

◇ 研究报告 ◇

a 系语气词特殊发音基频的情感表现*

舒炎昕 王军锋[†] 拓江敏 李悦

(西南科技大学制造科学与工程学院 绵阳 621010)

摘要 语气词在日常对话中频繁使用,能表达特定的情感。a 系语气词在生活广泛使用,口语中的某些语气词的特殊音调并不是普通话四个标准声调。该文从发音的基频角度出发,认为特殊音调能适用于一类情感的表达,提取特殊音调的基频数据,发现其在基频上的线性变化特征以及发音的规律。从基频上合成特殊音调语气词并应用到句子的语气词中,通过情感识别对比实验和情绪反馈对比实验表明,含有该种特殊音调语气词的情感语句使听音人产生较强烈的正向情绪,感到更加愉悦。

关键词 特殊音调,语气词,基频研究,汉语发音

中图分类号: TN912.33

文献标识码: A

文章编号: 1000-310X(2018)06-0947-09

DOI: 10.11684/j.issn.1000-310X.2018.06.017

The emotional expression of fundamental frequency of special tone in 'a' series statement label designator

SHU Yanxin WANG Junfeng TUO Jiangmin LI Yue

(School of the Institute of Manufacturing Science and Engineering, Southwest University of Science and Technology, Mianyang 621010, China)

Abstract Statement label designator is frequently used in daily conversations, it can express a particular emotion. 'a' series statement label designator are widely used in life, the special tones of some mood words in spoken language are not four standard tones in mandarin. This article starts from the fundamental frequency of pronunciation, it is considered that special tones can be applied to the expression of a class of emotions. The fundamental data of special tones is extracted, the characteristics of the linear variation on the fundamental frequency and the law of pronunciation are found. The synthesis of a special tone word is from the fundamental frequency and applied to the statement label designator of a sentence word, through the contrast experiment of emotion recognition and contrast experiment of emotional feedback, it is shown that the emotional statements containing the special tone make the listeners have a stronger positive emotion, and feel more pleasurable.

Key words Special tone, Statement label designator, Research of fundamental frequency, Chinese pronunciation

2018-03-12 收稿; 2018-07-13 定稿

*基于团队模式的研究生专业课程教学实践项目 (14JGCX07), 国家科技支撑计划项目 (2015BAH38F00), 西南科技大学博士研究基金项目 (17zx7123)

作者简介: 舒炎昕 (1993-), 男, 陕西汉中人, 硕士研究生, 研究方向: 语音交互设计、用户体验设计。

[†]通讯作者 E-mail: wjfchi@foxmail.com

1 引言

日常对话中经常使用到语气词。语气词在情态语气范畴中有表义功能、完句功能^[1]。它能够表达说话人的情感和反映说话人的意愿、态度以及评价,是人类情感表达的具体方式^[2]。现代汉语中六个典型语气词“的、了、吗、吧、啊、呢”派生出许多新的形式,如“呀、啦、哈”等语气词^[3]。孙锡信^[4]指出,近代汉语文本中最常用的语气词都是a系语气词,即都带a音或者主要元音都是a的语气词。普通话是以北方话为基础方言,以典范的现代白话文著作作为语法规范的现代标准汉语,包含阴平、阳平、上声、去声四种声调。而在生活中受特定环境因素的影响,说话人为了表现主观情绪,语气词的发音结合两种声调形成特殊发音。例如演员为了增加表演的趣味性,在一句台词的末尾加上特殊发音的语气词,引起观众大笑。观众通过演员传递的语音信息已感受到台词文本内容的趣味性,而结尾处语气词因其特殊的发音增加了整句台词的趣味性,强化了表演的情感表达。

语音信号主要包含语义和声学信息,在声学研究中基频(F_0)是重要的参数^[5],决定了发音的音调。在常见的韵律特征中基频相关特征有基音频率及均值、变化范围、变化率和均方差等^[6]。情感语音传递的信息包含语言信息、副语言信息和非语言信息,副语言信息是说话人下意识控制的,非语言信息表现说话人的心理和物理状态^[7]。

在合成语音中,基频(F_0)是声学研究的一个重要参数,它决定了声音的高低^[8]。基频的提取方法主要有时域的算法、频域算法、统计方法等^[9]。由基频合成语音的方法有基于目标逼近特征的基频建模方法,使合成语音有更好的情感表现力;时域信号调制的方法,提高相位连续性^[10];改进的即时域同步叠加固定合成方法,提高合成语音韵律修改能力^[11]。以上方法较好地改善了合成语音的自然性以及提高了特定情感语音的表现能力,但对于语气词特殊发音的合成和情感表现没有研究。

