

纪念马大猷先生诞辰 110 周年

## 大五人格对相信与不信语音声学特征的影响\*

胡 涵<sup>1,2</sup> 吴思衡<sup>1</sup> 顾文涛<sup>1†</sup> 叶名扬<sup>1</sup> 管欣怡<sup>1</sup>

(1 南京师范大学 文学院 南京 210097)

(2 麦吉尔大学 沟通科学与障碍学院 加拿大蒙特利尔 H3A1G1)

2024 年 10 月 29 日收到

2024 年 12 月 12 日定稿

**摘要** 为了考察汉语表达相信/不信态度的语音声学表现,以及说话人的大五人格维度(神经质、外倾性、开放性、宜人性、尽责性)对声学特征的影响,研究设计了情景交际语料,诱导 47 名大学生分别用相信和不信的态度说出目标句,用大五人格量表 NEO-FFI 测量被试的人格特质。基于 23 个声学参数,构建 6 种有监督分类器,对相信和不信语音做自动分类,获得的最髙识别率为 0.76。根据边际贡献值选出 8 个重要声学特征做后续分析。冗余分析和线性混合模型的结果显示,不信比相信的语音有较高的基频均值、基频标准差、基频最小值、基频最大值、谐波比,以及较低的语速、谐波差 H1-H2、基频扰动。同时,相信和不信语音的声学特征差异受到人格维度的调节:神经质调节基频均值、基频最小值、基频标准差;外倾性调节基频均值和谐波比;开放性调节基频标准差。线性混合模型树的结果进一步显示,大五人格整体上显著影响了相信和不信语音在基频均值、基频标准差、语速上的差异。

**关键词** 大五人格, 相信, 态度语音, 声学特征

PACS: 43.66, 43.70

DOI: 10.12395/0371-0025.2024325

CSTR: 32049.14.11-2065.2024325

## Effects of the Big Five personality on acoustic characteristics of the speech conveying belief and disbelief

HU Han<sup>1,2</sup> WU Siheng<sup>1</sup> GU Wentao<sup>1†</sup> YE Mingyang<sup>1</sup> GUAN Xinyi<sup>1</sup>

(1 School of Chinese Language and Literature, Nanjing Normal University Nanjing 210097)

(2 School of Communication Sciences and Disorders, McGill University Montreal, Canada H3A1G1)

Received Oct. 29, 2024

Revised Dec. 12, 2024

**Abstract** To investigate the acoustic characteristics of Putonghua speech conveying belief and disbelief, and the effects of the Big Five personality dimensions (Neuroticism, Extroverts, Openness, Agreeableness, Conscientiousness) on acoustic distinctions between these two attitudes, the scenarios were designed to elicit utterances in both tones of belief and disbelief from 47 participants. Their personality dimensions were measured using the Neuroticism Extraversion Openness Five-Factor Inventory (NEO-FFI). With six supervised classifiers, speech conveying belief and disbelief was identified on the basis of 23 acoustic features, achieving a maximum recognition accuracy of 0.76. Based on the SHapley Additive exPlanation (Shap) values, eight important acoustic features were selected for subsequent analyses. Redundancy analysis and linear mixed model showed that, compared to belief speech, disbelief speech exhibited higher mean, standard deviation, minimum, and maximum values of fundamental frequency, as well as harmonic-to-noise ratio. In addition, disbelief speech showed lower speech rate, harmonic difference (H1-H2), and Jitter. These acoustic distinctions, however, were moderated by the personality dimensions: Neuroticism affected the mean, minimum and standard deviation of fundamental frequency, while Extroversion influenced the mean fundamental frequency and harmonic-to-noise ratio, and Openness impacted the standard deviation of fundamental frequency. Linear mixed model trees further revealed that the

\* 国家社会科学基金重大项目(13&ZD189)、国家留学基金、江苏省研究生科研与实践创新计划项目(18120000302320)和南京师范大学博士学位论文优秀选题资助计划项目(18120000302418)资助

† 通讯作者: 顾文涛, wtgu@njnu.edu.cn

Five Big personality traits significantly modulated the differences in the mean fundamental frequency, the standard deviation of fundamental frequency, and the speech rate between the utterances conveying belief and disbelief.

**Keywords** Big Five personality, Belief, Attitudinal speech, Acoustic characteristics

## 引言

语音除了承载语言信息外,也传递了丰富的副语言信息,如说话人的态度、意图等<sup>[1-3]</sup>。很多学者研究了传递交际态度的语音。例如, Aubergé 团队先后定义并考察了 12 至 19 种态度,对英语、法语、日语、汉语等多种语言做了声学分析和感知实验<sup>[4,5]</sup>。另有一些学者则采用了正反对立的方法定义态度对: Pell 团队针对英语,系统考察了热情/冷漠、礼貌/粗鲁等若干对态度语音的声学关联物<sup>[6-9]</sup>; Gu 等针对汉语,着重考察了友好/敌对、礼貌/粗鲁、褒扬/贬讽等若干对态度语音的韵律特征<sup>[2,3]</sup>,发现不同态度的语音在时长和基频特征上有系统的差异,同时句重音的分布也会发生变化。

类似地,“相信”(belief)和“不信”(disbelief),也是言语交际中常用的相互对立的态度。Anderson<sup>[10]</sup>将人际互动过程中为了促使话者继续话题而向其提供反馈或激励的行为定义为“反馈行为”。Gardner<sup>[11]</sup>指出言语反馈行为不仅反映了听者对话者继续话题的激励,还反映了听者对话者说话内容的评价。表达相信和不信的态度是一种重要的言语反馈行为。“相信”指个体接受某项陈述,将其视作真值;“不信”指个体拒绝某项陈述,将其视作假值<sup>[12]</sup>。厘清相信和不信的态度语音的声学差异,不仅可以深入理解言语交际过程,还有助于完善人机智能语音交互系统。

目前仅有少量的研究考察了相信/不信的态度语音。Manusov 和 Trees<sup>[13]</sup>基于主观听感发现,被试倾向于用讽刺语调表达不信。Levitan 和 Hirschberg<sup>[14]</sup>通过分析英文语料库中采访者对被采访者陈述的相信和不信回应发现,不信比相信的语音语速更快、音强中值更高。Yu 和 Levitan<sup>[15]</sup>考察了西班牙语和英语,发现两种语言均表现为,相信比不信语音有较长的时长、较低的音强中值、较小的基频斜率。Gander 等<sup>[16]</sup>基于 16 个声学特征构建随机森林模型,对相信和不信语音的辨识率高达 0.76。此外,Armstrong 等<sup>[17]</sup>发现,5 岁儿童基于句末音高升降的语音线索,判断发音人传达的态度是相信还是不信,识别率高达 0.80。以上研究文献表明,相信/不信语音在声学特征上存在差异。

人格是影响言语交际行为的重要因素<sup>[18-21]</sup>。人格指个体稳定的行为模式, Costa 和 McCare<sup>[22,23]</sup>在 5 个维度上定义人格特质,称为“大五人格”: (1) 神经质 (Neuroticism, N): 高神经质人群情绪波动较大,倾向于体验消极情绪(如愤怒、敌意、焦虑、脆弱、抑郁)<sup>[20,23,24]</sup>; (2) 外倾性 (Extraverts, E): 高外倾性人群善于交际,喜欢与他人相处,经常体验到积极情绪,有温暖、社交性、自信、活力等特征<sup>[20,23,24]</sup>; (3) 开放性 (Openness, O): 高开放性人群表现出想象力丰富、有创造力、脚踏实地等特征,他们乐于探索世界,体验新鲜事物<sup>[20,22]</sup>; (4) 宜人性 (Agreeableness, A): 高宜人性人群乐于接受他人的意见、重视他人的观点,有信任、直率、谦虚、善良和顺从等特征<sup>[20,22,24]</sup>; (5) 尽责性 (Conscientiousness, C): 高尽责性人群习惯于控制、调节、指导自身的冲动,表现出有能力、有秩序、责任感强、自律和深思熟虑等特征<sup>[20,22,24]</sup>。

