科学通报 2019 年 第 64 卷 第 7 期: 705 ~ 714
论 文
《中国科学》杂志社
SCIENCE CHINA PRESS

# 视觉和听觉通道情绪知觉加工的神经机制及其异同

张瑶[1,2†], 张帆[1,2†], 毕泰勇[1,2,3*], 邱江[1,2*]

1. 西南大学心理学部, 重庆 400715;
2. 认知与人格教育部重点实验室, 重庆 400715;
3. 遵义医学院管理学院, 遵义 563000

† 同等贡献

* 联系人, E-mail: bitaiyong@126.com; qiuj318@swu.edu.cn

2018-11-12 收稿, 2018-11-27 修回, 2018-11-28 接受, 2019-01-23 网络版发表

国家自然科学基金(31600878, 31771231)、重庆自然科学基金(cstc2015jcyjA10106)、遵义医学院博士科研启动资金(F-875)和遵义医学院学术新苗培养创新探索专项(黔科合平台人才[2017]5733-014)资助

**摘要** 通过功能核磁共振(functional magnetic resonance imaging, fMRI)技术, 采用情绪知觉的判断任务, 探究了视觉与听觉情绪知觉加工的神经机制及其异同, 进一步揭示跨通道情绪知觉加工的多重感觉皮层的脑区定位. 我们选用含有恐惧、高兴或悲伤情绪的面孔和声音刺激作为实验材料, 使用事件相关设计考察每种刺激诱发的激活. 激活强度分析的结果显示: 情绪类型主效应显著的区域主要体现在左侧额下回岛盖部, 表明其作为视-听跨通道情绪知觉的皮层, 是视觉和听觉情绪知觉的共同通路, 且对于负性情绪, 尤其是恐惧情绪刺激最为敏感. 基于多体素模式分析(multi-voxel pattern analysis, MVPA)的结果发现: 右侧颞上沟的激活模式能区分高兴、悲伤、恐惧的面孔, 表明其对不同情绪效价的面孔知觉有重要作用; 右侧梭状回面孔区的激活模式能够区分高兴和悲伤的面孔, 表明其对于积极-消极情绪面孔知觉具有重要意义. 本研究为深入理解人类不同通道情绪知觉的加工过程及跨通道情绪知觉的多感觉皮层提供了重要的证据和支持.

**关键词** 功能核磁共振, 视觉通道, 听觉通道, 情绪知觉, 神经机制, 多重感觉皮层

在生活中, 来自面孔和声音的情绪信息往往是复杂而多变的, 如果无法感知理解这些非语言的情绪信号往往会导致沟通障碍. 在传统模型中, 大脑被认为具有独立的脑区分别处理面孔和声音信息, 多重感觉脑区整合不同感觉通道的信息[1]. 在此观点的影响下, 研究者分别探究了大脑对于面孔和声音信息的处理机制. 双侧梭状回(fusiform face area, FFA)、右侧额下回(inferior frontal gyrus, IFG)和颞上沟(superior temporal sulcus, STS)对于面孔加工有重要作用[2~4]. 大脑对不同情绪面孔的加工存在模式和倾向性的差异. Blair等人[5]的结果表明, 观看悲伤的面部表情时, 悲伤强度的增加与左侧杏仁核和右侧颞叶活动的增强有关. 也有研究显示杏仁核与高兴和恐惧的情绪面孔加工有关[6]. Zhu等人[7]和Corcoran等人[8]的研究证实颞上沟对于情绪面孔的识别有重要的作用. 情绪声音也是情绪处理机制研究的常见的刺激材料, 人类听觉的神经机制已经得到了较为广泛的研究, 除了初级听觉皮层(auditory cortex, AC)外, 脑影像方面的研究表明颞上沟和人类发声的特异性有关[9]. 在此基础上, 情绪声音的研究指出颞叶和额叶皮层对听觉情绪知觉有重要作用. 例如, 2000年的一项研究表明, 听到情绪声音刺激时, 和情绪声音有关的脑区位于颞上回后侧至初级听觉皮层[10]. 另有研究指出, 当被试聆听情绪声音时会激活外侧颞叶[11]. Ethofer

**引用格式**: 张瑶, 张帆, 毕泰勇, 等. 视觉和听觉通道情绪知觉加工的神经机制及其异同. 科学通报, 2019, 64: 705–714
Zhang Y, Zhang F, Bi T Y, et al. Similar and distinct neural mechanisms of visual and auditory emotion perception (in Chinese). Chin Sci Bull, 2019, 64: 705–714, doi: 10.1360/N972018-00721

© 2019《中国科学》杂志社

等人[11]和Imaizumi等人[12]发现与中性声音刺激相比, 呈现愤怒的声音刺激时, 颞上沟的激活强度增加. 还有一些研究表明在情绪声音的刺激下, 右侧前额叶皮层被显著激活[13].

随着研究的深入, 许多研究者认为不同感觉通道信息的神经系统之间并不是相互分离的, 而是存在相互整合的现象. 影像学研究表明, 颞上沟区域具有作为多感觉处理脑区的可能性. 例如, 猴子颞上回多感觉区部分神经元对视觉和听觉刺激都可以响应[14,15]. 较少的研究直接探讨前额叶皮层的多感觉加工, 在视觉和听觉任务中, 猴子的前额叶神经元对听觉和视觉刺激都有响应[16]. 但是, 解剖学研究指出多感觉整合的脑区也可能存在于初级皮层[17,18]. 例如, 研究表明, 单一感觉刺激处理的感觉皮层亦可以接受其他感觉刺激的输入并直接参与多感觉信息的整合[18]. 关于情绪知觉的视听整合是出现在单一感觉脑区还是多重感觉皮层, 没有十分明确的结论. 一些研究表明, 面孔和声音情绪的整合多发生在多感觉区域, 如眶额叶皮层、后颞上沟[19,20]. 但也有研究结果指出颞上回在情绪刺激的多模态整合中具有重要的作用, 这一发现支持传统上被认为是听觉皮层的颞上回区域也可能是多感觉处理的脑区的假设[21]. 而且2009年的一项研究表明, 纹状体视觉皮层不仅仅处理视觉信息, 还是多重感觉的处理的区域[22]. 这些研究证明, 跨通道的感觉信息不仅在大脑多感觉区域中整合, 还可能在更早的初级感觉脑区进行整合.