为了研究语气词特殊发音的情感表现能力,论文首先建立了特殊发音语料库,提取特殊发音基频特征进行研究,分析了特殊发音的基频随时间变化的关系规律,获得一种语气词特殊发音的方式。之后,以阴平发音为基础,将该特殊发音合成到不同语气词的发音中,生成含有特殊发音语气词的语音,并

进行了听辨对比。最后,利用主观数据验证了特殊发音的情感表现力,并总结了特殊发音在语气词中的应用条件。将这种方法应用到合成语音中,可以弱化用户对合成语音的乏味感,取悦用户,让用户获得良好的语音交互体验。论文的研究结果可以为语音交互的情感设计提供指导思路。

2 语气词特殊发音的基频分析

2.1 建立普通话特殊发音语料库

为了分析语气词特殊发音的基频信号,获取语气词特殊发音的方式,首先建立中文语气词特殊发音数据库。目前世界上获取语音数据的方式主要有表演数据和激励数据,由于表演数据易获取、可控制、情感表达较明显,因此选择表演数据。挑选十名发音清晰、普通话标准的人员录制特殊发音。男女各五名,录制语料文本取自节目台词:“人的一生很短暂,有的时候跟睡觉是一样的,眼睛一闭,一睁,一天过去了hao”。由于特殊发音的声调不在四个声调范围内,不能界定是哪一个汉字,故取汉字拼音作为发音语料,且设定语气词为一个字。要求训练语料录制人员模仿原本节目中演员的发音方式进行语料库的录制。每人录制十次,为了让语料录制人员模仿得更像,要求说出整段台词,该段台词的前半部分用于酝酿情感,特殊发音语气词在后半部分,为使特殊发音具有一定代表性,提取台词后半部分的语气词特殊发音作为研究对象,建立特殊发音语料库。

设定平常说话的情感为中立情感,用普通话以中立情感说出整段台词,作为研究对象的语气词也以中立情感的说话方式进行发音,只提取该段话的语气词发音建立中立发音语音库;录制作为研究对象的语气词的阴平发音作为特殊发音的基频参考语料。

在安静的实验室环境下进行数据库的录制,采用Adobe Audition CS6录音软件、SGC-598语音传输设备,录音采用Conexant SmartAudio HD声卡,声音的格式为16 kHz,单声道,16 bit。得到小规模的特殊发音数据库,共包含单个字特殊发音200个,单个字中立发音200个,阴平发音200个,最终获得样本数据600个。

2.2 发音基频分析

在时域内, 语音信号的特征是随时间变化的, 在浊音段表现出周期信号的特征, 可以检测到基频信号; 而在清音段, 高速的气流强制通过声道导致气流紊乱, 更像高斯白噪声, 不能检测到基频^[12]。因此主要研究浊音段基频线的特征, 提取语气词阴平发音的样本。praat 软件采用时域算法提取实验语料的基频数据和每个字的浊音段部分, 规定时长 400 ms, 提取基频数值 20 个样点, 获取所有人单个字发音的基频值。使用 OriginPro 8 软件分析数据, 并绘制阴平发音理想参考基频线, 作为合成特殊发音的基础数据。

处理参考语料数据, 结果如图 1 所示。图中两条曲线分别为单个字的男性和女性的平均阴平发音基频线, 女性平均基频普遍高于男性, 是因为男性声带宽厚, 振动频率低, 而女性声带窄而薄, 振动频率偏高。

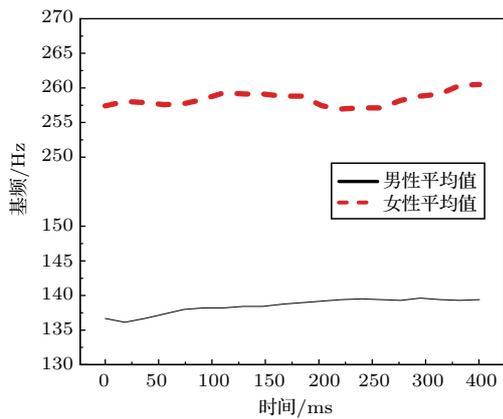


图 1 单个字阴平发音平均基频线

Fig. 1 Single word of the rising tone pronunciation average fundamental frequency line

从基频线性变化上发现阴平发音的基频线类似于一条平直线, 求出每段基频线的基频平均值 \bar{F} , 取每段基频线上的基频绝对值 $|F_0| = |F_i - \bar{F}|$, i 为时刻, 得到所有基频线上最大差值 $|F_{0MAX}| = 7.80$ Hz。每个人的基频线中呈现较小的上下浮动, 是因为人在发声时, 声带拉紧声门裂变窄, 在喉内肌肉协调下, 气流从气管和肺部冲出不断冲击声带, 引起嗓带规律性振动。但喉内肌肉不能一直保持固定的周期性运动, 肺部和气管的气流不能一直保持在固定的流速, 导致声带振动频率出现轻微的变化, 基频线产生浮动。

