



对自然语音和人工语音所诱发的情绪感知差异的 ERP 研究

杨 潇

(武汉体育学院,湖北 武汉 430000)

摘 要:采用事件相关电位对人脑在相同情绪效价的人工语音和自然语音(中立)下情绪处理进行探索,以及研究人工语音与其他情绪效价的自然语音(欢乐、生气)在脑电信号上有何区别。对 29 名大学生采用过德国 Brain Products 公司的脑电记录系统采集,按照 10-20 系统扩展的 64 导电极帽记录在带有情绪的语音启动下的 ERP,脑电系统记录其在情绪语音的条件下的事件相关电位和行为学结果。在 700~800ms 时间段内,人工语音比自然语音(欢乐、中立、生气)引发了更正走向的波幅($P<0.05$);大学生对相同效价的人工语音和自然语音确实存在一定的感知差异,在 700~800ms 的时间窗口里,人脑正在对人工语音进行处理,对于人工语音的资源分配显著多与其他语音的资源分配。

关键词:情绪;人工语音;事件相关电位

中图分类号: R749.3

文献标识码: A

文章编号: 1671-931X (2017) 05-0084-07

一、绪论

随着科学技术的不断发展,在计算机信息领域里,能处理包含在媒体中的情感信息语音交流的媒体系统的研究越来越引起人们的兴趣。因为语音信号不仅仅是多媒体人机交互的主要交流沟通方式,也是承载情感信息的重要媒体,所以包含在语音信号中的情感信息对于人机互动的研究就显得尤为重要。如今人机交互日益频繁,许多人工语音发音清晰,音色悦耳,但是否人工语音与自然语音给人们所带来的情绪感受是相同的呢?为了更好的了解有情感背景的语音与普通人工语音对人情绪感知的差异,特别运用高时间分辨率的事件相关电位,通过多次呈现情绪刺激材料,诱发情绪,记录脑电信号,通过处理得到事件相关电位,分析各个情绪效价的语音与 ERP 各成分之间的关系,在与行为实验的数据

相比较,从而探讨人脑对人工语音与自然语音情绪感知差异,希望能给将来更智能的人工语音的研究有所帮助。

(一)自然语音与人工语音材料的选用

从研究角度来讲,情绪化语音的语音材料一般可分为三类语音—自然语音、表演语音与诱导语音^[1],而本文将着重强调自然语音。一般定义中的自然语音是指说话人在真实情况下情绪的自然流露。而更为严苛的定义是说话人事先并不知道其语音会被用于研究,因而语音材料中携带的情绪信息是非常逼真和直接。但获取此类语音材料不易,样本少,没有与之相对应的对比组作为对照,录音效果差且有杂音,可使用性不强^[2]。所以研究人员一般以在实验室的条件下说话人模拟特定的情绪状态发音的语料作为自然语音方面的研究材料,因此本文采用的是中科院自动化所研制的 CASIA 汉语情感语音库。该语

收稿日期:2017-08-28

作者简介:杨潇(1993-),女,湖北武汉人,武汉体育学院研究生,研究方向:运动心理学。

音库中语音是由专业人士朗诵，语料都是单句并且长度适中，在语义上没有显性的情感倾向，最为重要的是该句子中语义的情感可塑性很大，不同语境可以诱导出不同的情感^[3]。而人工语音是指运用人类声道的发声机理、共振频率或是产生的共振峰所制造的“声音”来模拟人类发音的声音^[4]。现有可以将文本信息转换成语音信息的文本朗读技术，该技术旨在模拟人声，给用户带来多方面的体验。

（二）情绪诱发与声音的关系

情绪是一种持续变化的心理状态，并且具有组织性、深刻性、内涵性等特点。人们可以通过感知他人的情绪来捕捉到身边人的感情变化，而这样一次的交互过程被称为情绪感染。一般来说，自然状态下的非言语音节、声音录音以及音乐都可以诱发出特定的情绪。还有研究表明可以佐证有研究表明在感受他人情绪时是和自身体验同种情绪时的生理变化往往是相关一致的。例如，Oberman, Winkelman 和 Ramachandran 的研究发现，识别他人厌恶表情与自身体验厌恶情绪时，都激活了和厌恶相关的内脏反应^[5]。