有学者指出,内在的人格特质、外在的文化规范与情境等共同决定了个体的交际风格<sup>[19,25]</sup>。已有大量研究考察了人格维度与言语交际风格的相关关系。Leung 等<sup>[18]</sup>发现高外倾性和高开放性人群的言语表达更加直接、开放和准确。Ahmed 等<sup>[20]</sup>使用交际风格量表 (Communication Style Inventory) 测量了大学生群体交际风格的 6 个潜在维度: 表达性 (健谈、幽默、随意), 准确性 (考虑周到、语言简洁、内容充实、条理清晰), 语言攻击性 (易怒、强硬、使用贬低或攻击言辞、缺乏情感支持), 好奇性 (思维开放、充满好奇、善于辩论), 情绪性 (沟通中容易流露感伤、焦虑、紧张情绪, 采取防御性态度), 印象操控性 (奉承、有魅力、难以捉摸、隐瞒信息)。结果发现, 外倾性与表达性呈正相关, 神经质与语言攻击性、情绪性、印象操控性呈正相关, 尽责性与准确性呈正相关, 开放性与准确性、语言攻击性、好奇性呈正相关, 而宜人性与交际风格之间的相关性不显著<sup>[20]</sup>。

语音是言语交际的物质外壳,已有大量研究考察了人格维度对语音声学特征的影响<sup>[26]</sup>。Aronovitch<sup>[27]</sup>要求被试仅凭语音线索对话者的人格维度评分,发现外倾性与男性发音人的基频和能量范围呈正相关,与女性发音人的基频和能量均值呈正相关。Song 等<sup>[28]</sup>通过考察韩国人在正式交际场合中各人格维度的声学关联物发现: 神经质与基频均值仅在女性发音人中呈正相关; 外倾性与基频范围呈正相关, 并且在女

性发音人中与基频均值正相关、与共振峰间距负相关;开放性与基频范围呈正相关,并且在女性发音人中与共振峰间距负相关。Aronovitch<sup>[27]</sup>和Michalsky等<sup>[29]</sup>都揭示了女性发音人的语速和外倾性之间的关系,但结论相反:前者发现语速和外倾性呈正相关,而后者发现两者呈负相关。Gocsál<sup>[30]</sup>考察了男性发音人的语音声学特征与人格维度间的关系,指出语速与外倾性和开放性均呈正相关。Michalsky等<sup>[29]</sup>还发现语速与神经质之间的负相关关系、短语时长与尽责性之间的正相关关系。音质特征与人格维度之间也有相关性,如嘎裂声与较低的外倾性、宜人性、尽责性、神经质有稳定的联系<sup>[31]</sup>。此外,基于声学特征对5个人格维度(以中值为界分为高、低两个水平)做自动分类发现,外倾性(0.73)和尽责性(0.73)的识别率最高,其次为神经质(0.68)和宜人性(0.63),开放性的识别率最低(0.60)<sup>[32]</sup>。

此外,近年有研究开始探讨人格特质与情感类型对语音的综合作用。例如,盛晨<sup>[33]</sup>探索了外倾性与宜人性对汉语态度语音(包括积极态度与消极态度)声学特征的影响;此外,胡涵和顾文涛<sup>[34,35]</sup>揭示了依恋风格对亲密态度语音以及四种基本情绪语音的调节作用。为了深入研究人格特质对态度语音的作用规律,本文从大五人格出发,探究表达相信/不信态度的汉语语音声学特征,考察外倾性、神经质、宜人性、尽责性、开放性这5个人格维度对相信/不信语音声学特征差异模式的调节作用。

## 1 方法

### 1.1 实验材料

**人格测量问卷:**使用大五人格量表(Neuroticism Extraversion Openness Five-Factor Inventory, NEO-FFI)中文版<sup>[36]</sup>测量被试的人格维度。NEO-FFI由5个分量表构成,分别对应神经质(如:有时候我感到愤怒,充满怨恨)、外倾性(如:我喜欢周围有很多朋友)、开放性(如:我喜欢培养和发展新的爱好)、宜人性(如:我尽量对每一个遇到的人彬彬有礼,非常客气)、尽责性(如:我比较擅长为自己安排好做事进度,以便按时完成任务),每个分量表包含12个题项。每个题项均采用5度制计分(1:非常不符合,5:非常符合)。最终,计算12个题项的总分作为人格维度分数。用Cronbach's  $\alpha$ 系数检验样本量表题项的同质性,结果表明神经质的Cronbach's  $\alpha$ 系数为0.86,外倾性为0.84,开放性为0.85,宜人性为0.73,尽责性为0.84。

**产出语料:**采用情景诱导式语料。设计11个语义中性的陈述句作为目标句(音节数:均值 $M=8.58$ ,标准差 $SD=1.83$ ,范围 $range=6\sim 12$ )。每个目标句设计两个语境,分别诱发相信和不信态度。诱发相信态度的语境强调话者已知信息与传递信息相符,诱发不信态度的语境强调话者已知信息与传递信息不符。为了贴近日常言语交际,设计了3~4个交际话轮,由实验助理充当被试的交际对象,共同完成对话,并引导被试自然地说出目标句。以下是目标句为“外面在下大雨”的语境和对话脚本,目标句由被试A说出。

**【相信】**昨天晚上看天气预报说今天是晴天,但早上A打开窗户看到外面在下大雨,舍友B在收拾东西准备一会儿出门。

B:等一下我还要和小李去剪头发。

A:剪头发?那你们只能在附近了,记得带伞。

B:带伞?不用吧,天气预报说今天是晴天啊。

A:不是,我早上打开窗户看到的,外面在下大雨。

**【不信】**A昨天晚上看天气预报说今天是晴天。傍晚,舍友B和A聊起晚饭吃什么。

B:等会儿晚饭咱们只能点外卖吃了。

A:怎么了?为什么只能点外卖吃啊?

B:你也不看看,今天外面下大雨呢!

A:不是吧,外面在下大雨,天气预报不是说晴天吗?

### 1.2 被试选取

使用仿真算法,估计在显著水平为0.05(双尾)、功效值大于80%时态度(相信/不信)效应显著所需的样本量。基于预实验10名被试的数据,分别以每个声学特征为因变量构建线性混合模型(Linear Mixed Model, LMM),态度为固定效应,被试和目标句为随机效应,做1000次仿真模拟。计算所有声学特征上态度因子对应功效值的算术平均后发现,样本量为40的功效值为93%。用R 4.3.1<sup>[37]</sup>的MixedPower程序包<sup>[38]</sup>做仿真分析。

共招募47名无言语听力疾病史且普通话流利的被试(22男25女),平均年龄 $M=21.55$ 岁( $SD=2.59$ , $range=18\sim 27$ ),均为在校大学生或研究生。被试的背景信息和NEO-FFI测量的人格维度分数见表1。采用Wilcoxon秩和检验比较男性和女性在年龄和人格维度分数上的差异,结果显示性别效应均不显著( $ps>0.1$ ),表明年龄和人格维度分数在两性间分布均衡,可排除性别对后续统计结果的影响。



