心理科学进展 2014, Vol. 22, No. 5, 802–809
Advances in Psychological Science DOI: 10.3724/SP.J.1042.2014.00802

# 声音情绪的跨文化识别

孙继民 姚温青 胡学平

(西北师范大学心理学院, 兰州 730070)

**摘 要** 声音情绪跨文化识别是在跨文化前提下对无语义声音所传递情绪信息的识别。目前, 对声音情绪跨文化识别的研究热点主要集中于: 跨文化一致性、群内优势效应、性别和年龄问题、自发声音情绪识别、语言材料的选取等 5 个方面的研究。今后, 声音情绪跨文化识别的研究应在声音情绪的材料获取、影响因素、双语研究这三个方面作进一步的探索和完善。

**关键词** 声音情绪; 跨文化研究; 群内优势效应; 自发情绪

**分类号** B842

## 1 引言

情感信号对识别信息具有关键性作用, 例如, 警示人类危险。1872 年, 达尔文在《人类和动物的表情》中提出：人类对情绪情感的识别具有种族文化特征。为此许多心理学家进行了大量的实践验证和理论探索(Albas, McCluskey, & Albas, 1976; Jack, Garrod, Yu, Caldara, & Schyns, 2012; Scherer, Banse, & Wallbott,2001)。Elfenbein 和 Ambady (2002, 2003)研究不仅发现人们在跨文化前提下具有基本情绪的“普遍情感系统”, 还进一步阐明“普遍情感系统”由面部表情、语言和肢体等组成。单就语言线索而言, 人类的语言情感信息通过语义和非语义两种形式传递(Banse & Scherer, 1996; Hammerschmidt & Jürgens, 2007; Juslin & Laukka, 2003; Sauter, Eisner, Ekman, & Scott, 2010)。其中本文将要探讨的非语义情绪则包含两种形式(Laukka et al., 2013), 情绪韵律(emotional prosody) (Scherer, 1986)和非语言发声(non-linguistic vocalizations) (Gobl & Ni Chasaide, 2003)。对情绪韵律的定义是：不考虑词汇和语法结构, 通过重音和语调的变化来传达不同形式的情绪, 或者说独立于语言内容而传递的情感信息(Kitayama & Ishii, 2002; Monrad-Kohn, 1947)。Bostanov 和 Kotchoubey (2004)研究发现, 人们通过感知情绪语音中的声学线索, 从中提取出所携载的情绪意义以方便与其它个体的交流。因此在交流的过程中, 听者参与到语音、节奏和情感韵律的变化中, 并结合语言的解码方式识别语言的情感信息。对情绪韵律的研究发现, 韵律识别的模式具有普遍性(Frick, 1985), 在跨文化条件下被试能够超过随机猜测水平识别情绪韵律内容(Albas et al., 1976; Beier & Zautra, 1972; Bryant & Barrett, 2008; Kramer, 1964; Pell, Monetta, Paulmann, & Kotz, 2009; Scherer et al., 2001; Thompson & Balkwill, 2006)。而非语言发声仅仅是一种不用言语而发出的声音, 如叹息声、哭泣声或是性愉悦声等等。虽然非语言发声逐渐成为情绪跨文化普遍性(Ekman, 1992; Laukka et al., 2013; Russell, 1994)和情绪科学研究(Barrett, 2006; Ekman, 1992; Scherer, 1986 )的热点, 但是非语言发声的研究(Lima et al., 2013; Sauter et al., 2010; Sauter & Scott, 2007; Schröder, 2003; Simon-Thomas et al., 2009)比情绪韵律的研究相对较少。非语言发声研究也得出了与情绪韵律相似的结论, 在非语言发声条件下基本情绪也能够被识别, 甚至有时得到的识别率要高于情绪韵律方面的识别率(Hawk et al., 2009)。二者的区别则是, 情绪韵律具备语言规则和完整协调的发声器官, 而非语言发声不仅没有那些束缚并且比情绪韵律更能引发对情绪情境的识别(Scott et al., 2009)。因此, 本文将非语义情

收稿日期：2013-09-03
通讯作者：姚温青, E-mail: yaowenqinge@126.com

绪的跨文化研究对象(情绪韵律和非语言发声)命名为：声音情绪。综上所述，声音情绪跨文化识别就是在跨文化背景下对无语义声音所传递情绪信息的识别。纵观声音情绪的研究发现，其热点问题主要集中于 5 个方面：(1)是否具有跨文化的一致性；(2)是否具备群内优势效应；(3)声音情绪跨文化识别是否与性别和年龄相关；(4)是否自发声音情绪更易被识别；(5)实验材料的选取对研究的影响。