综上所述, 前人对于视-听通道情绪知觉的神经机制异同点并没有得到清晰的揭示, 且对情绪知觉的多感觉皮层的脑区定位还存在争议. 本研究以高兴、悲伤和恐惧的面孔和声音作为刺激材料, 采集被试在情绪知觉判断任务中影像数据. 使用多体素模式分析方法探究哪些脑区的激活模式可以显著地区分视觉和听觉通道的情绪信息; 通过激活强度分析, 找出与情绪类型、通道类型、情绪与通道类型交互作用显著的脑区, 进而确定情绪知觉的多重感觉皮层.

## 1 材料与方法

### 1.1 被试

19名来自西南大学的大学生参与实验(男生7名, 女生12名, 平均年龄22.3岁). 所有被试均为右利手, 视力或矫正视力正常, 签署知情同意书, 筛除具有精神疾病史、手术史的被试. 实验结束后给予被试一定的报酬.

### 1.2 实验材料与程序

本研究包含3个实验任务. 实验一是对低级视觉皮层(V1~V4)的定位使用经典的视网膜拓扑皮层定位材料和范式[23]; 实验二是对面孔区(FFA,STG)定位, 实验中屏幕上随机呈现0.3 s刺激材料(面孔图片/模糊图片), 图片间隔0.2 s, 24张图片作为一组连续呈现, 前后呈现的两张图片一致时, 被试进行按键反应. 该实验为组块设计, 每个组块包都含24个试次; 实验三为主实验, 刺激材料为情绪面孔与情绪声音. 面孔图片来自中国情绪面孔图库共60张, 高兴、悲伤、恐惧面孔各20张, 男女面孔各10张; 情绪声音来自互联网共60段, 其中高兴、悲伤、恐惧声音各20段, 男女声音各10段. 实验的流程如图1, 当注视点由红色变为绿色时提示被试实验刺激即将呈现, 绿色注视点持续1 s. 随后随机呈现听觉或视觉刺激, 如果呈现的是听觉刺激, 被试会听到时长为0.3 s的情绪声音, 被试通过按键判断情绪声音源于男性还是女性. 如果呈现的是视觉刺激, 被试看到一张情绪面孔图片, 持续1 s, 被试按键判断面孔是男性还是女性. 该实验是2×3(刺激通道类型: 视觉、听觉; 情绪类型: 高兴、悲伤、恐惧)的事件相关设计. 共有6种实验条件, 每种条件40个试次, 共有240个试次. 实验前招募了15名被试(未参与正式实验)对这些材料的效价和唤醒度进行了7点评分. 3种面孔材料的情绪效价和唤醒度分别为: 高兴面孔(效价: $M$=5.00, $SD$=0.74; 唤醒度: $M$=4.26, $SD$=0.99)、悲伤面孔(效价: $M$=2.75, $SD$=0.51; 唤醒度: $M$=3.81, $SD$=1.05)、恐惧面孔(效价: $M$=2.92, $SD$=0.59; 唤醒度: $M$=4.27, $SD$=1.06). 三者效价之间差异显著($F$(2,28)=73.52, $P$<0.001), 唤醒度之间差异不显著($F$(2,28)=2.43, $P$=0.11). 3种声音材料的情绪效价和唤醒度分别为: 高兴声音(效价: $M$=5.51, $SD$=0.62; 唤醒度: $M$=5.06, $SD$=1.07), 悲伤声音(效价: $M$=2.49, $SD$=0.49; 唤醒度: $M$=4.72, $SD$=1.05), 恐惧面孔(效价: $M$=1.55, $SD$= 0.44; 唤醒度: $M$=6.22, $SD$=0.84). 三者效价之间差异显著($F$(2,28)=198.41, $P$<0.001), 唤醒度之间差异显著($F$(2,28)=34.92, $P$<0.001).

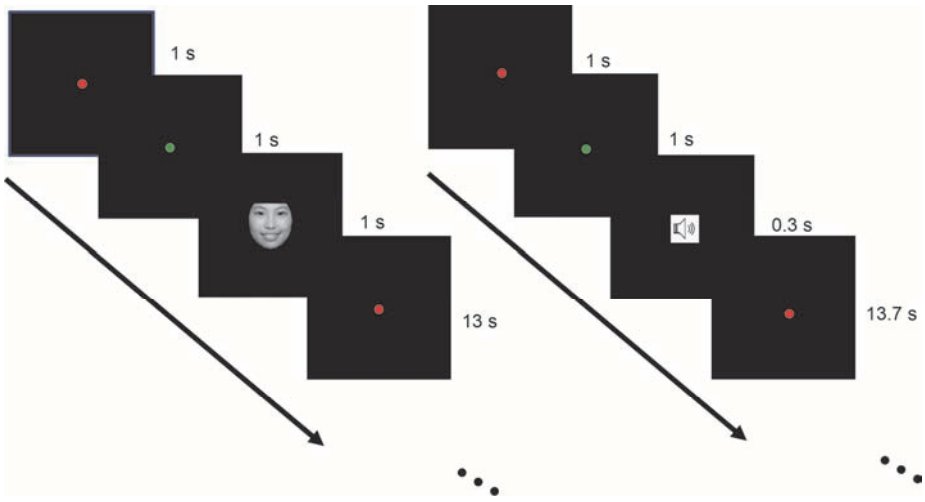



**图 1** (网络版彩色)实验流程图. 被试需要判断呈现的情绪面孔或情绪声音来自男性还是女性, 并进行按键反应

**Figure 1** (Color online) Experimental procedure. When the stimulus (an emotional face or voice) was presented, participants were required to make a gender judgement

