

doi: 10.3969/j.issn.1672-4933.2025.06.017

普通话人工耳蜗植入者对不同模式合成语音短句材料的识别分析

Intelligibility of Different Synthesized Mandarin Speech by Cochlear Implant Listeners

石颖 陈婧媛 陈彪 邹馨悦 李永新

SHI Ying, CHEN Jing-yuan, CHEN Biao, ZOU Xin-yue, LI Yong-xin

【摘要】目的 评估汉语普通话人工耳蜗植入(CI)者对不同模式合成语音短句的识别能力。**方法** 选取母语为汉语普通话的CI者17人,测试其安静环境中对自然语音与5种不同模式合成语音短句材料的言语识别率,包括不同讲话者性别、语速和讲话方式,分别记录被试的言语识别率后进行对比分析。**结果** 被试对自然语音识别率最高,对合成语音的B-Accent模式最差,短句的言语模式对言语识别率有显著影响 $[F(5,80)=53.8, P<0.001]$ 。多重比较分析发现,被试对自然语音的识别能力与其他所有合成语音的识别均有显著差异 $(P<0.001)$,F-Slow与F-Normal模式的识别率无显著差异,F-Normal与M-Normal模式识别率无显著差异,F-Normal模式的识别率显著高于F-Emotion和B-Accent模式,M-Normal模式的识别率显著高于F-Emotion和B-Accent模式;对F-Emotion模式的识别率显著高于B-Accent模式 $(P<0.05)$ 。受试者人口学变量均与其对合成语音短句的识别率无相关性 $(P>0.05)$ 。**结论** CI者对合成语音短句的识别能力易受合成语音的韵律信息影响,其变化越复杂,识别难度越大。

【关键词】 人工耳蜗植入;普通话;合成语音;言语识别

【中图分类号】 G762

【文献标识码】 A

【文章编号】 1672-4933(2025)06-0631-06

【Abstract】Objective To investigate the intelligibility of synthesized Mandarin speech by cochlear implant patients and analyze the influencing factors. **Methods** 17 Mandarin-speaking cochlear implant users participated in the test. All the sentences were naturally produced by a single female talker in slow clear speaking style (F-Nat) or were synthesized using the iFlyTek TTS engine. Five synthesized voices (F-Slow, F-Normal, M-Normal, F-Emotion and B-Accent) were selected to represent different talker genders, speaking rates, and speaking styles, speech recognition scores were analyzed. **Results** For CI users, the worst speech recognition scores were with the B-accent sentence. There was a significant effect of speech type $[F(5,80)=53.8, P<0.001]$. Post-hoc Bonferroni pair-wise comparisons showed that performance was significantly better with F-Nat than with any of the synthesized voices, significantly better with F-Norm than with F-Emot, or B-Acc, significantly better with F-Norm than with F-Emot or B-Acc, significantly better with M-Norm than with B-Acc or F-Emot, and significantly better with F-Emot than with B-Acc $(P<0.05)$. There was no significant difference between F-Slow and F-Normal $(P=0.97)$, so was F-Normal and M-Normal. There were no significant correlations among any of the demographic variables and speech types $(P>0.05)$. **Conclusions** CI users were susceptible to dynamic changes in speech, the more complex the variation is, the more difficult it is to identify.

【Key words】 Cochlear implant; Mandarin speech; Synthesized speech; Speech recognition

随着现代科学技术的快速发展,日常生活中的电子设备越来越智能化,经常会接触到合成语音(synthesize speech),即文语转换(text-to-speech, TTS),将任意文本转化为有声语言。TTS的应用包括各种人工语音服务(电话咨询、自动售票等)、汽车导航信息、电子读物、电子词典等。如今,大容量、高自然度的合成语音技术已经成熟,广泛应用于声讯服务、计算机、人工智能等各个领域,为人们的生活带来便利。