## 2　声音情绪识别的跨文化研究

### 2.1　跨文化一致性

声音情绪识别是否具有跨文化的一致性？大多数研究显示，对基本情绪的情绪韵律跨文化识别超过了随机猜测的水平而表现出了跨文化一致性 (Beier & Zautra, 1972;Jürgens, Drolet, Pirow, Scheiner, & Fischer, 2013; Kramer, 1964; McCluskey, Albas, Niemi, Cuevas, & Ferrer, 1975; Sauter et al., 2010; Scherer et al., 2001)。Kramer (1964)研究发现，美国被试能够超过随机猜测水平识别过滤掉日语语义内容的声音情绪。在 Beier 和 Zautra (1972)研究中录制了英语的 48 句不同句长的情绪(愉快、恐惧、悲伤、愤怒、冷漠)表达，并分别呈现给 52 名美国被试、55 名波兰被试和 54 名日语被试，结果显示他们对声音情绪材料的平均识别率几乎都高于 50%。而且，当随着材料语句的加长，其识别程度也会提高。这超出了随机猜测的可能性，得出跨文化一致性的结论。Ferrer 等(1975)给 6~11 岁的加拿大和墨西哥儿童各 10 人随机呈现由加拿大本土语言和墨西哥本土语言录制的愉快、悲伤、爱和愤怒 4 种词汇情绪材料，发现识别率超过猜测水平。随之，为了避免数据单薄而不能充分说明问题的嫌疑，Scherer 等(2001)让欧洲、美国和亚洲九国被试识别德语发声的情绪，验证了跨文化背景下声音情绪可以基本被识别的结论。为了进一步验证这一结论的可靠性，Thompson 和 Balkwill (2006)让 20 名英语被试对具备 4 种发声情绪(愉快、悲哀、愤怒、恐惧)的英语、德语、日语、中文和塔加拉族语的中性语义语句进行了情感识别。同样得出与先前研究一致的结论。继而, Pell 和 Skorup (2008)对声音情感加工的普适性和时间进程进行了验证。他们用英语和阿拉伯语作为语言材料，用相应的面部材料作为是非一致的判断。结果显示，两种语境下的被试能够判断到彼此韵律提供的情感信号。Pell (2009)比较了只懂自己语言的阿根廷西班牙人是怎样识别自己国家和其它三国非意义语言的情绪。结果显示，每种语言的情绪韵律都会超过随机猜测水平被识别。他们研究得出：尽管自身的语言和文化对识别者产生了约束，但是对他国语言情绪韵律的识别部分是独立于语言能力和普遍规则的。Sauter 和 Ekman 等(2010)在完全控制了语言因素的情况下进行了文化因素的单项研究，录制英语和辛巴语文化下的情绪“愤怒、厌恶、恐惧、愉快、悲哀、惊奇”发声，对英国被试和与之具有显著文化差异的辛巴族被试进行了“非语言发声识别”的研究。如，实验先给辛巴族被试用辛巴语讲授一个恐惧的故事“Fear Someone is suddenly faced with a dangerous animal and feels very scared.”，再呈现在英语语境下录制的恐惧和惊奇发声，让辛巴族被试做匹配判断，发现对发声情绪的识别超过随机猜测的水平。英语被试则进行了相对应的实验。由此，他们得出非语言情绪发声也可以通过跨文化识别而具备跨文化一致性。Sauter 等人还依据发声情感信号的一些积极情感状态(Sauter & Scott, 2007)，将愉快归结为积极情绪，而将研究中的其它基本情绪归结为消极情绪。发现对消极情感的识别具备完全的跨文化的一致性，而对积极情感的识别却包含有文化的特殊性。Sauter 和 Ekman 等人对此作了这样的解释：情感社会信号由组内成员所共享。而且，积极情感的交流更有助于群内社会的联系(Shiota, Keltner, Campos, & Hertenstein, 2004)。Laukka 等(2013)的非语言发声情绪识别研究得到与前面研究一致的结论，被试在跨文化背景下能够超过随机猜测水平识别积极和消极情绪。到目前为止，几乎所有的研究都支持了这样的假设：情绪能够通过语言声音识别(Ekman & Friesen, 1971)。因此可以得出结论：声音情绪能够在跨文化前提下超过随机猜测水平被识别，且呈现跨文化一致性效应。

### 2.2　群内优势效应

情绪表达者与情绪识别者如果来自同一种文化，那么他们彼此的情绪理解与交流更为精确，这种现象称为组内优势效应或者群内优势效应(in-group advantage)。这是由 Elfenbein 和 Ambady (2002)对 87 篇与跨文化有关的文献数据进行元分

析后提出的，并于 2003 年的研究中做了进一步的完善，以跨文化方式对群内优势效应进行研究，以正确率为指标进行现象解释。并得出以下结论：情绪表达者与情绪识别者来自相同文化时，他们的情绪理解与交流更为精准；情绪表达者与情绪识别者的文化暴露越大，群内优势效应越小；小群体成员识别大群体成员情绪的准确性要高于大群体成员识别小群体成员情绪的准确性；群体之间地理位置越近、联系越密切，群内优势效应越小；被试对呈现表情的人的文化背景越熟悉、联系越密切，群内优势效应越小，而群内优势效应与种族关联却很小(Elfenbein & Ambady, 2003)；如果个体习得了怎样理解其他文化的人的情绪，群内优势效应也会减小(Elfenbein & Ambady, 2003)。