表 1 被试的背景信息及人格维度分数

变量	描述统计量	总体, $N = 47$	女性, $N = 25$	男性, $N = 22$	$p$ -value
年龄	$M(SD)$	21.55(2.59)	21.92(2.45)	21.14(2.75)	0.395
	Min~Max	18.00~27.00	18.00~25.00	19.00~27.00	
神经质(N)	$M(SD)$	35.30(8.42)	34.88(8.09)	35.77(8.94)	0.790
	Min~Max	22.00~52.00	23.00~48.00	22.00~52.00	
外倾性(E)	$M(SD)$	36.81(7.63)	36.76(8.02)	36.86(7.36)	> 0.999
	Min~Max	24.00~50.00	24.00~50.00	24.00~49.00	
开放性(O)	$M(SD)$	44.23(8.35)	43.08(8.66)	45.55(7.99)	0.359
	Min~Max	23.00~59.00	23.00~59.00	30.00~57.00	
宜人性(A)	$M(SD)$	41.15(6.51)	40.92(6.27)	41.41(6.91)	0.661
	Min~Max	24.00~52.00	28.00~52.00	24.00~50.00	
尽责性(C)	$M(SD)$	44.83(7.15)	45.48(5.08)	44.09(9.02)	0.685
	Min~Max	27.00~57.00	37.00~55.00	27.00~57.00	

注:  $M$ 表示均值,  $SD$ 表示标准差, Min表示最小值, Max表示最大值,  $N$ 表示样本量。  $p$ -value为Wilcoxon秩和检验得到的概率值。

### 1.3 实验过程

录音在专业录音棚内进行。语音采集设备为 Apogee BOOM 声卡和 AKG 头戴式传声器,采集到的语音信号存为 WAV 文件(采样率 44.1 kHz,精度 16 bit)。使用 Eprime 3.0<sup>[39]</sup>设计实验程序,并在 Philips 220SW 显示屏(分辨率: 1280 × 1024)上呈现。实验分为三个阶段:问卷填写、实验准备、语音实验。实验持续约 1 小时。

问卷填写阶段。被试在阅读并签署知情同意后,填写个人基本信息和 NEO-FFI 线上问卷。人格维度得分对被试保密,防止被试根据某类人格的刻板印象做语音表达。

实验准备阶段。实验助理将被试引入录音棚,安置在距离显示屏约 70 cm 的座椅上,并调整座椅高度确保被试平视屏幕中央。用头戴式传声器采集语音,传声器距被试嘴部约 10 cm,实验过程中被试不得触碰传声器。被试有充分的时间熟悉语料。

语音实验阶段。按照角色扮演的对话脚本,被试感受角色的心理状态,与实验助理充当的对话者,以日常方式自然地交流。屏幕上首先呈现 500 ms 注视点,提示被试集中注意力。其次,呈现语境与交际话轮,目标句用粗体突出显示。被试先经练习阶段充分熟悉实验流程后,进入正式实验。如果被试有错读和漏读,或对话语表达效果不满意,可重新开始对话。所有目标句均由被试产出。不同目标句与其对应的交际情景随机呈现,同一目标句诱发相信和不信态度的情景顺次呈现。最终,共采集 47 被试 × 2 态度(相信/不信) × 11 目标句 = 1034 个语音文件。在去除 3 个有缺失值的样本后,得到 1031 个语音样本用于数据分析。

### 1.4 声学参数

参考以往有关情感<sup>[14-16]</sup>和人格<sup>[27-35]</sup>研究采用的声学特征,采用 Praat 6.3<sup>[40]</sup>,对每个目标句提取 11 个韵律参数和 12 个音质参数:

**基频 (f0):** 声带振动的频率。采用自相关算法获取基频的时变曲线,并手动修正异常值。对内插和平滑后基频曲线上的每个基频值做半音转换,公式为  $12 \cdot \log_2(f0/f0ref)$ ,其中  $f0ref$  为参考值(设置为 50 Hz)。基于转换后的基频曲线,计算均值 ( $f0mean$ )、标准差 ( $f0std$ )、最小值 ( $f0min$ )、最大值 ( $f0max$ )、范围 ( $f0range$ )。

**音强 (Int):** 语音的强度。计算语音幅值的均方根得到音强的时变曲线。基于内插和平滑后的音强曲线计算均值 ( $Intmean$ )、标准差 ( $Intstd$ )、最小值 ( $Intmin$ )、最大值 ( $Intmax$ )、范围 ( $Intrange$ )。单位为 dB。

**语速 (SpRate):** 发音人每秒产生的音节数。计算方法为全句的音节数除以全句的时长。单位为音节数/秒。

**谱矩 (Spec):** 频谱能量的分布,包括谱重心 (Spec-cog) 和谱离散度 (Spec-std)。前者为一阶原点矩,后者为二阶中心矩。单位为 Hz。

**谐波差 (HD):** 不同频段谐波能量的差异,反映声门的收紧程度。提取经过共振峰带宽修正的低频段谐波差 H1-H2、H2-H4,高频段谐波差 H2K-H5K (2000 Hz 和 5000 Hz 谐波能量的差值),以及 H1 与三个共振峰能量间的差值 H1-A1、H1-A2、H1-A3。谐波差值越大,声门的收紧度越低<sup>[41]</sup>。

**嗓音规则性参数:** 包括基频扰动 (Jitter)、振幅扰动 (Shimmer)、谐噪比 (HNR)、平滑倒谱峰值系数 (CPPS)。Jitter 和 Shimmer 分别表示相邻周期基频和

振幅的变异度,用百分数表示,数值越小,嗓音越规则。HNR表示谐波成分与噪声成分的能量比值,单位为dB,数值越大,嗓音越规则。CPPS指平滑倒谱能量的峰值与峰值下回归线之间的距离,单位为dB,数值越大,嗓音越规则。

## 1.5 统计分析

首先,用有监督分类器对相信/不信的态度语音做自动分类,考察两类态度语音在声学空间中的分布差异。基于分类模型计算每个声学特征对模型输出的边际贡献(SHapley Additive exPlanation, Shap),选取贡献高于均值的特征做后续分析。

其次,基于选取的重要声学特征做冗余分析(Redundancy Analysis, RDA),揭示解释变量矩阵(人格维度)对响应变量矩阵(声学特征)方差的解释率。具体分为两步:(1)基于原始声学特征做RDA,分别考察人格维度和相信/不信态度对声学特征的影响。(2)基于相信/不信的态度语音声学特征的绝对差值做RDA,探究人格维度对两种态度语音声学模式差异的调节作用。同时,采用层次划分法<sup>[42]</sup>计算每个人格维度对多维声学特征方差解释率的独立贡献。

然后,由于RDA未对人格维度的调节作用做假设检验,所以对选取的重要声学特征逐一拟合线性混合模型(LMM),考察每个人格维度与相信/不信态度的二阶交互效应。

最后,由于受到样本量的限制,LMM只能分别考察单个人格维度对相信/不信语音声学特征的调节作用。因此,进一步采用线性混合模型树(LMMT),综合考察5个人格维度整体上对两种态度语音的调节作用。LMMT由两部分组成:(1)利用决策树算法,根据人格维度将声学数据分为多个同质的亚组,作为叶节点;(2)对每个叶节点上的数据做LMM,计算固定效应和随机效应。

使用R 4.3.1<sup>[37]</sup>完成上述统计分析。完整的语料与实验数据均存储在Open Science Framework(OSF)平台([https://osf.io/365hc/?view\\_only=273159dc8c53438a9cdb78d18f8389b6](https://osf.io/365hc/?view_only=273159dc8c53438a9cdb78d18f8389b6))。