处理中立发音语料数据, 结果如图 2 所示。获得每段基频线的基频平均值, 求出该段线上每个点与平均值的差值, 并将所有基频线相同点差值平均后绘制平均基频线 l_0 和理想阴平发音基频参考线 l_1 , 得到阴平理想参考平均基频 $F_1 = 0.18$ Hz。

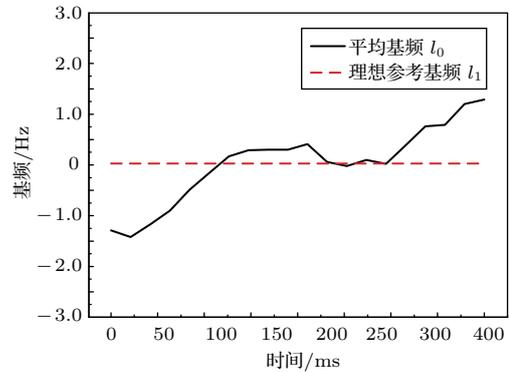


图 2 平均基频线和理想参考基频线

Fig. 2 Average fundamental frequency line and ideal reference base frequency line

对比中立发音语料和特殊发音语料数据, 结果如图 3 所示。“嚎、呵、哈”属于语气词的变体, 寻求认可、舒缓气氛是这类语气词的基本功能^[13]。语气词的形式发音, 包含该语气词的整段话语有表认同的疑问含义。无论男性还是女性, 特殊发音的浊音段时长长于中立音浊音段。受个体之间接受训练程度、模仿程度等因素影响, 不同个体在同一个浊音段时长上有略微差异, 在较短的时间周期里, 性别和发音时长没有较大关系。仅作为语气词的发音浊音段时长要短于特殊发音时长。

处理特殊发音语料数据, 提取特殊发音浊音段基频 50 个样点。以 l_1 为分界线, 大于 F_1 的值为

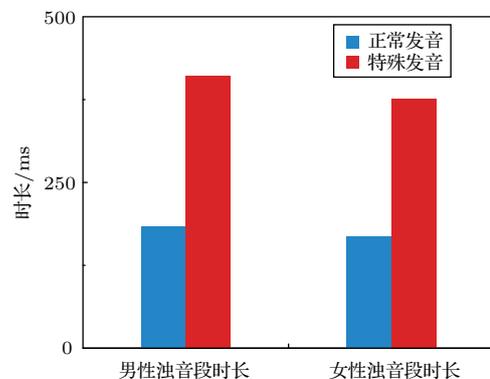


图 3 浊音段和语句平均时长

Fig. 3 The time length of the sonant segment and the sentence

正数,反之为负数,绘制特殊发音基频线 l_2 ,并把时间转换为百分制。图4是句子末尾语气词特殊发音基频变化线,平均 l_2 各点数据得到理想参考基频 $F_0 = 37.947$ Hz,并绘制参考基频线 l_3 。以 l_3 作为参考线研究特殊发音的基频线性变化,发现浊音段46%~100%的基频整体低于平均水平;0~6%的基频呈上升趋势,6%~68%基频逐渐下降,68%~100%基频上升,其中6%~24%基频急速下降,基频最低点在16%。0~6%基频值逐渐上升是因为要达到一定的声调,而声带振动频率是逐步达到的,整个基频线峰值点在6%。

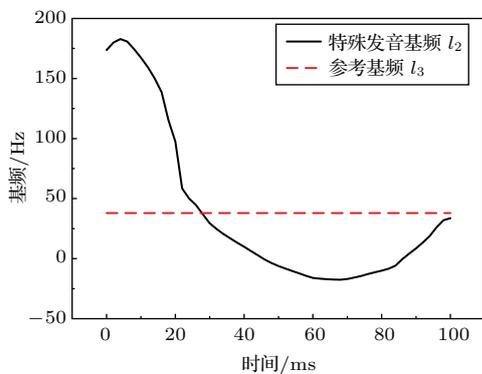


图4 句末语气词特殊发音基频线

Fig. 4 The fundamental frequency line of statement label designator special tone in the end of sentence

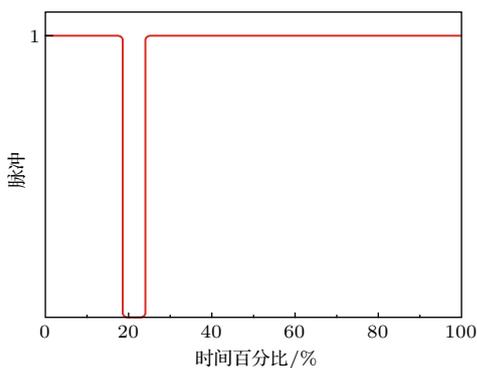


图5 句末语气词特殊发音脉冲图

Fig. 5 The pulse graph of statement label designator special tone in the end of sentence