（三）情绪的 ERP 成分

事件相关电位的时间分辨率可精确至微秒级，是刺激事件引起的脑电的实时波形，由于大量实验证明，事件相关电位可以敏感地反映出情绪活动在脑内的活动过程。研究者通过分析被试在不同情绪成分刺激下所引发的事件相关电位成分的波幅、分布区域等参数的变化来窥视心理活动内容，并且所得的数据便于与传统的行为测量指标—反应时有机地配合，进行过程性研究，事件相关电位技术具有无创性，可以精确地评价发生在脑内认知加工活动，因此该技术开始广泛应用于语言方面的研究^{[6][7]}。

根据相关研究所表明以下的 ERP（事件相关电位）成分是和情绪加工有关的。

1. 中期成分

在有新颖物体出现时，会有明显的痕迹的 N170，而该波会在刺激呈现 170ms 左右达到峰值。而潜伏期在 200ms 左右的 P200 是和目标刺激的早期识别相关连的，是与情绪有关，反应情绪的一个比较的重要指标^[8]。

2. 晚期成分

P300 是第三个正走向波，属于晚成分。在很多与情绪有关的研究中发现不管是正性情绪还是负性情绪都可以使 P300 发生变化。目前认为 P300 是测定大脑意愿活动最客观的手段。关于 P300 的产生机制，一般研究表明是刺激所具有的信息意义是引出 P300 所必需的条件，认为主动和被动模式潜伏期和幅值的差异表明 P300 的反应与听觉中枢的认知过程有关。

N400 是一个波峰潜伏期在 400 ms 左右的明显

负波。因其产生与言语刺激有关，故也被称为语言相关电位。N400 是研究人脑语言信息处理功能的高级认知电位，反映了言语加工及识别过程中大脑的电位变化，还有作为晚期成分的 LPC 成分，一般为正走向，该成分一般出现在刺激呈现 300ms–600ms 之后，相关研究表明，情绪刺激的自动化加工以及某些高级的认知活动如抑制加工、分类活动等都与该成分有关，此外，还有文献报告了情绪刺激的唤醒度也可能与 LPC 成分有关。由于 LPC 成分是晚成分，所以在该成分阶段基本上能够完成对情绪刺激意义的加工。LPP 成分被认为是与注意、动机有关的晚正成分，在 Lindon 的研究中发现该成分反映了视觉皮层活动的全局抑制，与情绪刺激加工活动的选择有关，可以作为情绪调节的指标^[9]。

LPC 成分(Late Positive Competent)，是晚期正成分，一般出现在刺激呈现后 300ms 到 600ms 左右，反映正性情绪词与负性情绪词的分化效应。它也可以被认为是一种与抑郁心境相关的情绪内容识别能力提高的现象，即情绪状态一致的情况下，其反应的波幅会增大。与中性词相比，情绪词会引起更大的正性波幅，尤其是正性中性词。因此，LPC 成分也可反应情绪启动效应中情绪的加工^[10]。

二、研究程序

（一）实验被试

在校大学生 29 人，均为右利手，听力正常没有受损，身体健康无重大病史也无相关心理疾病，参与实验时精神良好。被试总人数为 29 人，最终筛查只选取了 15 个被试的数据，其中男女数量为男 9 女 6，年龄 17–23，试验后均给了一定报酬。

（二）实验材料

本研究材料中的自然语音来自中文语言资源联盟(Chinese LDC)中科院自动化所的语音库，从中选取三个维度分别是开心、生气和中立，每个维度各有相同语义的 50 个句子，均由专业人员朗读，而人工语音则是用文字语音转换软件“能说会道”，均是采用相同的语句（见附录 1），语音的导出格式均为 mp3。用 Adobe Audition 软件进行降噪处理后，这 200 个句子的分贝量一致，比特率为 128bps。为了验证情感语料的有效性，制作了一个语音问卷，请来 35 名在校大学生进行评定，年龄在 19–21 岁，没有听力理解方面的障碍史。在 200 个句子中每个维度挑出 2 个句子做重复对比，一共 208 个五字到七字的短句子（其中 8 个句子作为练习），每种情绪的效价短句随机播放模式，并且每个句子会重复听 2 次，不同句子之间会有空白语音的提示。实验在教室内进行，通过“暴风影音”软件将语音文件播放给被试，要求被试在每个 5 秒间隔内对前面的目标句进行以下判断，对这些语音短句用这 7 个标准来评定，分别