## 2 结果

### 2.1 有监督分类器

基于1031个样本、23个声学特征,构建核K近邻模型(KKNNM)、灵活判别模型(FDM)、径向基支持向量机(RSVM)、随机森林(RF)、轻量级梯度提升

决策树(LGBDT),对语音表达的相信与否的态度做自动分类。为了提升模型的识别效果,将5个基模型3重10折交叉验证的预测值作为新特征,纳入线性惩罚模型,计算堆叠系数,构建堆叠模型(Stacks Model, SM)。为了消除个体差异的影响,每个声学特征以被试为单位转换为 $z$ 分数。数据集分为70%训练集(720个样本)和30%测试集(311个样本),训练集用于建模和调参,测试集用于评价模型的表现。采用模拟退火算法(迭代100次)和3重5折交叉验证法调参。使用准确率(Accuracy, ACC)和接受者工作特征曲线下面积(Area Under Curve, AUC)评估模型的识别效果。

如表2所示,5种基模型在测试集上的ACC都在0.71以上,显著高于0.50的机会水平(95%CI均不包含0.50),AUC都在0.79以上;其中,RSVM识别效果最优(ACC=0.76, AUC=0.84)。堆叠模型相比于除RSVM以外的基模型,AUC提升了2%~5%。

基于SM计算每个声学特征的Shap值,选出对相信/不信的判断有重要贡献的声学特征。如图1所示,贡献值高于绝对Shap均值的声学特征,由高至低排列为:f0mean, Jitter, HNR, f0std, f0max, f0min, SpRate, H2-H4。基于这8个声学特征做后续分析。

### 2.2 冗余分析

所有特征都在标准化(转换为 $z$ 分数)后做RDA。选取方差解释率占总体可解释方差的比例最高的两个维度(RDA1和RDA2)作三标图,如图2所示。图2(a)描述了人格维度和相信/不信语音在声学空间上的分布。不信比相信语音有较高的f0mean、f0min、f0max、HNR,较低的Jitter、H2-H4。人格维度对声学特征的方差解释率由高至低分别为:尽责性(37.12%)、外倾性(30.62%)、开放性(14.00%)、神经质(13.62%)、宜人性(5.12%)。尽责性与f0mean、f0std、f0max、f0min、SpRate呈正相关,与H2-H4、Jitter呈负相关。外倾性、开放性和神经质与H2-H4、Jitter呈正相关,与f0mean、f0min、f0max、HNR呈负相关。宜人性与HNR、f0min、f0mean呈正相关,与H2-H4、Jitter呈负相关。

图2(b)描述了人格维度对两种态度语音声学特征的绝对差值的影响。人格维度的方差解释率由高到低分别是:神经质(25.73%)、外倾性(25.48%)、开放性(22.58%)、尽责性(13.31%)、宜人性(12.98)。神经质与f0max、f0std、SpRate、Jitter呈正相关、与f0mean、f0min呈负相关。外倾性与f0mean、f0min、f0max、Jitter、HNR呈正相关,与SpRate、H2-H4的呈负相

表 2 有监督分类器的超参设置及分类效果

	KKNNM	FDM	RSVM	RF	LGBDT	SM
超参	kernel: inverse K: 9	nprune: 10 degree: 1	C: 15.7 sigma: 0.004	ntree: 1803 mtry: 3	ntree: 706, mtry: 9 tree_depth: 3 learn_rate: 0.004 lambda: 0.63, min_n: 10 loss_reduction: 0.003	lambda: 0.0001
ACC	0.71	0.75	0.76	0.76	0.75	0.75
95% CI	0.66~0.76	0.69~0.79	0.71~0.81	0.71~0.81	0.69~0.79	0.70~0.80
AUC	0.79	0.81	0.84	0.82	0.82	0.84
95% CI	0.74~0.84	0.77~0.86	0.79~0.88	0.77~0.87	0.77~0.87	0.79~0.88

注: kernel表示计算样本间距离权重的核函数, K表示用于预测类别的近邻点数量; nprune表示最终模型保留的特征数量, degree表示允许的最高交互阶数; C表示对误判样本的惩罚, sigma表示支持向量与决策边界的距离; ntree表示决策树的数量, mtry表示每次分裂基于的特征数量; tree\_depth表示树分裂的最大次数, learn\_rate表示提升算法的学习率, lambda表示L1正则化系数, min\_n表示节点所需的最小样本量, loss\_reduction表示分裂所需的最小损失。95% CI表示95%的置信区间。

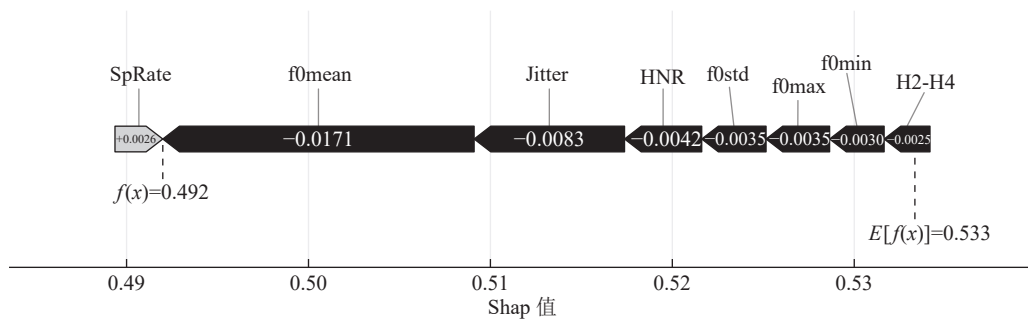


图 1 重要声学特征的 Shap 值 (仅展示 Shap 值高于绝对均值的 8 个声学特征。黑色底纹的箭头表示特征对模型输出的边际贡献为负, 灰色底纹的箭头表示特征对模型输出的边际贡献为正。  $f(x)$  为模型的预测值,  $E[f(x)]$  为基线模型的预测值。)

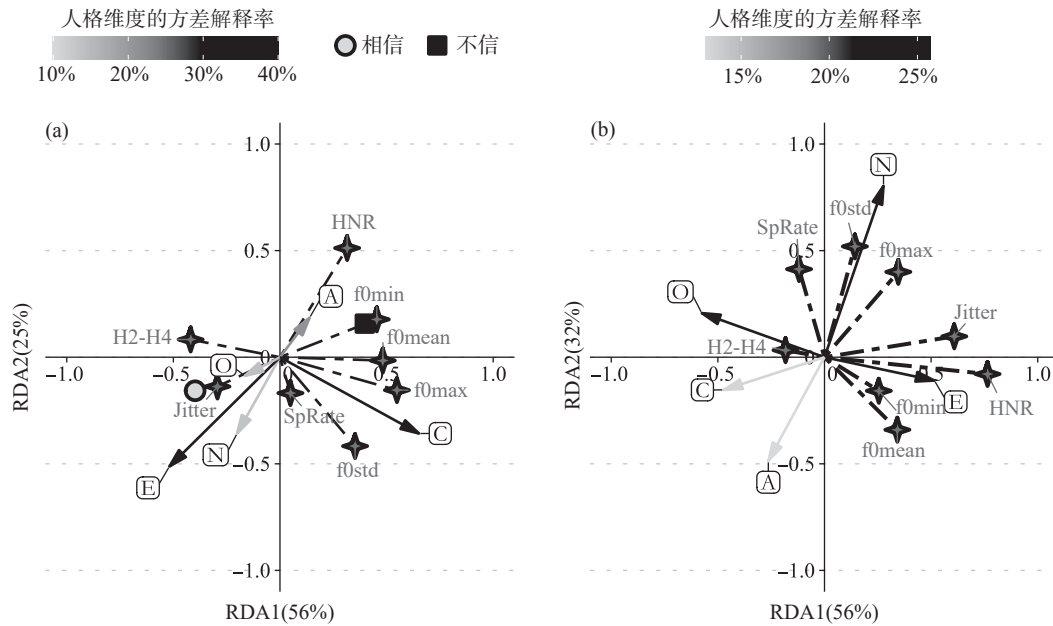


图 2 基于 RDA 的三标图 (a) 人格维度和相信/不信态度语音在声学空间中的分布; (b) 人格维度对两种态度语音声学特征的绝对差值的影响 (四角星号配以点划线表示声学特征向量; 实线箭头表示人格维度的向量, 箭头颜色的深浅表示每个人格维度对声学特征方差的解释率; 任意两个向量间夹角的余弦表示两个特征间的相关性; 相信和不信态度分别用灰色实心圆形和黑色实心方形表示, 它们在声学特征向量上的投影表示各自的声学特征取值; RDA1 和 RDA2 表示基于人格维度的线性组合构建的潜在维度, 百分数表示每个潜在维度的可解释方差占总体可解释方差的比例。)

