

# 常用中文言语测听材料在语后聋人工耳蜗植入者中的天花板效应\*

王焯<sup>1</sup> 石涯<sup>2</sup> 傅莹<sup>1</sup> 王倩<sup>1</sup> 符一飞<sup>1</sup> 郝昕<sup>1</sup>

**[摘要]** 目的:总结目前常用中文言语测试材料在语后聋人工耳蜗植入者言语识别上的效果。方法:采用临床常用的中文言语测试材料,对 32 例语后聋人工耳蜗植入者进行安静条件下单音节字、扬扬格词、中文 BKB 语句、中文 HINT 语句识别率以及嘈杂噪声下 HOPE 语句识别阈的测试,分析比较各类测试材料在语后聋人工耳蜗植入者康复成效评价时的临床适用性。结果:语后聋人工耳蜗植入者在安静条件下扬扬格词表、中文 BKB 句表和中文 HINT 句表的识别率达到 100% 的人数分别为 1 人、14 人和 4 人;同时,高分数段(≥85%)人数分别为 17 人、26 人和 14 人,均呈现出不同程度的天花板效应。反之,安静条件下单音节字识别率满分为 0 人,高分数段人数为 6 人;噪声下 HOPE 句表识别阈测试不受天花板效应干扰。结论:对于相当一部分语后聋患者,现有的中文言语测试材料已出现不同程度的天花板效应。安静条件下单音节字识别率和噪声下 HOPE 句表识别阈测试更能相对客观地评价患者的言语识别能力。

**[关键词]** 耳蜗植入术;语后聋;言语识别;天花板效应

doi:10.13201/j.issn.1001-1781.2015.04.003

**[中图分类号]** R764.5 **[文献标志码]** A

## Study of ceiling effect of commonly used Chinese recognition materials in post-lingual deafened patients with cochlear implant

WANG Ye<sup>1</sup> SHI Ya<sup>2</sup> FU Ying<sup>1</sup> WANG Qian<sup>1</sup> FU Yifei<sup>1</sup> XI Xin<sup>1</sup>

(<sup>1</sup>Department of Otorhinolaryngology Head and Neck Surgery, Chinese PLA General Hospital, Beijing, 100853, China; <sup>2</sup>Zhejiang Chinese Medical University, College of Medical Technology) Corresponding author: XI Xin, E-mail: xixin\_plagh@yeah.net

**Abstract Objective:** To evaluate and compare outcomes effects of Chinese recognition materials in post-lingually deafened patients with cochlear implant. **Method:** Thirty-two post-lingually deafened cochlear implant users participated in the study. Each one of them was given a series of speech recognition tests including the speech recognition score of monosyllable words, spondee words, Chinese BKB sentences and MHINT in quiet, as well as speech recognition threshold of HOPE corpus in babble noise. Speech recognition scores and thresholds were compared among these test materials using histograms, scatter diagrams and statistical methods. **Result:** Recognition scores for spondee words, Chinese BKB sentences and MHINT were affected by ceiling effects, with 1, 14 and 4 cases scored 100% respectively. Meanwhile, 17, 26 and 14 cases scored more than 85% correspondingly. On the other hand, speech recognition scores for monosyllable words in quiet and speech recognition threshold for HOPE corpus in babble noise were not affected. **Conclusion:** For a considerable part of post-lingually deafened patients using cochlear implant, some Chinese speech recognition materials have demonstrated different degrees of ceiling effects. While speech recognition scores for monosyllable words in quiet and recognition threshold for HOPE corpus in noise can evaluate the performance of speech recognition ability more objectively.

**Key words** cochlear implantation; post-lingually deafened; speech recognition; ceiling effect

人工耳蜗作为一种高科技生物医学装置,综合了多个学科的技术,能帮助众多双耳罹患重度或极重度感音神经性聋的成人及儿童获得听觉。目前全球人工耳蜗植入患者总数已超过 30 万<sup>[1]</sup>。从我国 1995 年首次实施人工耳蜗植入术以来,已有越来越多的患者通过此项手术重新获得了听觉,从一

开始的单通道人工耳蜗到如今的多通道人工耳蜗,术后效果不断进步,已远远超出人工耳蜗刚刚问世时人们的预期。随着人工耳蜗植入技术的日臻完善,语后聋人工耳蜗植入成效显著提高,中文言语测试也开始显露出天花板效应的趋势。以 MHINT 为例,既可在安静及噪声条件下测试患者特定言语级或固定信噪比的言语识别率,又可以自适应(adaptive)方式逐句调整每句话的信噪比来测试患者的信噪比识别阈。相对于英文版 HINT, MHINT 在编撰之初增加了语句的长度,其在噪声下进行识别阈测试的难度明显超过了成人耳蜗植

\* 基金项目:国家自然科学基金项目(No: 61370023, 81460099)

<sup>1</sup>解放军总医院耳鼻咽喉头颈外科(北京,100853)