通过深入研究发音语料、频谱、时长,对比图5处理后的脉冲图谱发现,在起初0~16%阶段肺部向外快速呼气,声带高速振动产生的基频信号高于参考基频线 l_3 水平;在18%~24%的位置脉冲信号微弱,这个短暂的时段声带张开振动减弱,产生的音

量微小,气流紊乱难以侦测到脉冲信号;声带经过短暂的张开立即闭合阻碍空气流动,产生较高频率的振动,同时发声音量增大,使34%~100%脉冲信号较强。整段特殊发音可以24%为分界线,可分为0~24%、24%~100%两大段:0~24%第一段基频值整体高于平均水平,且所占时长较短,对比发音时长数据发现,虽然不是完整的语气词发音段落并且时长较短,但是这一段基频已能感知并理解出字面意思和表达表语气的功能,该段发音类似语气词;24%~100%第二段基频值整体高于平均水平,时长较长,从主观感受上基频变化更加明显。整段发音第一段基频变化范围较大,但更多用于传递字面发音的信息,第二段时长较长,更多用于感知发音者的音调变化。

3 听辨实验

3.1 实验设计

为了让实验结果具有一定的普遍性,实验文本中包含常用的a系语气词,根据《现代汉语汉字频率表》,从使用频率较高(出现频率范围为4.8867%~0.015%)的前1000个字中提取a系语气词:吧(0.0447%)、吗(0.0431%)、啊(0.0264%)、哈(0.0163%)。将这4个字应用到实验语料中,每个语气词合成一类实验语料,共四类语气词语料对照组。每类语料对照组包含正常发音语料组和特殊发音语料组。

选十名声音清晰、普通话标准的人员,告诉听音人每段文本的情景演绎,将四类语气词分别以快乐、愤怒、悲伤、平静四种感情表达该段文本,作为四类语气词正常发音语料组。录制文本中四类语气词的阴平发音分别作为合成特殊发音的语料,将特殊发音语气词应用到快乐、愤怒、悲伤、平静四种语音作为四种特殊发音语料组。录制16 kHz采样率的实验语料。使用cool edit pro采用基频转化为半音^[14]的方法合成特殊发音,并矫正时间轨。

将合成好的语料制作成PPT演示文件,页面采用黑色字体、白色背景。每页为一个情景发音,例如第一页将“吧”语气词应用到快乐的正常发音中,首先文字叙述小明在参加派对特别开心的情景演绎,小明快乐地说道:“让我们从今晚嗨到天亮吧!”,然后说话文字内容和正常语音同时播放。第二页将“吧”语气词应用到快乐的特殊发音中,首先

给出相同的情景演绎,然后说话内容和特殊语音同时播放。第三页将“吧”语气词应用到愤怒的正常发音中,首先情景演绎一段小张愤怒的情景并说道“你赶紧滚吧!”,然后小张的说话内容和正常语音同时播放。第四页将“吧”语气词应用到愤怒的特殊发音中,同时播放说话内容和特殊语音。第五页和第六页为“吧”语料的悲伤语料对照组,第七页和第八页为“吧”语料的平静语料对照组。前八页为“吧”的语料对照组。每个语料对照组共8页,全PPT共计4个语料对照组、32句语料。

选择30名思维正常、没有语音学背景的人作为

听音人,要求听音人进入安静空旷的实验室并端坐正对27寸IPS显示器,采用DS-1338功放设备。操作人员负责播放PPT,每页语料播放完毕后,听音人进入答题阶段,听音人识别说话人所表达的情感,操作人员判断听音人听见说话后产生何种情绪反馈。答题结束后进入下一页实验内容,直到所有页面播放完毕。每次完成一组语料对照组,休息30s后,进行下一组语料对照组实验。为了避免因测试顺序出现的“顺序效应”,所测试的四类语气词语料对照组在PPT中的顺序采用拉丁方处理,具体方式如表1所示。

表1 听音人测试顺序

Table 1 The test order of listeners

	第一组	第二组	第三组	第四组
第一次	“吧”语料对照组	“吗”语料对照组	“啊”语料对照组	“哈”语料对照组
第二次	“哈”语料对照组	“吧”语料对照组	“吗”语料对照组	“啊”语料对照组
第三次	“啊”语料对照组	“哈”语料对照组	“吧”语料对照组	“吗”语料对照组
第四次	“吗”语料对照组	“啊”语料对照组	“哈”语料对照组	“吧”语料对照组