关。开放性与 SpRate、H2-H4 呈正相关, 与 f0mean、f0min、f0max、Jitter、HNR 的呈负相关。尽责性与 H2-H4 呈正相关, 与 f0mean、f0std、f0min、f0max、Jitter、HNR 呈负相关。宜人性与 f0std、f0max、SpRate、Jitter、HNR 呈负相关。

### 2.3 线性混合模型

为了检验 5 个人格维度对相信/不信态度语音的调节作用是否有统计学意义, 基于 8 个重要的声学特征逐一拟合 LMM。固定效应为每个人格维度 (连续变量) 和态度 (离散变量: 相信/不信) 间的二阶交互效应, 随机效应为被试和目标句。为平衡模型的复杂度和拟合优度, 采用重要性排序和逐步剔除法优化模型<sup>[43]</sup>。采用似然比检验, 估计最简模型的主效应和交互效应。对态度效应做事后检验时, 将人格维度以  $M \pm SD$  作为高、低两个水平, 用 Cohen's  $d$  表示效应量; 对人格效应做事后检验时, 分别计算人格维度在相信/不信条件下对声学特征的预测斜率。为了便于模型解读, 对人格维度做中心化处理。如表 3 所示, 除 f0max 和 Jitter 外的 6 个声学特征上, 均发现人格维度与态度间有显著的交互作用。具体结果呈现如下。

**f0mean:** 神经质 ( $\chi^2(1) = 9.24, p = 0.002$ )、外倾性 ( $\chi^2(1) = 4.72, p = 0.030$ )、尽责性 ( $\chi^2(1) = 6.70, p = 0.010$ ) 与态度间的交互效应显著。对态度做事后检验发现, 无论神经质、外倾性、尽责性的高低, 不信的 f0mean 均显著高于相信 (低神经质:  $b = 1.63, CI [1.37, 1.89], p < 0.001, d = 0.34$ ; 高神经质:  $b = 0.97, CI$

$[0.68, 1.25], p < 0.001, d = 0.20$ ; 低外倾性:  $b = 1.10, 95\% CI [0.83, 1.37], p < 0.001, d = 0.23$ ; 高外倾性:  $b = 1.51, CI [1.27, 1.75], p < 0.001, d = 0.32$ ; 低尽责性:  $b = 1.94, CI [1.44, 2.43], p < 0.001, d = 0.40$ ; 高尽责性:  $b = 1.39, CI [1.21, 1.57], p < 0.001, d = 0.29$ ), 但两种态度分别在低神经质、高外倾性、低尽责性人群中有较大的差异。对人格做事后检验未发现显著效应 ( $ps > 0.1$ )。

**f0std:** 神经质 ( $\chi^2(1) = 6.93, p = 0.008$ ) 和开放性 ( $\chi^2(1) = 7.21, p = 0.007$ ) 与态度间的交互效应显著。对于神经质与态度的交互, 态度的事后检验发现, 无论神经质高低, 不信的 f0std 均显著高于相信 (低神经质:  $b = 0.18, CI [0.00, 0.37], p = 0.050, d = 0.12$ ; 高神经质:  $b = 0.56, CI [0.36, 0.76], p < 0.001, d = 0.37$ ), 但高神经质人群在两种态度上的差异更大; 而神经质的事后检验发现, 神经质对不信条件下的 f0std 有显著的正向预测作用 ( $slope = 0.04, CI [0.01, 0.08], p = 0.014$ ), 但对相信条件下 f0std 的预测作用不显著 ( $slope = 0.02, CI [-0.01, 0.06], p = 0.211$ )。对于开放性与态度的交互, 态度的事后检验发现, 高开放性人群不信的 f0std 显著高于相信 ( $b = 0.36, CI [0.00, 0.49], p < 0.001, d = 0.24$ ), 但在低开放性人群中两种态度的差异不显著 ( $b = 0.02, CI [-0.28, 0.33], p = 0.890, d = 0.01$ ); 而开放性的事后检验未发现显著效应 ( $ps > 0.1$ )。

**f0min:** 神经质 ( $\chi^2(1) = 7.44, p = 0.006$ ) 和尽责性 ( $\chi^2(1) = 6.29, p = 0.012$ ) 与态度间的交互效应显著。对态度做事后检验发现, 低神经质人群不信的

表 3 每个人格维度与相信/不信态度在语音声学特征上的交互效应

	神经质(N)	外倾性(E)	开放性(O)	宜人性(A)	尽责性(C)
f0mean	低: 不信 > 相信 <sup>††</sup> 高: 不信 > 相信 <sup>†</sup>	低: 不信 > 相信 <sup>†</sup> 高: 不信 > 相信 <sup>††</sup>	n.s.	n.s.	低: 不信 > 相信 <sup>††</sup> 高: 不信 > 相信 <sup>†</sup>
f0std	低: 不信 > 相信 <sup>†</sup> 高: 不信 > 相信 <sup>††</sup> 相信: n.s.; 不信: +	n.s.	低: n.s. 高: 不信 > 相信	n.s.	n.s.
f0min	低: 不信 > 相信 高: n.s.	n.s.	n.s.	n.s.	低: 不信 > 相信 <sup>††</sup> 高: 不信 > 相信 <sup>†</sup>
SpRate	n.s.	n.s.	n.s.	低: 不信 < 相信 <sup>††</sup> 高: 不信 < 相信 <sup>†</sup>	n.s.
H2-H4	n.s.	n.s.	n.s.	低: n.s. 高: 不信 < 相信	n.s.
HNR	n.s.	低: 不信 > 相信 <sup>†</sup> 高: 不信 > 相信 <sup>††</sup> 相信: -; 不信: n.s.	n.s.	n.s.	n.s.

