

情绪信息的多通道整合

张亮^{1,2} 孙向红¹ 张侃¹

(1 中国科学院心理研究所, 北京 100101) (2 中国科学院研究生院, 北京 100049)

摘要 自然环境中人类情绪信息的交流是依靠多个感觉通道实现的, 多通道整合是情绪加工的基础。近年来的行为学、电生理学与神经成像的研究表明, 情绪信息的加工具有跨通道自动整合的特点, 它发生在认知加工的早期阶段, 与颞上回、颞中回、海马旁回和丘脑等脑区密切相关。不同情绪的整合既有共同的神经基础, 又有各自独特的加工区域。情绪信息的整合机制还可能与加工类型及注意资源有关。在未来研究中, 实验的标准化、动态化、自然化有助于提高研究的准确性和研究间的可比性, 而对特殊群体的研究, 以及综合研究情绪加工与注意等其他认知过程则有助于我们进一步探索多通道整合的神经机制。

关键词 情绪; 多通道整合; 面孔-声音的整合效应; McGurk 效应; 交互式分析

分类号 B842.6

对情绪准确、快速的觉知是人类进行正常社会交往所必须的技能。而人类和动物的情绪状态可以通过他们的面孔和声音表情来推断(达尔文, 1872)。事实上, 人类对情绪的判断既非单独依靠视觉(如面部表情), 也非完全依靠听觉(如语调语气), 而往往是两种感觉通道同时加工输入的结果(Spence & Driver, 2004)。

那么视觉信息和听觉信息是独立被加工的吗? 大量研究表明, 大脑会对来自不同感觉通道的信息进行多通道整合(Multisensory Integration, MSI)(Lugo, Doti, Wittich, & Faubert, 2008; McDonald, Teder-Salejarvi, & Ward, 2001), 情绪信息也不例外。新近的跨通道研究发现, 这种面孔-声音的整合效应(face-voice integration effects)广泛存在于成人(Collignon et al., 2008; de Gelder & Vroomen, 2000), 婴儿(Grossman, Striano, & Friederici, 2006), 病人(de Gelder, Vroomen, de Jong, Masthoff, Trompenaars, & Hodiament, 2005; Magnee, de Gelder, van Engeland, & Kemner, 2008)和动物(Ghazanfar & Logothetis, 2003)的情绪加工过程中。

本文综述了行为学、电生理学以及神经成像

研究的最新成果, 探讨了多通道情绪信息整合的神经基础和机制, 并对未来研究的发展方向提出了几点建议。

1 通道与多通道情绪信息加工

在自然环境中, 情绪信息的交流是通过多通道的形式实现的, 例如, 语气语调, 面部表情等等。大量研究表明, 这种多通道信息的加工并非是若干个单通道信息的简单叠加。生物体通过多个感觉通道接受信息, 而各个感觉系统的输入会依据一定的规则整合起来并进入知觉领域(McDonald et al., 2001; de Gelder et al., 2005)。多通道情绪信息在加工时间、神经脑区、以及发展阶段等多个方面都表现出其独特的特点。

大脑对单通道和多通道情绪信息的加工能力发展于大脑成熟的不同阶段(Barbara, Lucy, William, & Patricia, 2008; Flom & Bahrick, 2007)。婴儿研究表明, 对单通道和双通道的情绪认知有一个发展变化的过程: 对于搭配呈现的面孔和声音, 4个月大的婴儿就可以辨别情绪类型, 而对于单通道的情绪信息, 5个月大的婴儿才能判断语调的情绪, 7个月时才发展起对面孔情绪的识别(Flom et al., 2007)。

针对某些功能损伤病人的研究发现, 某些病症的患者在单通道情绪信息加工方面并无障碍, 但在多通道信息加工时却表现得异于常人。

收稿日期: 2009-04-17

通讯作者: 张侃, E-mail: zhangk@psych.ac.cn

Magnee (2008) 等人在对广泛性发展障碍 (Pervasive Developmental Disorder, PDD) 患者进行研究后发现, 在观看情绪性面孔时, 视觉刺激在 PDD 组和正常控制组中诱发了相同的 ERPs, PDD 患者表现出与常人无异的 P1 和 N170, 这说明 PDD 患者在快速知觉表情方面没有受损; 但如果在呈现恐惧表情的同时播放恐惧声音, 控制组能在顶-颞皮层看到明显的 N2, 而 PDD 患者则没有这种反应。这表明, PDD 患者在整合视觉和听觉的危险性信息方面与健康人的模式不同。由此, Magnee 等人认为, 虽然 PDD 患者对单通道情绪信息的加工正常, 但在多通道整合过程中出现了异常, 这很可能是导致患者在情绪知觉方面表现异常的主要原因。可见, 情绪知觉不仅要求大脑能准确快速的感知来自各个通道的信息, 还需要将这些信息有效的整合起来。任何一个环节出现问题都会影响个体正常的社会功能。