在语音合成过程中,文本分析、韵律处理非常重要,韵律(rhythm)是语音的超音段特征,是从语音抽象出的

语调、节奏和重音特性。在感知层面上表现为音高、语速以及音量的变化,从声学角度来说,表现为基频、音段时长和能量变化等^[1],韵律结构对语音合成的自然度和可懂度有重要影响。TTS系统中的韵律控制模块根据各项合成规则规划目标音高、音长、音强、停顿及语调等,将参数输入合成语音模块。对于汉语普通话而言,不同情感语言类型有相应的基频变化、能量变化、元音和辅音的变异规律,以及时长、重音等变化^[2,3]。当前合成语音无法很好地复制自然语言的韵律,其可懂度和自然度依然受限^[4]。

作者单位:首都医科大学附属北京同仁医院耳鼻咽喉头颈外科 北京 100730

作者简介:石颖 博士 主治医师;研究方向:耳科学

通讯作者:李永新, E-mail: entlyx@sina.com

在先前的研究中,笔者发现CI者的表现随着语速的加快而变差,不同的讲话方式及背景噪声会影响CI使用者对自然言语的感知^[5]。在合成语音中,语速是易控制和改变的,语音的基频信息、时长、能量等也可利用相应的软件进行调整,以合成各种带有不同特点的声音。本研究将对汉语普通话CI者对不同合成语音(语速、性别、感情色彩)短句的识别进行研究,探究其对合成语音的识别特点,分析其影响因素。

1 材料与方法

1.1 测试对象

招募17名母语为汉语普通话或可流利准确地使用普通话进行对话交流的CI者为被试,年龄15~64岁,平均23.82±13.53岁;人工耳蜗使用时间1.4~16.6年,平均8.29±5.13年;其中男10例、女7例;植入侧别左耳9例、右耳8例;有助听器配戴史7例,测试时配戴助听器(即双模式聆听)者1例;先天性前庭导水管扩大者8例(47%),颞骨骨折者2例(11.8%),其余7例(41.2%)病因不明;9例为先天性耳聋,最晚为57岁才发现耳聋;术前耳聋时间最短0.9年,最长19.4年。被试基本信息见表1。通过口语交流判断被试具有一定开放式言语识别能力,无其他影响聆听的疾病。本研究获得被试知情同意并通过医院伦理委员会批准(TREC2025-KYS202)。

1.2 测试材料

测试材料为开放式中文言语评估系统(mandarin speech perception test, MSP)中音位平衡(phonetically

balanced, PB)的短句材料。共10个句表,每表10个短句,每句7个汉字。表中各组声母、韵母、声调的分布频率与3500个常用汉字中的分布频率一致^[6]。该材料数字化录制后嵌入开放式中文言语评估系统软件中,标准化录制的发声者为女声专业播音员。分别进行安静环境与不同信噪比(signal-noise-ratio, SNR)下的言语识别率测试和稳态白噪声下的言语识别阈(speech recognition threshold, SRT)测试,该句表在健听人及CI者中进行了等价性验证,可应用于临床评估^[7]。测试时只需由测试者选择点击受试者正确识别的字词,软件将自动生成最终结果。

本实验过程中,除标准化录制的言语声,主要合成语音由科大讯飞InterSound™ 4.0嵌入式语音合成系统提供(包括1名女声、1名男声和1名儿童声)。科大讯飞的TTS引擎通过改变参数范围调节合成语音的速度(speaking rate)和音调(pitch)。测试语句有3种语速:慢语速(初始语速的1/2)、正常语速(录制或合成的初始语速)、快速语速(初始语速的2倍)。根据科大讯飞TTS合成系统的开发手册,直接在TTS合成系统中进行调整,获得实验所需的快速语速或慢语速,语速的调整值为+6(默认语速的双倍,快速)、0(默认语速)和-6(默认语速的一半,慢速)。整个测试共6种模式(1种自然语音,5种合成语音),自然语音记为F-Natural模式(标准化已验证录音材料),是MSP材料的标准化录制模式,为女声;5种合成语音涵盖不同性别、语速、讲话方式,具体情况如下:F-Slow模式:女声,语气模拟新闻播报形式,语速为慢速;F-Normal模式:女声(与1相同),语气模拟新闻播报形式,语速为常

