

Recognition of Emotional Prosody in the Proficient and the Non-Proficient Chinese-English Bilinguals

Mi Tang^{1,2}, Tong Ye², Ming Gu², Huijing Yu^{2,3}

¹Faculty of Psychology, Southwest University, Chongqing

²College of Psychology, Liaoning Normal University, Dalian Liaoning

³Faculty of Psychology, Tianjin Normal University, Tianjin

Email: demi_tang@foxmail.com

Received: Feb. 4th, 2020; accepted: Feb. 19th, 2020; published: Feb. 26th, 2020

Abstract

When we are communicating with others, it is important to perceive the emotion contained in the language correctly. The present study put emotional prosody recognition and language proficiency together, aiming at investigating whether language proficiency could affect the recognition of emotional prosody. To this end, some Chinese and English pseudo sentences were pronounced in 7 emotions (anger, disgust, fear, happiness, neutral, pleasant surprise, and sadness) by a native Chinese speaker and a native English speaker, respectively. These auditory stimuli were then presented to 20 proficient Chinese-English bilinguals and 20 non-proficient Chinese-English bilinguals. Recognition accuracy for each emotion was made into a confusion matrix. With the arcsine-transformed unbiased hit rate (*Hu* scores) serving as a dependent variable, a three-way repeated measure ANOVA was also performed. It is revealed by the confusion matrix that emotions shared the same valence were usually confounded by the participants. The results of repeated measures ANOVAs presented that except for the factor of language proficiency, which only had a marginally significant main effect, all the independent variables showed significant main effects. Also, there was a significant two-way interaction effect between emotions and language of materials. No other interaction effect was significant. In conclusion, language proficiency does not exist as an influence factor for emotional prosody recognition, which might be explained by the fact that recognition of emotional prosody can hardly be affected by the level of language ability we gained by nurture.

Keywords

Language Proficiency, Bilingual, Emotional Prosody, Prosody

熟练与非熟练中英双语者的情绪韵律识别

汤 宓^{1,2}, 叶 童², 顾 铭², 于慧婧^{2,3}

¹西南大学心理学部, 重庆

²辽宁师范大学心理学院, 辽宁 大连

³天津师范大学心理学部, 天津

Email: demi_tang@foxmail.com

收稿日期: 2020年2月4日; 录用日期: 2020年2月19日; 发布日期: 2020年2月26日

摘要

正确地感知口语中蕴含的情绪对言语交际至关重要。本研究将影响言语交流情绪感知的重要因素——情绪韵律, 与语言熟练度结合, 探究语言熟练度是否为影响情绪韵律识别的因素。为此, 熟练中英双语者与非熟练中英双语者分别被要求在听到了蕴含7种情绪之一的中英文假句后, 做出情绪分类判断。基于情绪识别正确率制作的混淆矩阵显示, 相同效价各情绪易被混淆。以经反正弦转换的无偏击中率(*Hu*分数)为因变量的方差分析结果表明, 虽然语言熟练度主效应边缘显著, 但该因素与其他自变量的交互作用均不显著。由此可推测, 语言熟练度不作为影响双语者情绪韵律识别的因素而存在, 人们识别第二语言情绪韵律的能力在发展过程中保持稳定趋势, 难以受到其他因素影响。

关键词

语言熟练度, 双语, 情绪韵律, 韵律

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在交流中能否正确地感知说话者蕴含在言语中的情绪情感, 能够直接地影响到交流的质量。例如, 假设我们从未学过英语, 当我们在街头遇到一名外国人向我们用英语焦急地说些什么时, 我们从他的语速、音调、声音的大小中或许就能推断出, 他正急着要去某个地方, 但却找不到路在哪儿。在这一过程中, 听话人对说话人声音中所带情绪的判断扮演着重要的角色。而这一判断, 实际上就是对于情绪韵律的识别。

韵律特征又被称为“超音段特征”。韵律特征是语言的一种音系结构, 与句法和语篇结构、信息结构等其他语言学结构密切相关。韵律特征可以分为三个主要方面: 语调、时域分布和重音, 这些韵律特征的方面就通过超音段特征实现。而超音段特征包括声音的音高, 强度以及时长。韵律是人类自然语言的一个典型特征, 也是语言和情绪表达的重要形式之一。

情绪韵律相比韵律特征, 则更加强了其中的情绪表达功能。在我们的口语表达中, 一些诸如声调、重音等的韵律特征也包含了一定的情绪线索(Min & Schirmer, 2011)。在语言理解中, 共有三个水平。其中初级水平是对单词的理解, 因为单词是组成语言材料的最小意义单位; 较高级水平是对短语和句子的理解, 从这一层次开始, 便涉及了对言语意义、语法关系的理解; 第三级水平是对说话人意图或动机的理解, 即明了说话人的真正的想法(黄希庭, 2011)。借助于情绪韵律提供的情绪线索, 语言理解才得以顺利进行, 尤其对于第三级水平的语言理解而言。

曾有研究发现,为了与其他个体更方便地进行交流,人们会从带有情绪的语音中感知声学线索,从而提取情绪意义(Bostanov & Kotchoubey, 2004)。由此不难推测,在交流的过程中,听话人会参与到对语音、节奏和情感韵律变化的感知中,并用一定的语言解码方式来识别语言中的情感信息。这种不考虑语言的词汇和语法结构,而只是通过重音和语调的变化来传达说话者情绪的韵律特征,即被称为情绪韵律(姚温青, 2014)。