3.2 情感识别实验

实验将观察特殊发音基频应用到不同语气词中听音人对情感辨别的辨认程度。播放PPT进行实验,答题阶段题目为选项评分题,每页一道题,给出快乐、愤怒、悲伤、平静四种基本情感评分项,听音人直接判断并给四种情感评分,分值为1~9分,认为表达的情感与评分项越相似,分值越高。

表2~表5为听音人听见四种情感语音后,对说话人所表达的情感进行的评价,所得分值均为平均分。数据中发现,表2~表5的正常发音组的四

种情感句对应的情感平均评分都较高,最高和最低分别达到7.77分和6.93分,听音人对说话人的评价与说话人自身表达的情感基本相符。而四类语气词特殊发音组中,四种情感语句相对应的情感平均评分都低于正常发音组,每种情感语句的四种情感评分较为接近,且特殊发音的标准偏差普遍高于正常发音,听音人听见特殊发音后不能做出明确的情感评价,还有部分听音人认为有评分项以外的情感,说明听音人对特殊发音的语句产生陌生感。

表2 语气词“吧”的四种情感语音平均分

Table 2 Four kinds of emotional speech average point table of statement label designator 'ba'

情感	语料							
	正常语料 快乐语句	特殊发音 快乐语句	正常语料 愤怒语句	特殊发音 愤怒语句	正常语料 悲伤语句	特殊发音 悲伤语句	正常语料 平静语句	特殊发音 平静语句
快乐	7.13	5.93	1.87	4.33	1.73	4.20	2.10	5.43
σ 快乐	1.40	1.01	1.04	1.49	1.01	1.71	1.44	1.24
愤怒	2.03	3.90	7.77	4.43	1.83	2.47	2.23	3.57
σ 愤怒	1.12	2.13	1.00	4.25	1.17	1.95	1.35	1.12
悲伤	1.90	4.06	2.40	3.13	7.27	3.00	2.53	2.76
σ 悲伤	1.29	1.92	1.43	1.40	1.31	1.20	1.54	1.36
平静	2.00	2.50	1.56	3.47	2.37	2.46	7.3	4.87
σ 平静	1.79	1.07	0.78	1.33	1.21	0.90	1.14	1.42

表3 语气词“吗”的四种情感语平均分

Table 3 Four kinds of emotional speech average point table of statement label designator 'ma'

情感	语料							
	正常语料 快乐语句	特殊发音 快乐语句	正常语料 愤怒语句	特殊发音 愤怒语句	正常语料 悲伤语句	特殊发音 悲伤语句	正常语料 平静语句	特殊发音 平静语句
快乐	7.00	6.10	2.80	3.90	1.78	3.13	1.90	5.43
σ 快乐	1.14	1.99	1.06	1.53	0.85	1.43	1.00	1.68
愤怒	2.23	3.00	7.67	5.80	1.67	2.43	1.80	3.57
σ 愤怒	1.14	1.41	0.95	1.71	0.76	1.04	0.85	1.22
悲伤	2.67	2.10	2.67	3.20	7.10	5.67	2.20	2.76
σ 悲伤	1.37	1.21	1.26	1.27	1.09	1.63	1.03	1.20
平静	3.00	3.67	1.97	4.00	2.47	3.67	7.53	4.87
σ 平静	1.50	1.60	0.81	2.10	1.19	1.40	1.22	1.50

表4 语气词“啊”的四种情感语平均分

Table 4 Four kinds of emotional average speech point table of statement label designator 'a'

情感	语料							
	正常语料 快乐语句	特殊发音 快乐语句	正常语料 愤怒语句	特殊发音 愤怒语句	正常语料 悲伤语句	特殊发音 悲伤语句	正常语料 平静语句	特殊发音 平静语句
快乐	7.77	6.33	2.33	4.20	2.10	3.53	2.23	4.47
σ 快乐	1.01	1.68	1.15	1.53	0.80	1.28	1.01	2.62
愤怒	1.43	2.57	7.33	5.57	3.00	2.57	2.00	2.13
σ 愤怒	0.50	1.19	1.47	1.72	1.22	0.94	1.02	1.14
悲伤	2.23	3.43	3.90	3.67	7.27	5.00	2.47	2.57
σ 悲伤	1.11	1.51	1.64	1.58	1.30	1.74	1.04	1.57
平静	2.90	4.00	1.43	2.66	2.67	3.56	7.00	3.87
σ 平静	1.39	1.26	0.57	1.06	1.12	1.33	1.17	2.03

表5 语气词“哈”的四种情感语平均分

Table 5 Four kinds of emotional speech average point table of statement label designator 'ha'