表1 被试基本信息

编号	性别	发现耳聋时年龄(岁)	术前耳聋时间(年)	测试时年龄(岁)	人工耳蜗使用时间(年)	侧别	病因	植入设备	现在是否配戴助听器	助听器使用时间(年)
P1	男	<1	2.1	15	12.9	左	不明	CI24C	否	0
P2	女	<1	3.6	15	11.4	左	LVAS	CI24RCS	否	0
P3	男	<1	6.5	15	8.5	右	LVAS	HR90K	否	3(双)
P4	男	2	1.7	16	14.3	左	不明	CI512	否	0
P5	女	12	1.0	17	4.0	右	不明	HR90K	否	0
P6	女	<1	11.9	17	5.1	左	LVAS	HR90K	否	3(双)
P7	男	<1	12.0	18	6.0	左	LVAS	CI24R	是	11(双)
P8	男	<1	10.9	18	7.1	左	LVAS	C40+	否	1.5(双)
P9	男	2	0.9	19	16.1	右	不明	CI24C	否	0
P10	女	2	1.1	19	15.9	右	不明	CI512	否	0
P11	女	<1	18.6	20	1.4	右	LVAS	Sonata	否	15(右)
P12	男	<1	17.0	21	4.0	左	LVAS	HR90K	否	17(双)
P13	女	4	1.4	22	16.6	右	不明	Sonata	否	0
P14	男	<1	19.4	23	3.6	右	LVAS	HR90K	否	0
P15	男	31	1.3	38	5.7	左	颞骨骨折	C40+	否	0
P16	男	43	2.0	48	3.0	左	颞骨骨折	HR90K	否	0
P17	女	57	1.7	64	5.3	右	不明	HR90K	否	1(左)

注:大前庭导水管综合征(large vestibular aqueduct syndrome, LVAS)

速; M-Normal 模式: 男声, 语气模拟新闻播报形式, 语速为常速; F-Emotion 模式: 女声, 语气模拟讲故事模式, 带有感情色彩, 语速常速; B-Accent 模式: 5 岁儿童声语气模拟日本动画人物蜡笔小新, 语速为常速。其中, F-Slow 模式的语速与 MSP 短句材料中的自然标准化语音 (F-Natural) 语速相近, 但略慢。6 种模式短句材料的具体基频和语速信息见表 2。F-Emot 模式的平均基频最高 (333 Hz), 其次为 B-Acc 模式 (284 Hz), M-Norm 的平均基频最低 (127 Hz)。

使用 AngelSimTV1.08.01 软件对所有不同模式的语音进行人工耳蜗仿真处理, 可模拟 CI 语音处理器, 将正常语音转换为类似 CI 者听到的声音。

1.3 测试环境

测试地点为北京市耳鼻喉科研究所测听中心标准双间隔声室, 背景噪声小于 20 dB(A)。在笔记本电脑 (Thinkpad T420 4180-N5C) 中安装开放式中文言语评估系统 (MSP, 版本 5.04.01) 软件, 其耳机插口经音频线连接至 MADSEN Orbiter 922 型听力计音频输入口, 再将听力计音频输出连接至配套扬声器。用软件配带的 1000 Hz 纯音作为校准信号, 按照国标 GB/T 7341.2-1998 校准声场, 测试设备 (丹麦 B&K Type 2209 声级计, 探头型号 B&K4145) 经北京市计量科学研究所校准。被试以 0° 入射角面向单个扬声器, 与扬声器相距 1 米, 测试过程中保证受试者双耳高度与扬声器喇叭中心一致, 给声强度为 65 dB SPL。

1.4 测试方法

所有被试测试前填写问卷, 收集基本信息。言语测听前均进行声场下的人工耳蜗助听听阈评估, 以保证人工耳蜗调试无异常, 测试音为啞音, 测试频率包括 0.25、0.5、1、2、4 kHz。测试过程中均使用被试日常生活中使用的程序, 不采用降噪模式、音乐模式等。

测试材料为开放式材料, 被试在声场中聆听, 采用听说复述法应答, 计算 70 个汉字的正确率记为言语识别率 (speech recognition perception, SRP), 每个汉字声母、韵母、声调全对时记为正确。测试开始前向受试者讲解测试要求, 询问疑惑并解答, 测试时不同模式与不同短句组别随机匹配, 每组句中短句出现的顺序随机生成, 每个模式播放顺序一致。测试过程中鼓励猜测, 给予正面评价, 但不提供正误反馈。每个测试项仅播放一次, 不予重复, 由于受试者口齿不清等原因, 测试者未听清复述时要求重复, 必要时以汉字或拼音书写确认。被试可随时要求休息, 整个测试中保证受试者至少休息 2 次。所有测试

