱,记为 M^s_{TFS x,i}(f).

$$M_{TFS, x, i}^{s}(f) = 20 \cdot \lg \left| \frac{1}{400} \left(\sum_{n \in POS_{i}} D_{n, x}^{s}(f) - \sum_{m \in NEG_{i}} D_{m, x}^{s}(f) \right) \right| (12)$$

d. 每位受试者重复步骤 a~c 共 100 次, 生成 FFR_{ITS} 原始对数幅度谱的分布.

e. 计算每位受试者 s,每种环境 x的 FFR_{TFS} ^{s, raw} 原始对数幅度谱 \hat{M}_{TFS} ,(f).

$$\hat{M}_{TFS, x}(f) = \frac{1}{100} \sum_{i=1}^{100} M_{TFS, x, i}^{s}(f)$$
(13)

f. 每位受试者的脑电数据,按照计算原始幅 度谱的方法步骤 a~e 重复 1000 次. 其中步骤 c 中 的频谱相位用[0, 2π]内随机分布的相位代替,将 负极性响应反相,计算 FFR_{TFS}本底噪声的对数幅 度谱.

$$\hat{N}_{TFS, x}(f) = \frac{1}{100} \cdot \sum_{i=1}^{100} 20 \cdot \lg\{\frac{1}{400} \middle| (\sum_{n \in POS_i} \Bigl| S_{n, x}^{s}(f) \Bigr| e^{j\omega_n} - \sum_{m \in NEG_i} \Bigl| S_{m, x}^{s}(f) \Bigr| e^{j\omega_n} \Bigr| \} (1 < k < 1000)$$
(14)

g. FFR_{TTS} 原始对数幅度谱相对于本底噪声做 归一化,即得到每位受试者、每种环境下的 FFR_{TTS}

幅度谱 SNR,记为 $\hat{M}_{TFS,x}(f)$.

$$\hat{M}_{TFS, x}^{s}(f) = \hat{M}_{TFS, x}(f) - \frac{1}{1000} \sum_{k=1}^{1000} \hat{N}_{TFS, x}(f)$$
(15)

3 实验结果

3.1 听觉选择性注意行为学实验结果

参与实验的受过音乐训练与未受过音乐训练的 受试者全部完成了行为学任务,实验结果高于随机 猜测水平(25%),且行为学统计量没有峰值和奇异 点,实验结果有效.图2显示了两类受试者完成空 间位置注意力实验和音高注意力实验的识别准确率 以及正确识别时的反应时间.



Fig. 2 Behavioral data showing effect of cue type (pitch *vs.* location) on accuracy (a) and reaction time (b) of performance in musicians and non-musicians

Error bars show standard error. *P < 0.05; **P < 0.01. \square : Non-musician; \blacksquare : Musician.

经多因素 ANOVA 及配对样本 t 检验统计,有以下结论.

a. 以人群差别(受过音乐训练或未受过音乐训 练)作为主因素,识别准确率存在显著差异[F(1,24)= 32.71; P < 0.001]. 因此,音乐训练可以改善声源 的识别准确率.

b. 以实验任务(空间位置或音高)作为主因素, 识别准确率存在显著差异[F(1, 24) = 26.97; P <
0.001]. 所以,通过空间位置比通过音高更容易准确识别声源.

c. 人群差别和实验任务之间存在交互效应 [*F*(1,24) = 8.69; *P* < 0.01]. 这意味着,音乐训练能 提高以音高为线索的声源识别准确率,并加快反应 时间,但是,并不能提高以空间位置为线索的声源 识别准确率.

d. 配对样本 t 检验的统计结果表明,受过音 乐训练的受试者其音高识别任务的准确率显著高于 未受过音乐训练的受试者[t(6) = 4.462, P < 0.01; musicians: (82.5 ± 2.5)%, SEM; non-musicians: (56.7± 3.9)%, SEM]. 两组人群的空间位置识别任 务的准确率相似,接近显著水平 0.05[t(6) = 2.255, P = 0.059; musicians: (89.1 ± 2)%, SEM; non-musicians: (80.9± 2.2)%, SEM].

e. 人群差别[*F*(1, 24) = 7.53; *P* < 0.05]和实验 任务[*F*(1, 24) = 9.76; *P* < 0.01]都会影响反应时间, 且这两个因素之间没有交互效应[F(1,24) = 0.02; P=0.889].

f. 受过音乐训练与未受过音乐训练的受试者 正确完成空间位置识别任务的反应时间没有明显差 别[t(6) = 1.869, P = 0.104; musicians: (0.832 ± 0.045) s, SEM; non-musicians: (0.927 ± 0.026) s, SEM], 而受过音乐训练的受试者完成音高识别任 务的反应时间低于未经过音乐训练的受试者, 接近 显著水平[t(6) = 2.275, P = 0.049; musicians: (0.941 ± 0.046) s, SEM; non-musicians: (1.045 ± 0.022) s, SEM]. 该结果排除了反应速度与准确率 之间存在权衡, 音乐训练带来的高准确率并不是以 反应时间长作为代价获得的. 所以, 音乐训练能提 高人对音高的识别能力, 但是不能明显提高人对声 源位置的识别能力.

3.2 音乐训练认知实验结果

利用文献[18]的方法计算出 7 名受过音乐训练 的受试者和 7 名未受过音乐训练的受试者的 FFR_{ENV} 长时 PLV,如图 3 所示.图 3a 和 b 分别表示安静 环境和噪声环境下的结果.统计两组人群在基频 *F*0 处 FFR_{ENV} 的 PLV,以及在所有谐波 H2~H8 FFR_{ENV} 的 PLV之和.两组人群对基频 *F*0的锁相能 力受外加噪声影响小.但是,受过音乐训练的受试 者无论是在安静还是噪声环境下,*F*0的锁相值均比 未受过音乐训练的受试者有所提高.