在关于情绪韵律的众多研究中,语料是一个被赋予大量关注的因素。有的实验使用了中性语义句子或词作为实验材料(Min & Schirmer, 2011; Paulmann, Furnes, Bøkenes, & Cozzolino, 2016; 黄贤军&张伟欣, 2014)。然而,考虑到在句子水平上,人们对语义的理解与对情绪韵律的识别存在交互作用。也就是说,语义会制约情绪韵律加工(Kotz & Paulmann, 2007)。出于此原因,有的实验便使用了假词、假句或经过过滤处理的词句声音材料进行研究(Dimoska, McDonald, Pell, Tate, & James, 2010; Uskul, Paulmann, & Weick, 2016)。以上语料的选取各有其利弊。使用具有语义的句子更有生态效度,但却避免不了语义对情绪韵律的干扰。去除语义的句子虽能更直观地体现情绪韵律,但也令人不免怀疑其真实性。除了词和句子外,还有实验会直接使用纯粹的情绪发声作为材料,如叹息声、欢呼声、哭泣声(Bostanov & Kotchoubey, 2004; Lima, Castro, & Scott, 2013)。不过使用这一材料的研究相对较少,目前得出的结论与其他使用完整词语、句子的实验结论都基本一致(Hawk, van Kleef, Fischer, & van der Schalk, 2009)。

此外,录制声音材料的录音者使用的是自然的自发声,还是刻意表达情绪的表演声,也会产生一定影响。有研究发现用自发声录制的声音材料其唤醒程度会更高(Bachorowski, 1999)。其内在的原因或许是,自发声由“产出优先”系统主导,这种系统与生理线索的改变有关;而表演声则由“感知优先”的系统主导(Owren, Amoss, & Rendall, 2011)。

跨文化情绪韵律识别则是该领域中的另一个研究热点。目前共有三种实验设计方法用以研究跨文化的情绪韵律识别:非平衡设计,不同文化下的被试识别某一种文化中的情绪韵律表达;反平衡设计,某一种文化下的被试识别不同文化中的情绪韵律表达;平衡设计,不同文化下的被试识别所有这些不同文化中的情绪韵律表达。

Bezooijen, Otto, and Heenan (1983)曾使用荷兰语短句材料对跨文化情绪韵律识别做出了研究。40名来自台湾、荷兰和日本的被试都能高于猜测水平识别情绪韵律,同时实验也发现了群内优势效应:荷兰被试在识别荷兰语的情绪韵律时,显著地比台湾和日本被试正确率更改。此研究便是一个使用了非平衡设计的范例。关于反平衡设计,Thompson and Balkwill (2006)曾做过一个实验,探究英语母语者能在多大程度上识别德语、英语、汉语、日语和他加禄语的情绪韵律。结果发现,尽管被试基本上都能成功识别其他非母语的情绪韵律,但是相比对其他语言情绪韵律的识别,当被试在识别英语的情绪韵律时有显著更高的正确率。使用非平衡设计与反平衡设计的实验得出的结论均具有较高的一致性:人们在识别不同文化中的情绪韵律时,会使用一种普遍性原则进行判断。同时,大部分的研究也发现了群内优势效应。即相比起其他文化,人们更擅长识别本民族文化中的情绪韵律。但长期以来,由于没有完全匹配情绪韵律表达者与感知者所处的文化之间的关系,非平衡设计与反平衡设计常受到“对跨文化情绪韵律识别探究不全面”的批判。因此,一种完全交叉设计,即平衡设计诞生了。Paulmann and Uskul (2013)将中文和英文的情绪韵律材料呈现给中国和英国的被试,要求其材料中的情绪韵律进行识别。通过对识别的错误率和正确率的分析,实验者得出了与非平衡设计与反平衡设计实验相一致的结论:被试在表现出跨文化情绪韵律识别的普遍性的同时,也表现出了对本文化情绪韵律的识别率显著更高的特殊性。

目前已有研究得出,熟练双语者在识别第二语言情绪词汇时的,表现水平与英语母语者一致(Ponari et al., 2015)。在口语韵律产出的相关研究方面,Chakraborty (2011)记录了熟练与非熟练孟加拉-英双语者在阅读孟加拉语和英语词汇时的唇部运动变化,发现第二语言的熟练程度不会对被试的第一语言和第二

语言的韵律产出产生影响。而关于语言的熟练程度是否也能影响情绪韵律的识别这一问题，却很少有研究涉及。

针对此研究空白，本研究旨在通过对比熟练与非熟练中英双语者在识别情绪韵律准确率上的差别，考察第二语言的熟练程度是否为情绪韵律识别的影响因素。本研究将探讨除目前已知的种种因素外，语言熟练度是否也会影响情绪韵律识别。当前国内对情绪韵律的研究较为滞后，从双语和语言熟练度角度开展的研究更为稀少。因此，本研究应能在深入探索情绪韵律的同时，为这一研究领域带来新的思考。

2. 方法

2.1. 预实验

预实验的目的为对声音材料进行筛选。邀请中文和英文母语被试对以各自的母语表达的声音情绪进行判断评价，从而保证接下来的正式实验中使用的实验材料，在情绪表达上具有一定的效度。

2.1.1. 被试

25名中文母语者参与了预实验(女性20名，男性5名)。被试的平均年龄为22.52岁，右利手，听力完好，视力或矫正视力正常。预实验结束后，每名被试获得了报酬。

2.1.2. 实验材料

本实验中使用的 all 情绪声音材料均来自于 Paulmann and Uskul (2013)录制的实验材料。预实验中被评价的中文声音材料的内容为40个中文假句(如：宁温鸡吐不不非糖哈)，由2名以中文为母语的被试制作。假句的句长与词长均已经过控制。1位具有表演经验的中文母语者用7种情绪(愤怒、厌恶、恐惧、愉快、中性、惊喜和悲伤)将假句朗读出来。虽然关于情绪的分类问题一直存有争议，但是根据 Ortony and Turner (1990)的统计，在众多情绪分类中有4种情绪最为普遍，即恐惧、愤怒、悲伤和愉快，其次便是厌恶和惊喜。因此，本研究使用了这6种最基本的独立情绪(以及作为基线参照的中性情绪)作为研究内容，最终生成了280个中文声音刺激(40个假句 × 7种情绪)。另外，Paulmann and Uskul (2013)通过同样的方法，还录制了一套以英文假句(如：Flotchderaded the downdarysnat)为内容的情绪声音材料。但是考虑到这些英文材料已经经过了英文母语者的评价(31名英语母语者，其中21名女性，10名男性，平均年龄24.09岁)，并且在中国很难找到较大的英语母语者被试样本，所以在预实验中没有再次对英文材料进行评价。