由同一人实施, 测试者在测试前已掌握纯音测听和言语测听相关理论和操作, 并熟悉言语测听的材料和操作。

1.5 统计学方法

应用 SPSS 21.0、sigmaplot 14.0 软件进行数据分析。检验方法使用方差分析 (ANOVA), T 检验及相关分析等。P<0.05 为存在统计学差异。

2 结果

2.1 被试声场下助听听阈评估结果

检测 17 名 CI 者声场下 0.25、0.5、1、2、4 kHz 的助听听阈, 其中 1 名测试者为双模式聆听, 具体助听听阈结果见表 3。

2.2 被试对不同模式短句的言语识别率比较

对经过人工耳蜗模拟处理后获取到的电谱图 (electrograms) 及原始声信号的语谱图 (spectrograms) 进行记录, 图 1 为短句“晚上一块去跳舞”的电谱图和语谱图, 可看出合成语音与自然语音 (F-Natural) 相比, 频域包络信息和基频信息存在不同程度的畸变, 尤其是 F-Emotion 和 B-Accent 模式。针对合成语音模拟耳蜗刺激得到的电谱图中, 频域包络信息和基频信息的失真更明显, 尤其是最后两个模式, 预示 CI 使用者识别合成语音的难度大于自然语音, 尤其是对最后两种合成语音 (F-Emotion 和 B-Accent 模式) 的识别能力可能更差。

图 2 为每种模式短句识别率箱图, 可见 CI 者对于 F-Natural 模式识别率最高, 其次为 F-Normal、F-Slow 和 M-Normal 模式, 对 F-Emotion 模式的识别能力较差, B-Accent 最差。采用单因素方差分析对比不同模式下的言语识别率, 发现短句的言语模式 (F-Natural、F-Slow、F-Normal、M-Normal、F-Emotion、B-Accent) 对言语识别率有显著影响 [$F(5, 80)=53.8, P<0.001$]。

2.3 被试对各种合成语音短句识别能力比较

经过多重比较分析发现, 被试对自然语音 (F-Natural) 的识别能力与其他所有合成语音的识别均有显著差异 ($P<0.001$), 表明语音质量 (即语音材料为自然语音还是合成语音) 对被试的识别率有显著影响; F-Slow 与 F-Normal 模式的识别率无显著差异 ($P=0.970$), 表明此两种合成语音的语速差异对 CI 使用者的识别无显著影响; F-Normal 与 M-Normal 模式识别率无显著差异 ($P=0.270$), 表明此两种合成语音的说话者性别因素对 CI 使用者的识别率无显著影响; F-Normal 模式的识别率显著高于 F-Emotion 和

表 2 每种模式语句材料的基频 (F_0 , Hz) 和语速 (字/秒) ($\bar{x}\pm s$)

	F-Nat	F-Slow	F-Norm	M-Norm	F-Emot	B-Acc
F_0 (Hz)	223.00±14.80	265.00±17.00	267.00±16.40	127.00±9.80	333.00±13.80	284.00±5.90
语速 (字/秒)	3.56±0.24	3.41±0.31	4.51±0.30	4.79±0.38	4.07±0.23	3.04±0.16

表3 被试声场下助听听阈(dB HL)

编号	频率(kHz)				
	0.25	0.5	1	2	4
S01	35	35	30	30	30
S02	35	30	30	35	35
S03	50	45	45	55	40
S04	55	50	45	50	40
S05	30	35	35	30	30
S06	35	35	35	30	35
S07	35	30	30	30	35
S08	25	30	20	20	30
S09	30	30	35	30	35
S10	40	35	35	35	30
S11	40	45	40	35	35
S12	45	45	50	50	45
S13	20	30	25	25	50
S14	30	30	25	30	30
S15	25	35	25	25	30
S16	55	50	50	45	40
S17	50	45	40	40	45