2 情绪信息的跨通道自动加工

行为学、神经成像与电生理学研究的結果均表明, 从一个通道输入的信息会改变对另一通道信息的觉知。视觉情绪信息能够影响大脑对听觉情绪信息的加工, 反之亦然。这种影响不仅发生在被试注意双通道的情况下, 也发生在被试仅仅注意一个通道, 甚至尽力忽视另一通道的情况下。这表明, 情绪信息的多通道整合具有自动加工的特点。

一个奇妙的例子是 McGurk 效应 (Mc Gurk & Mac Donald, 1976)。在屏幕上呈现 ba 的嘴型, 同时播放 ga 的声音, 人们却感觉自己听到的是 da。这一效应已经演变出很多实验范式, 它也适应于对多通道情绪信息整合的研究。de Gelder (2000) 等人采用这种范式探讨了视觉-听觉情绪信息的整合问题。实验者向被试随机呈现若干面孔, 各个面孔的表情从极度悲伤过渡到极度喜悦, 同时向被试播放声音, 声音只带有两种情绪, 或悲伤或喜悦。被试的任务是对情绪进行分类。结果表明, 被试可以有有效的整合来自两个通道的信息, 大脑对面孔情绪的判断受到同时呈现的语音情绪的影响。接下来, 研究者采用同样的范式, 但要求被试忽略声音, 仅判断面孔的情绪, 或者忽略面孔仅判断声音的情绪, 实验结果同样发现了

语音对判断面孔情绪的影响, 表明了视听情绪信息整合的自动加工特点。

同样的效应在对恐惧情绪的判断中表现得更为明显。在 Ethofer (2006) 等人的研究中, 实验采用了 3 种条件, 单通道视觉实验采用的是情绪度逐级变化的面孔表情 (从中性到 100% 恐惧以及从中性到 100% 喜悦); 单通道听觉实验采用喜悦或恐惧的语句作为刺激; 双通道实验则包含视听这两种刺激。单通道任务中, 被试的任务是判断刺激 (面孔或语调) 的情绪, 按 9 点量表打分; 双通道任务中被试仅对面孔的情绪打分。结果表明, 在伴随着恐惧声音的情况下, 被试对恐惧面孔的恐惧度评分更高, 甚至会将部分中性表情评定为恐惧表情。但愉悦的声音对表情评定并未产生任何影响。这很可能说明, 积极情绪与消极情绪的加工不同, 或者说大脑对一般情绪和同个体生存密切相关的情绪 (如恐惧) 的加工有很大不同。后者作为一种危险信号, 生物体对其快速而准确的觉察是生存所必须的。

3 多通道情绪信息整合的神经机制

近年来的认知神经研究表明, 可能有多个脑区参与多通道的信息整合, 包括皮层和皮层下组织, 如上丘 (superior colliculus, SC), 颞上沟 (superior temporal sulcus, STS) 顶叶以及前额叶的部分脑区等 (Driver & Noesselt, 2008)。那么多通道情绪信息整合的神经基础和机制又是怎样的呢? 整合发生在认知加工的哪个阶段, 又与哪些大脑结构相关呢?

最新的神经成像研究表明, 情绪性面孔-声音的整合区域与左侧颞中回 (middle temporal gyrus, MTG) (Ethofer et al., 2006; Park et al., 2008; Pourtois, deDelder, Bo, & Crommelinck, 2005), 左侧后颞上沟 (posterior superior temporal sulcus, pSTS) (Ethofer et al., 2006) 以及双侧的后颞上回 (posterior superior temporal gyrus, pSTG) (Kreifelts, Ethofer, Grodd, Erb, & Wildgruber, 2007) 密切相关。与单通道视觉和听觉情绪性刺激相比, 视听刺激还诱发了丘脑的活动 (Kreifelts et al., 2007)。对于恐惧性面孔-声音, 还可在杏仁核观测到这种多通道的整合 (Dolan, Morris, & de Gelder, 2001)。

