

符合普通话语境的宽动态范围压缩算法 低频释放时间研究*

金凤¹ 顾向阳¹ 孙晋² 陈友元²

[摘要] 目的:探讨在普通话语境下 WDRC(宽动态范围压缩算法)低频释放时间对助听器佩戴者言语识别率的影响。方法:将低频释放时间设置为50、100、200、400、800 ms五个组,分别在3种掩蔽噪声下进行实验,共形成15组实验条件。以听力障碍者作为实验对象,以普通话言语仿真声作为测试音,通过仿真平台输出,在特定的噪声条件下测试言语识别率。结果:WDRC低频释放时间设置为400 ms时,听障者言语识别率最佳;Babble Noise(多说话人噪声)对听力障碍者的噪声下言语识别率有显著影响。结论:在普通话语境下低频区间时间常数与噪声类型都会影响听力障碍者的普通话言语识别率。

[关键词] 宽动态范围压缩算法;普通话;释放时间;言语识别率;噪声

doi:10.13201/j.issn.1001-1781.2017.19.004

[中图分类号] R764 **[文献标志码]** A

Study of low frequency release time in WDRC on hearing-impaired individual in mandarin language environment

JIN Feng¹ GU Xiangyang¹ SUN Jin² CHEN Youyuan²

(¹Department of Otorhinolaryngology, Shanghai Xuhui Central Hospital, Zhongshan Hospital Xuhui Hospital, Shanghai, 200031, China; ²Shanghai Acoustics Laboratory, Chinese Academy of Sciences)

Corresponding author: GU Xiangyang, E-mail: gxy12341@163.com

Abstract Objective: To investigate the influence of low frequency release time of WDRC(wide dynamic range compression) to speech recognition score (SRS) of deaf subjects with hearing aids in mandarin language environment. **Method:** Release time constants are set differently from the values of 50 ms, 100 ms, 200 ms, 400 ms and 800 ms in low frequency (LF) channel, with 3 kinds of shaped noise, combining a certain LF release time constant with a certain shaped noise forms 15 different groups of experimental conditions. The testing material is mandarin sentences which are the outputs of simulation in the 15 different release time combinations. **Result:** If the release time is set the value of 400 ms, the patients with hearing aids can get the best mandarin speech intelligibility; the type of babble noise have significant effects on mandarin SRS. **Conclusion:** Statistical analysis of test results shows that the release time in LF channel or the type of shaped noise trends to influence mandarin SRS.

Key words wide dynamic range compression; mandarin; release time; speech recognition score; noise

听力障碍者能够感知音频信号的动态范围一般要比听力正常者窄,作为数字助听器听力响度补偿技术最常用的技术——宽动态范围压缩算法(WDRC),主要用于将自然言语动态范围按比例均匀地压缩到听障者残余听力所能感知的动态音频范围中,主要的技术参数包含压缩拐点、压缩比、增益、时间常数等,其中时间常数反映了助听器数字处理器对外界声音放大压缩处理速度的快慢,对助听器佩戴者的言语理解有重要影响。目前市售的助听器大部分为国外品牌,国外助听器厂家使用的WDRC时间常数,是根据英语等印欧语系的发声

特点来确定的,未考虑普通话语音特点,而汉语作为一种声调语言不同于英语语系,构成规律具有一定的特殊性。本课题通过对耳聋受试者作普通话语境下的句子测试,探究释放时间及环境噪声对普通话言语识别率的影响。

1 资料与方法

1.1 研究对象

受试者条件:年龄20~84岁;普通话为母语;0.25、0.5、1、2、4、8 kHz频率,各测听频率中最差阈值 ≤ 90 dBHL;平均听阈 ≥ 10 dB言语级强度下的裸耳最大言语识别率 $\geq 60\%$,单纯感音神经性聋,鼓膜完整,无耳科手术史。所有受试者均签署知情同意书,且自愿参加本试验。课题经上海市徐汇区中心医院伦理委员会批准实施。

2013-09—2014-10招募受试者30例,其中男18例,女12例;年龄20~84岁,平均65岁,其中

* 基金项目:上海市卫生和计划生育委员会青年科研项目(No:2013-154)

¹上海市徐汇区中心医院耳鼻咽喉科 复旦大学附属中山医院徐汇医院(上海,200031)