2.1.3. 实验程序

如图1所示，在预实验程序中，被试首先在电脑屏幕中央看到一个注视点“+”，持续1000ms。接着耳机中播放一句语音(假句)，声音的持续时间为1~6秒。被试需判断声音中蕴含的情绪，并用鼠标从7个情绪选项中选择自己所判断的情绪类别。提供给被试的情绪选项均为中文。被试在进行预实验前被告知，语音所说的内容是没有意义的，且每段语音只呈现一次，因此请又快又好地做出判断。实验程序通过E-Prime 2.0运行。为了避免被试过于疲劳，整个预实验被分为4块，共穿插了3次休息。预实验的持续时间约为30分钟。

2.1.4. 结果

以被试对每句语音做出判断的正确率为指标，识别正确率高于猜测水平14.2%三倍(即42.6%)的声音刺激均被筛选为合格的刺激。按照此标准，并在平衡7种情绪的前提下，最后共筛选出了101个中文声音刺激作为正式实验材料。中文材料各情绪的识别正确率为：愤怒 = 92%，厌恶 = 66%，恐惧 = 63%，愉快 = 82%，中性 = 100%，惊喜 = 80%，悲伤 = 93%。为了与中文材料相匹配，从已有的196个英文声音刺激中，随机筛选出了105个英文声音刺激为正式实验材料。原196个英文材料中，各情绪的识别

正确率为：愤怒 = 91%，厌恶 = 83%，恐惧 = 64%，愉快 = 55%，中性 = 91%，惊喜 = 69%，悲伤 = 91%。

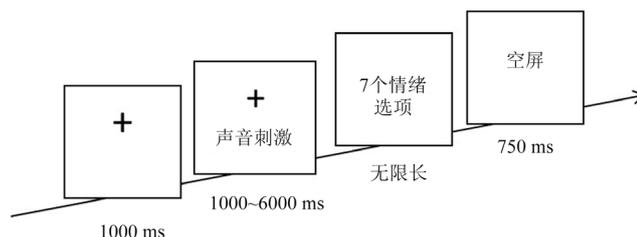


Figure 1. The procedure of the pre-experiment
图 1. 预实验流程图

使用 Praat 软件对被筛选出的声音刺激进行初步分析后，得到了如表 1 中所示的声学参数信息。为了明确这些声音材料的声学参数信息是否能够有效地预测其中蕴含的情绪，这些信息数据接受了一次 SPSS 中的判别分析。判别分析以声音的时长、基频、响度数据为自变量，7 种不同的情绪为因变量。结果显示，对于所有被筛选的中文材料而言，71.3% 的材料可被其声学参数信息正确预测。其中愤怒 = 86.7%，厌恶 = 69.2%，恐惧 = 61.5%，愉快 = 53.3%，中性 = 80%，惊喜 = 66.7%，悲伤 = 80%；所有被筛选的英文声音材料中，共有 65.7% 被正确预测：愤怒 = 86.7%，厌恶 = 46.7%，恐惧 = 73.3%，愉快 = 46.7%，中性 = 80%，惊喜 = 33.3%，悲伤 = 93.3%。综合被试判断的正确率及判别分析的结果，可以说明预实验中筛选的材料具有良好的效度。尽管英文材料中的惊喜情绪在判别分析中，没有达到前文规定的标准，但也高于了两倍的猜测水平。另外考虑到判别分析的效度并不如被试的正确率判断，且英文惊喜情绪的正确率已达到标准，因此所有的被筛选材料均被使用在了正式实验中。

Table 1. The acoustic parameters of the qualified emotional voices

表 1. 合格的情绪声音材料的基本声学参数信息

材料	情绪	平均时长(ms)	平均基频(Hz)	平均响度(dB)
中文	愤怒	1923.70	337.23	74.12
	厌恶	3231.90	254.87	63.00
	恐惧	3242.50	305.19	61.56
	愉快	2419.30	301.72	66.78
	中性	2903.30	220.55	56.84
	惊喜	2239.40	347.88	71.52
	悲伤	4998.10	286.99	62.81
英文	愤怒	2213.30	253.57	70.14
	厌恶	2579.50	225.55	63.28
	恐惧	2137.80	302.59	65.76
	愉快	2037.70	254.89	65.85
	中性	2270.00	202.04	62.84
	惊喜	1985.70	280.77	66.78
	悲伤	2294.70	252.88	62.47

2.2. 正式实验

2.2.1. 被试

共 40 名被试(女性 38 名，男性 2 名)参与了正式实验，平均年龄为 21.38 岁。其中熟练双语者 20 名，

平均年龄 21.2 岁；非熟练双语者 20 名(女性 18 名, 男性 2 名), 平均年龄 21.55 岁。熟练双语者均为英语专业, 且已通过专业英语四级(TEM-4)的学生(卢植&涂柳, 2010)。开始学习英语的平均年龄为 9.67 岁；非熟练双语者为非英语专业, 且未通过大学英语四级(CET-4)的学生(毛伟宾, 杨治良, 王林松, &袁建伟, 2008)。开始学习英语的平均年龄为 8.76 岁。所有被试的母语为中文, 英语是其掌握的第二语言。所有参与正式实验的被试均为右利手, 听力完好, 视力或矫正视力正常。被试在实验结束后获得少量报酬。