## 2 核磁共振扫描参数及预处理

采集被试结构像数据, 以便于数据分析时将结构像和功能像进行对齐. 影像数据的采集使用西门子3.0 T核磁共振扫描仪(Siemens Medical, Erlangen, 德国). 功能像的数据采集采用的是12通道线圈, 序列采用T2加权梯度回波平面成像序列技术. 扫描参数设置为: TR/TE=2000/30 ms, 层厚=3 mm, 翻转角=90°, FOV=192 mm×192 mm, 矩阵大小=64×64. 低级视觉皮层定位任务中扫描156层覆盖全脑; 面孔兴趣区定位任务中扫描180层覆盖全脑; 主实验中扫描132层覆盖全脑. 在感兴趣区分析中, 功能像的数据使用BrainVoyager QX分析软件包进行预处理和兴趣区的选定. 预处理的步骤为: 结构像与功能像的对齐、头动矫正、去线性漂移以及高通滤波(0.015 Hz)等过程. 在全脑分析中, 采用SPM8软件对功能像数据进行预处理和统计分析. 4名被试由于头动过大(头动偏移大于3 mm), 在后续的数据分析中剔除. 由于本研究感兴趣区的总数为9个(V1~V4, 双侧FFA, STS, AC), 因此统计分析中Bonferroni矫正后的显著性水平是0.05/9=0.0056.

## 3 数据分析结果

### 3.1 感兴趣区及其激活强度分析结果

将预处理之后的功能像数据拟合成一般线性模型(general linear model, GLM), 在定位低级视觉皮层时, 首先利用定位扫描的功能像数据将每个视觉区域的边界勾画出来. 然后, 利用T统计量建立每个体素对定位任务反应的强度. 将每个边界内激活最显著($P$<0.001)的区域定义为V1~V4感兴趣区. 面孔区则被定义为对面孔图片的反应比模糊图片更强烈的区域($P$<0.001), Talairach坐标见表1. 本研究中听觉兴趣区为听觉皮层(AC), 因为听觉皮层的解剖结构相对复杂且体积较小[24], 因此在本研究中未进一步区分. 根据神经解剖学理论本研究将听觉皮层定位为对声音刺激的反应比视觉刺激的反应更强的区域(包括布鲁德曼41和42区).

激活强度分析是针对上述兴趣区, 将其拟合成为一般线性模型. 每个区域的激活强度值为GLM估计的$\beta$值. 利用重复测量方差分析(repeated measures ANOVA)方法计算兴趣区在6种实验条件下的通道类型、情绪类型主效应和通道与情绪类型的交互效应. 结果显示: 初级视觉皮层和面孔选择区对情绪面孔的激活强度显著高于对情绪声音的激活强度(V1: $F$(1,14)=142.15, $P$<0.001; V2: $F$(1,14)=117.02, $P$<0.001; V3: $F$(1,14)=127.64, $P$<0.001; V4: $F$(1,14)=102.41, $P$<0.001; lFFA: $F$(1,14)=91.92, $P$<0.001; rFFA: $F$(1,14)=96.63, $P$<0.001; lSTS: $F$(1,14)=33.75, $P$<0.001; rSTS: $F$(1,14)=9.76, $P$<0.05). 听觉皮层对情绪声音的激活强度显著高于对情绪面孔的激活强度($F$(1,14)=186.59, $P$<0.001). 情绪类型的主效应, 交互效应均不显著.

**表 1 面孔兴趣区和全脑分析结果在Talairach坐标体系中的位置**

**Table 1** The coordinates of the region of interest and whole brain analysis result in the Talairach coordinate system

| | | 脑区 | *X* | *Y* | *Z* |
|---|---|---|---|---|---|
| 面孔兴趣区 | | 左侧梭状回 | −41.93 | −46.90 | −16.13 |
| | | 右侧梭状回 | 39.01 | −48.76 | −16.45 |
| | | 左侧颞上沟 | −50.48 | −59.17 | 7.56 |
| | | 右侧颞上沟 | 47.01 | −51.80 | 7.94 |
| 全脑分析结果 | 情绪通道主效应 | 左侧额下回岛盖部 | −54 | 18 | 33 |
| | 情绪与感觉通道交互效应 | 左侧楔前叶 | 6 | −63 | 30 |
| | | 左侧背外侧额上回 | −21 | 54 | 36 |
| | | 左侧顶上小叶 | 21 | −45 | 75 |
| | | 左侧额中回 | −42 | −78 | 30 |
| | | 右内侧额上回 | 12 | 62 | 18 |

### 3.2 多体素模式分析结果

使用多体素模式分析方法(multi-voxel pattern analysis, MVPA)来探究不同情绪类型的面孔和声音刺激是否可以改变兴趣区内所有体素在任务过程中激活的空间分布模式. 多体素模式分析包括以下几个步骤: 特征选择; 创建样本; 训练分类器; 泛化测试. 使用libsvm工具包[25]对兴趣区在6种不同条件下的情绪知觉过程进行激活模式分析, 即每个兴趣区对从两种通道知觉到的3种情绪的区分准确率. 结果显示: 右侧的STS能够区分被试知觉到的是高兴、悲伤还是恐惧的面孔, 区分准确率为40%($P$<0.05, Bonferroni矫正); 右侧的FFA能够区分被试知觉到的是高兴还是悲伤的面孔, 区分准确率为57%($P$<0.05, Bonferroni矫正). AC对视觉-听觉通道知觉到的3种情绪的区分正确率达到38%($P$=0.02, 未矫正), 结果如图2所示.

### 3.3 全脑分析结果

上述兴趣区的分析中, 没有发现直接体现情绪类型主效应的脑区. 所以, 利用全脑分析来探究是否存在与面孔和声音情绪知觉相关的脑区. 首先, 将6种条件对应的功能核磁数据拟合建成一般线性模型. 然后, 通过全脑分析找到与每个条件对应的激活显著的区域. 最后, 分别对6种条件进行全脑激活强度分析. 采重复测量的方差分析, 分别计算情绪类型、感觉通道类型的主效应以及情绪类型与感觉通道类型的交互效应(FDR矫正, $P$<0.05, 体素>200). 结果显示: 情绪类型的主效应体现在左侧额下回岛盖部; 刺激通道类型的主效应主要出现在枕叶、颞叶等经典的视觉和听觉信息加工的脑区, 并与激活强度分析结果相互印证; 交互作用体现在5个脑区, 分别是右侧楔前叶、左侧背外侧额上回、右侧顶上小叶、左侧枕中回、右内侧额上回, 具体结果见表1和图3.