Fig. 3 Long-time PLV of FFR_{ENV} in quiet (a) and noise (b) in response to fundamental (F0) and harmonic frequencies (H2~H8)

The inset bar graphs represent the F0 and the summed representation of H2 \sim H8 in quiet and noise, respectively. Error bars show standard error. * $P \le 0.05$. \square : Non-musician; \blacksquare : Musician.

两因素 ANOVA 及配对样本 t 检验统计结果 表明:

a. 人群差别会导致听神经元锁相能力差异显 著[*F*(1, 24) = 8.91, *P* < 0.01],表明经过音乐训 练会提高人的听觉神经的锁相能力.

b. 以环境因素(安静或噪声)作为主因,不会明显影响锁相能力 PLV[F(1,24)=0.05, P=0.821].

c. 人群差别和环境因素之间不存在交互效应 [F(1, 24) = 1.12, P = 0.303].

d. 音乐训练可以提高人脑对基频的认知. 无 论是在安静还是噪声环境下,受过音乐训练的受试 者对基频 FO(100 Hz)包络的 PLV 均显著高于未受 过音乐训练的受试者[安静: t(6) = 2.985, P < 0.05; musicians: 0.120 ±0.017, SEM; non-musicians: 0.092±0.02, SEM. 噪声: t(6) = 2.574, P < 0.05; musicians: 0.113 ±0.016, SEM; non-musicians: 0.094±0.015, SEM].

e. 音乐训练可以改善噪声环境中对谐波包络的认知能力. 两组人群对谐波的锁相能力均受到噪声的影响. 噪声环境中,受过音乐训练的受试者对谐波 H2~H8(200~800 Hz)的包络神经元锁相编码能力明显强于未受过音乐训练的受试者. 人群差异 [F(1, 24) = 3.13, P = 0.042]和环境因素[F(1, 24) = 35.649, P < 0.001]都显著地影响神经元锁相能力, 两者未见交互效应[F(1, 24) = 0.004, P = 0.949]. 尽管在安静环境中,两组人群对谐波的神经元锁相 能力相当 [t(6) = 1.303, P = 0.199; musicians: 0.396±0.039, SEM; non-musicians: 0.356±0.046, SEM], 但是在噪声影响下, 受过音乐训练的受试 者对谐波包络的神经元锁相能力明显高于未受过音 乐训练的受试者[*t*(6)=2.266, *P* = 0.028; musicians: 0.236±0.031, SEM; non-musicians: 0.189±0.011, SEM].

图 4 是 7 名受过音乐训练的受试者和 7 名未受 过音乐训练的受试者的平均短时 PLV 谱图,时间 分辨率为 20 ms,频率分辨率为 8 Hz.刺激声音的 基频 100 Hz 及谐波 H2~H8(200~800 Hz)均是计 算频率的中心点.由图可知,神经元稳态锁相的持 续时间受音乐训练和声源频率影响.具体结论如下.

a. 在响应时间内两组人群的神经元锁相值基本稳定.

b. 安静环境下,受过音乐训练的受试者对 100 Hz 的锁相值显著增强;频率越高,持续锁相 时间越短.

c. 噪声环境下,受过音乐训练的受试者对 100 Hz 的锁相能力仍然比未受过音乐训练的受试 者更好,并且响应的潜伏期比未受过音乐训练的受 试者更短;受过音乐训练的受试者表现出更强的基 频抗噪声能力以及更好的对高频的持续锁相.

d. 噪声影响了两组人群对 200~800 Hz 谐波 的锁相能力. 谐波点的 PLV 逐渐减弱,甚至降低 至零假设.

图 5 显示了两组人群正 / 负极性响应的瞬时相 位差的极性图及相位差均值矢量.



Fig. 4 Grand average short-term PLV of FFR_{ENV} in response to the complex tone in quiet [musicians (a) and non-musicians (b)] and in noise [musicians (c) and non-musicians (d)]



Fig. 5 Polar of FFR's instantaneous phase difference [musicians (a) and non-musicians (b)] and vector of FFR's grand averaged phase difference [musicians (c) and non-musicians (d)]

由图 5a, b 可知, 在刺激声持续的间期 10~ 180 ms(10 ms 为潜伏期)受过音乐训练的受试者对 刺激声的神经元锁相能力更稳定. 受过音乐训练的 受试者的低频 FFR 响应相位差(如 100 Hz, 200 Hz) 基本一致, 接近 0; 高频(如 700Hz, 800 Hz)相位 差接近 180°. 相位差随时间变化的轨迹稳定. 未 受过音乐训练的受试者的 PLV 瞬时相位差所有频 率都接近 0. 当频率提高时相位差随时间变化的轨 迹混乱.图 5c,d显示了两组人群对包络和 TFS 的 持续锁相能力. 如果受试者的听觉神经元对刺激声 的某个频率的持续锁相能力差,则相位差均值 $\Theta(f)$ 小;反之,持续锁相能力强则Θ(f)大.若相位差均 值矢量 $\Theta(f)$ 接近 0°,则 FFR 锁相于刺激声的包络; $\overline{H}_{\Theta}(f)$ 接近 180°,则表示 FFR 锁相于时域细节结 构. 当声源频率低时, 仅对慢变的包络锁相可以大 大降低神经元放电所需的人体化学物质消耗,听觉 神经编码更快速有效, $\Theta(f)$ 应当接近 0° ; 当频率 提高时,声音的 TFS 丰富,如果要对声音进行准 确编码,就应当锁相于 TFS,对应 $\Theta(f)$ 应当接近 180°. 在低频(100~400 Hz)范围内,两组人群的 $\Theta(f)$ 均靠近 0°,均有良好的包络锁相能力;在高 频(500~800 Hz)范围内,受过音乐训练的受试者 的 $\Theta(f)$ 始终在 180°附近,而未受过音乐训练的受 试者的高频 $\Theta(f)$ 则远离 180°.因此,受过音乐训 练的受试者对声源谐波的 TFS 的持续锁相能力优 于未受过音乐训练的受试者.