2.2.2. 实验材料

正式实验中使用的实验材料即预实验中筛选出的 101 个中文声音刺激, 以及 105 个英文声音刺激。中文和英文的声音所表达的 7 种情绪已经过了平衡控制。各声音刺激的情绪识别率均在随机水平以上, 因此具有良好的效度。详细数据信息见 2.1.4。

2.2.3. 实验程序

正式实验的实验程序与预实验使用的程序一致(图 1), 由 E-Prime 2.0 运行。唯一的一点不同在于程序所呈现的实验材料。预实验的目的为评定中文声音材料, 所以程序中呈现了未经筛选的 280 个中文声音刺激。而正式实验中, 为了探究熟练与非熟练双语者对于中文和英文情绪韵律识别的特点, 呈现给每一个被试的则是经过筛选的 101 个中文声音刺激和 105 个英文声音刺激。被试同样需要在听到每个声音刺激后, 在 7 个情绪选项中选择自己感知到的情绪。实验开始前, 为了避免被试对实验目的进行猜测, 被试会被告知他们的英语水平与实验没有很大关系。只是因为实验中使用了中英文的材料, 所以需要了解他们的英语水平。实验结束后, 被试将被告知真正的实验目的。

2.2.4. 结果

表 2 为正式实验中被试对每一种情绪做判断的正确率与错误率。从表中可看出熟练双语者对所有材料的识别正确率介于 36%~100%, 非熟练双语者对所有材料的识别正确率在 38%~100%之间。中文材料被正确识别的概率在 50%~100%之间, 英文材料被正确识别的概率为 36%~90%。

从识别的错误率中可以看出, 被试在识别中文句子时, 倾向于将厌恶和恐惧情绪识别为悲伤, 也有倾向将愤怒情绪感知为厌恶。在识别英文句子时, 被试产生了更多的混淆: 被试会倾向于将愤怒判断为厌恶, 惊喜判断为愉快。而在在恐惧、悲伤两种情绪之间, 则容易产生相互混淆。另外, 被试也有一定的可能性将中性情绪与悲伤、愉快、厌恶混淆起来。

对比熟练与非熟练双语者的结果发现, 两类被试识别中英文情绪韵律的模式大致相同。两类被试在识别中文时犯的的错误均少于英文, 且对于所有情绪的识别均高于猜测水平。但对于熟练双语者而言, 出现了一个意料之外的结果: 熟练双语者将英文的愉快情绪混淆为中性情绪的概率高于其识别愉快情绪的正确率。但配对样本 t 检验的结果显示, 两者之间差异不显著 $t(19) = -1.613, p = 0.123$ 。因此熟练双语者在实验中将愉快错认为中性的概率并不高, 或许他们只是将英文的愉快和中性情绪相互混淆了起来。而非熟练双语者识别英文愉快情绪的正确率均高于其他错误率。

综合混淆矩阵的结果可以发现, 在 7 种情绪种, 被试在 4 种消极情绪(愤怒、厌恶、恐惧、悲伤)之间经常存在混淆。同样的特点也存在于 2 种积极情绪(愉快、惊喜)之间。而对于中性情绪的识别, 往往都是最好的。这与被试在实验结束后的口头反馈一致。

Wagner (1993)曾提出, 在以类别判断为任务的实验研究中, 单纯的正确率并不是一个好的因变量指标, 它受到的误差影响使得其不能完全代表被试的行为水平。如此, 作为比正确率更有效的指标, Hu 分数(无偏击中率, unbiased hit rate)被提出。但是若要将 Hu 分数用于推论统计分析, 还需将其进行反正弦转换。

Table 2. The confusion matrix of the recognition accuracy of each emotion
表 2. 各情绪识别正确率的混淆矩阵

被试	材料	情绪	反应						
			愤怒	厌恶	恐惧	愉快	中性	惊喜	悲伤
熟练双语者	中文	愤怒	0.92	0.08	0.00	0.00	0.00	0.00	0.00
		厌恶	0.07	0.50	0.07	0.03	0.08	0.03	0.22
		恐惧	0.00	0.02	0.62	0.01	0.02	0.02	0.31
		愉快	0.00	0.04	0.00	0.79	0.06	0.09	0.01
		中性	0.00	0.00	0.00	0.00	1.00	0.00	0.00
		惊喜	0.07	0.01	0.02	0.04	0.00	0.86	0.01
		悲伤	0.00	0.00	0.00	0.00	0.00	0.00	0.99
	英文	愤怒	0.79	0.16	0.03	0.00	0.01	0.00	0.00
		厌恶	0.04	0.60	0.04	0.10	0.10	0.05	0.07
		恐惧	0.04	0.05	0.48	0.10	0.08	0.08	0.16
		愉快	0.00	0.02	0.02	0.36	0.46	0.09	0.04
		中性	0.00	0.05	0.00	0.02	0.90	0.00	0.03
		惊喜	0.00	0.01	0.00	0.33	0.01	0.63	0.00
		悲伤	0.00	0.00	0.11	0.00	0.15	0.00	0.73
非熟练双语者	中文	愤怒	0.84	0.14	0.00	0.00	0.00	0.01	0.00
		厌恶	0.08	0.51	0.11	0.02	0.02	0.02	0.24
		恐惧	0.00	0.02	0.60	0.00	0.01	0.02	0.33
		愉快	0.02	0.03	0.01	0.74	0.10	0.10	0.00
		中性	0.00	0.00	0.00	0.00	1.00	0.00	0.00
		惊喜	0.10	0.03	0.02	0.05	0.00	0.78	0.01
		悲伤	0.01	0.00	0.07	0.01	0.00	0.00	0.92
	英文	愤怒	0.73	0.21	0.02	0.01	0.01	0.01	0.01
		厌恶	0.08	0.55	0.06	0.05	0.16	0.04	0.06
		恐惧	0.05	0.06	0.38	0.10	0.10	0.07	0.24
		愉快	0.01	0.07	0.04	0.39	0.04	0.10	0.05
		中性	0.00	0.15	0.01	0.03	0.78	0.00	0.02
		惊喜	0.04	0.06	0.01	0.30	0.00	0.59	0.01
		悲伤	0.00	0.02	0.12	0.02	0.11	0.01	0.71