情感	语料							
	正常语料 快乐语句	特殊发音 快乐语句	正常语料 愤怒语句	特殊发音 愤怒语句	正常语料 悲伤语句	特殊发音 悲伤语句	正常语料 平静语句	特殊发音 平静语句
快乐	7.00	5.33	2.81	3.67	2.17	3.19	3.57	3.76
σ 快乐	1.41	2.02	0.95	2.05	0.76	2.15	1.10	1.96
愤怒	1.19	2.1	6.93	4.05	2.05	2.29	1.57	1.67
σ 愤怒	0.48	1.27	0.86	2.11	0.84	1.29	0.81	0.84
悲伤	1.19	2.29	4.05	3.76	7.20	3.95	2.33	2.71
σ 悲伤	0.61	1.21	0.87	2.06	1.11	2.11	1.15	1.17
平静	3.36	3.33	3.14	2.81	3.59	4.29	7.13	3.71
σ 平静	0.93	2.22	0.93	1.72	1.04	1.88	1.03	1.76

图6表示同一情感下正常发音和特殊发音的平均分绝对差值,绝对差值越小表示同种情绪下两种语料所表达的情感更加相似,反之更加迥异。其中“吧”的愤怒情感语句中愤怒的绝对差值为3.34分,说明在愤怒情感下的特殊发音感受最不明确,听音人很难辨别说话人是否在表示愤怒。四类语气词中特殊发音愤怒语句的愤怒评分都要低于快乐的评分,并且听音人判断出有选项以外的情绪,说明特殊发音弱化了愤怒的表现。“吧”的悲伤情感语句中悲伤的绝对差值为4.27分,在四种情感语句中,悲伤语句相对应的悲伤情感绝对差值最高,说明悲伤的特殊发音的情感最难辨认。“吗”、“啊”、“哈”的平静情感语句相对应的平静绝对差值分别为2.66分、3.13分、3.42分,在四种情感语句相应的情感评分中都处于最高,且四类语气词悲伤的特殊发音对快乐的评分都高于正常发音,说明特殊发音不能让听音人认可平静的情感,听音人给出说话人是在表达开心的判断,说明四类语气词中平静的特殊发音对平静情感的认同度很低。在四类语气词中愤怒、悲伤、平静三种情感语句中,相应的情感评价绝对差值都高于快乐的情感语句,说明特殊发音对愤怒、悲伤、平静的认同度都较低。

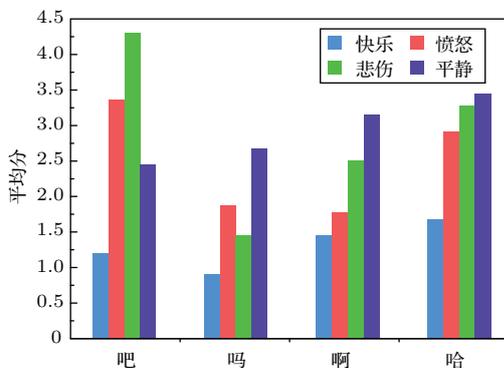


图6 四类语气词中四种情感语句相对应的情感绝对差值

Fig. 6 The absolute difference of emotion corresponding to four kinds of emotional sentences in four kinds of statement label designator

结合表2~表5和图6发现,四类语气词中快乐语句相对应的快乐的绝对差值分别为1.2分、0.9分、1.44分、1.67分,在四种情感语句相应的情感评分中最低,说明快乐的特殊发音所表达的情感在四种情感语句中认可度最高,虽然四类语气词特殊发音快乐的评分都低于正常发音,可能弱化了情感的表达,

但是都高于其他三种特殊发音的情感评分,听音人认为说话人可能想通过特殊的方式表达自己更加高兴的状态,提高了对快乐的认同度。说明该种特殊发音方式适用于表达快乐的情感。

3.3 情绪反馈实验

实验将观察特殊发音应用到不同情感语句中对听音人主观感受的影响,判断听音人自身产生的情绪反馈。当前广泛应用二维激活-效价情感空间模型描述人的情感^[15-17]。纵轴激活维度(Arousal)反映说话人的生理反应的激烈度,如听见情感语句后情绪激动的程度。横轴效价维度(Valence)反映对事物的正负面评价,如听见情感语句后是否感到有趣。播放PPT进行实验,答题阶段题目为评分题,分值为1~9分,听音人从激活维度和效价维度上对两种语料的主观感受进行评分。

图7为四类语气词的四情感语音的激活-效价空间的平均值,评价听音人的自身感受。黑色正方形为听音人正常发音组的评价,红色三角形为特殊发音组的评价。正常发音组中评价大都集中在第二、第三、第四象限,听音人在听到语音后更多地表示出焦虑、厌烦、乏味、难过等低激活度和低效价度的情感,说明听音人比较认同正常语音情感的表现,没有什么有趣的地方可以提高关注度。