进一步提取出情绪类型主效应所在脑区的信号值, 用配对样本$T$检验计算同一情绪类型下不同感觉通道间激活强度的差异, 同一感觉通道不同情绪类型间激活强度的差异. 结果表明: 知觉恐惧面孔的激活强度显著高于知觉恐惧声音的激活强度($T$=2.552, $P$<0.05); 知觉恐惧面孔的激活强度显著高于知觉悲伤面孔的激活强度($T$=2.765, $P$<0.05); 知觉恐惧面孔的激活强度显著高于知觉高兴面孔的激活强度($T$=4.075, $P$<0.01); 知觉恐惧声音的激活强度显著高于知觉高兴声音的激活强度($T$=2.806, $P$<0.05).

进一步进行MVPA的全脑搜索分析, 为了找出所有有可能存在对情绪信息区分性的脑区, 因此采用比较宽泛的显著性水平(AlphaSim, $P$<0.05, 体素个数>30). 结果表明左侧枕下回、枕中回、额中回、脑岛和右侧楔前叶、额中回、尾状核、额下回的激活模式能够能够区分快乐、悲伤和恐惧的面孔刺激; 左侧颞上回、丘脑和右侧的颞上回、中央后回等脑区的激活模式可以区分快乐、悲伤和恐惧的情绪声音(表2).

## 4 讨论

本研究使用功能性磁共振成像技术来探究视-听通道情绪知觉的神经机制及其异同, 并进一步揭示跨通道情绪知觉的多重感觉皮层的脑区定位. 多体

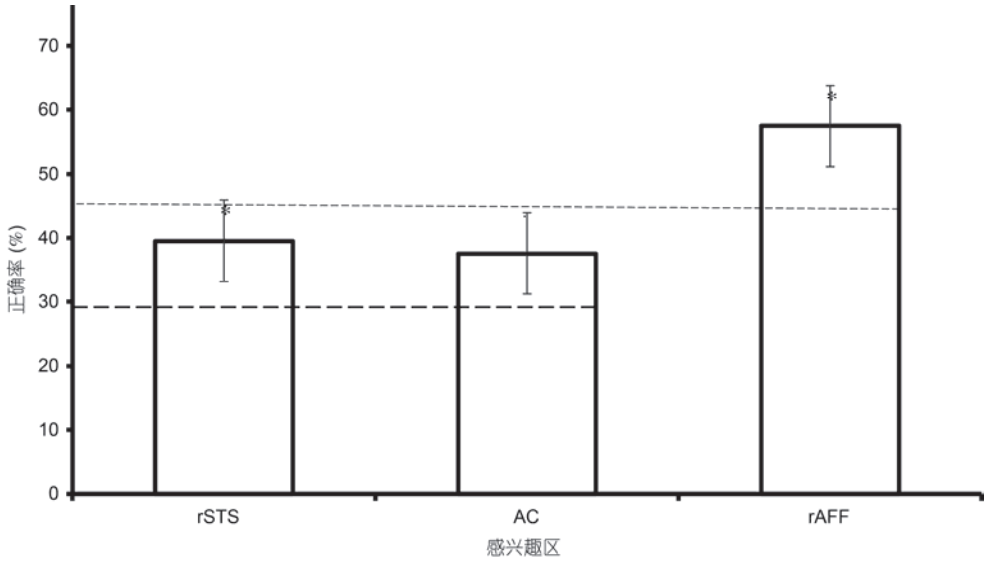



**图 2** 多体素模式分析结果. 右侧STS对被试知觉到的高兴、悲伤还是恐惧面孔的区分正确率与AC对于两种感觉通道的高兴、悲伤和恐惧情绪的区分正确率显著高于随机水平 33%; 右侧FFA对高兴、悲伤的情绪面孔区分正确率显著高于随机水平 50%(*: $P$<0.05, Bonferroni矫正; ·: $P$=0.02, 未矫正)

**Figure 2** Multivariate pattern analysis results. The activity pattern of rSTS was able to characterize the perception of human faces with emotional valence (happy, sad and fear); the activity pattern of AC could discriminate perception of face and voice with emotional valence (happy, sad, fear), and their decoding accuracy was significantly higher than 33%. The activity pattern of rFFA can characterize happy and sad faces and its decoding accuracy was significantly higher than 50% (*: $P$<0.05, Bonferroni multiple comparison; ·: $P$=0.02, uncorrected)

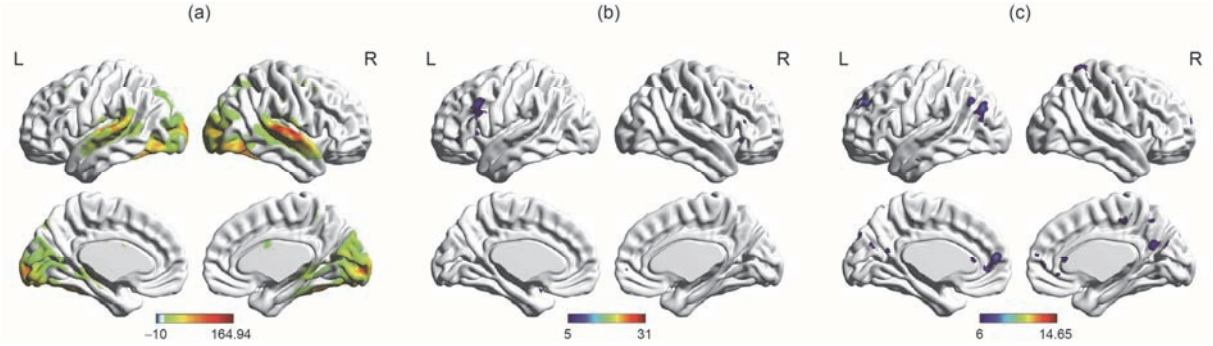



**图 3** (网络版彩色)全脑分析结果. (a) 刺激通道类型主效应主要出现在枕叶、颞叶、额叶等经典的视觉和听觉处理脑区; (b) 情绪类型的主效应出现在左侧的额下回岛盖部; (c) 交互作用体现在五处脑区, 分别是右侧楔前叶、右侧顶上小叶、右内侧额上回、左侧背外侧额上回、左侧枕中回(FDR矫正, $P$ =0.05, 体素个数>200)