²中国科学院声学研究所东海研究站

通信作者:顾向阳, E-mail: gxy12341@163.com

<60岁 9 例, ≥60岁 21 例。进行纯音气导听阈检查,平均听阈分级:极重度聋(>90 dB) 0 例,重度聋(71~90 dB) 2 例,中重度聋(56~70 dB) 16 例,中度聋(41~55 dB) 12 例,轻度聋(26~40 dB) 0 例。所测得的平均纯音气导听阈见图 1。

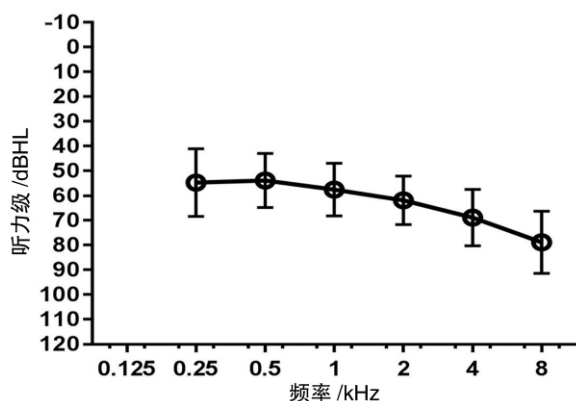


图 1 受试耳的平均纯音气导听阈

1.2 设备及材料

实验地点:隔声室(本底噪声 ≤ 30 dBA SPL,隔声水平 ≥ 40 dBA SPL);外置声卡:TASCAM US-144MK II;诊断型听力计:Entomed SA 204TM;插入式耳机(E. A. R. Tone 3A 50 Ω);测试助听器:斯达克 Wi series i100 RIC 机[苏食药监械(准)字 2012 第 2460383 号];测试平台:基于 MATLAB 开发的普通话句子测试平台;声级计:杭州爱华 AWA6228 型多功能声级计;扬声器,电脑等。

以 Xi 等^[1]研究的普通话句子语音库作为测试材料,由男性播音员录制,且该播音员母语为普通话,一共包括 32 个句表,每个句表有 10 个普通话句子,7 个汉字组成一句话。

模拟环境噪声的掩蔽噪声分别为 Car Noise(汽车噪声),Babble Noise(多说话人噪声)和 Speech-Shaped Noise(语谱噪声)。

1.3 方法

1.3.1 预处理 将受试者测得的纯音气导听阈输入 NAL-NL2 处方公式软件^[2],得出助听器在不同输入声压级下的增益目标曲线,涵盖 16 通道增益、压缩拐点及压缩比。宽动态范围压缩算法时间常数设定为:启动时间选取为常用的短启动时间 10 ms;释放时间低频段分别选取为 50、100、200、400、800 ms,高频段根据赵建平等^[3]的结论统一选取为 100 ms。由操作人员将上述参数编程至受试者佩戴助听器,并根据受试者的描述情况进行微调。在受试者适应调试助听器 20 min 后进行言语测试。

1.3.2 言语测试 在隔声室内进行普通话句子言语测试,测试前用声级计对声场进行校准。受试者测试耳佩戴助听器,非测试耳用耳印泥封堵。基于

MATLAB 开发的普通话句子测试平台以 Xi 等^[1]研究的普通话句子语音库作为测试材料,分别以 Car Noise、Babble Noise 和 Speech-Shaped Noise 为掩蔽噪声,并可调整信噪比。测试材料经扬声器播放,取扬声器与受试者 0°角方位。正式实验开始之前,受试者需进行预实验,预实验中测试材料与正式实验测试材料一致。通过预实验,使受试者熟悉测试材料和测试流程,避免学习效应;同时可测得 Speech-Shaped Noise 下的普通话句子言语识别率 60% 的信噪比值,作为正式实验时使用,避免产生天花板效应,30 位测试者的平均信噪比值为 8.97 dB。

以上述 5 个不同释放时间和 3 个噪声类型共得出 15 组测试条件,其中测试条件的顺序以及每组测试条件下播放的测试材料普通话语表顺序均由基于 MATLAB 开发的普通话句子测试平台控制随机。每种条件测试一个语表,每个语表包含 10 句普通话,每句话含有 7 个左右的汉字。受试者复述所听噪声下普通话句子,实验操作人员记录测试结果,每隔 20 min 由实验操作员观测受试者情绪状态,予以适当休息。

1.4 统计学分析

所有数据采用 SPSS 23.0 软件作单变量多因素方差分析(GLM Univariate),考虑年龄因素可能对言语识别率的影响,将受试者分为中青年组(<60岁)9 例和老年组(≥60岁)21 例,进行多因素方差分析。