注: 高、低表示各人格维度的高、低两个水平, 分别设为  $M + SD$ 、 $M - SD$ 。当相信和不信语音的声学特征在人格维度高、低两个水平上都有显著差异时, <sup>††</sup>与<sup>†</sup>分别表示其中差异值较大的和较小的。+ 和 - 分别表示人格维度在相信或不信条件下对声学特征的正向与负向预测作用。n.s.表示无显著效应。另两个声学特征f0max和Jitter, 由于人格维度与相信/不信态度之间的交互效应都不显著, 因此未在表中列出。



f0min 显著高于相信 ( $b = 1.44, CI [0.87,2.01], p < 0.001, d = 0.29$ ), 但高神经质人群在两种态度上的差异不显著 ( $b = 0.16, CI [-0.79,0.47], p = 0.612, d = 0.03$ )。无论尽责性高低, 不信的 f0min 都显著高于相信 (低尽责性:  $b = 2.17, CI [1.07,3.27], p < 0.001, d = 0.43$ ; 高尽责性:  $b = 0.99, CI [0.59,1.39], p < 0.001, d = 0.20$ ), 但低尽责性人群在两种态度上的差异更大。对神经质和尽责性做事后检验均未发现显著效应 ( $ps > 0.1$ )。

**SpRate:** 宜人性与态度间的交互效应显著 ( $\chi^2(1) = 5.50, p = 0.019$ )。对态度做事后检验发现, 无论宜人性高低, 不信的 SpRate 显著慢于相信 (低宜人性:  $b = -0.36, CI [-0.49, -0.24], p < 0.001, d = -0.46$ ; 高宜人性:  $b = -0.21, CI [-0.28, 0.15], p < 0.001, d = -0.27$ ), 但低宜人性在两种态度上的差异更大。对宜人性做事后检验未发现显著效应 ( $ps > 0.1$ )。

**H2-H4:** 宜人性与态度间的交互效应显著 ( $\chi^2(1) = 5.09, p = 0.024$ )。对态度做事后检验发现, 高宜人性人群不信的 H2-H4 显著小于相信 ( $b = -0.47, CI [-0.74,-0.19], p = 0.002, d = -0.17$ ), 但低宜人性人群在两种态度上的差异不显著 ( $b = -0.16, CI [-0.72,0.40], p = 0.577, d = -0.06$ )。对宜人性做事后检验未发现显著效应 ( $ps > 0.1$ )。

**HNR:** 外倾性与态度间的交互效应显著 ( $\chi^2(1) = 4.92, p = 0.027$ )。对态度做事后检验发现, 无论外倾性高低, 不信的 HNR 都显著高于相信 (低外倾性:  $b = 2.18, CI [0.41,3.95], p = 0.016, d = 0.18$ ; 高外倾性:  $b = 4.88, CI [3.29,6.47], p < 0.001, d = 0.41$ ), 但高外倾性人群在两种态度上的差异更大。对外倾性做事后检验发现, 外倾性对相信条件下的 HNR 有显著的负向预测作用 ( $slope = -0.39, CI [-0.66,-0.13], p = 0.004$ ),

但对不信条件下 HNR 的预测作用不显著 ( $slope = -0.22, CI [-0.48,0.05], p = 0.104$ )。

### 2.4 线性混合模型树

以不信/相信语音声学特征的差值作为结果变量, 5 个人格维度作为预测变量构建 LMMT, 进一步综合考察 5 个人格维度整体对相信/不信语音声学特征的影响。LMM 的固定效应为截距, 随机效应为被试和目标句。使用 T 检验比较叶节点之间的差异 (Benjamini-Hochberg 法修正  $p$  值), 效应量为 Cohen's  $d$ 。为了防止模型过拟合, 设置叶节点的最小样本量为 20, 并基于贝叶斯信息准则对树剪枝。如图 3 所示, 只在 f0mean、f0std、SpRate 上人格有显著效应。

**f0mean:** 所有叶节点的均值都显著大于 0 (Node3:  $M = 1.05, CI [0.46,1.63], p < 0.001$ ; Node5:  $M = 2.85, CI [2.14,3.56], p < 0.001$ ; Node6:  $M = 1.53, CI [0.77,2.30], p < 0.001$ ; Node7:  $M = 1.04, CI [0.60,1.48], p < 0.001$ ), 说明不信的 f0mean 显著高于相信。事后检验发现, Node5 (神经质  $\leq 32$ 、外倾性  $> 38$ 、尽责性  $\leq 50$ ) 的 f0mean 显著高于其余三个叶节点 (Node3:  $b = 1.80, CI [0.98,2.62], p < 0.001, d = 0.96$ ; Node6:  $b = 1.31, CI [0.35,2.28], p = 0.017, d = 0.70$ ; Node7:  $b = 1.80, CI [1.08,2.52], p < 0.001, d = 0.96$ )。

**f0std:** 除 Node3 ( $M = -0.31, CI [-0.71,0.09], p = 0.122$ ), 其余叶节点的均值都显著大于 0 (Node4:  $M = 0.58, CI [0.06,1.09], p = 0.028$ ; Node5:  $M = 0.51, CI [0.25,0.76], p < 0.001$ )。事后检验发现, Node3 (开放性  $\leq 39$ 、神经质  $\leq 41$ ) 的 f0std 显著小于其他叶节点 (Node4:  $b = -0.89, CI [-1.50,-0.28], p = 0.008, d = -0.62$ ; Node5:  $b = -0.82, CI [-1.23,-0.41], p < 0.001, d = -0.57$ )。

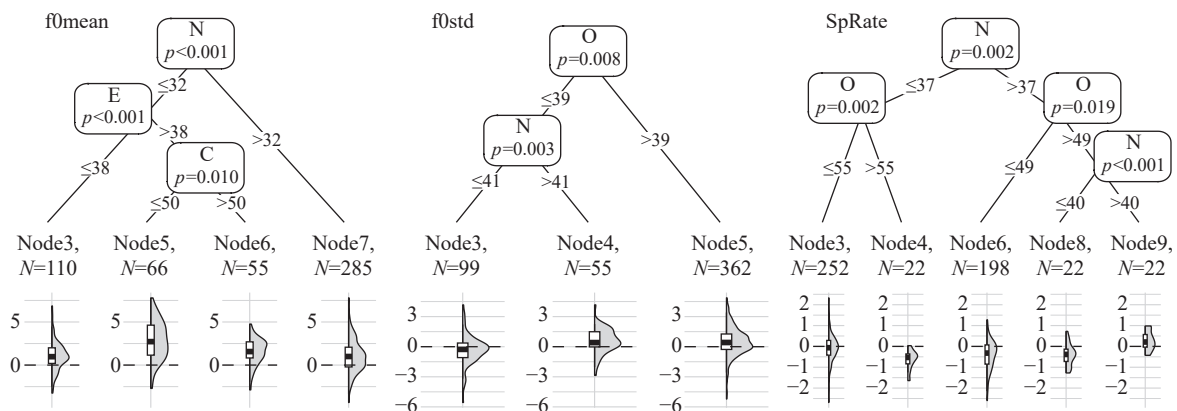


图 3 不同人格维度的组合对不信和相信语音声学特征差值的影响 (根节点和内部节点表示人格维度, 参数稳定性检验得到的  $p$  值标记在人格维度下方,  $p < 0.05$  表示节点变量与结果变量显著相关。叶节点显示不信和相信语音声学特征差值的分布, 每个叶节点的样本量用  $N$  表示。)



**SpRate:** Node4( $M = -0.64$ ,  $CI [-1.06, -0.22]$ ,  $p = 0.004$ ) 和 Node6( $M = -0.42$ ,  $CI [-0.58, -0.25]$ ,  $p < 0.001$ ) 的均值显著小于 0, Node8 的均值边缘显著小于 0( $M = -0.40$ ,  $CI [-0.82, 0.02]$ ,  $p = 0.063$ ), Node3( $M = -0.08$ ,  $CI [-0.23, 0.07]$ ,  $p = 0.301$ ) 和 Node9( $M = 0.28$ ,  $CI [-0.14, 0.70]$ ,  $p = 0.187$ ) 的均值与 0 没有显著差异。事后检验发现, Node3(神经质  $\leq 37$ 、开放性  $\leq 55$ ) 的 SpRate 显著高于 Node4( $b = 0.56$ ,  $CI [0.13, 0.99]$ ,  $p = 0.030$ ,  $d = 0.88$ ) 和 Node6( $b = 0.34$ ,  $CI [0.16, 0.52]$ ,  $p = 0.005$ ,  $d = 0.53$ )。Node9(神经质  $> 40$ 、开放性  $> 49$ ) 的 SpRate 显著高于 Node4( $b = 0.92$ ,  $CI [0.33, 1.50]$ ,  $p = 0.009$ ,  $d = 1.44$ )、Node6( $b = 0.70$ ,  $CI [0.26, 1.13]$ ,  $p = 0.009$ ,  $d = 1.10$ ) 和 Node8( $b = 0.68$ ,  $CI [0.10, 1.26]$ ,  $p = 0.046$ ,  $d = 1.07$ )。