**Figure 3** (Color online) Whole brain analysis results. (a) Main effect of modality: Classical visual and auditory cortex located in the occipital lobe, temporal lobe and frontal lobe; (b) main effect of emotional valence: Left opercular part of inferior frontal gyrus; (c) interactive effect of modality and emotional valence: Right precuneus, superior parietal lobule, medial superior frontal gyrus, left middle occipital gyrus, dorsolateral superior frontal gyrus (FDR correction, $P$<0.05, voxel>200)

素模式分析发现, 右侧STS的激活模式能够区分高兴、悲伤还是恐惧的情绪面孔, 右侧的FFA的激活模式来区分被试知觉到的是悲伤还是高兴的面孔, 表明右侧的STS和FFA在情绪面孔知觉中的重要作用; AC的激活模式对视觉-听觉通道的高兴、悲伤、恐惧情绪表现出可区分的倾向, 虽然未过矫正, 但其有可能是视-听通道情绪知觉的共同通路. 全脑分析发现情绪类型主效应体现在左侧额下回岛盖部, 表明其作为视-听情绪知觉的多重感觉皮层, 是视觉与听觉情绪知觉的共同通路, 进一步的分析表明其对负性的, 尤其是对恐惧情绪激活增强, 这些结果均指出其在情绪知觉中的具有重要作用.

多体素模式分析结果表明, 右侧FFA和STS的激活模式表现出这两个兴趣区可能参与面孔情绪知觉,

**表 2 MVPA的全脑分析的结果在Talairach坐标体系中的位置**[a)]

**Table 2** The coordinates of the Searchlight analysis results in the Talairach coordinate system

| MVPA全脑搜索分析 | 脑区 | *X* | *Y* | *Z* | 体素个数 |
|---|---|---|---|---|---|
| 面孔解码结果 | 左侧枕下回 | −22.42 | −87 | −8.1 | 83 |
| | 左侧枕中回 | −26.55 | −81.35 | 16.18 | 53 |
| | 左侧额中回 | −30.71 | 5.45 | 45.98 | 47 |
| | 左侧脑岛 | −32.04 | 23.64 | 6.09 | 62 |
| | 右侧楔前叶 | 29.54 | −64.61 | 36.28 | 55 |
| | 右侧额中回 | 45.81 | 15.91 | 37.84 | 52 |
| | 右侧尾状核 | 17.15 | 27.49 | −0.96 | 43 |
| | 右侧额下回 | 39.57 | 46.94 | 4.19 | 47 |
| 声音解码结果 | 左侧颞上回 | −52.54 | −27.1 | 12.63 | 179 |
| | 左侧丘脑 | −7.65 | −13.6 | 16.44 | 83 |
| | 右侧颞上回 | 50.02 | −17.4 | 5.83 | 203 |
| | 右侧中央后回 | 52.52 | −17.74 | 42.13 | 42 |

a) 左侧枕下回、枕中回、双侧额中回等脑区能够区分被试知觉的是高兴、悲伤还是恐惧的面孔; 双侧颞上回、左侧丘脑、右侧的中央后回的激活模式可以区分高兴、悲伤、恐惧的声音(AlphaSim, $P<0.05$, 体素个数>30)

有研究显示FFA, STS不仅会表现出对面孔刺激的偏好, 而且会对视觉通道的情绪信息进行处理[26,27]. STS作为经典的多感觉脑区, 参与视觉和听觉刺激的加工[28~30], 进一步的研究表明STS在情绪面孔加工中被显著激活[31]. Alaerts等人[32]发现STS的功能连接较弱可以预测孤独症患者情绪识别能力. 关于面孔知觉的研究还发现, STS通过整合面部动作的视觉空间信息进而对面部情绪知觉过程起重要作用[33]. 结合前人的研究结果, 我们认为STS可能是视觉情绪加工的重要脑区. 右侧的FFA的激活模式能够区分被试知觉到的是高兴还是悲伤的面孔, 2012年有研究表明右侧FFA在情绪信息的处理中起到了重要作用[34]. Geday等人[35]发现FFA和枕下回之间的神经网络在情绪线索识别过程中具有特殊的作用. AC在视觉-听觉通道知觉到的高兴, 悲伤和恐惧情绪的激活模式表明其作为跨通道情绪知觉的共同通路的倾向性. 2015年一项对于人工耳蜗植入者的研究表明, AC的激活强度与面孔识别能力呈显著正相关[36]. 关于恒河猴的研究指出, 在面孔和声音信息的整合过程中AC与STS起到了调节作用[37]. Koelsch等人[38]及Pantev和Herholz[39]关于情绪音乐的研究指出, AC在情绪声音加工过程起到关键作用. 我们的研究和前人结果的相同之处是均发现FFA和STS在面孔情绪信息加工中的作用, 而AC可能是视听跨通道情绪知觉的多重感觉皮层. 不同之处在于, 前人的研究指出STS在视听信息整合中具有重要作用, 我们的结果并未发现这一现象. 可能的原因在于, 我们的研究并未要求被试对视觉和听觉信息进行整合, 因此STS只体现出在视觉通道中的作用, 而缺乏其受到听觉信息调控的证据.