2 结果

本研究结果提示,WDRC 低频时间常数及噪声类别对听力障碍者的言语识别率均有显著影响($F=27.378, P<0.01$),无论老年组还是中青年组,其交互作用均有显著影响[老年组 $F=14.285, P<0.01, F(\text{释放时间})=3.000, P<0.05; F(\text{噪声类别})=36.856, P<0.01$;中青年组 $F=16.781, P<0.01; F(\text{释放时间})=0.825, P>0.05; F(\text{噪声类别})=48.693, P<0.01$]。在 Babble Noise 环境中,低频释放时间在 800 ms 时,受试者的言语识别率最高,低频释放时间在 50 ms 时,受试者的言语识别率最低;在 Car Noise 及 Speech-Shaped Noise 环境中,低频释放时间在 400 ms 时,受试者的言语识别率最高,低频释放时间在 50 ms 时,受试者的言语识别率最低。以言语识别率为应变变量,以噪声及低频释放时间为影响因子进行统计学分析,结果提示不同的释放时间及不同的噪声条件下,普通话语境下的言语识别率差异有统计学意义。随着释放时间的延长,受试者的言语识别率有明显提高,一般在 400 ms 时,言语识别率最高($F=2.434, P<0.05$)。不同的噪声条件对受试者的言语识别率影响明显,其中以 Babble 噪声的影响最为显著($F=$

77.266, $P < 0.01$), 这可能与 Babble 噪声同测试材料声学性质相近、容易混淆有关(表 1~2); 释放时间及噪声与言语识别率的关系见图 2。

3 讨论

WDRC 时间常数的长短反映了助听器对外界声音进行放大压缩处理的快慢能力, 它包含启动时间和释放时间 2 个部分。启动时间反映了 WDRC 对输入信号强度增加的反应速度, 释放时间反映了 WDRC 对输入信号强度减弱的反应速度。综合目前国内外的观点来看, WDRC 时间常数的设定尚未有统一标准。对于启动时间, 一般选用较短的时间常数, 比如 10 ms^[4]。诸多研究者对释放时间进行了探讨, 不同时间常数对于言语识别率^[5-8]的影响, 结论不一。Alexander 等^[5]探讨 WDRC 的释放时间和通道数的关系, 表明当使用 4 个通道时, 短的释放时间(40 ms)稍好, 当使用 16 个通道时, 长的释放时间(640 ms)较好。Neuman 等^[6]对压缩比及释放时间的研究发现, 长的释放时间可以获得较好的言语舒适度。Novick 等^[7]的研究提示无论在消声环境还是课堂等噪声环境, 释放时间对言语识别率无影响。Van Toor 等^[8]通过主观及客观方面对不同时间常数的设定, 对噪声环境下言语可懂度的影响, 发现不同的实验类别所产生的最佳释放时间不同。

国外有关助听器时间常数设定的研究均是基于印欧语系的特点来进行的, 结论不完全适合汉语普通话的声学特点。不同于英语等印欧语系, 汉语普通话有 22 个辅音音位, 每个音节都有音调, 且汉语普通话中一般存在重音。音节是普通话的基本

发声单位, 主要包括声母、韵母及声调; 一个音节包含 1~4 个音素; 音素有元音及辅音(清、浊)2 种, 音节中音素的排列结构一般有 4 种类型: 元、辅-元、元-辅、辅-元-辅。浊辅音和元音一样, 都是经声带发声, 故浊辅音的能量接近元音, 而清辅音不经声带发声, 能量相对较小。吴宗济等(1964)报道普通话语音的共振峰不同, 元音和浊辅音的共振峰主要集中在低中频区域, 多在 2 kHz 以下, 高频区域主要集中在清辅音的能量, 多在 2 kHz 以上。声调识别对于汉语等声调语言的语音理解具有重要意义, 调整普通话元音的共振峰轮廓, 有助于普通话的声调识别, 进一步加强对于汉语普通话的理解^[9]。有研究从语音知觉角度出发, 利用识别任务测试方法对轻-中度感音神经性聋患者的声调知觉模式特性进行分析, 发现该群体的声调知觉模式仍具有范畴知觉特征, 与听力正常人间无显著差异^[10]。同时, 普通话音节结构的特点, 有比较低的清辅音的能量, 在较长的语音中, 如果选取比较长时间常数, 会降低辅音增益, 降低言语识别度, 故在 WDRC 算法当中, 为了避免清辅音的增益不足, 需选择合理的时间常数。如: 词语“爆破(bao po)”, “p”为清辅音, 如果“p”增益不足, 易听成“抱哦(bao o)”, 出现理解差异。