注：表中加粗字体为正确率，未加粗字体为错误率。

Table 3. The confusion matrix for calculating the unbiased hit rate
表 3. 计算 *Hu* 分数(无偏击中率, unbiased hit rate)的混淆矩阵

刺激	反应			总数
	1	2	3	
1	a	b	c	a + b + c
2	d	e	f	d + e + f
3	g	h	i	g + h + i
总数	a + d + g	b + e + h	c + f + j	N

注：表中字母均为频数，其中 $N = a + b + c + d + e + f + g + h + i$ 。

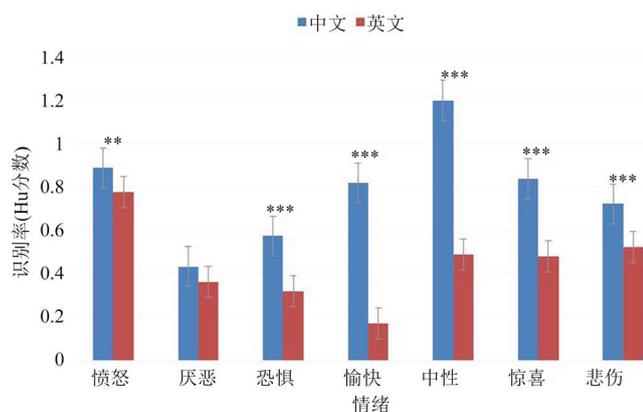
根据表 3 中的数据, 对于其中的刺激 1, Hu 分数的计算公式如下:

$$Hu = \frac{a}{a+b+c} \times \frac{a}{a+d+g}$$

以语言熟练度(熟练双语, 非熟练双语)为被试间自变量, 材料语言(中文, 英文)和情绪(愤怒、厌恶、恐惧、愉快、中性、惊喜和悲伤)为被试内自变量, 以经过反正弦转换的 Hu 分数为因变量, 三因素重复测量方差分析的结果显示, 材料语言自变量的主效应显著 $F(1,38) = 97.83, p < 0.001, \eta_p^2 = 0.759$, 被试对中文情绪韵律的识别($M = 0.78, SD = 0.03$)显著地比对英文的识别($M = 0.44, SD = 0.02$)更准确。情绪自变量的主效应显著 $F(6,228) = 61.836, p < 0.001, \eta_p^2 = 0.666$, 表明某些情绪类别在实验中被更好地识别。其中中性情绪($M = 0.84, SD = 0.03$)和愤怒情绪($M = 0.81, SD = 0.03$)的识别率最高, 其次便是惊喜($M = 0.66, SD = 0.03$), 悲伤($M = 0.62, SD = 0.02$), 愉快($M = 0.50, SD = 0.03$), 恐惧($M = 0.45, SD = 0.03$), 厌恶 ($M = 0.40, SD = 0.02$)。成对比较的结果显示, 除了愤怒与中性、恐惧与厌恶、恐惧与愉快、惊喜与悲伤情绪之间识别率差异不显著, 其他情绪之间的识别率均有显著差别 $ps < 0.001$ 。作为被试间自变量的语言熟练度虽然在 5% 的显著性水平上主效应不显著, 但根据其 p 值, 也可被判定为边缘显著 $F(1,38) = 3.55, p = 0.069, \eta_p^2 = 0.103$ 。说明熟练与非熟练中英双语者在识别中英文情绪韵律上, 存在一定差别。熟练中英双语者的表现($M = 0.64, SD = 0.03$)略好于非熟练中英双语者($M = 0.58, SD = 0.02$)。

材料语言和情绪的交互作用显著 $F(6,228) = 61.836, p < 0.001, \eta_p^2 = 0.502$, 表明被试分别在识别中文和英文的某些情绪韵律的表达上, 存在显著差别(图 2)。针对此交互作用的简单效应分析发现, 除了厌恶情绪 $F(1,39) = 1.63, p = 0.211$, 对于其他 6 种情绪, 被试对中文的识别率均显著高于英文(愤怒 $F(1,39) = 9.69, p < 0.05$; 恐惧 $F(1,39) = 28.57, p < 0.001$; 愉快 $F(1,39) = 166.02, p < 0.001$; 中性 $F(1,39) = 186.20, p < 0.001$; 惊喜 $F(1,39) = 38.97, p < 0.001$; 悲伤 $F(1,39) = 17.91, p < 0.001$)。除此之外, 语言熟练度与材料语言之间的交互作用, 语言熟练度与情绪之间的交互作用, 以及语言熟练度、材料语言与情绪之间的三因素交互作用均不显著 $Fs \leq 1.227, ps > 0.05$ 。

为了确定被试开始学习第二语言的年龄是否对实验结果产生了干扰, 在进行重复测量方差分析的同时, 也以初学第二语言的年龄为自变量, 进行了一次协方差分析。协方差分析的结果表明, 被试初学第二语言的年龄不影响本实验的结果。年龄的主效应不显著 $F(1,37) = 0.27, p = 0.609$; 年龄与其他自变量的交互作用也均不显著 $Fs \leq 1.2, ps > 0.05$ 。