声音情绪的群内优势效应研究也引发了很多研究者的兴趣。Albas 等(1976)在对加拿大的两组被试进行方言语言情感内容方面研究发现，印第安被试识别印第安文化下的语言情绪要显著的比白人被试识别的好，相反白人被试识别自己文化下的语言情绪要显著比印第安被试好。这与 Pell 等(2009)让阿根廷的西班牙被试识别本国的其它 3 种语言研究中得出的结论是一致的。阿根廷的西班牙被试对本族语言的识别具有显著高的精确度，而对本国的 3 种与之相似度极低的其它语言的情绪识别就比较低。随后, Scherer 等(2001)研究发现，被试所成长的区域若与德国相近的话，识别率也相对来说要高于其它不相近国家被试而显现出群内优势效应。为了进一步验证群内优势效应, Wickline, Bailey 和 Nowicki (2009)选取了非裔美国学生、欧洲裔美国学生，非洲留学生和欧洲留学生作为 4 组被试，要求被试通过面部表情和声音表情来识别非裔美国人和欧裔美国人的面部情绪。结果发现，非洲留学生识别美国人面部情绪和声音的正确率普遍低于其他 3 组，欧洲裔美国人识别美国人面部情绪和声音的准确性高于其他 3 组。Ekman 等人于 2010 年的发声研究也得出，英语被试对辛巴族被试的发声识别要低于辛巴族被试对辛巴族语的发声识别。研究发现：本土文化下声音情绪的表达和解读习惯(Elfenbein, Beaupré, Lévesque, & Hess, 2007)对非本土文化的声音情绪加工产生影响(Mesquita & Frijda, 1992; Pell & Skorup, 2008; Scherer et al., 2001)。

然而，也有一些差异存在, Jürgens 等(2013)让德国，罗马尼亚和印度尼西亚的被试识别德语情绪音频，发现德国被试在识别情绪上只具有一点微乎其微的优势效应。这与先前的研究略微不一致。这样的数据在声音情绪识别中的研究中不多，但在面部表情的研究中却有典型例子, Matsumoto, Olide 和 Willingham (2009)选取了在 2004 年雅典奥运会上美国运动员和日本运动员在获得奖牌后的面部表情照片，让美国被试和日本被试对这些照片进行识别，进行是否会得奖的判断，却发现群内优势效应在现实生活情境下不存在了。于是, Matsumoto (2002)认为组内优势效应是否存在，从跨文化的研究内容来看，可能跟研究方法和材料的不同有关。而且，有研究表明，情绪识别的准确性与人格特质(Matsumoto et al., 2000)、社交性人格类型(Mill, 1984)和社会调节能力(Leppaenen & Hietanen, 2001)有关，这些因素是否也会作为调节因素对群内优势效应造成影响呢？考虑到对声音情绪的跨文化研究还不是很成熟，有必要对群内优势效应进行进一步的探索研究。

**2.3** 性别因素和年龄因素

就声音情绪跨文化识别年龄因素的研究，最早是由 McCluskey 等人在 1975 年进行的，他们让加拿大和墨西哥的 6~11 岁男孩与男性本科生用愉快、悲伤、愤怒和爱等情绪说出任何他们喜欢的词汇，然后将过滤掉语义的情绪声音词汇随机呈现给他们，发现随着年龄的上升，识别精确度也是上升的。这也仅是将男性作为被试进行的实验，如果是女性被试是否也会得出同样的结论，这需要进一步的研究。而 Albas 等(1976)做的实验也是基于白人和印第安男性被试进行的，由于并未将性别作为研究因素，也仅证明了跨文化识别的一致性结论。DePaulo 和 Rosenthal (1978, 1979)让 632 名本土被试识别本土文化的非语义语言情绪，经研究发现，在识别精确度方面女性被试明显高于男性被试，高年龄段被试明显高于低年龄段被试。McCluskey 和 Albas (1981)对墨西哥和加拿大 5~65 岁年龄段的被试呈现情绪(愉快、悲伤、爱、愤怒)声音，发现随着年龄的增加，识别情绪韵律的能力也是增加的。2007 年 Paulmann 等人在跨文化前提下对不同性别或年龄的被试进行了基本情绪(愤怒、厌恶、恐惧、愉快、惊奇、悲伤)识别的探索。他们要求被试尽可能的通过声音韵

律识别情绪。结果发现，情绪识别的正确率体现在不同的情绪方面，这些模式变化与年龄相关，但与性别无关。此外, Tracy 和 Robins (2008)探索了是否骄傲的表现会与种族和性别有一定的关系，发现它二者对骄傲声音情绪的跨文化识别无影响。