本文采用在低频区选取 5 种释放时间, 并选取 3 种不同类型的掩蔽噪声, 对听力障碍者进行普通话句子测试, 探究不同释放时间及噪声组合对于普通话言语可懂度的影响。结果提示 WDRC 低频时间常数中释放时间设置为 400 ms 时, 不同年龄段的听障者言语识别率均最佳, 特别是老年听障者优势

表 1 不同掩蔽噪声及释放时间下的受试者普通话言语识别率及分析

%, $\bar{x} \pm s$

噪声类型	释放时间/ms					F	P
	50	100	200	400	800		
多说话人噪声	50.11±26.74	56.11±19.28	53.64±24.06	55.00±23.61	57.41±21.65	27.378	<0.01
汽车噪声	70.04±19.25	75.10±14.04	77.10±15.25	78.20±15.20	73.25±17.52		
语谱噪声	67.02±20.44	70.92±15.22	68.76±18.03	73.31±13.49	71.46±18.13		

表 2 不同掩蔽噪声及释放时间下的中青年、老年受试者普通话言语识别率及分析

%, $\bar{x} \pm s$

组别	释放时间/ms					F	P
	50	100	200	400	800		
中青年组							
多说话人噪声	41.50±23.50	49.91±19.27	45.74±27.12	43.30±19.28	46.64±26.64	16.781	<0.01
汽车噪声	76.03±15.71	80.63±14.96	78.70±15.37	79.09±16.08	67.30±27.87		
语谱噪声	63.78±19.15	68.96±17.47	66.41±25.75	67.70±17.07	68.57±20.45		
老年组							
多说话人噪声	53.79±29.92	58.76±20.85	57.02±23.78	60.00±26.16	62.02±19.19	14.285	<0.01
汽车噪声	67.48±23.05	72.73±16.44	76.41±17.04	77.82±17.89	75.80±16.06		
语谱噪声	68.41±22.10	71.76±16.47	69.77±15.85	75.71±15.08	73.80±16.28		

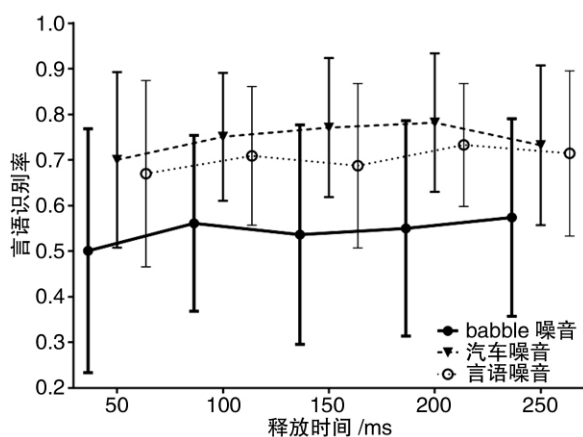


图 2 噪声、释放时间及言语识别率关系图

显著;噪声类型对听障者言语理解度有显著影响,其中 Babble Noise 的影响最大。本实验表明,对于汉语普通话,助听器低频区间释放时间不宜设置过短,一般选取 400 ms 较为合适;Babble Noise 严重影响助听器的作用,可改善相关技术,提高助听器在 Babble Noise 环境下的效用。

(致谢:感谢参加本课题的所有志愿者;特别感谢中国科学院声学研究所东海站及斯达克实验室提供的技术支持。)