3.3 行为学实验结果与认知实验结果的相关性

3.3.1 在谐波处,听觉神经主要是对 TFS 的锁相. 我们计算全部 8 名受过音乐训练的受试者的 FFR_{TFS} 幅度谱 SNR,并分析其与行为学数据的关系,结论如下.

a. 如图 6 所示,受过音乐训练的受试者的 FFR_{TFS}幅度谱 SNR 在谐波 H2~H8(200~800 Hz)的 均值与音高识别任务的识别准确率线性相关(相关 系数 r = 0.7231, P < 0.05).因此,经过音乐训练, 增强了人的听觉神经系统对 TFS 的神经元锁相能 力,其行为学上表现为提高了音高的识别准确率.

b. 受过音乐训练的受试者在基频处的 FFR、 未受过音乐训练的受试者在基频和谐波处的 FFR 均与对应的行为学结果无关(*P*>0.1).

c. 空间位置识别的行为学结果与 FFR在各个

频率的响应也不相关(P>0.1).





A higher phase locking of brainstem response to the TFS of harmonics in quiet was found to correspond to a better behavioral.

3.3.2 在基频处,听觉神经主要是对包络的锁相. 由于心理声学的音高感知与生理声学的基频认知存 在直接对应性,因此,由 3.2 节可知,音乐训练增 强了基频处 FFR_{ENV} 的锁相能力,以及基频处 FFR_{ENV} 的抗噪声能力.因此,音乐训练提高了基频 的神经元锁相能力,在心理感知上直接地增强了人 对音高的感知,从而提高了基于音高感知的听觉选 择性注意的感知.

3.3.3 由于音乐训练引起大脑的特异性可能与训练时长和训练开始年龄相关,我们进一步分析了音乐训练时间时长(以年为单位,精确到小数点后一位表示月份)与12组听觉脑干频率跟随响应的关系.均未发现音乐训练时长与FFR有显著关联.结论如下.

a. 在安静环境下, 音乐训练时长与 F0 处的 FFR_{ENV} 的锁相值(P = 0.75322 > 0.05)和幅度谱信噪 比(P = 0.89001 > 0.05)均无显著相关性.

b. 在噪声环境下, 音乐训练时长与 F0 处的 FFR_{ENV} 的锁相值(P = 0.81037 > 0.05)和幅度谱信噪 比(P = 0.84520 > 0.05)均无显著相关性.

c. 在安静环境下,音乐训练时长与 H2~H8 处的 FFR_{ENV} 的锁相值(P = 0.87771 > 0.05)和幅度谱 信噪比(P = 0.86359 > 0.05)均无显著相关性.

d. 在噪声环境下, 音乐训练时长与 H2~H8 处的 FFR_{ENV} 的锁相值(P = 0.45783 > 0.05)和幅度谱 信噪比(P = 0.16949 > 0.05)均无显著相关性.

e. 在安静环境下, 音乐训练时长与 H2~H8
 处的 FFR_{TFS} 的锁相值(P = 0.83987 > 0.05)和幅度谱
 信噪比(P = 0.73323 > 0.05)均无显著相关性.

f. 在噪声环境下,音乐训练时长与 H2~H8 处的 FFR_{TFS} 的锁相值(P = 0.85180 > 0.05)和幅度谱 信噪比(P = 0.86002 > 0.05)均无显著相关性.

4 讨 论

4.1 音乐训练对听觉选择性注意的可塑性神经机制探讨

音乐训练对人脑可塑性的神经机制非常复杂, 现有的研究尚无定论.我们从宏观的神经系统和微 观的神经物质分别讨论本文结论所对应的可能的神 经机制.

4.1.1 音乐训练对神经系统的可塑性

已有的受过音乐训练与未受过音乐训练的受试 者区别音高的实验结果认为,受过音乐训练的受试 者能更敏锐而准确地发现音高的变化^[20].本文是将 已有的研究延伸至探究音乐训练对高级选择性听觉 注意任务的作用.结果表明,音乐训练有助于提高 听觉神经系统对基频的神经元相位同步,从而提高 基于音高的听觉选择性注意力的识别准确率,并缩 短响应时间. Song 等四选用了中心位于基频和谐 波的 40 Hz 频率分析窗,计算了共振峰过渡区和稳 态区 FFR 的平均幅度谱,指出 FFR 的强基频响应 与听觉感知的提高相关.与之相反,Munte 等四通 过记录脑电也发现钢琴家的声源空间位置识别能力 与受过音乐训练的受试者没有显著差别,而专业指 挥家的空间位置识别准确率才有明显提高. 这与本 文的结论一致,即并未发现音乐训练可以明显提高 空间位置感知能力.

本文认为音乐训练能够提高行为学上的音高识 别能力,对应的神经机制之一是听觉神经系统中神 经元的 FFR_{TFS} 幅度谱 SNR 的提高. 文献[19]指出, FFR_{TES} 具备抗噪声和竞争语音能力强的特征,这也 意味着音乐训练可以提高 FFR_{TFS} 幅度谱 SNR,从 而增强人在噪声声场中的听觉认知能力. 噪声对谐 波处神经元同步锁相能力的影响强于其对基频处的 影响. 在噪声环境下, 受过音乐训练的受试者的谐 波处神经元同步锁相能力优于未受过音乐训练的受 试者,然而,在安静环境中并不存在这种优势.现 有研究表明,音乐训练会提高噪声中的语音感知能 力,并且其对噪声敏感谐波信号的神经元同步能力 增强,二者线性相关[2].本文结果与以上相关研究 结果一致. 受过音乐训练的受试者在基于音高的听 觉选择性注意力的增强将有助于提高人的声音流分 离和声源对象形成,二者是听觉系统完成复杂环境

中语音感知的主要过程四.