是1(非常生气),2(比较生气),3(有些生气),4(中性情绪),5(有些高兴),6(比较高兴),7(非常高兴),需要判断句子的情绪成分并按照所给定的标准在下面打勾。一共35名被试,筛选出33分有效问卷,经过SPSS11.5的单因素方差分析评定,人工语音与自然语音中的生气效价差异显著 $P=0.032<0.05$ 、与开心效价显著差异, $P=0.027<0.05$,与自然语音中的中立效价差异不显著, $P=0.827>0.05$ 。该语音材料有效,可以使用。

(三)实验过程

实验采用了汉语情感语音库中的150个句子(有三个维度欢乐、中立、生气,每个维度50个句子,句子含义相同)和50个人工语音(与上述句子含义相同),将其导出mp3格式编写到E-prime2.0中。实验中一共有416个试次(其中16个试次作为练习,其余分为8个区组,每个区组里面有50个句子),每个试次会出现一句语音,效价随机。刺激呈现的显示器屏幕为17寸LED液晶显示器(刷新频率:70Hz)和立体声音箱,隔被试距离为80cm左右。实验室为隔音并且光线比较暗的环境,被试靠在斜椅上,在佩戴完电极帽后,主试会说出指导语。被试任务为判断语音的效价,用右手按数字键“1”、“2”、“3”、“4”、“5”给出评分。实验过程会先在屏幕中央呈现“+”注视点200ms,之后便马上出现1000–2000 ms的语音,语音播放完毕后需要被试对此进行反应,按键结束后便有500ms的空白,之后便是下一个试次。正式实验过程需要半个小时到40分钟左右,在每个区组间隙被试会进行适当休息。

实验指导语:欢迎您来参与本次实验!您将会听到来自语音库的一系列句子,在您听到语音之后,您要用数字键“1”(非常生气)、“2”(比较生气)、“3”(中性情绪)、“4”(比较开心)、“5”(非常开心)来对其进行评价,注意要忽略该语音的语义只要对其情感成分进行评价,现在您可以用右手在数字键盘上进行训练。准备好后,按“Q”键进入实验。感谢您的参与!

(四)脑电数据记录与处理

脑电信号通过德国Brain Products公司的脑电记录系统采集,按照10–20系统扩展的64导easy cap电极帽记录EEG,采集脑电数据时以头顶的FCz作为参考点,在右眼上下记录眼电(IO)。采样频率为500Hz,头皮电阻调制为5k Ω 以下,带通滤波频段为0.05至100Hz。参考电极置于双侧乳突连线,接地电极Ground安放于前额,右眼下1cm安放电极记录眼电(IO)。滤波带通为0.05–100Hz,采样频率为500Hz,以刺激呈现前100ms作为基线,将记录参考转换为平均参考,手动剔除伪迹明显的试次,然后采用独立成分分析(ICA)来剔除眼电和明显的肌电活动,因此对这四种实验条件诱发的EEG分别进行叠加和平均,基线校正后,剔除了波幅超过 $\pm 50\mu V$ 的

伪迹。仅对反应正确的听觉刺激诱发的EEG进行叠加,每个条件的ERP波形的平均叠加试次为104。分析时程为200–2000ms,用刺激呈现前100ms的脑电为基线。通过平均波形及每个被试和一个个电极的数据排查,确定ERP数据分析的7个时窗为500–550ms,550–600ms,600–650ms,650–700ms,700–750ms,750–800ms,900–950ms。分析的电极分别为:额区(F3/FC1/F4/FC2/FP2/FZ),中央顶区(C4/CZ/FCZ/P8),枕区(O1)和颞叶区(FC6/F8/F7)。由于分别有三个因素,第一个因素是条件(开心、中立、生气、人工语音),第二个因素是电极(C4、CZ、F3、F4、F7、F8、FC1、FC2、FC6、FP2、FZ、O1、P8、FCZ),第三个因素是时间窗口(500–600ms,550–600ms,600–650ms,650–700ms,700–750ms,750–800ms,900–950ms)。