### 3 讨论

#### 3.1 相信/不信语音在声学空间中的分布差异

基于 23 个声学特征构建的有监督分类器发现, 表达相信/不信的语音在声学空间中有清晰的决策边界, 模型的 ACC 和 AUC 最高可达 0.76 和 0.84。根据 Shap 值, f0mean、Jitter、HNR、f0std、f0max、f0min、SpRate、H2-H4 对区分相信/不信有较大贡献。基于上述 8 个重要特征做 RDA(图 2(a)) 和 LMM 发现, 不信比相信语音有更高的 f0mean、f0std、f0min、f0max、HNR, 以及更低的 SpRate、H2-H4、Jitter。

汉语的不信比相信语音有更高的 f0mean 和 f0std, 这与西班牙语和英语的研究结果一致<sup>[15]</sup>。较高的 f0mean 与疑问、怀疑、惊讶有关, 例如: Liu 和 Xu<sup>[44]</sup> 发现句末音高上升是疑问语气的重要特征; Jiang 和 Pell<sup>[9]</sup> 指出, 怀疑比确信的语音有更高的 f0mean; Liu 等<sup>[45]</sup> 发现惊讶比中性的语音有更高的 f0mean。较大的 f0std, 根据努力编码 (Effort Code) 假说<sup>[46]</sup>, 表明话者对信息有更多的强调。

表达不信时, 语音有较小的 H1-H2 和 Jitter、较大的 HNR, 说明嗓音更加规则和清晰<sup>[41,47]</sup>, 这与讽刺语音的嗓音表现一致<sup>[48]</sup>, 支持“个体倾向于使用讽刺语调表达不信态度”的观点<sup>[13]</sup>。使用清晰且规则的嗓音, 有利于提高语音的可懂度, 从而让否定更加突显。

#### 3.2 人格维度对相信/不信语音声学特征差异的调节作用

RDA(图 2(b)) 和 LMM 发现, 神经质、外倾性、开放性对相信/不信语音之间的声学特征差异有重要的调节作用。神经质的调节作用表现在 f0mean、f0min

和 f0std, 外倾性的调节作用表现在 f0mean 和 HNR, 而开放性的调节作用表现在 f0std。

相比于低神经质人群, 高神经质人群在表达相信和不信态度时 f0mean 和 f0min 的差异较小, 这与高神经质人群有较弱的情绪表达能力有关<sup>[20,23,24]</sup>。而且, 表达不信态度时, 高神经质人群比低神经质人群有更高的 f0std, 表明更高的情绪激活<sup>[49]</sup>, 符合高神经质人群易情绪化的性格特征<sup>[20,23,24]</sup>。

相比于低外倾性人群, 高外倾性人群表达相信和不信态度时 f0mean 和 HNR 的差异较大, 这与高外倾性人群有较强的交际能力和情绪表达能力相符<sup>[20,23,24,50]</sup>。而且, 表达相信态度时, 高外倾性人群比低外倾性人群对声门控制的规则性减弱 (较小的 HNR), 说明他们在言语表达时较为放松和随意, 与文献<sup>[20]</sup>的结果一致。

相比于低开放性人群, 高开放性人群表达相信和不信态度时 f0std 的差异较大, 这可能受到开放性人群好奇心交际风格 (思维开放、充满好奇、善于辩论等) 的影响<sup>[20]</sup>。好奇心促使高开放性人群表达不信时倾向于加强疑问语气, 导致不信和相信的 f0std 有较大的差异。但是, 开放性对相信/不信态度 f0std 的预测作用都不显著, 该假设仍需进一步验证。

LMMT 综合考察了大五人格整体上对相信/不信语音之间声学特征差异的影响。相比于其他人群, 同时具有低神经质、高外倾性、低尽责性的人群, f0mean 的差异较大; 同时具有低神经质、低开放性的人群, f0std 的差异较小; 同时具有低神经质、低开放性的人群, 以及同时具有高神经质、高开放性的人群, SpRate 的差异较小。

#### 3.3 不足与展望

虽然研究的样本量超过了 80% 统计功效所需的最小样本量, 但是毕竟样本量较小, 后续研究可通过增加样本量获得更可靠的结论。同时, 人格维度是普适的心理学概念、与语言无关, 因此可以预期, 人格维度对相信/不信语音声学特征的影响, 有很多语言共通的性质, 后续研究可将被试人群拓展至其他语言, 获得更具普遍性的结论。此外, 态度在语音上的表现不会在话语中均匀分布, 某些局部的声学特征更能体现态度上的区别<sup>[2,3,9]</sup>。例如, Gu 等<sup>[2,3]</sup> 发现, 基频特征在不同态度间的对比, 在句重音部位表现更加强烈, Jiang 和 Pell<sup>[9]</sup> 发现态度对比的声学差异随音节在句中的位置而变化。本文只分析了全句的整体声学参数, 后续研究可以通过提取句子不同位置的声学参数做更细致的局部分析。

## 4 结论

采用有监督机器学习、冗余分析、线性混合模型和线性混合模型树四种分析方法, 揭示了相信和不信语音的声学特征的差异, 以及 5 个人格维度 (神经质、外倾性、开放性、宜人性、尽责性) 对这些声学特征的影响。研究发现, 相信/不信的态度语音在声学空间中有清晰的聚类边界, 不信比相信语音有更高的基频均值、基频标准差、基频最小值、基频最大值、谐波比, 以及更低的语速、谐波差 H1-H2、基频扰动; 大五人格维度, 尤其是神经质、外倾性、开放性, 对相信和不信语音的声学特征差异有显著的调节作用。未来的智能人机语音交互系统, 离不开个性化的需求, 而人格特质正是个性化的重要体现。本文以相信与不信语音为例, 揭示了大五人格对态度语音声学特征的影响, 为个性化人机语音交互的发展提供了数据支持。