由此总结，在声音情绪的跨文化识别中年龄是一个影响因素，随着年龄的上升，识别率是增大的。但是否是一个直线上升还是在一定阶段会平缓或是下降趋势，就需要更多的研究对其做进一步的探究。就目前研究而言，性别对声音情绪的跨文化识别几乎没什么影响，但又不能完全排除。Chun, Park, Park 和 Kim (2012)在相同文化中进行了性别是否影响情绪识别的研究。他们采用不同性别制作哭、笑和中性三种材料，让不同性别的被试识别，发现同性之间更能识别愉快，而异性之间更易识别悲伤。研究还发现，在听到笑时同性之间比异性之间的在海马回上有更多激活，可能反映了更多的情绪加工和注意倾向。而在听到哭时，异性之间比同性之间在后扣带回产生更多的激活，这可能反映了更多的同理心加工。Chun 等人认为可能的原因是，男女在识别方面的大脑反应不同，男性更多是情感的表现，而女性更多为同理心的关注。虽然对这样的解释还没有定论，但是已开始涉及到脑电研究的领域。且该研究将会对跨文化声音情绪识别研究产生积极影响。

**2.4** 自发声音情绪

对面部表情的观察研究发现，来自不同文化的成员都能可靠地判断出自发面部表情(Matsumoto, Olide, & Schug, 2009)。那么来自不同文化的成员能否正确识别自发或是表演的声音情绪？也逐渐成为了跨文化研究的热点。Drolet, Schubotz 和 Fischer (2012)研究发现，声音的真实性会影响情感意义的识别。Scheiner 和 Fischer (2011)将 80 句无语义德语的自发声音情绪和表演声音情绪语句呈现给德国、罗马尼亚、印度尼西亚三国被试。经研究发现，三国被试的跨文化识别模式很相似，而且对自发和表演的判断识别不是很显著(刚刚超过随机水平)，表演情绪的“愤怒”较容易被识别，自发情绪的“悲伤”较容易被识别。声音情绪对被识别者行为的影响众所周知(Buchanan et al., 2001; Pell, M. D., Jaywant, A., Monetta, L., & Kotz, S. A., 2011; Scherer, 1991)，大脑激活以及相关概念已成了现在的研究热点( Vigneau, 2006; Wildgruber et al.,2002)。Drolet 等(2012)通过 fMRI 的脑成像实验研究，进一步探索发现在声音情绪的识别方面自发刺激比表演刺激具备特殊的真实性。研究发现，表演刺激比自发刺激在 ToM (theory of mind)网络中引发了更大的激活，而自发刺激则更能提高 ToM 网络的激活和促进社会情境的加工。并认为声音的自发或表演将是影响人类识别情绪的重要影响因素。Drolet, Schubotz 和 Fischer (2013)进一步研究发现不管是表演的还是自发的声音刺激都能被被试识别，且都会促进被试的内侧前额叶、颞皮层和胼胝体皮质对工作记忆和任务决策的加工。尽管表演的声音情绪表达比自发情绪表达具备更高的精确度、更加典型和激烈(Jürgens et al., 2011; Laukka et al., 2012; Scherer, 2013; Wilting, Krahmer, & Swerts, 2006)、更具社交性(Hunt, 1941; Matsumoto et al., 2009)。但自发声音情绪要比表演声音情绪具备真实性而更容易被识别，并对大脑的激活具有很显著的影响(Drolet et al., 2012)。这也许是因为在现实生活中，自发情绪具备更多的社会显著性，或自发情绪比表演情绪在激活自然心智化能力时具备更大的效力。基于这些研究，2013 年 Rebecca 等人从德国音频档案卷内获取了描述 4 种情绪(愤怒、悲伤、愉快、恐惧)的语句，让专业演员进行录制评定后，呈现给德国、罗马尼亚和印度尼西亚的 120 名被试。发现来自 3 个国家的被试几乎不能很好的区分表演和自发语言情绪。而印度尼西亚被试和罗马尼亚被试更容易识别自发情绪的“悲伤”；德国被试倾向于选择自发和表演的“愤怒”。于是 Rebecca 等人推测，是否集体主义民族倾向于加工消极情感，而个体主义民族倾向于加工与之平时生活相一致的愤怒情绪？不同文化结构产生的自我构建在声音情绪跨文化识别方面的差异问题？或者，是自发情绪与表演情绪的跨文化差异导致了这样的现象？这些都需要作进一步的探索。因此，今后的研究应该更加慎重的选用自发声音情绪和表演声音情绪。