音乐训练 FFR 实验与文献[19]FFR 实验的结论 一样,即外加噪声主要影响谐波 FFR 的神经元同 步功能,对基频 FFR 影响不明显.进一步分析显 示,噪声环境下受过音乐训练的受试者的听觉神经 元对噪声的锁相能力高于未受过音乐训练的受试 者,两组人群在安静环境下没有差别.因此,音乐 训练可以提高噪声环境下的听觉认知能力,提高受 噪声干扰的谐波刺激声源的响应,提高噪声中的语 音识别能力^[24].此外,音乐训练能有效地提高选择 性注意音高的能力,有助于增强声音分离和声源目 标形成能力,这两部分正是复杂环境中的人感知语 音的主要过程.因而,音乐训练可提高噪声环境中 的听觉认知能力.

音乐训练不仅会影响与音乐相关的脑区,而且 会改善整个听觉神经生理系统的功能.本文发现音 乐训练明显提高了 TFS 的编码能力.这一现象可 能与神经系统功能和结构的改变相关.经过音乐训 练后,在区分声音细微差别所诱发的大脑皮层电位 更强^[3].受过音乐训练人群的大脑听觉皮层和感觉 运动皮层的灰质均比与未受过音乐训练人群更多^[3], 出现了白质增多的痕迹.音乐训练可能改变海马区 的功能和结构^[3],反映了神经的新生能力.音乐训 练对神经元的存活与连接有显著影响,能提高听觉 神经回路的活性,这些被激活的神经回路在成年后 仍能持续保持稳定^[3].TFS 的分辨需要神经回路网 络的协同作用.因此,音乐训练很可能会改变大脑 物质,改善脑网络连接功能^[3],从而增强 TFS 的编 码能力.

本文中音乐训练带来的快响应时间和强锁相能 力,一部分原因是训练使得脑网络的控制和平衡能 力增强.音乐训练经常要锻炼高阶脑区,很可能增 强听觉神经系统自上而下的控制能力.结合已有的 研究我们可以推测,音乐训练受试者的 ABR/FFR 的增强很可能降低了脑皮层 - 丘脑、脑皮层 - 中脑 和脑皮层 - 耳蜗核的投射,从而增强了感知系统的 自上而下的控制能力^[30].投射的降低决定了训练的 行为学特性和细胞特性的相关度,影响听觉相关的 下丘脑和内侧膝状体.

由于年龄会明显影响人脑的听觉认知响应,为 避免年龄因素和训练时长干扰实验结果,本文受试 者选择的是年龄相仿且音乐训练年限相似的青年 人.研究表明:对接受过音乐训练的成年人而言, 音乐训练时长和训练开始的年龄均与听觉脑干响应 ABR 反映的声音辨别能力无关,随着年龄的增长, 音乐训练的时长对人脑听觉神经的可塑性不显著, 此结论与本文的结论一致.与此同时,大部分讨论 音乐训练与大脑关系的文献中未分析训练时长的影 响,若受试者同为成年人,只对比接受音乐训练和 未接受音乐训练组间差异,而非组内差异.值得指 出的是,学龄前儿童音乐训练时长与其对/ba/和 /ga/的共振峰过度区听觉脑干响应中的相移相关. 此外,训练和发育引起的人脑变化在抑制和兴奋网 络中达到平衡:从儿童至成年抑制网络增强,而在 成年大脑的抑制网络可塑性提高^[1].训练可以调谐 兴奋与抑制网络之间的平衡^[2].即使在成年之后接 受训练,丰富的感知环境也能降低抑制功能.因 此,虽然成年人的大脑可塑性受音乐训练时长影响

此,虽然成年八的八脑內垄住又自示所综时长影响 不显著,但是,音乐训练提供的丰富听觉环境仍能 提高成年人的神经网络功能,从而实现听觉认知的 可塑性.

本文的实验范式保证了实验结果能说明音乐训 练的神经机制.本文的刺激声为周期性复合音,因 此音乐训练提高了听觉神经对声音规律的敏感性, 从而增强认知神经回路,并长久保持^[3],认知回路 能力的增强会表现出听觉行为能力的提升^[34].本文 的受试者为被动听取声音,而选择性注意为主动听 觉认知,从结果可知,主动与被动的听觉认知之间 存在相互作用关系.近来,也有研究发现被动的 ABR 与主动的注意力、工作记忆能力相关^[35].音 乐训练提高了听觉选择性主动注意能力,同时也增 强了被动诱发电位.

此外,本文在 FFR 分析方法上与已有方法有 所差异.a. 已有研究中分析脑电时频信息的方法 是先计算若干试验的时域平均波形,再计算时域平 均 STFT 的幅度谱^[36].本文计算每个独立试验的 FFR 的 STFT, 通过计算瞬时相位的一致性获得短 时 PLV 谱图,进而可以分析 FFR 的随时间变化的 锁相特性.b. 现有研究中,时域平均波形 STFT 的方法旨在估计脑干 FFR 是否能对刺激声锁相, 其谱图与潜伏期和频率有关. 这种方法的前提假设 是时域平均会提高由刺激激发的锁相与刺激起始段 的 FFR. c. 若每个试验的 FFR 瞬时相位是高度一 致的,此时时域平均波形 STFT 的方法与本文短时 PLV 谱图方法等效. 但是, 一般情况下, 每个试 验的响应相位及幅度都存在差异, PLV 能更直接 地估计 FFR 的时间信息,不需要考虑因伪迹等带 来的试验之间响应相位与幅度的变化.

4.1.2 音乐训练对突触、细胞及基因表达的元可 塑性

元可塑性(metaplasticity)指的是训练或经验引 起突触、神经和大脑的改变^[37].元可塑性从训练诱 发的神经和突触的结构改变出发,提供了分析神经 细胞机制的框架.某些胞间信号分子(如神经调质、 儿茶酚胺、荷尔蒙)具有实时而短暂的可塑性.元 可塑性则能够使神经生理和生物化学特性的持久变 化,提高人脑的后续学习潜质.