注: 图中*表示 $p < 0.05$, **表示 $p < 0.01$, ***表示 $p < 0.001$ 。

Figure 2. The analysis of simple effect for the interaction between language of materials and emotion

图 2. 材料语言与情绪交互作用的简单效应分析

3. 讨论

本研究的主要目的为探索语言的熟练程度是否为影响情绪韵律识别的因素。语言的熟练程度通过控制被试的英文水平实现。对情绪韵律的识别体现为，被试在听到用特定情绪表达出来的中文和英文假句后，做出判断的 *Hu* 分数。

首先，本研究以对每个情绪声音做出判断的正确率为指标，分析了被试反应的混淆矩阵。从混淆矩阵中呈现的数据模式发现，被试对情绪韵律的判断很明显受到了情绪效价的影响。即同一效价内情绪较易被混淆，主要表现为：中文的厌恶和恐惧情绪容易被感知为悲伤；英文中愤怒易被判断为厌恶，惊喜易被判断为愉快。如此的识别模式，或许与被混淆的情绪间存在着类似的声学参数模式有关。已有研究证明过，声音刺激的物理模式会在一定程度上增强或减弱情绪韵律识别的效果(Cornew, Carver, & Love, 2009)。从表 1 呈现的情绪声音声学参数信息可以发现，同一效价的几种情绪，其时长、基频、响度之间的差别不大。为进一步证实这一观点，以表 1 中的数据为基础，将情绪作为自变量、三种声学参数分别作为因变量，进行单因素方差分析后，事后分析的结果表明：在中文情绪中，易被混淆的几种情绪除了在时长上存在显著差异外 $ps < 0.05$ ，在基频和响度上的差异均不显著 $ps > 0.05$ ；在英文情绪中，除了愤怒和厌恶、恐惧和悲伤这两对情绪在基频和响度上存在显著差异外 $ps < 0.01$ ，其他几种易混淆情绪在三个指标上均无显著差异 $ps > 0.05$ 。此结果可以说明，声音情绪刺激的声学参数模式相近，可以部分地解释被试在识别这些刺激时表现出的混淆偏向。

其次，根据本实验的方差分析结果，可以推测语言的熟练程度并不是一个能够影响到情绪韵律识别的因素。不仅自变量的主效应不显著，语言熟练度与其他自变量的交互作用也不显著。虽然将语言熟练程度与情绪韵律识别结合的研究并不多，但是根据一些相关研究，本研究的结果与前人的探究结论基本一致。Chakraborty (2011)曾对第二语言的熟练程度是否会影响到韵律产出进行过研究，结果发现第二语言的熟练程度不会影响韵律产出。Min and Schirmer (2011)通过研究发现，第二语言言语理解和情绪韵律的整合与母语的情绪韵律整合无发展先后的差异，且这种言语整合也不会随着熟练度而变化。本研究则对语言熟练程度和情绪韵律识别的直接关系进行研究，也没有发现语言熟练程度在情绪韵律识别中的影响。此外，又考虑到熟练与非熟练中英双语者开始学习第二语言的年龄大致相同 $t(38) = 1.139, p > 0.05$ ，因此可以推测，在语言习得的早期，就如同对母语情绪韵律的掌握，人们识别第二语言情绪韵律的能力已被获得，且在发展的过程中保持较稳定趋势，不受到后天学习中对这门外语的熟练程度的影响。当然，这一推论需要通过发展心理学的研究加以验证。

虽然本研究没有发现语言熟练度的显著主效应，但依然得到了该自变量边缘显著的结果。这说明熟练与非熟练中英双语者在两种语言的情绪韵律识别上依然存在一定差异。从均值的比较中可以发现熟练中英双语者表现地更好。另外，尽管语言熟练度与材料语言的交互作用不显著，但从两个自变量相互结合后，各水平对应的均值来看，熟练双语者在识别英文的情绪韵律时($M = 0.48, SD = 0.02$)，准确率略高于非熟练的中英双语者($M = 0.40, SD = 0.02$)。同样的模式也可见于对中文的情绪韵律识别中(熟练双语者： $M = 0.81, SD = 0.05$ ；非熟练双语者： $M = 0.75, SD = 0.04$)。此边缘显著的结果或许由两种原因所导致：第一，本研究中的熟练中英双语者为来自外语专业且通过专业四级的学生，平时对于英语的接触频率要明显高于非熟练双语者的被试。Shtyrov (2011)曾发现，通过不断地接触与学习口语新异词汇，被试将形成特定的记忆痕迹。而这种表现为接触效应的记忆痕迹，使得人们对新异词汇的表征接近于对真实、熟悉词汇的表征。因此，在本研究中，或许语言的熟练程度通过接触效应对被试的情绪韵律识别造成了影响，而且这种影响也调节了被试对母语的情绪韵律识别，只是影响的效应不大，尚未达到 5% 的统计显著水平；第二，语言的熟练程度本应产生显著的影响，但在本研究中由于自变量的操纵不够有效，因此只

产生了边缘显著的效应。在往常针对语言熟练度的研究中,通常将通过专业英语八级的英语专业学生定义为熟练双语者(张积家&王悦, 2012),但在本研究实施的过程中,专八分数尚未公布,已考过专八的研究生数量又不足以构成大样本,因此本研究才缩小了自变量水平之间的差距,只得到了边缘显著的结果。

最后,对于材料语言和情绪交互作用显著的结果,则可以从跨文化研究的角度解释。若除去语言熟练度这一因素,本研究便是一个使用反平衡设计进行的跨文化情绪韵律识别研究。而根据前人跨文化研究的结果,当被试识别本文化和异文化背景下的情绪韵律时,总会呈现出跨文化一致性和群内优势效应共存的结果(Bezooijen et al., 1983; Paulmann & Uskul, 2013; Thompson & Balkwill, 2006)。本研究的结果与前人结果一致:相比英文,所有被试对中文的情绪韵律识别地更好。同时,所有被试对大部分声音刺激的识别均高于猜测水平。