**2.5** 语言材料的选取

根据跨文化研究范式，声音情绪识别材料的制作与选取主要有 3 种方式：(1)平衡设计，录制各文化下所有参与实验的被试的声音材料，然后

全部呈现给所有被试识别(Albas et al., 1976; Kramer, 1964; McCluskey et al., 1975; Paulmann & Uskul, 2013; Sauter et al, 2010; Tracy & Robins, 2008); (2)非平衡设计，采用单一文化背景下的声音材料，饰以基本情绪情感后呈现给多国文化背景下的语言被试(Beier & Zautra, 1972; Jürgens et al., 2013; Scherer et al., 2001); (3)反非平衡设计，采用多国文化下的非语义声音，呈现给一国文化下被试识别(Laukka et al., 2013; Pell et al., 2009; Thompson & Balkwill, 2006)。在这里，有些事实不得不阐明，几乎绝大部分声音材料为了避免语义的干扰而做过一些处理，如过滤、假词和假句等。Beier 和 Zautra (1972)研究材料是典型的第二种方法，他们用录音带制作了一个 48 句不同句长(包含 4 种情绪)的语库。让 100 名美国被试进行等级评定，达到将近 60%~80%的同意度。再用录音呈现给各国语言被试。McCluskey 等(1975)用第一种方法进行了材料选取，在研究中他们让加拿大和墨西哥被试用愉快、悲伤、爱和愤怒 4 种情绪表达各自本土的语言词汇。将经过情感过滤器处理后的情绪语言样本随机呈现给两国被试，发现随着年龄的上升识别精确度也逐渐提高了。Sauter 等(2010)情绪发声研究也用了第一种方法。第三种材料选取的典型研究是 Petri Laukka 等(2013)进行的非语言发声的研究。他们录制印度、肯尼亚、新加坡和美国情绪声音，只呈现给瑞典被试识别。不难发现, 4 种方法各具利弊。当进行平衡材料设计时，对声音情绪的研究是可对比的。其说服力要相对好些。但考虑到被试的负荷量以及研究的局限，方法二和方法三，就一语多国被试和多语一国被试的研究也的确是个很好的验证。而基于此的方法还有侧重于阅读韵律情绪的研究和探索。将逐渐成为声音情绪研究的方向。而对语义的处理更是体现了对声音情绪的研究过程中的额外因素的控制。语义不加改动时，是最原始和真实的表现，因此让被试去识别时所得到的也是真实的记录，但又免不了会怀疑被试是否会依据词汇发生猜测。而若是改动了语义词汇，也许会消除被试对发声的特别猜测和顾虑，但又会考虑其真实性。或是，自发的非语言发声，更多侧重生物特征而忽略文化作用。可见，声音材料的选择对研究目的和结论起到不可小觑的影响。虽然对刺激的选择很慎重，但是识别者的个体差异还是会影响情绪模式(Banse & Scherer, 1996; Scherer, Banse, Wallbott, & Goldbeck, 1991)。因此，声音情绪跨文化识别研究中的材料设计和录制将成为未来完善实验研究的重点。

## 3 小结与展望

相对面部表情识别的研究而言，对声音情绪的研究显得势单力薄。这与人类自身复杂的生物学特性有关。人们通过感知情绪语音中的声学线索，从中提取出所携载的情绪意义(Bostanov & Kotchoubey, 2004)。不同的声学线索能够反映情绪的不同维度，如愉悦度(Bänziger & Scherer, 2005; Juslin & Scherer, 2008)。人们不仅可以通过口语的音段成分传递的语义信息来理解说话人所表达的情绪，还可以通过口语声音中音高、时长、强度等声学线索的变化，感知说话人的不同情绪状态(Besson, Magne, & Schön, 2002)。因此，对声音情绪的跨文化研究将会对今后的跨文化交流以及学习产生重要的影响和作用。现如今对声音情绪识别的跨文化研究主要集中于跨文化一致性、群内优势效应以及自发声音等热点问题，今后的研究应该更加关注发声材料的获取、影响因素以及双语研究三个方面。

首先，发声材料的获取就是考虑材料录制设计以及自发和表演差异。不同的研究目的需要不同的材料设计，不同的材料设计需要不同的实验模式，不同的实验模式得出的研究结果也有差异。尤其声音材料的自发与表演差异性与文化的交互问题，对文化的一致性与群内效应的影响也是显而易见。如前文提到的“个体主义文化倾向于识别表演和自发的愤怒情绪，而集体主义的被试倾向于识别自发的悲哀情绪”等。这些都需要在研究之前，对研究目的以及材料获取做充足的认识和理解，以免获得无效结果。

其次，在影响因素方面，声音的跨文化特性是情绪发声以及识别的重要影响因素。被试作为独立具备自我情绪模式的个体，不管在性别方面或在年龄方面都将会对声音情绪的跨文化研究产生影响。Schirmer 和 Kotz (2003)的 ERP 研究中发现：性别对情绪韵律的加工有影响，女性比男性在情绪的理解上更敏锐；在非语义情绪识别方面，女性相对男性识别率高。然而，这方面的研究也只是寥寥数篇，因此需要更多的实证研究和理论