3.1 多通道情绪信息的整合阶段——ERPs 研究

情绪信息的整合发生在认知加工的早期还是晚期? 根据已有的行为实验结果来看, 多通道信息的整合是极其快速而自动化的, 那么视觉和听觉情绪信息的交互很可能发生在早期阶段 (Eimer, 2001)。事件相关电位 (ERPs) 等电生理学手段初步确定了视听信息整合的时间进程。de Gelder 等人 (1999) 的研究证明, 同时呈现情绪相冲突的面孔和声音引发了一个早期的失匹配波 (Mismatch-negativity, MMN), MMN 约出现在刺激后 180ms。进而, 又一项 ERPs 研究发现, 与单独呈现情绪性声音相比, 伴随着表情的声音诱发出的 N1 振幅升高, 这一变化发生在刺激后 110ms 左右 (Pourtois, de Gelder, Vroomen, Rossion, & Crommelinck, 2000)。还有研究发现, 情绪不匹配的视听刺激在刺激后 220ms 发生了 N1 到 P1 的正向偏转 (positive deflection), 但情绪匹配的视听刺激发生正向偏转的时间有所延迟。亦有电生理学的研究表明视听信息整合发生在刺激后 200ms 内 (Barbara et al., 2008)。总之, 这些研究均说明, 视觉与听觉情绪信息的交互发生在早期, 即早期知觉阶段而不是晚期决策阶段。

值得注意的是, 有研究者提出, 视听整合要求空间与时间上的一致性, 即要求声音与面孔同时出现在同位置才会发生跨通道的捆绑现象。而研究也发现, 跨通道的捆绑 (Intermodal Binding, IB) 不仅对时间和空间有所要求, 对两种通道的信息内容也有要求。捆绑仅仅发生在面孔与声音情绪匹配的情况下, 对于倒置的面孔或者面孔与声音情绪不匹配的情况则不存在这种整合效应。在 Pourtois 等人 (2000) 的实验中, 研究者采用正置与倒置的面孔, 同时呈现情绪匹配或不匹配的声音, 要求被试对面孔进行分类, 结果发现, 匹配的视听刺激比不匹配的视听刺激诱发出振幅更强的 N1。并且, 这种听觉情绪的促进作用仅发生在面孔正置的情况下, 如果面孔是倒置的, 则无论任何情绪的声音都没有差别。可见, 听觉 ERP 的差别并不是初级图片特征的差异导致的。总之, 这些电生理学研究发现情绪信息的多通道整合发生在大脑加工的早期 (约刺激后 110~220ms)。

对婴儿的 ERPs 研究初步揭示了人类获得这

种整合能力的年龄阶段。Grossmann 等人 (2006) 对婴儿的视听情绪整合能力进行了研究。他们给 7 个月大的婴儿呈现面孔表情 (愉悦或愤怒) 和声音, 声音的情绪与面孔匹配或不匹配。结果发现, 情绪不匹配的声音诱发了较强的负波, 而情绪匹配的声音则诱发了较强的正波。根据已有的脑电研究结果, 负波的衰减和晚期正波 (late positive component, LPC) 的加强反映了婴儿对物体的识别加工过程。也就是说, 匹配声音的出现促进了婴儿对表情的识别, 而不匹配的声音则没有起到促进作用。说明 7 个月大的婴儿已经能够进行多通道的情绪信息整合, 识别面孔和声音中共有的情绪。

3.2 多通道情绪信息整合的神经基础——神经成像研究

神经成像技术为研究者提供了探讨信息整合的神经基础的手段。通过比较单通道刺激与多通道刺激的激活脑区, 研究发现多通道刺激的加工区域并非是各个单通道刺激加工区域的算术叠加。交互式分析 (Interaction analyses) 的方法是: 如果 $Bimodal > Unimodal1 + Unimodal2$, 则超出部分被定义为正性交互 (Positive Interaction), 表明这部分脑区与多通道信息的整合有关 (Ethofer et al., 2006; Wildgruber, Ackermann, Kreifelts, & Ethofer, 2006; Park et al., 2008)。