此外,在对交互作用的简单效应分析中,发现被试对厌恶情绪的识别并不存在语言上的差异。这或许和中英文厌恶情绪声音的物理参数模式相似有关。但是,通过对中英文厌恶情绪声音的声学参数进行独立样本 t 检验发现,两者只在响度上差异不显著 $t(26) = -0.29, p > 0.05$,在时长和基频上却存在显著差别 $ps < 0.01$ 。但是当再对其他 6 种情绪声学参数的中英文差别进行 t 检验后,却仍然发现,那些在简单效应检验中的识别正确率存有显著差异的情绪,在声学参数上也大多是存在显著差别的。这样的结果表明,在解释情绪韵律识别的特征时,在一定程度上,可以将情绪韵律识别的过程类比为单纯的加工声学参数的过程。但在对具体的结果进行解释时,这一观点则不一定完全适用。陈熙海(2011)曾提出,关于情绪韵律的认知机制研究,应在考虑到听觉加工的大背景下进行,同时也需要考虑到背景信息、任务要求、声学参数的局部变化等因素对情绪韵律加工的影响。可以说,识别情绪韵律是一个自下而上的认知过程,但其中也必然会有自上而下加工的参与,这就决定了人们在感知声音中的情绪语调时,不仅仅只是在进行单纯的听觉加工。那么基于怎样的机制,使得情绪韵律加工拥有不同于低级听觉加工的特殊性?对此,有研究通过事件相关电位(ERP)的方法,发现声学线索是通过自下而上的驱动效应调节着情绪韵律的早期加工;而在加工的全过程中,实验任务的要求通过自上而下的方式调节着情绪韵律的加工(江爱世, 2013)。

至于本研究中发现厌恶情绪不存在群内优势效应这一结果,更具体地体现为“对中文的厌恶情绪识别率较低”这一特点。从表 2 呈现的数据来看,被试对中文厌恶情绪的识别正确率是所有中文情绪中最低的。在前人的相关研究中也出现过类似的结果:在对动态情绪面孔的做识别是,相比西方人,东方人更频繁地将厌恶情绪与恐惧和惊讶混淆(高敏, 2016)。刘晓燕(2008)的研究中以 QQ 表情为材料,发现被试对厌恶表情的识别正确率最低。而这一结果应该与情绪表达的强度有关,因为被试对厌恶情绪强弱程度的评分也是最低的。此观点或许也可解释本研究的结果,但本研究未曾对材料的情绪强度进行评价,因此这一解释还需进一步验证。

本研究存在的不足之处主要有以下几点:1) 被试的性别数量不平衡。尤其在性别被作为情绪韵律识别的影响因素的前提下。已有研究发现,女性对情绪韵律的感受性比男性更高(van den Brink et al., 2012)。因此本研究的结果或许会略高于总体水平,因为女性在全体被试中占了很大的比例。2) 样本数量不足以构成标准的大样本。标准的大样本容量为 30 人,但是由于时间和实验条件的限制,本实验只在各个条件下安排了 20 个被试。若容量能进一步扩大,结果将更具有说服力。3) 语言熟练度的两个水平间的间距有待拉大。即增强自变量操纵的有效性。关于这一点已在上文展开过讨论,在此便不再赘述。4) 预实验中对情绪韵律进行评价的维度过少。预实验中只对情绪类别进行了评价,若需要更全面的匹配刺激,还可以从愉悦度、唤醒度、熟悉度、优势度这几个维度进行评价(王一牛, 周立明, & 罗跃嘉, 2008)。通过此种评价方式,对于厌恶情绪为何在本研究中不存在群内优势效应这一结果也能进行更好的分析。5) 研究层面过浅,只在行为层面对这一课题进行探究。将来有望从生理水平,如通过事件相关电位的方法,从 Schirmer and Kotz (2006)提出的情绪韵律加工三阶段模型着手开展研究。情绪韵律加工的三阶段模型

认为, 情绪韵律的加工可以大致分为: 对声学线索的识别分析; 从声学线索中提取并识别情绪意义; 整合韵律情绪到更高水平, 这三个阶段。以这一模型为基础的研究, 可以从每个阶段分别涉及的脑电成分入手。如第一阶段发生较早, 与 N1 有关, 而第三阶段发生较晚, 且涉及到语义整合, 因此与 N300、N400 等晚期成分有关。以这些具体的脑电成分为指标, 对比熟练与非熟练中英双语者的差异, 会得到比本研究更为全面的结果。