三、数据处理

(一)行为实验统计结果

运用SPSS11.5统计软件进行行为实验的数据分析,结果如表1:对行为实验的数据进行多重比较可得:关于不同效价语音之间的差异和问卷调查所得出的结论一致。

表1 对不同情绪语音的评分以及被试的平均反应时进行单因素方差分析

效价	评分	平均反应时
欢乐	3.94 ± 0.572	$2940.14 \pm 1207.037ms$
中立	2.97 ± 0.326	$3080.55 \pm 1217.723ms$
生气	2.13 ± 0.649	$3061.53 \pm 1002.853ms$
人工语音	2.98 ± 0.138	$2343.80 \pm 884.535ms$

(二)三因素统计结果

对数据先进行三因素统计可得结果如下:对条件(欢乐、中立、生气、人工语音) \times 电极点(C4、CZ、F3、F4、F7、F8、FC1、FC2、FC6、FP2、FZ、O1、P8、FCZ) \times 时间窗口(500–600ms,550–600ms,600–650ms,650–700ms,700–750ms,750–800ms,900–950ms)之间的三因素方差分析,Levene 齐质性检验大于0.05,得到条件的主效应为 $F(3,78)=56.570, P=0.000<0.05$,主效应显著;电极的主效应为 $F(13,18)=60.492, P=0.000<0.05$,主效应显著;时间的主效应为 $F(6,39)=13.123, P=0.000<0.05$,时间主效应显著。条件和电极交互效应, $F(39,234)=4.772, P=0.000<0.05$,交互作用显著;条件和时间交互作用 $F(18,234)=87.377, P=0.000<0.05$,交互作用显著,时间和电极的交互作用为 $F(78,234)=1.712, P=0.000<0.05$,交互作用显著;条件、电极、时间三种交互作用显著, $F(234,5488)=6.496, P=0.000<0.05$ 。随后按照电极来进行下一步的两因素方差分析。

(三)二因素统计结果

对数据进行二因素(条件、时间)检验结果如表2。

(四)单因素统计结果

表 2 对数据进行二因素统计结果(条件、时间)

电极	Levene 检验	时间主效应	条件和时间交互作用
C4	sig.>0.05	$F(6, 18)=1.470, P=.187 > .05$	$F(18, 392)=5.731, P=.000<.05$
CZ	sig.>0.05	$F(6, 18)=2.378, P=.029<.05$	$F(18, 392)=11.965, P=.000<.05$
F3	sig.>0.05	$F(6, 18)=3.763, P=.001<.05$	$F(18, 392)=17.769, P=.000<.05$
F4	sig.>0.05	$F(6, 18)=2.607, P=.017<.05$	$F(18, 392)=13.467, P=.000<.05$
F7	sig.>0.05	$F(6, 18)=3.334, P=.003<.05$	$F(18, 392)=9.445, P=.000<.05$
F8	sig.>0.05	$F(6, 18)=2.295, P=.034<.05$	$F(18, 392)=7.872, P=.000<.05$
FC1	sig.>0.05	$F(6, 18)=2.826, P=.011<.05$	$F(18, 392)=17.997, P=.000<.05$
FC2	sig.>0.05	$F(6, 18)=2.499, P=.022<.05$	$F(18, 392)=16.009, P=.000<.05$
FC6	sig.>0.05	$F(6, 18)=2.194, P=.043<.05$	$F(18, 392)=11.911, P=.000<.05$
FP2	sig.>0.05	$F(6, 18)=1.002, P=.424 > .05$	$F(18, 392)=4.524, P=.000<.05$
FZ	sig.>0.05	$F(6, 18)=3.409, P=.003<.05$	$F(18, 392)=17.469, P=.000<.05$
O1	sig.>0.05	$F(6, 18)=2.498, P=.022<.05$	$F(18, 392)=9.775, P=.000<.05$
P8	sig.>0.05	$F(6, 18)=2.942, P=.008<.05$	$F(18, 392)=8.530, P=.000<.05$
FCZ	sig.>0.05	$F(6, 18)=4.018, P=.001<.05$	$F(18, 392)=23.523, P=.000<.05$