结合表2~表5、图6和图7发现,特殊发音组快乐和宁静语句的激活-效价空间评价主要集中在第一象限,听音人听见快乐的语音原本感到很平常,突然末尾语气词发出特殊发音,产生陌生感,通过一定的心理作用,认为说话人表达快乐的特殊方法很有趣,从陌生转为趣味,从而使该段语音让听音人产生有趣、开心、更加愉悦等正向情绪。快乐的特殊发音组的正向评价普遍高于其他三组特殊发音,通过特殊发音的表现,听音人认为说话人想表达出更加愉悦、得意的情感,增加了语句内容的趣味性,使听音人产生乐趣。愤怒原本是负面情绪的一种,在特殊发音的表现下,原本负向的情绪让听音人瞬间产生陌生感,认为一个愤怒的人更像是在搞怪,对负面情绪的认知产生较大的反差,让听音人同样产生较高的正向情绪反应。从搞怪、取悦听音人的角度来看,愤怒的特殊发音能明显改变愤怒的负面情感表现,从而达到取悦听众的效果。平静的语句本身表达的情感是中性的,在评价中不易被其他情感因素所干扰,特殊发音对语句情感的表达也会有一定的影响,说话人同样对特殊发音语句产生一种陌

生感,通过心理作用使听音人产生愉悦的感觉。而悲伤的特殊发音语句对听音人产生更多的如平静、满足等激活度较低情感,该种音调的特殊发音不适合于表达悲伤和平静两种情感,虽然音调的陌生感会让人产生小的兴趣,但更多的只会让听音人感到奇怪、乏味。实验可以说明人对特殊发音的理解是

通过特定情感引导,从对该音调产生陌生感逐渐转向认可的过程。两组试验中快乐的语音评价表中都能较明确判断出相应的情感,并且在特殊发音语句的激活-效价空间中听音人产生比较强烈的正向情绪,说明该种特殊发音可能比较适用于表现高兴的情感。

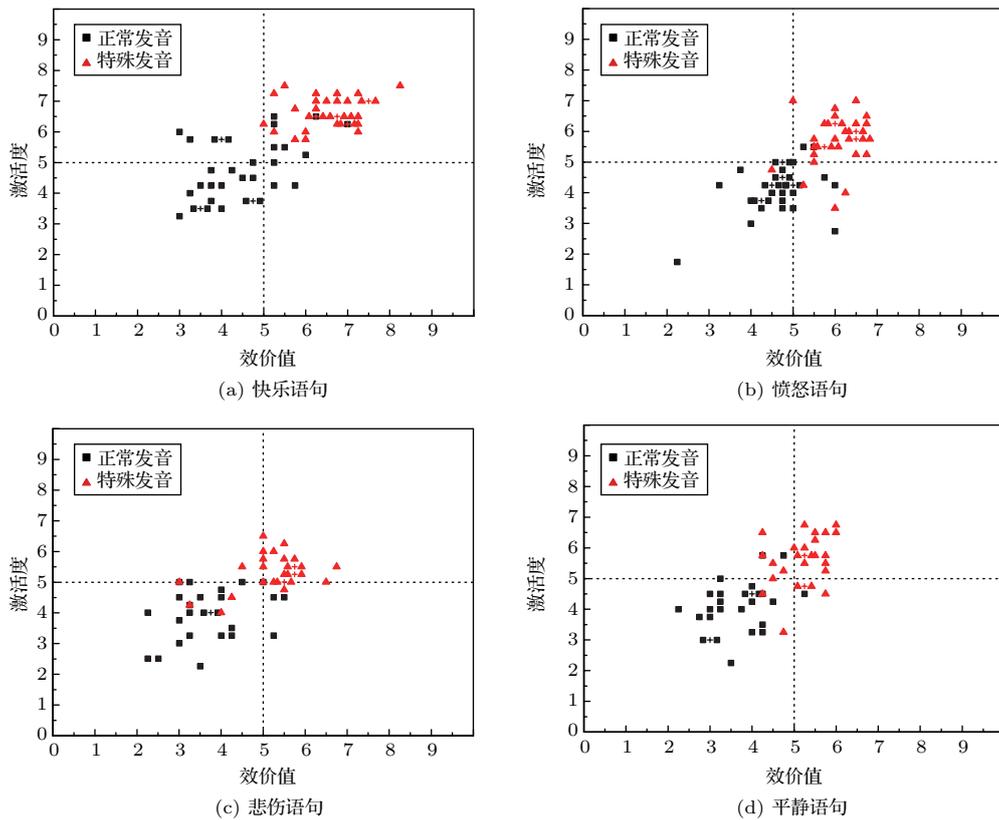


图7 四类语气词的激活效价空间评分

Fig. 7 Four kinds of statement label designator emotional speech arousal-valence space average