综合本文和已有研究的结论,可以推测音乐训 练在细胞层面的可塑性主要源自两方面: a. 训练 引起神经突触的可塑性增强; b. 训练使得谷氨酸 盐和脑源性神经营养因子(brain-derived neurotrophic factor, BDNF)受体与听觉系统的交互作用增强, 细胞表达能力增强.蛋白转录因子可以调控基因的 表达,继而控制突触的传导.目前,已识别 2000 余种转录因子, 而训练与转录因子的关系研究尚处 于起步阶段,本文认为,音乐训练引起的大脑结构 和功能的可塑性可能与两类基因表达相关:N甲基 D 天门冬氨酸受体(N-methyl-D-aspartate receptor, NMDAR)亚单元表达和酪氨酸激酶 B 受体 (tyrosine kinase receptor B, TrkB)表达. TrkB 与影响学习和 记忆的脑源性神经营养因子 (brain-derived neurotrophic factor, BDNF)受体有关. 经过训练 后,NMDA 谷氨酸盐受体会促进突触可塑性,提 高突触的效能[38].

本文实验采用具有强规律性(周期性)的复合音 作为听觉刺激.有研究表明,音乐训练能增强人对 声音规律的敏感度,并且相应的 NMDAR 亚单位 和 BDNF/TrkB 表达增强^[39].这与本文结论一致. 训练会增强突触的可塑性,促进突触新生和神经新 生,增强细胞的活化度,从而影响神经营养因子 BDNF 的表达和受体 TrkB,最终诱发的 NMDAR 神经可塑性^[49]. NMDAR 会修饰听觉刺激,调控 BDNF 和 TrKB 的表达,增强听觉区域外的脑区的 学习能力^[41].所以,音乐训练能促进基因表达,提 高突触的可塑性和细胞的活化度,最终提高听觉认 知水平.

本文的研究集中于听觉脑干区域的 FFR/ABR, 听觉训练能刺激该区域的 NMDAR 调谐神经响应 的延展性^[42]. FFR/ABR 能选择性地估计神经系统 对的声音特征编码,包括本文分析听觉神经对声音 的基频和谐波的编码,或者一个单刺激声的不同区 域等均能体现在由 ABR/FFR 诱发的电位中.听觉

脑干和中脑神经核团的 TFS 和高频锁相能准确地 反映在 ABR 中,我们记录到的 FFR 也准确地复现 了锁相能力. 有研究表明, 经过音乐训练后, 人会 选择性地提高对谐波的神经元编码能力. 将大鼠置 于音乐环境中,它们脑干的 NMDAR 蛋白表达比 对照组更稠密^[43]. NMDAR 和 TrkB/BDNF 的可塑 性机制有一定的交互作用,包括激活代谢型谷氨酸 受体,增强邻近突触的可塑性变化,继而促使突触 传导发生持续性变化^[44]. NMDAR 对听觉感知(如 声源获取、时长区分等)提升有明显的作用[49,并 且受外界噪声等环境影响.因此,本文在安静和噪 声环境下的音乐训练引起的 FFR 增强均受 NMDAR 调控.本文发现在背景噪声存在的情况 下,音乐训练受试者的听觉神经编码能力更稳定. 很可能是由于音乐训练加强了神经元同步性, NMDAR 调控能力提高,最终信噪比增强.

4.2 听觉 FFR 的起源神经核团的探讨

本文进一步通过分析可分辨复合音 FFR_{ENV}的 锁相上限频率探讨 FFR 的起源.大部分生物实验 结果普遍认为 FFR 体现的是来自听觉脑干的神经 元群体的锁相活动.但是,要准确找到产生人类头 皮采集的 FFR 神经源是非常困难的.一些冷冻切 除实验和认知比较实验认为,脑干的下丘脑 IC、 外侧丘系 LL 以及耳蜗核 CN 很可能是 FFR 的神经 元^[46-47].特别地,下丘脑被认为最有可能是 FFR 的 主要神经源.例如,早期 Smith 等^[48]对猫的下丘实 施了低温手术,发现此时猫的 FFR 幅度下降,而 下丘温度回升后 FFR 的幅度恢复.Gardi 等^[49]切除 了猫的耳蜗核,发现 FFR 幅度有近 50%的大幅下 降.在人体实验中,破坏下丘也发现了 FFR 消失 的现象.

利用文献[18]和[19]的方法,计算了全部 16 位 受试者 FFR_{ENV} 的 PLV,以及各个频率下具有显著 有效响应的受试者人数,如图7所示.在安静环境 下, FFR_{EW}的 PLV 从中频 500 Hz 开始持续下降, 并且从高频 600 Hz 开始 PLV 和有效受试者人数均 迅速下降. 低次谐波 200~400 Hz, FFR_{ENV} 的 PLV 几乎相等,有效受试者人数均等于或接近全部受试 者人数 16. FFR_{ENV} 在 300 Hz、400 Hz 的 PLV 与 200 Hz的PLV百分比分别为112.39%、101.34%. 但是中频 500 Hz 相比 200 Hz 的 PLV 下降了 76.30%. 高次谐波 600 Hz、700 Hz 和 800 Hz 的锁 相能力比 200 Hz 下降了近一半, 百分比分别为 56.62%、54.21%和47.52%. 在噪声环境下, FFR_{ENV}的 PLV 主要集中在基频和低频,高频的神 经元锁相能力受到严重影响.因此,以上实验结果 表明, FFR_{ENV} 对可分辨复合音的包络锁相响应的上 限频率为 600 Hz.