参考文献

- 陈熙海(2011). *刺激特征和任务要求对情绪韵律加工的调制——来自电生理的证据*. 博士论文, 北京: 中国科学院.
- 高敏(2016). *基于动态面部表情的东方人情绪识别的研究*. 硕士论文, 成都: 电子科技大学.
- 黄希庭(2011). *心理学导论*. 北京: 人民教育出版社.
- 黄贤军, 张伟欣(2014). 两种任务下情绪韵律的加工进程. *心理科学*, 37(4), 851-856.
- 江爱世(2013). *情绪韵律的动态加工机制*. 博士论文, 北京: 中国科学院大学.
- 刘晓燕(2008). *QQ 面部表情图片的识别研究*. 硕士论文, 长沙: 湖南师范大学.
- 卢植, 涂柳(2010). 二语习得年龄与高熟练度中英文双语者心理词典表征. *外国语(上海外国语大学学报)*, 4, 47-56.
- 毛伟宾, 杨治良, 王林松, 袁建伟(2008). 非熟练中-英双语者跨语言的错误记忆通道效应. *心理学报*, 40(3), 274-282.
- 王一牛, 周立明, 罗跃嘉(2008). 汉语情感词系统的初步编制及评定. *中国心理卫生杂志*, 22(8), 608-612.
- 姚温青(2014). *声音情绪的跨文化识别研究——以汉语、藏语和维语为例*. 硕士论文, 兰州: 西北师范大学.
- 张积家, 王悦(2012). 熟练汉-英双语者的语码切换机制——来自短语水平的证据. *心理学报*, 44(2), 166-178.
- Bachorowski, J. A. (1999). Vocal Expression and Perception of Emotion. *Current Directions in Psychological Science*, 8, 53-57. <https://doi.org/10.1111/1467-8721.00013>
- Bezooijen, R. V., Otto, S. A., & Heenan, T. A. (1983). Recognition of Vocal Expressions of Emotion. *Journal of Cross-Cultural Psychology*, 14, 387-406. <https://doi.org/10.1177/0022002183014004001>
- Bostanov, V., & Kotchoubey, B. (2004). Recognition of Affective Prosody: Continuous Wavelet Measures of Event-Related Brain Potentials to Emotional Exclamations. *Psychophysiology*, 41, 259-268. <https://doi.org/10.1111/j.1469-8986.2003.00142.x>
- Chakraborty, R. (2011). Influence of L2 Proficiency on Speech Movement Variability: Production of Prosodic Contrasts by Bengali-English Speakers. *Bilingualism: Language and Cognition*, 14, 489-505. <https://doi.org/10.1017/S1366728910000441>
- Cornew, L., Carver, L., & Love, T. (2009). There's More to Emotion than Meets the Eye: A Processing Bias for Neutral Content in the Domain of Emotional Prosody. *Cognition and Emotion*, 24, 1133-1152. <https://doi.org/10.1080/02699930903247492>
- Dimoska, A., McDonald, S., Pell, M. C., Tate, R. L., & James, C. M. (2010). Recognizing Vocal Expressions of Emotion in Patients with Social Skills Deficits Following Traumatic Brain Injury. *Journal of the International Neuropsychological Society*, 16, 369-382. <https://doi.org/10.1017/S1355617709991445>
- 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. <https://doi.org/10.1037/a0015178>
- Kotz, S. A., & Paulmann, S. (2007). When Emotional Prosody and Semantics Dance Cheek to Cheek: ERP Evidence. *Brain Research*, 1151, 107-118. <https://doi.org/10.1016/j.brainres.2007.03.015>
- 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. <https://doi.org/10.3758/s13428-013-0324-3>
- Min, C. S., & Schirmer, A. (2011). Perceiving Verbal and Vocal Emotions in a Second Language. *Cognition and Emotion*, 25, 1376-1392. <https://doi.org/10.1080/02699931.2010.544865>
- Ortony, A., & Tumer, T. J. (1990). What's Basic about Basic Emotions? *Psychological Review*, 3, 315-331. <https://doi.org/10.1037/0033-295X.97.3.315>
- Owren, M. J., Amoss, R. T., & Rendall, D. (2011). Two Organizing Principles of Vocal Production: Implications for Non-human and Human Primates. *American Journal of Primatology*, 73, 530-544. <https://doi.org/10.1002/ajp.20913>
- Paulmann, S., & Uskul, A. K. (2013). Cross-Cultural Emotional Prosody Recognition: Evidence from Chinese and British Listeners. *Cognition and Emotion*, 28, 230-244. <https://doi.org/10.1080/02699931.2013.812033>

-
- Paulmann, S., Furnes, D., Bøkenes, A. M., & Cozzolino, P. J. (2016). How Psychological Stress Affects Emotional Prosody. *PLoS ONE*, *11*, e0165022. <https://doi.org/10.1371/journal.pone.0165022>
- Ponari, M., Rodríguez-Cuadrado, S., Vinson, D., Fox, N., Costa, A., & Vigliocco, G. (2015). Processing Advantage for Emotional Words in Bilingual Speakers. *Emotion*, *15*, 644-652. <https://doi.org/10.1037/emo0000061>
- Schirmer, A., & Kotz, S. A. (2006). Beyond the Right Hemisphere: Brain Mechanisms Mediating Vocal Emotional Processing. *Trends in Cognitive Sciences*, *10*, 24-30. <https://doi.org/10.1016/j.tics.2005.11.009>
- Shtyrov, Y. (2011). Fast Mapping of Novel Word Forms Traced Neurophysiologically. *Frontiers in Psychology*, *2*, 340. <https://doi.org/10.3389/fpsyg.2011.00340>
- Thompson, W. F., & Balkwill, L.-L. (2006). Decoding Speech Prosody in Five Languages. *Semiotica*, *158*, 407-424. <https://doi.org/10.1515/SEM.2006.017>
- Uskul, A. K., Paulmann, S., & Weick, M. (2016). Social Power and Recognition of Emotional Prosody: High Power Is Associated with Lower Recognition Accuracy than Low Power. *Emotion*, *16*, 11-15. <https://doi.org/10.1037/emo0000110>
- van den Brink, D., Van Berkum, J. J. A., Bastiaansen, M. C. M., Tesink, C. M. J. Y., Kos, M., Buitelaar, J. K., & Hagoort, P. (2012). Empathy Matters: ERP Evidence for Inter-Individual Differences in Social Language Processing. *Social Cognitive and Affective Neuroscience*, *7*, 173-183. <https://doi.org/10.1093/scan/nsq094>
- Wagner, H. L. (1993). On Measuring Performance in Category Judgment Studies of Nonverbal Behavior. *Journal of Nonverbal Behavior*, *17*, 3-28. <https://doi.org/10.1007/BF00987006>