探索，来验证年龄的作用程度，以及性别的存在差异与否。所以，实证研究和理论研究的数量也对研究结论的考究具有重要的作用。有研究发现，识别率的高低显示其个体识别倾向性，如个体主义被试对积极情绪的识别要好些，而集体主义被试对消极情绪的识别要好些。因此，当文化作为一项重要影响因素，如自我建构，那么在这方面的研究就需要进一步的探究。

最后，近期研究将跨文化声音情绪研究逐渐转向了双语教学和学习的研究。如 Zhao (2013)在情感语调维度进行了研究，当将英语材料分别呈现给中文被试、中文为母语的中英双语被试、英语被试后发现英语被试可以很好的识别英语材料情感。而比起中文单语被试而言，中英双语被试对英文材料的识别更加接近英语被试的水平。但是，双语被试在英语识别的调型变化上受母语迁移的影响。因此，在双语角度上的声音情绪跨文化研究，将会使得双语人更加准确的掌握两个语言之间的差异，进而促进交流和学习。

## 参考文献


江爱世, 陈煦海, 杨玉芳. (2009). 言语情绪韵律加工的时间进程. *心理科学进展, 17*(6), 1109–1115.

张秋颖, 陈建文, 于全磊, 辛鹏. (2011). 情绪识别中的群内优势效应. *心理科学进展, 19*(2), 209–216.

Albas, D. C., McCluskey, K. W., & Albas, C. A. (1976). Perception of the emotional content of speech: A comparison of two Canadian groups. *Journal of Cross-Cultural Psychology, 7*(4), 481–489.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*(3), 614–636.

Bänziger, T., & Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech Communication, 46*(3-4), 252–267.

Barrett, L. F. (2006). Are emotions natural kinds? *Perpectives on Psychological Science*, 1, 28–58.

Beier, E., & Zautra, A. J. (1972). The identification of vocal expressions of emotion across cultures. *Journal of Consulting and Clinical Psychology, 40*(4), 560.

Besson, M., Magne, C., & Schön, D. (2002). Emotional prosody: Sex differences in sensitivity to speech melody. *Trends in Cognitive Sciences, 6*(10), 405–407.

Bostanov, V., & Kotchoubey, B. (2004). Recognition of affective prosody: Continuous wavelet measures of event-related brain potentials to emotional exclamations. *Psychophysiology, 41*(2), 259–268.

Bryant, G. A., & Barrett, H. C. (2008). Vocal emotion recognition across disparate cultures. *Journal of Cognition and Culture, 8*(1), 135–148.

Buchanan, T. W., & Lovallo, W. R. (2001). Enhanced memory for emotional material following stress-level cortisol treatment in humans. *Psychoneuroendocrinology, 26*(3), 307–317.

Chun, J. W., Park, H. J., Park, I. H., & Kim, J. J. (2012). Common and differential brain responses in men and women to nonverbal emotional vocalizations by the same and opposite sex. *Neuroscience Letters, 515*(2), 157–161.

Darwin. C. (1872). *The Expression of Emotions in Man and Animal*. London: John Murray.

DePaulo, B. M., & Rosenthal, R. (1978). Age changes in nonverbal decoding as a function of increasing amounts of information. *Journal of Experimental Child Psychology, 26*(2), 280–287.

DePaulo, B. M., & Rosenthal, R. (1979). Age changes in nonverbal decoding skills: Evidence for increasing differentiation. *Merrill-Palmer Quarterly of Behavior and Development, 25*(2), 145–150.

Drolet, M., Schubotz, R. I., & Fischer, J. (2012). Authenticity affects the recognition of emotions in speech: Behavioral and fMRI evidence. *Cognitive, Affective, & Behavioral Neuroscience, 12*(1), 140–150.

Drolet, M., Schubotz, R. I., & Fischer, J. (2013). Explicit authenticity and stimulus features interact to modulate BOLD response induced by emotional speech. *Cognitive, Affective, & Behavioral Neuroscience, 13*(2), 318–329.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*, 169–200.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*(2), 124–129.

Ekman, P., Freisen, W. V., & Ancoli, S. (1980). Facial signs of emotional experience. *Journal of Personality and Social Psychology, 39*(6), 1125–1134.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin, 128*(2), 203–235.

Elfenbein, H. A., & Ambady, N. (2003). Universals and cultural differences in recognizing emotions. *Current Directions in Psychological Science, 12*(5), 159–164.

Elfenbein, H. A., Beaupré, M., Lévesque, M., & Hess, U. (2007). Toward a dialect theory: Cultural differences in the expression and recognition of posed facial expressions. *Emotion, 7*(1), 131–146.

Frick, R. W. (1985). Communicating emotion: The role of prosodic features. *Psychological Bulletin, 97*(3), 412–429.