### 参 考 文 献

- Xu Y. Speech prosody: A methodological review. *J. Speech Sci.*, 2011; **1**(1): 85-115
- Gu W, Zhang T, Fujisaki H. Prosodic analysis and perception of Mandarin utterances conveying attitudes. The 12th INTER-SPEECH, Florence, Italy, 2011: 1069-1072
- Gu W, Fujisaki H. Data acquisition and prosodic analysis for Mandarin attitudinal speech. In: Peng G, Shi F (Eds.), East flows the great river: Festschrift in honor of Prof. William S.-Y. Wang's 80th birthday. Hong Kong: City University of Hong Kong Press, 2013: 483-500
- Shochi T, Rilliard A, Aubergé V, et al. Intercultural perception of English, French and Japanese social affective prosody. In: Hancil S. (Ed.), The role of prosody in affective speech (Linguistic Insights vol. 97). Bern: Peter Lang AG, 2009: 31-59
- Lu Y, Aubergé V, Rilliard A. Prosodic profiles of social affects in Mandarin Chinese. The 7th International Conference on Speech Prosody, Dublin, Ireland, 2014: 125-129
- Truesdale D M, Pell M D. The sound of passion and indifference. *Speech Commun.*, 2018; **99**: 124-134
- Caballero J A, Vergis N, Jiang X, et al. The sound of im/politeness. *Speech Commun.*, 2018; **102**: 39-53
- Fish K, Rothermich K, Pell M D. The sound of (in)sincerity. *J. Pragmat.*, 2017; **121**: 147-161
- Jiang X, Pell M D. The sound of confidence and doubt. *Speech Commun.*, 2017; **88**: 106-126
- Andersen P A. Nonverbal communication: Forms and functions. Mountain View, CA: Mayfield, 1999
- Gardner R. Between speaking and listening: The vocalisation of understandings. *Appl. Linguist.*, 1998; **19**(2): 204-224
- Harris S, Sheth S A, Cohen M S. Functional neuroimaging of belief, disbelief, and uncertainty. *Ann. Neurol.*, 2008; **63**(2): 141-147
- Manusov V, Trees A R. "Are you kidding me?": The role of non-verbal cues in the verbal accounting process. *J. Commun.*, 2002; **52**(3): 640-656
- Levitan S I, Hirschberg J. Believe it or not: Acoustic-prosodic cues to trust and mistrust in spoken dialogue. The 11th International Conference on Speech Prosody, Portugal, Lisbon, 2022: 610-614
- Yu Y, Levitan S I. Acoustic-prosodic cues to trust and mistrust in Spanish and English dialogues. The 12th International Conference on Speech Prosody, Leiden, The Netherlands, 2024: 965-969
- Gauder L, Pepino L, Riera P, et al. A study on the manifestation of trust in speech. ArXiv Preprint: 2102.09370, 2021
- Armstrong M, Esteve Gibert N, Hübscher I, et al. Developmental and cognitive aspects of children's disbelief comprehension through intonation and facial gesture. *First Lang.*, 2018; **38**(6): 596-616
- Leung S K, Bond M H. Interpersonal communication and personality: Self and other perspectives. *Asian J. Soc. Psychol.*, 2001; **4**(1): 69-86
- Waldherr A, Muck P M. Towards an integrative approach to communication styles: The Interpersonal Circumplex and the Five-Factor Theory of personality as frames of reference. *Communications*, 2011; **36**(1): 1-27
- Ahmed J, Naqvi I. Personality traits and communication styles among university students. *Pak. J. Soc. Clin. Psychol.*, 2015; **13**(2): 53-59
- Solaja O M, Idowu F E, James A E. Exploring the relationship between leadership communication style, personality trait and organizational productivity. *Serb. J. Manag.*, 2016; **11**(1): 99-117
- Costa Jr P T, McCrae R R. From catalog to classification: Murray's needs and the five-factor model. *J. Pers. Soc. Psycho.*, 1988; **55**(2): 258-265
- Costa Jr P T, McCrae R R. Four ways five factors are basic. *Pers. Individ. Differ.*, 1992; **13**(6): 653-665
- Matthews G, Deary I J, Whiteman M C. Personality traits. Cambridge: Cambridge University Press, 2003
- McCrae R R, Costa Jr P T, Ostendorf F, et al. Nature over nurture: Temperament, personality, and life span development. *J. Pers. Soc. Psycho.*, 2000; **78**(1): 173-186
- Lee S, Park J, Um D. Speech characteristics as indicators of personality traits. *Appl. Sci.*, 2021; **11**(18): 8776
- Aronovitch C D. The voice of personality: Stereotyped judgments and their relation to voice quality and sex of speaker. *J. Soc. Psychol.*, 1976; **99**(2): 207-220
- Song J, Kim M, Park J. Acoustic correlates of perceived personality from Korean utterances in a formal communicative setting. *Plos One*, 2023; **18**(10): e0293222
- Michalsky J, Niebuhr O, Penke L. Do charismatic people produce charismatic speech: On the relationship between the Big Five personality traits and prosodic features of speaker charisma in female speakers. The 10th International Conference on Speech Prosody, Tokyo, Japan, 2020: 700-704
- Gocsál Á. Female listeners' personality attributions to male speakers: The role of acoustic parameters of speech. *Pollack Periodica*, 2009; **4**(3): 155-165
- Pearsell S, Pape D. The effects of different voice qualities on the perceived personality of a speaker. *Front. Commun.*, 2023; **7**: 909427

- 32 Mohammadi G, Vinciarelli A. Automatic attribution of personality traits based on prosodic features. The International Conference on Affective Computing and Intelligent Interaction, Xi'an, China, 2015: 29–32
- 33 盛晨. 外倾性与宜人性格特质的态度语音实验研究. 硕士学位论文, 南京: 南京师范大学, 2019
- 34 胡涵, 顾文涛. 个体依恋风格对亲密话语韵律及嗓音特征的影响. *声学学报*, 2022; **47**(2): 276–286
- 35 胡涵, 顾文涛. 成人依恋风格对情绪语音声学特征的作用. *声学学报*, 2024; **49**(6): 1274–1283
- 36 姚若松, 梁乐瑶. 大五人格量表简化版 (NEO-FFI) 在大学生人群的应用分析. *中国临床心理学杂志*, 2010; **18**(4): 457–459
- 37 R Core Team. R: A language and environment for statistical computing. <https://www.R-project.org/>, 2023
- 38 Kumle L, Vö M L H, Draschkow D. Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behav. Res. Methods*, 2021; **53**(6): 2528–2543
- 39 Psychology Software Tools, Inc. E-Prime 3.0. <https://support.pst-net.com/>, 2016
- 40 Boersma P, Weenink D. Praat: Doing phonetics by computer. Version 6.3. <http://www.praat.org/>, 2022
- 41 Garellek M. The phonetics of voice 1. In: Katz W, Assmann P (Eds.), *The Routledge handbook of phonetics*. Routledge, 2019: 75–106
- 42 Lai J, Zou Y, Zhang J, *et al.* Generalizing hierarchical and variation partitioning in multiple regression and canonical analyses using the rdacca. hp R package. *Methods Ecol. Evol.*, 2022; **13**(4): 782–788
- 43 Voeten C C. Buildmer: Stepwise elimination and term reordering for mixed-effects regression. R package version 2.11, <https://CRAN.R-project.org/package=buildmer>, 2023
- 44 Liu F, Xu Y. Parallel encoding of focus and interrogative meaning in Mandarin intonation. *Phonetica*, 2005; **62**(2–4): 70–87
- 45 Liu X, Xu Y, Zhang W, *et al.* Multiple prosodic meanings are conveyed through separate pitch ranges: Evidence from perception of focus and surprise in Mandarin Chinese. *Cogn. Affect. Behav. Neurosci.*, 2021; **21**(6): 1164–1175
- 46 Gussenhoven C. Intonation and interpretation: Phonetics and phonology. The 1st International Conference on Speech Prosody, Aix-en-Provence, France, 2002: 47–57
- 47 Teixeira J P, Gonçalves A. Accuracy of jitter and shimmer measurements. *Procedia Technol.*, 2014; **16**: 1190–1199
- 48 Li S, Gu W, Liu L, *et al.* The role of voice quality in Mandarin sarcastic speech: An acoustic and electroglottographic study. *J. Speech Hear. Res.*, 2020; **63**(8): 2578–2588
- 49 Laukka P, Juslin P, Bresin R. A dimensional approach to vocal expression of emotion. *Cogn. Emot.*, 2005; **19**(5): 633–653
- 50 Boone R T, Buck R. Emotional expressivity and trustworthiness: The role of nonverbal behavior in the evolution of cooperation. *J. Nonverbal Behav.*, 2003; **27**: 163–182