通过对 6 种最基本的人类情绪研究, Kreifelts 等人 (2007) 发现, 人对面孔-声音的情绪识别率显著高于单独呈现面孔或声音。而且无论是积极情绪 (兴奋、愉悦), 还是消极情绪 (厌恶、愤怒、悲伤、恐惧) 都表现出了这种优势。进一步的 fMRI 结果显示, 与单通道视觉和听觉情绪刺激相比, 视听结合的刺激激活了双侧的后颞上回 (pSTG) 以及右侧丘脑。其中, 左侧 pSTG 的激活最为显著, 其 BOLD 反应水平与伴随声音情绪带来的行为增益基本成线性关系。右侧丘脑和右侧 pSTG 也表现出这种趋势, 但没有达到显著水平, 表明左侧颞叶的部分结构对多通道情绪信息整合非常关键。

另外还有研究发现, 与单通道刺激相比, 恐惧的声音-面孔激活了左侧后颞上沟 (pSTS) 和左侧杏仁核, 并延展到杏仁核周围灰质 (periamygdaloid), 表明这些结构很可能参与了多通道情绪信息的整合。而且, 恐惧的声音比愉

悦的声音更强烈地激活了右侧梭状回腹部 (midsection of right fusiform gyrus), 研究者认为, 这说明该区域的活动与警觉信息的加工密切相关 (Ethofer et al., 2006)。

正电子发射断层扫描研究 (positron emission tomography, PET) 也发现 (Pourtois et al., 2005), 与单独的视觉和听觉刺激相比, 双通道刺激激活了左侧颞中回 (MTG) 和双侧颞上回 (STG)。研究还发现, 恐惧的表情以及面孔-声音激活了杏仁核的反应, 但单独呈现恐惧声音并没有激活杏仁核。更为有趣的发现是, 喜悦的面孔-声音激活区主要位于左半球前部, 而右侧半球对恐惧的面孔-声音更为敏感。这表明, 正性情绪与负性情绪整合有着不同的神经基础 (Armony & Dolan, 2002; Pourtois et al., 2005)。

继而, Park 等人 (2008) 又在实验中引入了中性刺激, 通过单通道刺激与双通道刺激比较发现视听刺激的整合区域位于左侧颞中回 (MTG), 这一结论印证了 Pourtois 等人的 PET 研究结果。Park 的研究同时发现, 3 种类型的视听刺激 (喜悦、愤怒、中性) 均激活了颞上回 (STG), 额下回 (inferior frontal gyrus, IFG) 和海马旁回 (parahippocampal gyrus), 包括杏仁核。这表明, 这些视听刺激虽然情绪类型不同, 却在整合过程共享了同一个神经网络。另一方面, 不同的情绪又有其不同的加工机制。通过情绪刺激与中性刺激比较, 研究者发现特定的恐惧整合加工脑区位于扣带回后部, 梭状回和小脑; 而颞中回, 海马, 扣带回前部, 颞下回与颞中回等则与喜悦相关。特定的情绪激活区域暗示各类情绪信息既有共同的神经基础, 又有其独特的整合特点。

另外, 各研究采用的范式和情绪类型与研究结果有着密切的关系。在前两个研究中 (Ethofer et al., 2006; Kreifelt et al., 2007), 两者均发现了颞上回在整合过程中的重要作用。而这两个实验均采用外显的情绪分类任务, 亦即自上而下的信息加工, 由注意驱动。但是, 在后两个研究中 (Pourtois et al., 2007; Park et al., 2008), 实验采用的是内隐任务, 如对性别进行分类等, 情绪加工是内隐的、自下而上的加工, 这种加工过程是自动化的, 与注意因素无关。如果比较外显与内隐的情绪整合过程, 可以发现, 颞上回更多的参

与了外显的情绪信息整合。可见, 情绪加工类型以及注意在其中的作用是研究情绪信息整合不可忽视的因素。

4 研究展望

情绪信息的多通道整合是一个十分复杂的过程, 尽管近年来国外对此已开展了大量研究, 但对于多通道整合的机制与神经基础还没有明确的答案。综合这些研究结果, 尚存在以下几点问题: 一是不同研究者采用的实验材料有所差异, 在情绪效价上 (如恐惧的程度, 愉悦的程度等) 没有统一的衡量标准, 导致不同的实验结果之间难以对比, 无法排除材料差异对实验结果的影响; 二是目前研究大部分采用静态表情图片作为视觉刺激 (Eimer et al. 2007; Holmes, Vuilleumier, & Eimer, 2003; Vuilleumier, Armony, Driver, & Dolan, 2001), 这与现实生活中人们常见的动态表情差异较大。早有研究表明, 大脑对静态信息与动态信息的加工机理有很大差异。因此, 要想将实验结论推广到自然环境中去, 还需要对实验材料和范式进行较大的改进, 使其更符合人们的日常生活。