Hammerschmidt, K., & Jürgens, U. (2007). Acoustical correlates of affective prosody. *Journal of Voice, 21*(5), 531–540.

Hawk, S. T., Van Kleef , G. A., Fischer A. H., & Van der Schalk J. (2009). "Worth a thousand words": absolute and

relative decoding of nonlinguistic affect vocalizations. *Emotion , 9*, 293–305.

Hunt, W. A. (1941). Recent developments in the field of emotion. *Psychological Bulletin, 38*(5), 249–276.

Gobl, C., & Ní Chasaide, A. (2003) The role of voice quality in communicating emotion, mood and attitude. *Speech Communication*, 40, 189–212.

Izard, C. E. (1971). *The face of emotion*. New York, NY: Appleton-Century-Crofts.

Jack, R. E., Garrod, O. G. B., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences of the United States of America, 109*(19), 7241–7244.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129*(5), 770–814.

Juslin, P. N., & Scherer, K. R. (2008). Speech emotion analysis. *Scholarpedia, 3*(10), 4240.

Jürgens, R., Drolet, M., Pirow, R., Scheiner, E., & Fischer, J. (2013). Encoding conditions affect recognition of vocally expressed emotions across cultures. *Frontiers in Psychology, 4*, 111.

Jürgens R., Hammerschmidt K., & Fischer J. (2011). Authentic and play-acted vocal emotion Expressions reveal acoustic differences. Frontiers in Emotion Science, 2,180.

Kitayama, S., & Ishii, K. (2002). Word and voice: Spontaneous attention to emotional utterances in two languages. *Cognition & Emotion, 16*(1), 29–59.

Kramer, E. (1964). Elimination of verbal cues in judgments of emotion from voice. *The Journal of Abnormal and Social Psychology, 68*(4), 390–396.

Laukka, P., Audibert, N., & Aubergé, V. (2012). Exploring the determinants of the graded structure of vocal emotion expressions. *Cognition & Emotion, 26*, 710–719.

Laukka, P., Elfenbein, H. A., Söder, N., Nordström, H., Althoff, J., Chui, W., & Thingujam, N. S. (2013). Cross-cultural decoding of positive and negative non-linguistic emotion vocalizations. *Frontiers in Psychology, 4*, 353.

Leppaenen, J. M., & Hietanen, J. K. (2001). Emotion recognition and social adjustment in school-aged girls and boys. *Scandinavian Journal of Psychology, 42*, 429–435.

Lima, C. F., Castro, S. L., & Scott, S. K. (2013). When voices get emotional: A corpus of nonverbal vocalizations for research on emotion processing. *Behavior Research Methods, 45*, 1234–1245.

Matsumoto, D. (2002). Methodological requirements to test a possible in-group advantage in judging emotions across cultures: Comment on Elfenbein and Ambady (2002) and evidence. *Psychological Bulletin, 128*, 236–242.

Matsumoto, D., LeRoux, J. A., Wilson-Cohn, C., Raroque, J., Kooken, K., Ekman, P., ... Goh, A. (2000). A new test to measure emotion recognition ability: Matsumoto and Ekman's Japanese and Caucasian Brief Affect Recognition Test (JACBART). *Journal of Nonverbal Behavior, 24*, 179–209.

Matsumoto, D., Olide, A., & Willingham, B. (2009). Is there an ingroup advantage in recognizing spontaneously expressed emotions? *Journal of Nonverbal Behavior, 33*(3), 181–191.

McCluskey, K. W., Albas, D. C., Niemi, R. R., Cuevas, C., & Ferrer, C. A. (1975). Cross-cultural differences in the perception of the emotional content of speech: A study of the development of sensitivity in Canadian and Mexican children. *Developmental Psychology, 11*(5), 551–555.

McCluskey, K. W., & Albas, D. C. (1981). Perception of the emotional content of speech by Canadian and Mexican children, adolescents, and adults. *International Journal of Psychology, 16*(1-4), 119–132.

Mesquita, B., & Frijda, N. H. (1992). Cultural variations in emotions: A review. *Psychological Bulletin, 112*(2), 179–204.

Mill, J. (1984). High and low self-monitoring individuals: Their decoding skills and empathic expression. *Journal of Personality, 52*, 815–822.

Monrad-Kohn, G. H. (1947). The prosodic quality of speech and its disorders (a) brief survey from a neurologist's point of view. *Acta Psychiatrica Neurologica Scandinavica, 22*, 255–269.

Paulmann, S., & Uskul, A. K. (2013). Cross-cultural emotional prosody recognition: Evidence from Chinese and British listeners. *Cognition & Emotion, 28*, 230–244.

Pell, M. D., Jaywant, A., Monetta, L., & Kotz, S. A. (2011). Emotional speech processing: Disentangling the effects of prosody and semantic cues. *Cognition and Emotion, 25*, 834–853.