全脑分析的结果表明, 左侧额下回岛盖部是视觉和听觉通道情绪知觉的多重感觉皮层. 有关于情绪面孔的研究表明观看厌恶情绪的面孔图片时, 激活了额叶的岛盖部[40]. 2010年的一项研究指出在不同表情加工整合过程中前额叶被显著激活, 而不同的面部表情加工中又存在不同通路, 识别高兴情绪时左侧的额下回岛盖部显著激活[41]. 关于病人的研究发现额下回岛盖部对情绪认知有重要作用[42,43]. 例如, 抑郁症患者左侧的额下回岛盖部, 左侧后扣带回, 双侧颞下回等与情绪控制相关的脑区活动异常[42]. 在情绪面孔识别任务中, 与正常被试相比, 精神分裂症患者岛盖部额下回白质纤维束具有显著损伤[44]. 本研究进一步发现额下回的岛盖部对恐惧情绪, 尤其是恐惧面孔的激活强度最强. 这可能与情绪知觉过程中的负性偏向和视觉信息处理具有优先权有关. 情绪认知是具有负性偏向的, 这是由于人类需要对负性事件做出更快的反应, 从而逃避危险. 这种优先加工体现在大脑对负性情绪事件的敏感性, 使其在加工中优先于中性事件和正性事件[45]. 情绪认知的负性偏向在情绪的知觉阶段同样有所体现, 在知觉阶段和认知评价阶段中出现了恐惧表情加工的偏向, 与高兴和中性表情相比, 恐惧表情诱发更大的波幅[46]. 被试知觉恐惧面孔时岛盖部额下回的激活

强度最强, 这可能是因为与听觉刺激相比视觉信息具有优先处理的优势[47]. 杏仁核在情绪反应中具有重要的作用, 但在情绪类型主效应分析中并未发现其显著激活, 这可能是由于杏仁核仅对突发性的情绪刺激反应敏感, 而对实验中诱发的持续出现的情绪状态不太敏感[48]. 本研究为负性情绪, 尤其是恐惧情绪加工的偏向性提供了证据支持. 全脑分析结果指出通道与情绪类型交互作用体现在楔前叶和额中回等脑区, 进一步的MVPA全脑搜索分析结果表明左侧枕下回、枕中回、脑岛和右侧楔前叶、额中回、尾状核、双侧额下回的激活模式可以区分不同类型情绪面孔刺激. 较多的研究指出枕下回, 枕中回、楔前叶和额中回在情绪面孔知觉中显著激活[49,50]; 2008年的一项研究表明右侧的颞上回在情绪声音韵律加工中起到重要作用[51], 还有实验表明双侧颞上回对高兴, 恐惧和愤怒情绪声音激活强烈[52]. 结合前人研究结果, 我们认为枕下回、枕中回、楔前叶和额中回等脑区视觉情绪加工中有重要作用; 颞上回是听觉情绪加工的重要脑区.

2017年一项脑影像研究表明右侧颞上沟在情绪面孔和情绪声音加工的多感觉整合过程中具有重要作用, 并且与情绪面孔与情绪声音一致时(恐惧面孔-恐惧声音)相比当呈现的信息不一致时(恐惧面孔-中性声音)该脑区显著激活[53]. 本研究从视觉通道和听觉通道信息在皮层中的激活强度的角度出发, 发现左侧额下回岛盖部是视听跨通道情绪知觉的多重感觉皮层. 我们的研究可能缺乏视听信息整合的过程, 因此并不能说明STS在多通道情绪信息整合中的作用. 未来的研究可进一步探索额下回在视听情绪信息的整合中是否也发挥了重要作用.

## 5 结论

本研究以情绪面孔和情绪声音作为实验刺激, 探究了视觉通道和听觉通道的情绪知觉的神经机制的异同. 研究结果表明: 右侧STS的激活模式可以有效地识别高兴、悲伤和恐惧的情绪面孔, 右侧FFA的激活模式可以有效地识别悲伤和高兴的面孔; AC对视-听通道的高兴、悲伤、恐惧情绪有一定的区分作用. 全脑分析的结果表明, 视-听通道具有共同的处理情绪知觉信息的多感觉皮层即左侧的额下回岛盖部, 且该脑区对于负性情绪, 尤其是恐惧情绪的激活强度更强, 也表明其在情绪加工中具有偏向性.

## 参考文献


1 Felleman D J, Van Essen D C. Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex, 1991, 1: 1–47
2 Haxby J V, Horwitz B, Ungerleider L G, et al. The functional organization of human extrastriate cortex: A PET-rCBF study of selective attention to faces and locations. J Neurosci, 1994, 5: 6336–6353
3 Kanwisher N, Mcdermott J H, Chun M M. The fusiform face area: A module in human extrastriate cortex specialized for face perception. J Neurosci, 1997, 17: 4302–4311
4 Golarai G, Ghahremani D G, Eberhardt J L, et al. Representation of parts and canonical face configuration in the amygdala, superior temporal sulcus (STG) and the fusiform “face area” (FFA). J Vis, 2004, 4: 131
5 Blair R J R, Morris J S, Frith C D, et al. Dissociable neural responses to facial expressions of sadness and anger. Brain, 1999, 122 (Pt 5): 883–893
6 Breiter H C, Etcoff N L, Whalen P J, et al. Response and habituation of the human amygdala during visual processing of facial expression. Neuron, 1996, 17: 875–887
7 Zhu Q, Nelissen K, Van d S J, et al. Dissimilar processing of emotional facial expressions in human and monkey temporal cortex. Neuroimage, 2013, 66: 402–411
8 Corcoran C M, Grinband J, Gowatsky J L, et al. Event-related repetitive TMS to right posterior STS (but not OFA) in healthy volunteers (HV) briefly recapitulates face emotion recognition (FER) deficits of schizophrenia. Brain Stimul, 2017, 10: e44–e45
9 Fecteau S, Armony J L, Joanette Y, et al. Is voice processing species-specific in human auditory cortex? An fMRI study. Neuroimage, 2004, 23: 840–848
10 Buchanan T W, Lutz K, Mirzazade S, et al. Recognition of emotional prosody and verbal components of spoken language: An fMRI study. Cognit Brain Res, 2000, 9: 227–238
11 Ethofer T, Anders S, Erb M, et al. Cerebral pathways in processing of affective prosody: A dynamic causal modeling study. Neuroimage, 2006, 30: 580–587
12 Imaizumi S, Mori K, Kiritani S, et al. Vocal identification of speaker and emotion activates different brain regions. Neuroreport, 1997, 8: 2809–2812