Neural response frequency/Hz

Fig. 7 Grand average long-term PLV of FFR_{ENV} in response to the complex tone in quiet (a) and in noise (b), and the corresponding numbers of subject with significant phasing locking response

沿着上行听觉通路,各个阶段的锁相能力不同,锁相上限频率也会下降.现有的大部分研究主要关注下丘脑的锁相.下丘被认为是脑皮层记录到的神经元锁相脑干电位最有可能的起源部位. Kuwada 等^[50]研究了猫的下丘脑内的神经单元,发现 600 Hz 以上锁相响应很少出现,而 1 200 Hz 以上不存在任何锁相响应下丘内的神经元对纯音的锁 相仅限于几百赫兹,只有极少数的神经元能够锁相 至1kHz.而神经元对复合音的锁相大多数都是局 限于100~200 Hz,有少数神经元对复合音的锁相 频率高达400 Hz^[51].Liu等^[32]发现豚鼠的26%单位 细胞的特征频率小于300 Hz,55%单位细胞的特征 频率小于600 Hz.本文认为,FFR 对包络锁相的 上限频率约为600 Hz,且从500 Hz 开始锁相能力 即迅速下降.这一个结果与现有研究关于下丘脑对 声音包络锁相的上限频率一致.因此,我们认为 FFR 的起源(或大部分神经核团)很可能位于下丘脑.

5 结 论

听觉选择性注意是人在复杂环境下重要的感知 和认知能力.本文通过行为学感知实验和听觉脑干 频率跟随响应认知实验,研究了音乐训练对基于音 高和基于空间位置的两种听觉选择性注意的可塑 性. 感知实验结果表明, 经过音乐训练后, 人的听 觉选择性注意的能力提高,是完成以音高为特征的 听觉选择性任务中,受过音乐训练的受试者的行为 学准确率明显提高,且响应时间加快.因此,音乐 训练对基于音高的听觉选择性感知具有可塑性. 另 一方面,本文提出分析 FFR_{ENV} 和 FFR_{TTS} 的四个新 方法,用于分析认知实验数据,结果表明,受过音 乐训练的受试者在基频处的 FFR_{EW} 神经元锁相能 力明显强于未受过音乐训练的受试者,且表现出更 强的 FFR_{EW} 基频抗噪声能力; 受过音乐训练的受试 者在谐波处的 FFR_{TFS} 的持续锁相能力优于未受过 音乐训练的受试者,且受过音乐训练的受试者在谐 波处的 FFR_{ITS} 幅度谱信噪比与其音高行为学实验 的准确率呈线性正相关.综上所述,音乐训练对基 于音高的听觉选择性注意的行为感知和神经认知能 力均具有可塑性, 音乐训练提高了听觉认知神经元 在基频的 FFR_{ENV} 锁相能力、抗噪声能力,并提高 了神经元在谐波的 FFR_{TFS} 的持续锁相能力、抗噪 声能力和幅度谱信噪比,进而提高了基于音高的听 觉选择性注意的行为学准确率.最后,本文通过分 析 FFR_{EW} 锁相的上限频率,认为听觉 FFR 可能起 源于下丘脑.

虽然人脑训练与学习对应的生物机制己有一定 的研究成果,但是在人日常生活的实际意义却知之 甚少.音乐训练可进一步应用于改善人的学习和记 忆,是认知神经可塑性的很好的实际应用.音乐训 练引起大脑的特异性可能与训练时长和训练开始年 龄相关^[53].与此同时,神经突触的生发、以及树突 棘的膨胀与繁殖等贯穿整个生命过程,受发育和训 练共同控制.突触的可塑性随发育而降低,随着训 练的增强而提高^[54].儿童时期是发育敏感期,该阶 段接受音乐训练后脑皮层的响应会增强,音乐家脑 皮层的特异性源自训练而非先天因素^[55].因此,本 文下一步工作是研究不同训练时长、成年人与儿童 和不同乐器训练对听觉认知的可塑性,并进一步研 究从下丘脑开始的听觉神经上行通路的听觉认知可 塑性影响.

致谢 感谢 Center for Computational Neuroscience and Neural Technology, Boston University 的 Barbara G Shinn-Cunningham 教授, Jing Xia 博士和 Hari Bharadwaj 博士,以及清华大学生物医学工程 系的胡广书教授对本文的实验设计及方法的指导.

参考文献

- Schafer R J, Moore T. Selective attention from voluntary control of neurons in prefrontal cortex. Science, 2011, 332(6037): 1568–1571
- [2] Shinn-Cunningham B G, Best V. Selective attention in normal and impaired hearing. Trends in Amplification, 2008, 12(4): 283–299
- [3] Lehmann A, Schönwiesner M. Selective attention modulates human auditory brainstem responses: relative contributions of frequency and spatial cues. PloS One, 2014, 9(1): e85442
- [4] 宣 宾, 张达人. 时间选择性注意的认知神经机制. 生物化学与 生物物理进展, 2009, 36(6): 663-667
- Xuan B, Zhang D R. Prog Biochem Biophys, 2009, 36(6): 663–667
 [5] Micheyl C, Delhommeau K, Perrot X, *et al.* Influence of musical
- and psychoacoustical training on pitch discrimination. Hearing Research, 2006, **219**(1): 36–47
- [6] Zendel B R, Alain C. Concurrent sound segregation is enhanced in musicians. Journal of Cognitive Neuroscience, 2009, 21(8): 1488– 1498
- [7] Parbery-Clark A, Skoe E, Kraus N. Musical experience limits the degradative effects of background noise on the neural processing of sound. The Journal of Neuroscience, 2009, 29(45): 14100–14107
- [8] Wong P C, Skoe E, Russo N M, et al. Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nature Neuroscience, 2007, 10(4): 420–422
- [9] Carcagno S, Plack C J. Subcortical plasticity following perceptual learning in a pitch discrimination task. Journal of the Association for Research in Otolaryngology, 2011, 12(1): 89–100
- [10] Greenberg S, Marsh J T, Brown W S, *et al.* Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones. Hearing Research, 1987, 25(2): 91–114
- [11] Meddis R, O'Mard L. A unitary model of pitch perception. The Journal of the Acoustical Society of America, 1997, 102(3): 1811– 1820
- [12] Jeffress L A . A place theory of sound localization. Journal of Comparative and Physiological Psychology, 1948, 41(1): 35–39
- [13] Du Y, Kong L, Wang Q, et al. Auditory frequency-following response: A neurophysiological measure for studying the "cocktailparty problem". Neuroscience and Biobehavioral Reviews, 2011, 35(10): 2046–2057
- [14] Marsh J T, Worden F G, Smith J C. Auditory frequency-following response: neural or artifact? Science, 1970, 169(3951): 1222–1223
- [15] Krishnan A. Human frequency-following responses: representation of steady-state synthetic vowels. Hearing Research, 2002, 166(1): 192–201
- [16] Chandrasekaran B, Kraus N. The scalp-recorded brainstem response to speech: Neural origins and plasticity. Psychophysiology, 2009,