B-Accent 模式 ($P<0.05$), 表明合成语音的语气、重音等情绪因素(vocal emotion)对CI使用者的识别能力有显著影响;此外,M-Normal模式的识别率显著高于F-Emotion和B-Accent模式 ($P<0.05$);对F-Emotion模式的识别率显著高于B-Accent模式 ($P<0.05$)。

2.4 被试人口学变量与合成语音短句识别率的相关性分析

对不同合成语音模式的短句识别率与CI组受试者与

测试时年龄、术前耳聋时间、CI使用时间、中低频助听听阈(0.25、0.5、1 kHz 均值)、高频助听听阈(2、4 kHz 均值)进行线性回归分析,未发现与任何因素具有相关性($P>0.05$)。

3 讨论

人工耳蜗植入(CI)技术是目前重度和极重度感音神经性聋患者听力重建的重要方法,CI者在日常生活中会遇到需要聆听合成语音的情况。健听者和CI者对自然语音的识别能力均明显高于合成语音,CI者对合成语音识别的难度更大,尤其是对语速快的合成语音^[8]。分析原因,可能是由于CI的频率分辨率有限,使用者能获取的有效频域和时域信息相对局限,对于各种非典型语言(针对于清晰、语速缓慢、中性情绪的言语而言),如带有情绪的语言、真实环境中的对话、耳语、方言等,其识别难度会明显高于健听者。汉语普通话是一种声调语言,通常可分为即阴平(一声)、阳平(二声)、上声(三声)、去声(四声)和轻声5个声调,声调变化实质上由基频、发音时长、音强决定,对汉语普通话的识别很大程度依赖于基频信息的获取。当语句复杂或非典型的动态变化多时,会使基频信息的捕捉难度加大,因此,汉语普通话CI者与其他母语为非声调语言的CI者相比,对合成语音的识别能力可能会更差。

3.1 合成语音语速对CI者言语识别的影响

自然语音的声学特性随着语速变化而变化,如元音及辅音时长,相邻音节及音素的间隔时间,以及基频变化。合成语音的语速调整方法与自然语音存在区别。自

晚上一块去跳舞 Wǎnshàng yīkuài qù tiàowǔ (Let's go dancing tonight)