4 结论

论文从声音的基频角度合成特殊发音并应用到句子的语气词中,进行了情感语音的听辨实验,采用主观量表评价特殊语音的情感表现能力。通过对照实验,分析实验数据之后发现,特殊发音适用于表达高兴的情感,并能使听音人产生比较强烈的正向情绪反应。在常规语句中合成特殊发音语气词的方法能显著提高听音人的愉悦度。这为特殊发音应用于情感语音交互领域提供了可能。

本研究虽达到了预期效果,但仍存在一些不足。本文仅从基频的角度研究了特殊发音的情感表现,就情感表达而言,时长、能量变化也是韵律的重要

方面,它们同样对情感表现有很大的影响,后续工作将从多维特征空间研究语音信号,提高实验结论的可靠性。

参 考 文 献

- [1] 齐沪扬. 情态语气范畴中语气词的功能分析[J]. 南京师范大学文学院学报, 2002(3): 141-152.
Qi Huiyang. Analysis of the function of modal particles in modal category[J]. Journal of School of Chinese Language and Culture Nanjing Normal University, 2002(3): 141-152.
- [2] 谢群. 语气词的主观意义研究——语气词主观意义研究系列之一[J]. 外语学刊, 2015(6): 75-79.
Xie Qun. A study on the subjective meaning of modal words—A serial study on the subjective meaning of

- modal words(I)[J]. *Foreign Language Research*, 2015(6): 75-79.
- [3] 李云凤. 现代汉语语气词“哦、哟、咯”的话语功能及其主观性研究[D]. 合肥: 安徽大学, 2017.
- [4] 孙锡信. 近代汉语语气词[M]. 北京: 语文出版社, 1999.
- [5] 惠琳, 俞一彪. 短时频谱通用背景模型群联合韵律的年龄语音转换[J]. *声学学报*, 2017, 42(6): 762-768.
Hui Lin, Yu Yibiao. Voice conversion of different ages using universal background model groups of short-time spectra and prosodic features[J]. *Acta Acustica*, 2017, 42(6): 762-768.
- [6] 刘振焘, 徐建平, 吴敏, 等. 语音情感特征提取及其降维方法综述[J]. *计算机学报*, 2017, 40: 1-22.
Liu Zhentao, Xu Jianping, Wu Min, et al. Review of emotional feature extraction and dimension reduction method for speech emotion recognition[J]. *Chinese Journal of Computers*, 2017, 40: 1-22.
- [7] Fujisaki H. Prosody, models, and spontaneous speech[C]. *Computing prosody*, New York, USA: Springer-Verlag, 1997: 27-42.
- [8] Titze I R. Principles of voice production[M]. Iowa City, IA, USA: National Center for Voice and Speech, 2000.
- [9] 张杰, 龙子夜, 张博, 等. 语音信号处理中基频提取算法综述[J]. *电子科技大学学报*, 2010, 39(S1): 99-102, 126.
Zhang Jie, Long Ziye, Zhang Bo, et al. A summarize of pitch detection algorithmic in speech signals processing[J]. *Journal of University of Electronic Science and Technology of China*, 2010, 39(S1): 99-102, 126.
- [10] 高丽. 统计参数语音合成中的基频建模与生成方法研究[D]. 合肥: 中国科学技术大学, 2015.
- [11] 汤国祥. 机器人语音变调系统设计与实现[D]. 广州: 广东工业大学, 2013.
- [12] 张晓蕊. 语音变调算法研究及其在语音合成中的应用[D]. 济南: 山东大学, 2011.
- [13] 王治平, 赵力, 邹采荣. 基于基音参数规整及统计分布模型距离的语音情感识别[J]. *声学学报*, 2006, 31(1): 28-34.
Wang Zhiping, Zhao Li, Zou Cairong. Emotional speech recognition based on modified parameter and distance of statistical model of pitch[J]. *Acta Acustica*, 2006, 31(1): 28-34.
- [14] 高和. 基频与音分的通向公式[J]. *乐器*, 1982, 44(5): 21.
- [15] Tato R, Santos R, Kompe R, et al. Emotional space improves emotion recognition[C]. *ICSLP-INTERSPEECH*, Denver, USA, 2002: 2029-2032.
- [16] Cowie R, Douglas-Cowie E, Tsapatsoulis N, et al. Emotion recognition in human-computer interaction[J]. *IEEE Signal Processing Magazine*, 2001, 18(1): 32-80.
- [17] 谢波. 普通话语音情感识别关键技术研究[D]. 杭州: 浙江大学, 2006.