47(2): 236–246

- [17] Lomber S G, Malhotra S. Double dissociation of "what" and "where" processing in auditory cortex. Nature Neuroscience, 2008, 11(5): 609–616
- [18] Zhu L, Bharadwaj H, Xia J, et al. A comparison of spectral magnitude and phase-locking value analyses of the frequency-following response to complex tones. The Journal of the Acoustical Society of America, 2013, 134(1): 384–395
- [19] 朱 莉. 噪声环境中的语音增强和听觉认知的研究[D]. 北京: 清 华大学医学院生物医学工程系, 2013

Zhu Li. Research on the speech enhancement and auditory Cognition in noisy environment [D]. Beijing: Department of Biomedical Engineering, Medical School, Tsinghua University, 2013

- [20] Kishon-Rabin L, Amir O, Vexler Y, et al. Pitch discrimination: are professional musicians better than non-musicians? Journal of Basic and Clinical Physiology and Pharmacology, 2001, 12(2): 125–143
- [21] Song J H, Skoe E, Banai K, et al. Perception of speech in noise: neural correlates. Journal of Cognitive Neuroscience, 2011, 23(9): 2268–2279
- [22] Munte T F, Kohlmetz C, Nager W, et al. Neuroperception: Superior auditory spatial tuning in conductors. Nature, 2001, 409(6820): 580
- [23] Parbery-Clark A, Skoe E, Kraus N. Musical experience limits the degradative effects of background noise on the neural processing of sound. The Journal of Neuroscience, 2009, 29(45): 14100–14107
- [24] Anderson S, Skoe E, Chandrasekaran B, et al. Brainstem correlates of speech-in-noise perception in children. Hearing Research, 2010, 270(1): 151–157
- [25] Chobert J, Marie C, Francois C, et al. Enhanced passive and active processing of syllables in musician children. Journal of Cognitive Neuroscience, 2011, 23(12): 3874–3887
- [26] Bermudez P, Lerch J P, Evans A C, et al. Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. Cerebral Cortex, 2009, **19** (7): 1583– 1596
- [27] Groussard M, La Joie R, Rauchs G, et al. When music and long-term memory interact: effects of musical expertise on functional and structural plasticity in the hippocampus. Plos One, 2010, 5(10)
- [28] Alladi P A, Roy T, Singh N, et al. Prenatal auditory enrichment with species-specific calls and sitar music modulates expression of Bcl-2 and Bax to alter programmed cell death in developing chick auditory nuclei. International Journal of Developmental Neuroscience, 2005, 23(4): 363–373
- [29] Halwani G F, Loui P, Ruber T, et al. Effects of practice and experience on the arcuate fasciculus: comparing singers, instrumentalists, and non-musicians. Frontiers in Psychology, 2011, 2(6): 156
- [30] Strait D L, Parbery-Clark A, O'Connell S, et al. Biological impact of preschool music classes on processing speech in noise. Developmental Cognitive Neuroscience, 2013, 6(4): 51–60
- [31] Bavelier D, Levi D M, Li R W, et al. Removing brakes on adult brain plasticity: from molecular to behavioral interventions. Journal of Neuroscience, 2010, 30(1): 14964–14971
- [32] Takesian A E, Kotak V C, Sanes D H. Presynaptic GABA (B) receptors regulate experience-dependent development of inhibitory short-term plasticity. Journal of Neuroscience, 2010, 30(45): 2716-