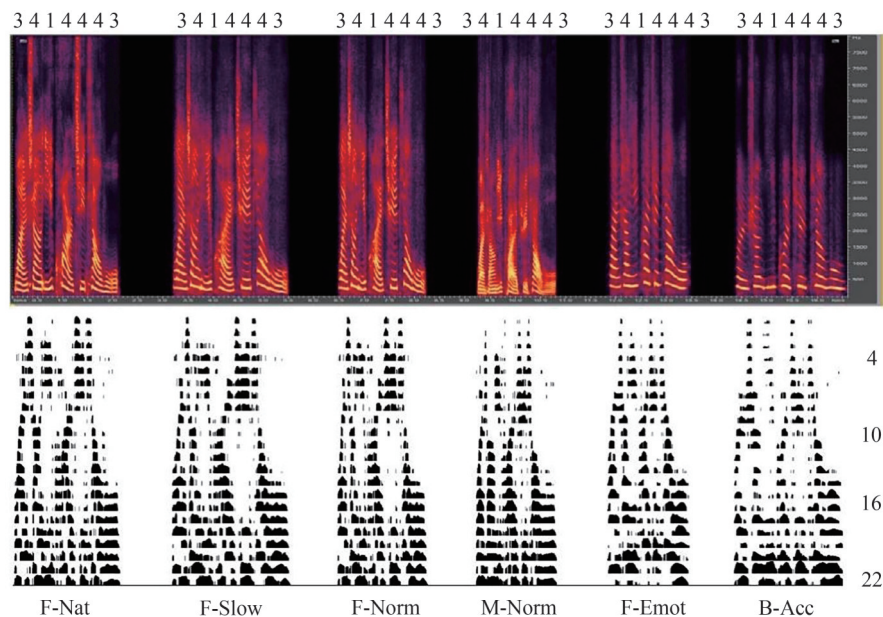


图1 “晚上一块去跳舞”的电谱图和语谱图

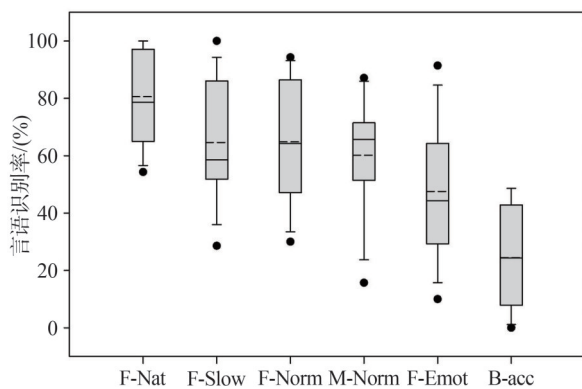


图2 CI组受试者对不同模式短句言语识别率(注:图中矩形的上下边缘分别为上、下四分位数,矩形中实线表示的为此组数据的中位数,虚线表示的为此组数据的均值)

然语音中的语速调整一般有两种方法,一种是利用自然产生的快速语句和慢速语句直接当作目标语句进行录音和识别,这种方法可最大程度保留自然语音中的动态声学特性;另一种是通过均匀的时间缩放和时间扩展调整,保留语句中的基频和共振峰信息,而元音持续时间、声音释放时间等都会被改变,会有一定程度的失真。合成语音是通过调整TTS中的语速调节参数实现,根据不同的选择方式可将单音素、双音素或者三音素混合实现目标语速,同样造成声音信号中声学信息的失真,尤其是时域信息,即时域失真,且合成语音的失真程度可能要超过时间调整方法实现的自然语音。这可能是快语速的合成语音识别难度较大的根本原因之一。不管是健听者还是CI者,对快语速的合成语音识别能力都比常速合成语音及自然语音差^[8]。

在先前研究中,笔者发现CI者对快速语句的识别能力差于常速和慢速语句^[9]。对合成语音语速影响CI者言语识别效果的研究结果显示,常速与快速语句的言语识别率有显著差异,表明将合成语音的语速从快速减慢为常速可有效提高CI者的言语识别能力。因此,推断其对合成语音的识别也有类似效果,本实验把合成语音的语速设定为慢速和常速,探究将合成语音语速减慢是否能帮助CI者提高识别效果。F-Slow与F-Norm两种短句模式的语速相差1.1字/秒,CI者对这两种模式合成语音短句的言语识别率无显著差异,这与之前对常速和慢速自然语音识别率的研究结果类似^[10],意味着将常速合成语音短句的语速减慢无法有效提高CI者的言语识别能力。因此,推断合成语音的语速在某个范围内减慢(目前粗略认为从快速减到常速)可帮助CI者提高识别能力,但当减慢到这个范围之外后,不再有显著影响。实验中所说的常速是以清晰缓慢的言语测试材料为标准进行对比,这种语速比日常生活中对话交流的真实语速慢。

3.2 合成语音的感情色彩对CI者言语识别的影响

本实验中,CI者对F-Emo模式的识别率显著低于其他语音模式,表明TTS系统在合成带有感情色彩的语音(即情绪语音)时,无法保留对汉语普通话识别的关键基频信息、振幅信息和时长信息等。从几种合成语音参数的对比及语音波形、语谱图分析可以看出,F-Emot模式的平均基频最高(333Hz),与其他模式相比,其三声和四声的基频变化范围相对较窄,谐波分布也不同;B-Acc模式与谱图也存在类似畸变,其音节的振幅与其他模式不同。因此,从语谱图畸变的程度可预测F-Emot模式和B-Acc模式的识别相对困难,这与本实验结果一致。CI者对普通自然语音和带有感情色彩的自然语音识别率比较无显著差别^[10],当聆听的内容为带有感情色彩的自然语音时,CI者易受其基频信息变异影响。