此外, 大量的神经成像和电生理研究已经探讨了多通道整合的时程与相关脑区, 初步研究了情绪信息整合的神经基础, 但这些研究多将关注点聚焦于情绪信息的简单加工。而事实上, 情绪的加工与其他认知加工过程密切相关, 比如注意在情绪加工中扮演的角色 (Mazza, Turatto, Rossi, Turatto, & Umilt, 2007)。情绪信息的整合是一个快速而自动化的过程, 但注意资源是否能起到一定的调控作用? 已有研究曾试图探讨单通道情绪信息的自动加工问题, 但在注意所扮演的角色方面仍存在争议 (Holmes et al., 2003; de Gelder et al., 2005; Pauli & Roeder, 2008)。而研究多通道情绪信息加工与注意的关系则有助于更深入地探索情绪信息的整合机制。

另外, 特殊群体的研究也有助于我们弄清跨通道信息加工的神经机制。推广实验对象有两个思路和缘由。其一, 纵向研究。对发育中的个体, 如婴儿、儿童的研究有助于我们弄清人类情绪认知发展的过程以及多通道整合的发生发展阶段 (如 Grossman, Neil 等人的研究); 其二, 横向研究。某些特殊人群在情绪加工方面有其独特的特

点,如脑损伤病人(如 Morris, DeGelder, Dolan 等人的研究)、精神疾病患者(如 de Gelder, Vroomen 等人的研究)、广泛性发展障碍(PDD)患者或注意力缺陷多动障碍(ADHD)的儿童等等,比较他们与常人在情绪认知时的脑活动差异,有助于我们了解情绪加工的神经基础;还有一些特殊人群的特点表现在跨通道加工方面,如听觉或视觉缺失的残疾人,很多研究者通过研究盲人的代偿行为来探索跨通道加工的过程,在不同发育期损伤视力的盲人(先天盲人、早期盲人、后期盲人)表现出的代偿行为和相关脑机制存在很大不同(Bavelier & Neville, 2002),但关于情绪认知的跨通道发展与代偿还有待研究。

多通道整合是人类知觉的重要方式,而情绪加工是人类最基本的一种认知活动,它对于每一个生活在社会中的个体来说都有着极其重要的意义,是个体生存与交往的必需。因此,情绪信息的多通道整合研究是探索人类认知活动的神经基础,是促进神经康复医学的基础课题。希望上述研究成果与几点建议能对相关领域的研究者有所启发。

参考文献

- Armony, J. L., & Dolan, R. J. (2002). Modulation of spatial attention by fear-conditioned stimuli: an event-related fMRI study. *Neuropsychologia*, *40*, 817–826.
- Barbara, A. B., Lucy J. M., William J. G., & Patricia, L. D. (2008). Multisensory integration in children: A preliminary ERP study. *Brain Research*, *1242*, 283–290.
- Bavelier, D., Neville, H. J. (2002). Cross-modal Plasticity: where and how. *Nature Reviews Neuroscience*, *3*, 443–452.
- Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Research*, *1242*, 126–135.
- de Gelder, B., Boecker, K., Tuomainen, J., Hensen, M., & Vroomen, J. (1999). The combined perception of emotion from voice and face: early interaction revealed by human electric brain responses. *Neuroscience Letters*, *260*, 133–136.
- de Gelder, B., & Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, *14*, 289–311.
- de Gelder, B., Vroomen, J., de Jong, S. J., Masthoff, E. D., Trompenaars, F. J., & Hodiament, P. (2005). Multisensory integration of emotional faces and voices in schizophrenics. *Schizophrenia Research*, *72*, 195–203.
- de Gelder, B., Morris, J. S., & Dolan, R. J. (2005). Unconscious fear influences emotional awareness of faces and voices. *PNAS*, *102*, 18682–87.
- Dolan, R.J., Morris, J.S., & de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proceeding of National Academy Science USA*, *98*, 10006–10010.
- Driver, J., & Noesselt, T. (2008). Multisensory Interplay Reveals Crossmodal Influences on 'Sensory-Specific' Brain Regions. *Neuron*, *57*, 11–20.
- Eimer, M. (2001). Crossmodal links in spatial attention between vision, audition, and touch: evidence from event-related brain potentials. *Neuropsychologia*, *39*, 1292–1303.
- Eimer, M., & Holmes, A. (2007). Event-related brain potential correlates of emotional face processing. *Neuropsychologia*, *45*, 15–31.
- Ethofer, T., Pourtois, G., & Wildgruber, D. (2006). Investigating audiovisual integration of emotional signals in the human brain. *Progress in Brain Research*, *156*, 345–361.
- Flom, R., & Bahrick, L.E. (2007). The development of infant discrimination of affect in multimodal and unimodal stimulation: The role of intersensory redundancy. *Developmental Psychology*, *43*, 238–252.
- Ghazanfar, A. A., & Logothetis, N. K. (2003). Neuroperception: Facial expressions linked to monkey calls. *Nature*, *423*, 937–938.
- Grossmann, T., Striano, T., & Friederici, A.D. (2006). Crossmodal integration of emotional information from face and voice in the infant brain. *Developmental Science*, *9*, 309–315.
- Holmes, A., Vuilleumier, P., & Eimer, M. (2003). The processing of emotional facial expression is gated by spatialattention: evidence from event-related brain potentials. *Cognitive Brain Research*, *16*, 174–184.
- Kreifelts, B., Ethofer, T., Grodd, W., Erb, M., & Wildgruber, D. (2007). Audiovisual integration of emotional signals in voice and face: An event-related fMRI study. *Neuroimage*, *37*, 1445–1456.
- Lugo, J. E., Doti, R., Wittich, W., & Faubert, J. (2008). Multisensory Integration: Central Processing Modifies Peripheral System. *Psychological Science*, *19*, 989–997.
- Magnée, M., de Gelder, B., van Engeland, H., & Kemner, C. (2008). A typical processing of fearful face-voice pairs in Pervasive Developmental Disorder: An ERP study. *Clinical Neurophysiology*, *119*, 2004–2010.
- Mazza, V., Turatto, M., Rossi, Turatto., & Umilt, C. (2007). How automatic are audiovisual links in exogenous spatial attention? *Neuropsychologia*, *45*, 514–522.
- Mc Donald, J., Teder-Salejarvi, W., & Ward, L. M. Multisensory integration and crossmodal attention effects