由以上二因素结果可得这 14 个电极均可进行单因素方差分析,以下数据是从这 14 个数据进行单因素方差检验过后,所筛选出来的新数据。

表 3 单因素方差分析结果

电极	700–750ms	750–800ms	900–950ms
CZ	人工 > 自然(*)		人工 > 自然(*)
F3	人工 > 自然(*) 中立 > 生气(*)	人工 > 自然(*) 生气 > 中立(*)	
F4	人工 > 自然(*) 欢乐 > 中立(*)		
FC6		人工 > 自然(*) 欢乐 > 中立(*)	
FZ		人工 > 自然(*)	
P8	人工 > 自然(*)		
FCZ	人工 > 自然(*) 欢乐 > 中立(*) 生气 > 中立(*)	人工 > 自然(*) 生气 > 中立(*) 生气 > 欢乐(*)	

注:* 表示在 0.05 水平上显著。

四、实验结果讨论与总结

(一)实验结果分析与讨论

1.地形图分析

由图 1 可知,人脑对不同效价的语音处理的时间是不同的,比如额叶、顶叶和颞叶对欢乐语音处理的时间是在 600–700ms,在 700–800ms 时只有额叶进行不太明显的处理,对于中立效价的自然语音来说除开枕叶外,额叶、顶叶以及颞叶在 500–600ms 的时间里都对其进行了加工处理,在 600–700ms 只有额叶对其进行微弱的处理。对于生气效价的自然语音来说,只有额叶对其进行处理,时间段为 800–1000ms。但是对于人工语音来说,除枕叶外,700–800ms,额叶、顶叶包括部分颞叶都对此进行明显加工。

2.脑电图分析

由图 2 可知,大部分电极在 50ms 左右,各个效价均出现微弱的 P50,说明被试能够听到各个效价的语音。而在 450–600ms 这个时间段,条件中立的脑波最先出现晚期成分 LPP,诱发出比其他条件更大的正走向波幅,持续时间较长。此外在 500–700ms 时间窗口中,条件为欢乐的脑波,诱发出晚期成分,其走向为正,并且 LPP 的波幅明显大于其他条件诱发的波幅。在 700–800ms 的时间窗口中,条件为人工语音的 LPP 波幅与其他条件差异显著,而条件为生气