2727

- [33] Francois C, Schon D. Neural sensitivity to statistical regularities as a fundamental biological process that underlies auditory learning: the role of musical practice. Hearing Research, 2014, 308 (2): 122-128
- [34] Kraus N, Strait D L, Parbery-Clark A. Cognitive factors shape brain networks for auditory skills: spotlight on auditory working memory. Annals of the New York Academy of Sciences, 2012, 1252 (2): 100–107
- [35] Strait D L, O'Connell S, Parbery-Clark A, et al. Musicians' enhanced neural differentiation of speech sounds arises early in life: developmental evidence from ages three to thirty. Cerebral Cortex, 2014, 24(9): 2512–2521
- [36] Bidelman G M, Krishnan A. Effects of reverberation on brainstem representation of speech in musicians and non-musicians. Brain Research, 2010(1355): 112–125
- [37] Abraham W C, Bear M F. Metaplasticity: the plasticity of synaptic plasticity. Trends in Neurosciences, 1996, 19(4): 126–130
- [38] Lau C G, Zukin R S. NMDA receptor trafficking in synaptic plasticity and neuropsychiatric disorders. Nature Reviews Neuroscience, 2007, 8(6): 413-426
- [39] Herholz S C, Boh B, Pantev C. Musical training modulates encoding of higher-order regularities in the auditory cortex. European Journal of Neuroscience, 2011, 34(3): 524–529
- [40] Cunha C, Brambilla R, Thomas K L. A simple role for BDNF in learning and memory? Frontiers in Molecular Neuroscience, 2010, 3: 1-14
- [41] Chikahisa S, Sei H, Morishima M, et al. Exposure to music in the perinatal period enhances learning performance and alters BDNF/TrkB signaling in mice as adults. Behavioural Brain Research, 2006, 169(2): 312–319
- [42] Feldman D E, Brainard M S, Knudsen E I. Newly learned auditory responses mediated by NMDA receptors in the owl inferior colliculus. Science, 1996, 271(5248): 525–528
- [43] Xu J, Yu L, Cai R, et al. Early auditory enrichment with music enhances auditory discrimination learning and alters NR2B protein expression in rat auditory cortex. Behavioural Brain Research, 2009, 196(1): 49–54
- [44] Abraham W C. Metaplasticity: tuning synapses and networks for plasticity. Nature Reviews Neuroscience, 2008, 9(5): 387–399
- [45] Tang Y P, Shimizu E, Dube G R, et al. Genetic enhancement of learning and memory in mice. Nature, 1999, 401(6748): 63–69
- [46] Chandrasekaran B, Kraus N. The scalp-recorded brainstem response to speech: Neural origins and plasticity. Psychophysiology, 2009, 47(2): 236–246
- [47] Kraus N, Nicol T. Brainstem origins for cortical "what" and "where" pathways in the auditory system. Trends in Neuroscience, 2005, 28(4): 176–181
- [48] Smith J C, Marsh J T, Brown W S. Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. Electroencephalography and Clinical Neurophysiology, 1975, 39(5): 465–472
- [49] Gardi J, Merzenich M, McKean C. Origins of the scalp recorded frequency following response in the cat. Audiology, 1979, 18 (5): 358-381
- [50] Kuwada S, Yin T C, Syka J, *et al.* Binaural interaction in low-frequency neurons in inferior colliculus of the cat. \mathbb{IV} .

Comparison of monaural and binaural response properties. Journal of Neurophysiology, 1984, **51**(6): 1306–1325

- [51] McKinney M F, Tramo M J, Delgutte B. Neural correlates of the dissonance of musical intervals in the inferior colliculus. In Physiological and Psychophysical Bases of Auditory Function, eds Breebaart D J, Houtsma A J M, Kohlrausch A, Prijs V F, Schoonhoven R. 2001, (Maastricht, Shaker), pp 83–89
- [52] Liu L F, Palmer A R, Wallace M N. Phase-locked responses to pure tones in the inferior colliculus. Journal of Neurophysiology, 2006,

95(3): 1926-1935

- [53] Gaser C, Schlaug G. Brain structures differ between musicians and non-musicians. Journal of Neuroscience, 2003, 23(27): 9240–9245
- [54] Bednarek E, Caroni P. Beta-Adducin is required for stable assembly of new synapses and improved memory upon environmental enrichment. Neuron, 2011, 69(6): 1132–1146
- [55] Steele C J, Bailey J A, Zatorre R J, et al. Early musical training and white-matter plasticity in the corpus callosum: evidence for a sensitive period. Journal of Neuroscience, 2013, 33(3): 1282–1290

Research on Neural Mechanisms of Plasticity in Auditory Selective Attention^{*}

ZHU Li¹), ZHENG Fen¹), DENG Juan²), YAN Zheng³), XU Xiao-Ling^{1)**}

(¹⁾ College of Information Engineering, Nanchang University, Nanchang 330031, China;
 ²⁾ Institute of Biomedical Engineering, The Chinese Academy of Sciences & Peking Union Medical College, Tianjin 300192, China;
 ³⁾ College of Information Science and Engineering, Huaqiao University, Xiamen 361021, China)

Abstract This study explores whether musical training is able to enhance the perception of pitch- and locationbased selective auditory attention, then analyzes the corresponding neural mechanism of the auditory plasticity through musical training. In the auditory perception experiment, listeners were instructed to select one of two simultaneous digits based on either the target digit's pitch or location. In the auditory cognitive experiment, the listeners' brainstem frequency-following responses (FFRs)were recorded in response to the complex tones with various frequency resolutions in both quiet and in noisy acoustic environments. Further, four approaches of FFR analysis were proposed, including short-term phasing locking value of envelope-related frequency-following response (FFR_{ENV}), polar of FFR's instantaneous phase difference, vector of FFR's grand averaged phase difference, and spectrum signal-to-noise ratio of temporal-fine-structure-related frequency-following response (FFR_{TS}). Results show that compared to non-musicians, listeners with musical training were both more accurate at reporting the target digit and had a shorter reaction time, especially for attend-pitch trials. For both groups, noise had little effect on the neural encoding of the fundamental frequency (F0) but significantly degraded the neural encoding of harmonics. Compared to non-musicians, musicians showed enhanced phase locking of F0 as well as more robust phase locking to stimulus harmonics in the noise condition. Neural spectrum SNRs of FFR_{ITS} to harmonics were correlated with listeners' behavioral correction rate on the pitch-based auditory attention, for which robustness of musicians' FFR out performed the non-musicians'. These findings suggest musician's better pitch perception in auditory selective attention is relevant to their enhanced neural cognition. Musical training is assumed to be able to enhance the FFR_{ENV} phase locking to the F0, the robustness and continuous phase locking of FFR_{TFS} to the harmonics, and spectrum SNR FFR_{IFS} to the harmonics. In conclusion, musical training has a significant effect on listeners' plasticity in auditory selective attention.

Key words auditory plasticity, frequency-following response, auditory selective attention, neural mechanism, pitch, musical training

DOI: 10.16476/j.pibb.2015.0125

^{*}This work was supported by grants from The National Natural Science Foundation of China (61463035), The Natural Science Foundation of Jiangxi Province (20142BAB217022, 20142BAB207004), and The Natural Science Foundation of Jiangxi Education Commission (GJJ14193). **Corresponding author.

Tel: 86-791-83969680, E-mail: xuxiaoling98@163.com

Received: June 2, 2015 Accepted: June 25, 2015