参考文献

- [1] XI X, CHING T Y, JI F, et al. Development of a corpus of Mandarin sentences in babble with homogeneity optimized viapsychometric evaluation[J]. Int J Audiol, 2012, 51: 399—404.
- [2] KEIDSER G, DILLON H, FLAX M, et al. The NAL-NL2 Prescription Procedure[J]. Audiol Res, 2011, 23: e24.
- [3] 赵建平, 原猛, 冯海泓. WDRC 释放时间对普通话言语理解度的影响研究[J]. 声学技术, 2013, 32(3): 217—221.
- [4] MOORE B C. The choice of compression speed in hearing AIDS: theoretical and practical considerations and the role of individual differences[J]. Trends Amplif, 2008, 12: 103—112.
- [5] ALEXANDER J M, MASTERSON K. Effects of WDRC release time and number of channels on output SNR and speech recognition[J]. Ear Hear, 2015, 36: e35—e49.
- [6] NEUMAN A C, BAKKE M H, MACKERSIE C, et al. The effect of compression ratio and release time on the categorical rating of sound quality[J]. J Acoust Soc Am, 1998, 103(5 Pt 1): 2273—2281.
- [7] NOVICK M L, BENTLER R A, DITTBERNER A, et al. Effects of release time and directionality on unilateral and bilateral hearing aid fittings in complex sound fields[J]. J Am Acad Audiol, 2001, 12: 534—544.
- [8] VAN TOOR T, VERSCHUURE H. Effects of high-frequency emphasis and compression time constants on speech intelligibility in noise[J]. Int J Audiol, 2002, 41: 379—394.
- [9] LUO X, FU Q J. Enhancing Chinese tone recognition by manipulating amplitude envelope: implications for cochlear implants[J]. J Acoust Soc Am, 2004, 116: 3659—3667.
- [10] 元贝尔, 刘鹏, 古鑫, 等. 轻中度感音神经性聋患者汉语普通话声调知觉模式研究[J]. 临床耳鼻咽喉头颈外科杂志, 2017, 31(1): 38—40.
- [11] 元贝尔, 刘鹏, 古鑫, 等. 轻中度感音神经性聋患者汉语普通话声调知觉模式研究[J]. 临床耳鼻咽喉头颈外科杂志, 2017, 31(1): 38—40.
- [12] 元贝尔, 刘鹏, 古鑫, 等. 轻中度感音神经性聋患者汉语普通话声调知觉模式研究[J]. 临床耳鼻咽喉头颈外科杂志, 2017, 31(1): 38—40.
- [13] 元贝尔, 刘鹏, 古鑫, 等. 轻中度感音神经性聋患者汉语普通话声调知觉模式研究[J]. 临床耳鼻咽喉头颈外科杂志, 2017, 31(1): 38—40.
- [14] 元贝尔, 刘鹏, 古鑫, 等. 轻中度感音神经性聋患者汉语普通话声调知觉模式研究[J]. 临床耳鼻咽喉头颈外科杂志, 2017, 31(1): 38—40.
- [15] ALBU S, TRABALZINI F, AMADORI M. Usefulness of cortical mastoidectomy in myringoplasty[J]. Otol Neurotol, 2012, 33: 604—609.
- [16] TOROS S Z, HABESOGLU T E, HABESOGLU M, et al. Do patients with sclerotic mastoids require aeration to improve success of tympanoplasty[J]. Acta Otolaryngol, 2010, 130: 909—912.
- [17] MISHIRO Y, SAKAGAMI M, TAKAHASHI Y, et al. Tympanoplasty with and without mastoidectomy for non-cholesteatomatous chronic otitis media[J]. Eur Arch Otorhinolaryngol, 2001, 258: 13—15.
- [18] MISHIRO Y, SAKAGAMI M, KONDOH K, et al. Long term outcomes after tympanoplasty with and without mastoidectomy for perforated chronic otitis media[J]. Eur Arch Otorhinolaryngol, 2009, 266: 819—822.
- [19] 方敏, 杨贵军, 李朝军, 等. 乳突病变切除加鼓室成形术与单纯鼓室成形术治疗中耳炎疗效的 Meta 分析[J]. 听力学及言语疾病杂志, 2016, 24(1): 75—80.
- [20] 何圆圆, 肖红俊. 鼓室及乳突病变状况与鼓膜成形术[J]. 临床耳鼻咽喉头颈外科杂志, 2011, 25(1): 45—47.
- [21] MCCOUL E D, ANAND V K, CHRISTOS P J. Validating the clinical assessment of eustachian tube dysfunction: the Eustachian Tube Dysfunction Questionnaire (ETDQ-7)[J]. Laryngoscope, 2012, 122: 1137—1141.
- [22] DOYLE W J, SWARTS J D, BANKS J, et al. Sensitivity and specificity of eustachian tube function tests in adults[J]. JAMA Otolaryngol Head Neck Surg, 2013, 139: 719—727.
- [23] SCHRÖDER S, LEHMANN M, SUDHOFF H, et al. [Assessment of chronic obstructive eustachian tube dysfunction: Evaluation of the German version of the Eustachian Tube Dysfunction Questionnaire]. HNO, 2014, 62: 160, 162—164.
- [24] 肖社平, 谭代忠, 廖敏清, 等. 声导抗测试对慢性化脓性中耳炎咽鼓管功能的评价[J]. 临床耳鼻咽喉头颈外科杂志, 2006, 20(7): 306—311.

(收稿日期: 2017-07-19)

(收稿日期: 2017-06-05)