Pell, M. D., Monetta, L., Paulmann, S., & Kotz, S. A. (2009). Recognizing emotions in a foreign language. *Journal of Nonverbal Behavior, 33*(2), 107–120.

Pell, M. D., Paulmann, S., Dara, C., Alasseri, A., & Kotz, S. A. (2009). Factors in the recognition of vocally expressed emotions: A comparison of four languages. *Journal of Phonetics, 37*(4), 417–435.

Pell, M. D., & Skorup, V. (2008). Implicit processing of emotional prosody in a foreign versus native language. *Speech Communication, 50*(6), 519–530.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin, 115*, 102–141.

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the*

*National Academy of Sciences of the United States of America, 107*(6), 2408–2412.

Sauter, D. A., & Scott, S. K. (2007). More than one kind of happiness: Can we recognize vocal expressions of different positive states? *Motivation and Emotion, 31*(3), 192–199.

Scheiner, E., & Fisher, J. (2011). Emotion expression: The evolutionary heritage in the human voice. In *Interdisciplinary Anthropology* (pp. 105–129). Heidelberg & New York: Springer.

Scherer. F. M. (1991). Changing Perspectives on the Firm Size Problem. in Zoltan J.Acs and David B.Audretsch (eds.). *Innovation and Technological Change: An International Comparison* (pp. 24–38*),* Ann Arbor: University of Michigan Press,

Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, *99*, 143–165.

Scherer, K. R. (2013). Vocal markers of emotion: comparing induction and acting elicitation. *Computer Speech and Language, 27*, 40–58.

Scherer, K. R., Banse, R., Wallbott, H. G., & Goldbeck, T. (1991). Vocal cues in emotion encoding and decoding. *Motivation and Emotion, 15*(2), 123–148.

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology, 32*(1), 76–92.

Schirmer, A., & Kotz, S. A. (2003). ERP evidence for a sex-specific Stroop effect in emotional speech. *Journal of Cognitive Neuroscience, 15*(8), 1135–1148.

Schröder, M. (2003). Experimental s tudy of affect bursts. *Speech Commun, 40*, 99–116.

Scott, S. K., Sauter, D., & McGettigan, C. (2009). Brain mechanisms for processing perceived emotional vocalizations in humans. In Brudzynski SM (Eds.), *Handbook of Mammalian Vocalizations* (pp 187– 198). Academic Press, Oxford.

Shiota, M. N., Campos, B., Keltner, D., & Hertenstein, M. J. (2004). Positive emotion and the regulation of interpersonal relationships. In P. Philippot & R. S. Feldman (Eds.), *The regulation of emotion* (pp. 127–155). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Simon-Thomas,E.R., Keltner, D. J., Sauter, D., Sinicropi-Yao, L., & Abramson, A. (2009). The voice conveys specific emotions: Evidence from vocal burst displays. *Emotion, 9* , 838–846.

Thompson, W. F., & Balkwill, L. L. (2006). Decoding speech prosody in five languages. *Semiotica, 158*, 407–424.

Tracy, J. L., & Robins, R. W. (2008). The nonverbal expression of pride: Evidence for cross-cultural recognition. *Journal of Personality and Social Psychology*, *94*(3), 516–530.

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., ... Tzourio-Mazoyera, N. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *Neuroimage, 30*(4), 1414–1432.

Wickline, V. B., Bailey, W., & Nowicki, S. (2009). Cultural in-group advantage: Emotion recognition in African American and European American faces and voices. *The Journal of Genetic Psychology, 170*(1), 5–29. NeuroImage 15, 856–869

Wildgruber, D., Pihan, H., Ackermann, H., Erb, M., & Grodd., W. (2002). Dynamic brain activation during processing of emotional intonation: influence of acoustic parameters, emotional valence and sex. *NeuroImage, 15*, 856–869.

Wilting, J., Krahmer, E., & Swerts, M. (2006). Real vs. acted emotional speech. In *Proceedings of the Ninth International Conference on Spoken Language Processing* (pp. 805–808). Pittsburgh, PA, USA.

Zhao, W. J. (2013). *A study on the production and perception of English emotional prosody by mandarin-English bilingual adults* (Unpublished master's thesis). The Chinese University of Hong Kong.

## The Recognition of Emotional Vocally Across Cultures

SUN Jimin; YAO Wenqing; HU Xueping

(*School of Psychology, Northwest Normal University, Lanzhou* 730070, *China*)

**Abstract:** The recognition of vocal emotions across cultures means understand the emotional information through voice in cross-culture situations. Whereas existing researches mainly focus on: cross-cultural consistency; the effect of in-group advantage; the study of gender and age; the recognition of spontaneous voice emotions; the selection of language materials. So our future research should further explore the aspects of materials obtaining, the influence factors and bilingualism.

**Key words:** vocally emotions; cross-cultural study; in-group advantage effect; spontaneous emotions