in the human brain. (2001). *Science*, 292, 1791.

Mc Gurk, & Mac Donald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746–748.

Morris, J. S., de Gelder, B., Weiskrantz, L., & Dolan, R. J. (2001). Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain*, 124, 1241–1252.

Neil, P.A., Chee-Ruiter, C., Scheier, C., Lewkowicz, D. J., & Shimojo, S. (2006). Development of multisensory spatial integration and perception in humans. *Developmental Science*, 9, 454–464.

Park, J. Y., Gu, B. M., Kang, D. H., Shin, Y. W., Choi, C. H., Lee, J. M., & Kwon, J. S. (2008). Integration of cross-modal emotional information in the human brain: An fMRI study. *Cortex*, In Press, Available online 30 July 2009.

Pauli, W. M., & Roeder, B. (2008). Emotional salience changes the focus of spatial attention. *Brain Research*, 1214, 94–104.

Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B., & Crommelinck, M. (2000). The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport*, 11, 1329–1333.

Pourtois, G., de Gelder, B., Bo, A., & Crommelinck, M. (2005). Perception of facial expressions and voices and of their combination in the human brain. *Cortex*, 41, 49–59.

Spence, C., & Driver, J. (2004). *Crossmodal space and crossmodal attention*. Oxford, New York, USA: Oxford University Press.

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of Attention and Emotion on Face Processing in the Human Brain: An Event-Related fMRI Study. *Neuron*, 30, 829–841.

Wildgruber, D., Ackermann, H., Kreifelts, B., & Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. *Progress in Brain Research*, 156, 249–268.

The Multisensory Integration of Emotional Information

ZHANG Liang^{1,2}, SUN Xiang-Hong¹, ZHANG Kan¹

(1 Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China)

(2 Graduate University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: In the natural environment, human receive emotional information through multiple sensory modals. The Multisensory Integration (MSI) is the foundation of emotion processing. Plenty of behavioral, electrophysiological and neuroimaging researches show that the multisensory integration of emotional information is automatically and integrates in the early stage of processing. The STG, MTG, parahippocampal gyrus and thalamus are involved in the integration of emotion processing. Each emotion uses a separate network and shares a common network. The process is also affected by the task and the attention. Further researches in this field can focus on the neural mechanism by standardizing and comparing researches, widening specific populations. The attention and the related cognitive processing also play key roles in the multisensory integration of emotion information.

Key words: emotion; multisensory integration; face-voice integration effects; McGurk effect; interaction analyses