13 George M S, Parekh P I, Rosinsky N, et al. Understanding emotional prosody activates right hemisphere regions. JAMA Neurol, 1996, 53: 665–670
14 Beauchamp M S. See me, hear me, touch me: Multisensory integration in lateral occipital-temporal cortex. Curr Opin Neurobiol, 2005, 15: 145–153
15 Barraclough N E, Xiao D, Baker C I, et al. Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. J Cogn Neurosci, 2005, 17: 377–391
16 Watanabe M. Frontal units of the monkey coding the associative significance of visual and auditory stimuli. Exp Brain Res, 1992, 89: 233–247
17 Céline C, Pascal B. Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. Eur J Neurosci, 2006, 22: 2886–2902
18 Shi M X, Zhou J, Zhang H, et al. Progress of studies on multisensory integration in the mammalian sensory cortex (in Chinese). J Biol, 2011, 28: 77–80 [施夏明, 周俊, 张辉, 等. 哺乳动物大脑感觉皮层对多种感觉信息整合的研究进展. 生物学杂志, 2011, 28: 77–80]
19 Watson R, Latinus M, Charest I, et al. People-selectivity, audiovisual integration and heteromodality in the superior temporal sulcus. Cortex, 2014, 50: 125–136
20 Romanski L M. Representation and integration of auditory and visual stimuli in the primate ventral lateral prefrontal cortex. Cereb Cortex, 2007, 17(Suppl 1): i61–i69
21 Robins D L, Hunyadi E, Schultz R T. Superior temporal activation in response to dynamic audio-visual emotional cues. Brain Cogn, 2009, 69: 269–278
22 Zhao Y L. An functional MRI study of sensation and sensorimotor integration deficits in early Parkinson's disease (in Chinese). Doctor Dissertation. Hangzhou: Zhejiang University, 2009 [赵艺蕾. 脑功能成像对早期帕金森病感觉功能损害和感觉-运动整合缺陷的神经机制研究. 博士学位论文. 杭州: 浙江大学, 2009]
23 Engel S A, Glover G H, Wandell B A. Retinotopic organization in human visual cortex and the spatial precision of functional MRI. Cereb Cortex, 1997, 7: 181–192
24 Penhune V B, Zatorre R J, MacDonald J D, et al. Interhemispheric anatomical differences in human primary auditory cortex: Probablistic mapping and volume measurement form magnetic resonance scans. Cereb Cortex, 1996, 6: 661–672
25 Chang C, Lin C. Libsvm: A library for support vector machines. ACM Intel Syst Tec, 2011, 2: 1–27
26 Grossman E D, Blake R. Brain areas active during visual perception of biological motion. Neuron, 2002, 35: 1167–1175
27 Schultz R T, Grelotti D J, Klin A, et al. The role of the fusiform face area in social cognition: Implications for the pathobiology of autism. Philos Trans: Biol Sci, 2003, 358: 415–427
28 Hikosaka K. The polysensory region in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. Biomed Res, 1993, 14: 41–45
29 Bruce C, Desimone R, Gross C G. Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. J Neurophysiol, 1981, 46: 369–384
30 Beauchamp M S, Lee K E, Argall B D, et al. Integration of auditory and visual information about objects in superior temporal sulcus. Neuron, 2004, 41: 809–823
31 Hoffman E A, Haxby J V. Distinct representations of eye gaze and identity in the distributed human neural system for face perception. Nat Neurosci, 2000, 3: 80–84
32 Alaerts K, Woolley D G, Steyaert J, et al. Underconnectivity of the superior temporal sulcus predicts emotion recognition deficits in autism. Soc Cogn Affect Neurosci, 2014, 9: 1589–1600
33 Calder A J, Young A W. Understanding the recognition of facial identity and facial expression. Nat Rev Neurosci, 2005, 6: 641–651
34 Atkinson A P, Vuong Q C, Smithson H E. Modulation of the face- and body-selective visual regions by the motion and emotion of point-light face and body stimuli. Neuroimage, 2012, 59: 1700–1712
35 Geday J, Gjedde A, Boldsen A S, et al. Emotional valence modulates activity in the posterior fusiform gyrus and inferior medial prefrontal cortex in social perception. Neuroimage, 2003, 18: 675–684
36 Stropahl M, Plotz K, Schonfeld R, et al. Cross-modal reorganization in cochlear implant users: Auditory cortex contributes to visual face processing. Neuroimage, 2015, 121: 159–170
37 Ghazanfar A A, Chandrasekaran C, Logothetis N K. Interactions between the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. J Neurosci, 2008, 28: 4457–4469
38 Koelsch S, Skouras S, Lohmann G. The auditory cortex hosts network nodes influential for emotion processing: An fMRI study on music-evoked fear and joy. PLoS One, 2018, 13: e0190057
39 Pantev C, Herholz S C. Plasticity of the human auditory cortex related to musical training. Neurosci Biobehav Rev, 2011, 35: 2140–2154
40 Zhang X J. Two types of disgust materials: Disgust sensitivity related fMRI research (in Chinese). Master Thesis. Chongqing: Southwest