3.3 合成语音说话者性别对CI者言语识别的影响

健听者易分辨讲话者的性别,因为男声和女声的基频不同,女声在50~250 Hz,男声在100~500 Hz,不同语境中特有的语速、音调、语气等声学信息,从人群中辨别出讲话者并不困难^[11]。这一部分与说话者的性别、年龄和方言等因素有关的信息被叫做索引信息(indexical information),主要体现为不同说话者间的声学差异,会影响CI使用者的言语识别能力^[12]。当CI者感知不同说话者的言语时,不同音高(与基频有关)、声道长度、CI处理器相关的因素(如通道数,频率分布)和患者自身因素(如电极神经之间的匹配程度及耳聋时间)等,均会影响CI者对不同说话者语言的理解程度。

目前,国内还没有针对CI者对索引信息识别的研究,国外研究相对较多。当男女声的基频差异为100 Hz时,CI者的语音性别识别率达94%,但当男女声的基频差异为10 Hz时,CI者的语音识别率降为68%,同时发现目前的人工耳蜗处理策略只能传递有限的索引信息^[13,14]。Vongphoe等^[15]通过对健听人与CI者进行多人身份识别测试,发现健听者的鉴别正确率达84%,而CI者的鉴别正确率仅为20%。虽然目前还没有针对CI者对合成语音中索引信息对言语感知影响的研究,但是基于相似的感知机理,索引信息的改变也会影响CI者的言语识别,可推测其影响会大于自然语音。在CI者对合成语音识别率的相关研究中发现,说话者的性别因素会影响CI者($P=0.021$)的言语识别,但与以往研究不同,实验中说话者的性别因素虽然证明对言语识别率有显著影响,在数值差别不大(2.71%)^[8]。本研究中,CI被试对F-Normal与M-Normal模式短句的言语识别率无显著差异($P=0.27$),表明合成语音的讲话者性别差异未影响CI组被试的言语识别率。本研究使用的合成语音中,F-Normal与M-Normal模式的基

频相差 140 Hz, 识别率分别为 64.82% 和 60.29%, 未出现难度差异, 但测试过程中询问被试是否感觉到两种声音存在差异、易于区分时, 被试普遍反应可以很好区分。可见, CI 者对合成语音的性别差异可能不会对其识别难度造成影响, 在区分度较好的情况下, 易于区分。今后需要将不同性别合成语音的基频信息进行调整, 以探究区分两种声音性别的最小基频差, 此外, 本研究的两种声音对比在中性情绪语句基础上, 今后可增加感情色彩等不同韵律特点, 以探究感情色彩、重音等声学信息与索引信息是否会产生交互作用, 影响 CI 者对合成语音的识别。

参考文献

[1] 王永鑫, 蔡莲红. 语法信息与韵律结构的分析与预测[J]. 中文信息学报, 2010, 24(1): 65-70.
 [2] 陶建华, 蔡莲红, 赵晨. 汉语语音合成中的文本分析和韵律处理[C]. 辉煌二十年——中国中文信息学会二十周年学术会议论文集. 2001. 279-286.
 [3] 冯哲, 孙吉贵, 张长胜, 等. 汉语语音合成的研究进展[J]. 吉林大学学报(信息科学版), 2007, 25(2): 198-206.
 [4] Paris CR, Thomas MH, Gilson RD, et al. Linguistic Cues and Memory for Synthetic and Natural Speech[J]. Human Factors: The Journal of the Human Factors and Ergonomics Society, 2000, 42(3): 421-431.
 [5] Ji C, Galvin JJ, Xu A, et al. Effect of Speaking Rate on Recognition of Synthetic and Natural Speech by Normal-Hearing and Cochlear Implant Listeners[J]. Ear and hearing, 2012, 34(3): 313-323.
 [6] Shi Y, Peng KA, Chen B, et al. Interaction between Speech Variations and Background Noise on Speech Intelligibility by Mandarin-Speaking