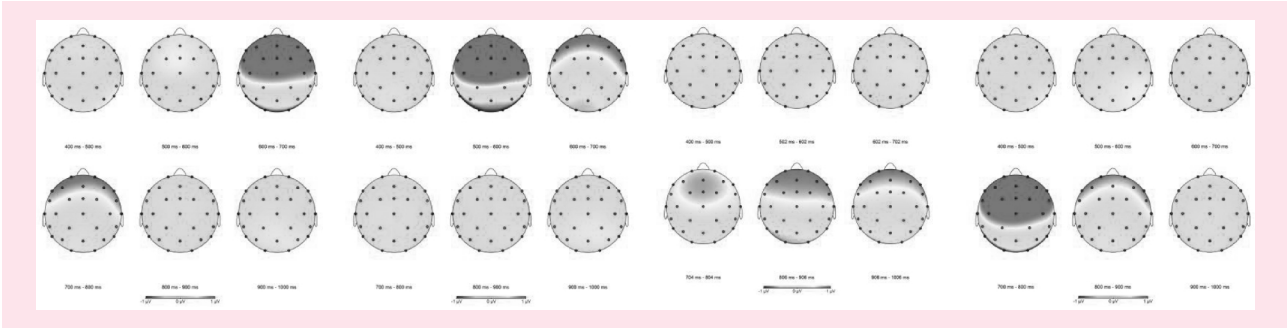
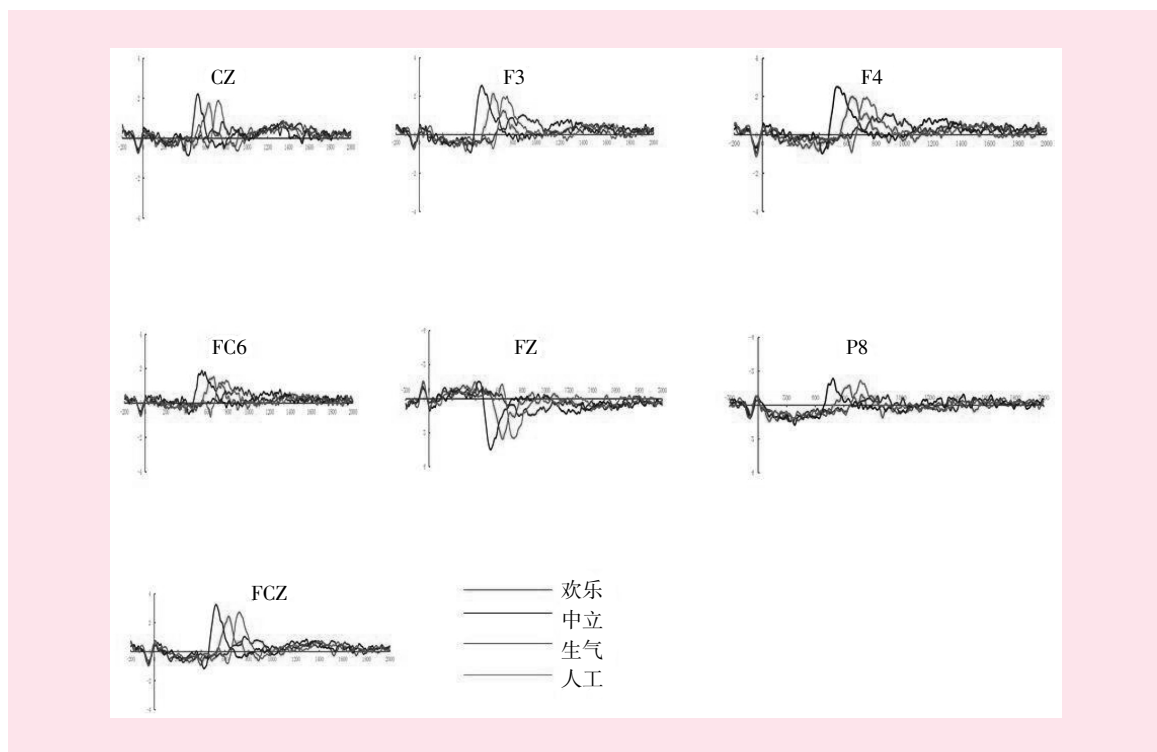


图 1 从左到右依次为欢乐、中立、人工语音和生气的差异地形图(μv)

杨潇：对自然语音和人工语音所诱发的情绪感知差异的 ERP 研究

图2 左侧总平均波幅(横坐标单位为 ms,纵坐标单位为 μV)

注:*表示在 0.05 水平上显著。

的波幅潜伏期长,在 850–1000ms 之间诱发出较为明显的 LPP,持续时间最长。

3.综合讨论

由以上的行为数据、统计结果、脑电图以及地形图可知,额叶、顶叶以及部分颞叶对语音处理的情况。所选的电极 FCZ、FZ、F3、F4、FC6 代表额叶,而 P8 代表颞叶,CZ 代表顶叶,一般来说额叶与听觉关系紧密,而顶叶是负责躯体感觉,额叶是具有高级认知活动(注意、记忆、问题解决)的调节和控制运动的能力比如筹划、决策和目标设定有关,因此我们可以了解到在听到语音之后的 700–800ms,大学生的额叶会对人工语音进行处理以及分析,人工语音与自然语音的处理时间段是有差别,对相同效价的人工语音和自然语音是有不同之处的。

从行为数据的统计结果来看,除中立效价的自然语音与人工语音的评分差异不显著之外,其他条件两两比较差异显著,这一结果和问卷评分的结果相似。而对于反应时来说,人工语音的反应时与其他条件的自然语音反应时差异显著,人工语音的反应时明显短与自然语音的反应时,猜想是因为实验需要被试听完语音再进行评分,并且一旦辨认出是人工语音,可以直接选“3”,而不像自然语音还有衡量是哪个维度的情绪,情绪强度等。

从事件相关电位的成分来看,每个电极的脑电图里均出现 P50 这个和听觉有关的外源性成分,说明被试能够听到语音^[1]。而出现的 P100 和 N100 也代表了额颞上回听觉区域里可以识别外源性成分^[12],一