University, 2011 [张晓静. 两类厌恶材料: 厌恶敏感性相关的 fMRI 研究. 硕士学位论文. 重庆: 西南大学, 2011]
41 Gao P. The preliminary study of facial expression recognition in normal adults by fMRI (in Chinese). Master Thesis. Guangzhou: Jinan University, 2010 [高鹏. 正常成人面部表情情绪识别的脑功能磁共振初步研究. 硕士学位论文. 广州: 暨南大学, 2010]
42 Guo M Y, Zhang S X, Xu M Z, et al. Resting-state functional magnetic resonance imaging of first-onset major depressive disorder (in Chinese). J Southern Med Univ, 2014, 34: 1277–1281 [郭牟莹, 张水兴, 许明智, 等. 首发抑郁症患者的静息态功能成像. 南方医科大学学报, 2014, 34: 1277–1281]
43 Huang G Q. Functional MRI study of gender-related differences of emotional memory in Major depressive disorder (in Chinese). Master Thesis. Hefei: Anhui Medical University, 2013 [黄国权. 抑郁症情绪记忆性别差异 fMRI 研究. 硕士学位论文. 合肥: 安徽医科大学, 2013]
44 Yao J J, Lü Y D, Zhao X X, et al. Facial emotion cognition deficits and the damage of white matter fiber tracts in the frontal lobe by diffusion tensor imaging in patients with first-episode schizophrenia (in Chinese). Chin J Psychiatry, 2017, 50: 127–132 [姚晶晶, 吕一丁, 赵晓鑫, 等. 精神分裂症首次发病患者面孔情绪认知与额叶白质纤维功能的相关性. 中华精神科杂志, 2017, 50: 127–132]
45 Luo Y, Huang Y, Li X, et al. Effects of emotion on cognitive processing: Series of event-related potentials study. Adv Psychol Sci, 2006, 14: 505–510
46 Luo W B, Luo Y J. The brain mechanism of facial expression processing in rapid serial visual presentation (in Chinese). The National Academic Congress of Psychology, 2009 [罗文波, 罗跃嘉. 快速序列视觉呈现中面孔表情加工的脑机制. 全国心理学学术大会, 2009]
47 Chen X Y, Wang M X, Gao Y T, et al. Interaction of visual and auditory emotional information during their relationship evaluation (in Chinese). Psychol Sci, 2016, (4): 842–848 [陈晓宇, 王美馨, 高雅婷, 等. 视觉与听觉情绪信息关系判断中的交互作用. 心理科学, 2016, (4): 842–848]
48 Damasio A R, Grabowski T J, Bechara A, et al. Subcortical and cortical brain activity during the feeling of self-generated emotions. Nat Neurosci, 2000, 3: 1049–1056
49 Sato W, Kochiyama T, Yoshikawa S, et al. Enhanced neural activity in response to dynamic facial expressions of emotion: An fMRI study. Cogn Brain Res, 2004, 20: 81–91
50 Fusar-Poli P, Placentino A, Carletti F P, et al. Laterality effect on emotional faces processing: ALE meta-analysis of evidence. Neurosci Lett, 2009, 452: 262–267
51 Hoekert M, Bais L, Kahn R S, et al. Time course of the involvement of the right anterior superior temporal gyrus and the right fronto-parietal operculum in emotional prosody perception. PLoS One, 2008, 3: e2244
52 Ethofer T, Bretscher J, Gschwind M, et al. Emotional voice areas: Anatomic location, functional properties, and structural connections revealed by combined fMRI/DTI. Cereb Cortex, 2012, 22: 191–200
53 Davies-Thompson J, Elli G V, Rezk M, et al. Hierarchical brain network for face and voice integration of emotion expression. Cereb Cortex, 2018, bhy240

Summary for “视觉和听觉通道情绪知觉加工的神经机制及其异同”

# Similar and distinct neural mechanisms of visual and auditory emotion perception

Yao Zhang[1,2†], Fan Zhang[1,2†], Taiyong Bi[1,2,3*] & Jiang Qiu[1,2*]

[1] *Faculty of Psychology, Southwest University, Chongqing 400715, China;*
[2] *Key Laboratory of Cognition and Personality of Ministry of Education, Chongqing 400715, China;*
[3] *School of Management, Zunyi Medical University, Zunyi 563000, China*
† Equally contributed to this work
* Corresponding authors, E-mail: bitaiyong@126.com; qiuj318@swu.edu.cn

Emotion is a rapidly changing psychological and physical phenomenon. In daily life, people need to use the information of various sensory modalities (visual, auditory, etc.) to perceive non-verbal emotional information. Non-verbal emotion from faces and voices are often complex and varied. Previous studies have revealed that there are common and distinct neural networks underlying perception of human faces and voices. However, the neural mechanisms underlying visual and auditory emotional perception have not been well studied. Furthermore, despite researches on audiovisual integration of cross-modal emotional information, the multisensory cortex of the visual and auditory emotional information remains elusive. Therefore, it is necessary to study the similarities and differences between the neural mechanisms of emotion perception in visual and auditory modalities, and to explore the multiple sensory cortex of cross-modal emotion perception.

The present fMRI (functional magnetic resonance imaging) study adopted a 2*3 (stimulus presentation modality: Visual, auditory; emotional valence: Happy, sad, fear) event-related design to investigate the neural mechanisms of emotion perception in visual and auditory modalities. When the stimulus (an emotional face or voice) was visually or aurally presented, participants were required to make a gender judgement. The results showed that the activation intensity of emotional faces in V1−V4, bilateral fusiform gyrus and bilateral superior temporal sulcus (STS) was significantly higher than that of emotional voices. Conversely, the activation intensity of emotional voices in auditory cortex (AC) was significantly higher than that of emotional faces. The results from multivoxel pattern analysis (MVPA) showed that the activation patterns of the right STS could discriminate the perception of human faces with emotional valence (happy, sad and fear face), indicating that the rSTG plays important role in perception of faces with different emotional valence; the activation patterns of the right FFA were different for happy and sad faces, indicating that the rFFA is crucial for positive and negative emotional face perception. A voxel-based whole brain analysis was further performed to examine the cortical areas that modulated perception of emotional valence. The whole brain analysis showed that the main effects for emotional valence was significant in the left opercular part of inferior frontal gyrus, indicating that this region might be a multisensory cortex of visual-auditory emotional perception. In summary, our study provided important evidence for further understanding the processing of emotion perception in different modalities and multisensory cortex of cross-channel emotion perception.

**functional magnetic resonance imaging, visual modality, auditory modality, emotion perception, neural mechanism, multisensory cortex**

doi: 10.1360/N972018-00721