Cochlear Implant Patients[J]. Speech Communication, 2018, 104(11): 89-94.
 [7] Fu QJ, Zhu M, Wang X. Development and validation of the Mandarin speech perception test[J]. J Acoust Soc Am, 2011, 129(6): L267-L273.
 [8] 王顺成, 石颖, 李永新, 等. 开放式言语评估系统短句材料在人工耳蜗植入者中等价性的初步分析[J]. 听力学及言语疾病杂志, 2013, 21(3): 74-77.
 [9] Li Y, Zhang G. Effects of speaking style on speech intelligibility for Mandarin-speaking cochlear implant users[J]. Journal of the Acoustical Society of America, 2011, 129(6): EL242.
 [10] Su Q, Galvin JJ, Zhang G, et al. Effects of Within-Talker Variability on Speech Intelligibility in Mandarin-Speaking Adult and Pediatric Cochlear Implant Patients[J]. Trends Hear, 2016, 20: 2331216516654022.
 [11] Liu C, Galvin J, Fu QJ, et al. Effect of spectral normalization on different talker speech recognition by cochlear implant users[J]. The Journal of the Acoustical Society of America, 2008, 123(5): 2836.
 [12] Borrie S, McAuliffe M, Liss J, et al. The role of linguistic and indexical information in improved recognition of dysarthric speech[J]. J acoust soc am, 2011, 130(4): 2447-2447.
 [13] Fu QJ, Chinchilla S, Galvin JJ. The role of spectral and temporal cues in voice gender discrimination by normal-hearing listeners and cochlear implant users[J]. J Assoc Res Otolaryngol, 2004, 5(3): 253-260.
 [14] Fu QJ, Chinchilla S, Nogaki G, et al. Voice gender identification by cochlear implant users: the role of spectral and temporal resolution[J]. J acoust soc am, 2005, 118(3 Pt 1): 1711-1718.
 [15] Vongphoe M, Zeng FG. Speaker recognition with temporal cues in acoustic and electric hearing[J]. J Acoust Soc Am, 2005, 118(2): 1055-1061.

收稿日期 2023-12-14
 责任编辑 赵 倩

(上接 618 页)

[6] 李小娟, 时盼姣. 引导式教育和感统训练联合治疗对自闭症患儿的效果[J]. 国际精神病学杂志, 2019, 46(6): 1036-1038.
 [7] 何怡, 庞子建. 听觉辨别训练对儿童构音障碍的疗效[J]. 中国康复理论与实践, 2020, 26(11): 1344-1348.
 [8] 李胜利. 语言治疗学[M]. 第 2 版. 北京: 人民卫生出版社, 2013. 138-139.
 [9] 李胜利. 构音障碍的评价[J]. 中国康复, 1993, 8(2): 84-84.
 [10] 韩知娟. 普通话言语的发展: 言语清晰度音位对比及声学特征[D]. 上海: 华东师范大学, 2005.
 [11] 缪鸿石. 康复医学理论与实践[M]. 上海: 上海科学技术出版社, 2000. 424-430, 892.
 [12] 张玲, 吴浪龙, 康怀鑫, 等. 口部定位疗法联合引导式教育治疗言语障碍儿童的疗效观察[J]. 听力学及言语疾病杂志, 2023, 31(1): 26-30.
 [13] 李树森. 儿童康复医学[M]. 北京: 人民卫生出版社, 2006. 187-189.
 [14] 王阔, 李健, 王昕. 早期言语-认知功能训练联合家庭参与听觉训练在语言发育迟缓患儿中的应用[J]. 检验医学与临床, 2023, 20(22): 3375-

3378.
 [15] 尹敏敏, 葛胜男, 邱卓英, 等. 运用世界卫生组织国际分类家族构建儿童构音障碍诊断、评估和整体康复方案[J]. 中国康复理论与实践, 2020, 26(1): 28-36.
 [16] 吴晓璐, 王丽, 郝京萍, 等. 儿童功能性构音障碍辅音错误特点分析[J]. 北京口腔医学, 2023, 31(5): 321-323.
 [17] 冯彬, 林银萍, 王燕红, 等. 引导式健康教育改善髋部骨折手术患者术后功能、负性情绪及并发症的效果[J]. 国际护理学杂志, 2022, 41(3): 431-434.
 [18] 刘巧云, 王丹, 赵航, 等. 听障儿童听觉分辨能力训练[J]. 中国听力语言康复科学杂志, 2015, (1): 66-68.
 [19] 夏小玲, 王修敏, 曹晓光, 等. 引导式健康教育对吞咽障碍合并构音障碍患者吞咽、语言康复训练的效果分析[J]. 中国听力语言康复科学杂志, 2023, 21(3): 306-309.

收稿日期 2024-12-13
 责任编辑 赵 倩