般认为,可能是颞上回的听觉区域产生的 P100、N100,Tarkka 等 132 通过脑磁波描记术和脑电源分析研究表明 N1 和 P2 起源于颞叶^[13]。与额叶有关的电极中,会出现 N200,这个成分与执行任务的矛盾冲突的神经相关,也与不恰当的反应时认知有关。这可能是因为实验要求被试听完语音之后,再作反应,因此抑制冲突而导致的 N200^[14]。比较有意思的是,没有出现 P200,这可能表示负性情绪偏向不明显^[15]。由于有研究表明给予被试中性刺激时,它的 P300 的波幅最大,而本研究中中立效价的语音的波幅较大,由于 P300 有可能会延迟到 1000ms^[16],这与本研究符合。由上文分析可以看出,中立效价的语音会引起 N400,而 N400 是语言信息处理功能的高级认知电位,反映了言语加工及识别过程中大脑的电位变化^[17]。这可能是由于实验只让被试注重情绪背景,忽略字词的意思,但中立效价的语音情绪背景不太明显,因而被试才会自动注重中性效价语音的言语处理。每个效价的语音均引发出了正晚期成分,正晚期成分与情绪刺激的加工和一些高级认知活动也有关系^[18],是代表着对情绪刺激意义的加工,表示人脑对各个效价的语音均做了加工处理。

从时间窗口来看,时间大约是 500–1000ms,是处理语音的时间。大约 500–600ms 是处理条件为中立的语音,其次是处理条件开心的语音,处理时间为 600–700ms,之后才是人工语音,处理时间为 700–800ms,最后在 800–1000ms 加工的是条件为生气的语音。一般来说 500ms 之后的都属于晚期成分,晚期

成分与工作记忆有关表示对情绪刺激意义的加工^[19],且晚期成分是和情绪强度成正比的。由以上数据分析可以看出,中立效价的自然语音最先出现晚期成分,可能是因为中立效价的语音较好辨认,不用对其情绪背景进一步的评判,导致工作记忆对其的信息量较少,造成在记忆保持上占用较少的认知资源,因而最先诱发出正走向大波幅的晚期成分。其后是欢乐效价的自然语音诱发出晚期成分,可能是因为人脑对正性情绪有种特殊的偏好^[20],所以处理欢乐效价的自然语音要比带有其他情绪色彩的语音要快,更敏感,激起情绪反应的速度较快。人工语音诱发出晚期成分的时间较晚,可能是因为人工语音的情绪表达不太明显,加上音色和一般人类的口音有出入,所以就调动更多的大脑内部资源参与处理和加工。最后是生气效价的自然语音诱发晚期成分,一般来说,人们对于负性情绪有着不同的敏感性,在一些情绪任务中,人们对负性情绪的反应速度会变慢^[21]。这和本研究结果相似,猜想是由于生气的情绪引发了更多的集中性加工,被分配了更多的注意资源进行额外的编码,影响了同一时间段的其他语音,因此在处理生气效价的语音时,波幅都不大。

(二)建议

本实验的结果显示大学生对自然语音的中立条件最为敏感,次之是自然语音的开心条件,而对人工语音的处理时间相对靠后,最后是负性语音。这说明人脑在处理中立效价的自然语音最快,消耗认知资源最少,因此在一些紧急情况下,或者给工作记忆和注意力有损伤的人,尽可能采用中立效价的自然语音进行交流和处理,这样效率会高,效果会好。

五、实验结论

由以上分析可得出以下结论:大学生对相同效价的人工语音和自然语音确实存在一定的感知差异,这个差异不仅是统计的差异也是处理情绪时所反映出的差异,在700-800ms的时间窗口里,人工语音与其他效价的自然语音在统计检验上有显著性差异,而此时人工语音引起晚期正成分,人脑正在对人工语音进行处理,对于人工语音的资源分配显著多与其他语音的资源分配。

参考文献:

- [1] 高慧,苏广川,陈善广.不同情绪状态下汉语语音的声学特征分析[J].航天医学与医学工程,2005,(18).
- [2] 赵力,钱向民,邹采荣,等.语音信号中的情感识别研究

[J].软件学报,2001,(12).

- [3] 吕庆莉,年玮.明日的播音明星几可乱真的人工语音问世[J].国外科技动态,2003,(3).
- [4] 章婷,顾文涛.汉语情感语音的语料设计与韵律研究[J].南京师范大学文学院学报,2011,(3).
- [5] 孟昭兰.情绪心理学[M].北京:北京大学出版社,2005.
- [6] 郑志伟,黄贤军,张钦.情绪韵律调节情绪词识别的ERP研究[J].心理学报,2013,(45).
- [7] 张森泉.ERP成分分解方法研究[D].成都:电子科技大学,2008.
- [8] Rossion B, Curran T, Gauthier I. A defense of the subordinate-level expertise account for the N170 component[J].Cognition,2002,(12).
- [9] 屈南.视听记忆编码和提取的ERP研究[D].北京:首都师范大学,2005.
- [10] Mecklinger A, Pfeifer E. Event-related potentials reveal topographical and temporal distinct neuronal activation patterns for spatial and object working memory [J]. Cognitive Brain Research,1996,(1).
- [11] 周佳莉.跨视听通道的再认ERP研究[D].秦皇岛:燕山大学,2014.
- [12] 罗跃嘉,魏景汉,翁旭初.汉字视听再认的ERP效应与记忆提取脑机制[J].心理学报,2004,(33).
- [13] 宋阳.不同情绪下认知过程的事件相关电位研究[D].天津:河北工业大学,2007.
- [14] 梁勇,王正敏.听觉事件相关电位P300的主动与被动模式记录比较[J].临床耳鼻咽喉科杂志,2000,(07).
- [15] 张明岛,陈兴时.脑诱发电位学[M].上海:上海科技教育出版社,1995.
- [16] 王继军,张明岛,外间宏.事件相关电位N400神经发生源研究[J].上海精神医学,2006,(02).
- [17] 周昕,梁勇.听觉事件相关电位各亚成分的起源及意义[J].临床耳鼻咽喉头颈外科杂志,2008,(11).
- [18] 常翼,庞小梅,许晶,等.情绪语音信息自动加工的失匹配负波研究[J].医学与哲学,2013,(3).
- [19] Herrmann C S, Mecklinger A, Pfeifer E. Gamma responses and ERPs in a visual classification task [J].Clinical Neurophysiology,1999,(4).
- [20] 杨丽珠,董光恒,金欣俐.积极情绪和消极情绪的大脑反应差异研究综述[J].心理与行为研究,2007,(09).
- [21] 杨阳,张钦,刘旋.积极情绪调节的ERP研究[J].心理科学,2011,(27).

[责任编辑:许海燕]

An ERP Study of Emotional Perception Differentiated by Natural and Artificial Voice

YANG Xiao

(Wuhan Institute of Physical Education, Wuhan 430000, China)

Abstract: This study aimed to identify the differences in emotional processing between artificial voices and natural human voices (neutral) by using event related potential (ERP) and to clarify how different between artificial voices and natural human voices (happy, mad) in ERP. The ERP elicited under different emotional voices conditions were recorded of 29 university students and their behavioral performance was collected. The event-related EEG signals were recorded and analyzed to find out the appropriate time window of ERP. During the 700–800ms time window, artificial voices aroused more large positive amplitudes (LPP or LPC) than natural human voices ($P < 0.05$). Participants were most sensitive to neutral intonation of natural human voices. Artificial voices lead to more positive LPP or LPC, the emotional processing of artificial voices was slower than neutral intonation of natural human voices. The resource allocation of artificial voices was much more than other kinds of natural human voices when brains addressed voices. The speed of addressing neutral intonation of natural human voices for the brains was the fastest and it consumed the least cognitive resources.

Key words: emotion; artificial voice; event-related potential (ERP)

(上接第 75 页)

Research on the Influencing Factors of Film Box Office Income Based on VEC Model

BI Yan-cheng

(College of Economics and Business Administration, Central Normal University, Wuhan 430000, China)

Abstract: Box-office revenue is an important measure of the industry's development, and the factors that affect box-office earnings are numerous and complex. At first, this paper summarized the research achievements of predecessors; secondly, the paper analyzed the influence factors of domestic box-office receipts in 2000–2016 by establishing the co-integration and vector error correction model (VECM). Finally, some suggestions were given in order to make some references to the future development of the domestic film industry and relevant policies.

Key words: film industry; box-office revenue; VEC; law of development