<sup>2</sup>浙江中医药大学听力与言语科学学院

通信作者:郝昕, E-mail: xixin\_plagh@yeah.net

入者的水平<sup>[2]</sup>。故临床工作中,受制于客观时间和条件限制,大多数听力师采用的是安静条件下对人工耳蜗植入患者进行测试,但值得注意的是,语后聋人工耳蜗植入患者在安静条件下 MHINT 语句识别率却已开始出现天花板效应。国内相关文献报道,植入 1 年左右的成人患者 MHINT 识别率平均分也能够达到 70%<sup>[3]</sup> 以上。由此可见,找到最行之有效的评估人工耳蜗植入效果的方法,和人工耳蜗的技术革新一样难。有学者认为理想的言语测试材料的特点应当是:可靠、不同测试条件下具有高敏感性、同时还与现实听觉环境下的言语感知高度相关<sup>[4]</sup>。由于植入者的表现会受到诸如生理变化、调机改变、音量/敏感度调节、劳累等多种因素的影响,单纯依靠一种测试方法无法满足客观的评估需求。日常的临床工作中,听力师也会采取多种评估方法相结合的办法,通过不同测试间的相互印证判断测试结果的可靠性。因此本文的目的是:使用常用的中文言语测试材料对患者进行测试并分析结果;评估不同测试结果之间的相关性,观察是否可以将不同类型的测试组合起来用于人工耳蜗植入者的评估。

### 1 资料与方法

#### 1.1 研究对象

32 例语后聋人工耳蜗植入者参与本研究,其基本情况见表 1。在耳聋前均达到正常言语能力水平,术前诊断均为双侧重度感音神经性聋。32 例中,男 15 例,女 17 例;植入年龄 4~59 岁,平均 25.8 岁。所有受试者在术后 1 个月左右开机并适时调机。耳蜗装置使用时间 12 个月~10 年,平均 38 个月。

#### 1.2 测试材料

使用汉语普通话版 MHINT 句表进行安静条件下识别率测试。MHINT 语句由香港大学、美国 House 耳研所、北京市耳鼻咽喉科研究所联合开发,共 12 张句表,每张表 20 个短句,每句 10 个字。MHINT 语音文件已集成到美国 House 耳研所提供的 BLIMP(bilateral implant test system)软件中,可以控制扬声器给声强度及侧别,根据受试者正确复述的字的数量,自动计算每张表的识别率。

解放军总医院和清华大学编制的“心爱飞扬”言语测试软件,包括 22 张单音节字表(每张 25 个音节)、6 张扬扬格词表(每张 40 词)、12 张安静条件下 BKB 语句表(每张表 10 句共 50 个关键词)、14 张嘈杂噪声下 BKB 语句表(每张表 18 句 100 个关键词)。所有语音测试材料均经过等价性验证。软件可实现安静条件下单音节字识别率、扬扬格词识别率、安静条件下 BKB 语句识别率和嘈杂噪声下 HOPE 语句识别阈的测试,并自动归档测试结果。

#### 1.3 测试方法

测试在经校准满足准自由场条件的隔声室进行,本底噪声在 20 dBA 以下。使用 USB 音频转接头连接笔记本电脑和经校准的 GSI61 听力计,通过高精度扬声器播放语音,给声强度为 65 dB 言语听力级别。扬声器位于其植入侧 45°前方 1 m 处,与患者坐姿时耳部高度齐平。测试前均使用练习表进行练习,确保受试者熟悉测试流程并积极配合。受试者在练习和测试过程中不使用重复的测试材料,测试人员根据受试者复述的内容以字、词句中关键词为单位记录结果,软件自动计算语句识别率及识别阈。

表 1 32 例语后聋人工耳蜗植入者的基本情况

例序	年龄/岁	病因	佩戴时间/年	例序	年龄/岁	病因	佩戴时间/年
1	28	药物性聋	5	17	52	药物性聋	1
2	26	药物性聋	4	18	10	大前庭水管	8
3	27	药物性聋	4	19	11	不明原因	1.5
4	16	脑膜炎	4	20	26	突发性聋	3
5	37	不明原因	3	21	12	药物性聋	6
6	60	突发性聋	1	22	12	药物性聋	8
7	21	不明原因	4	23	15	不明原因	5
8	35	突发性聋	1	24	16	突发性聋	10
9	50	突发性聋	1	25	24	不明原因	4
10	39	突发性聋	1.5	26	19	药物性聋	3
11	46	突发性聋	1	27	29	突发性聋	1
12	18	不明原因	1	28	38	突发性聋	3
13	40	突发性聋	1	29	58	突发性聋	1
14	56	突发性聋	3	30	44	突发性聋	1
15	22	不明原因	1	31	19	不明原因	5
16	18	不明原因	1	32	59	突发性聋	3

### 1.4 统计学方法

使用 SPSS 18.0 统计软件,分析所有受试者安静条件下单音节字识别率、扬扬格词识别率、BKB 语句识别率、MHINT 语句识别率和嘈杂噪声下 BKB 语句识别率的分布规律,并对上述结果两两进行相关性分析。

## 2 结果

### 2.1 安静条件下人工耳蜗植入受试者言语识别率及嘈杂噪声下识别阈测试结果

所有受试者五项测试(安静条件下单音节字、扬扬格词、中文 BKB 语句、MHINT 语句识别率以及噪声下 HOPE 语句识别阈)的测试结果分布情况见图 1 及表 2。若患者对某项测试的识别率为 100%,则达到了该测试的“天花板”,定义为天花板效应。而嘈杂噪声下 HOPE 语句识别阈的测试,使用了自适应方法来得到患者识别率 50%所对应的信噪比,测试时的信噪比总是随患者应答的正误在识别阈附近上下波动,故不存在天花板效应。32 例受试者,有 4 例的安静下 MHINT 语句识别率、14 例的中文 BKB 语句识别率达到了天花板效应,而且高分段(识别率 $\geq 85\%$ )人数也较多。相比之下,安静条件下单音节字识别率的分布相对均匀,没有一个受试者的单音节字识别率达到 100%,安

静条件下扬扬格词测试中只有 1 例识别率得了满分。图 1e 显示了嘈杂噪声下 HOPE 语句识别阈的分布情况,与单音节词识别率测试的结果十分类似,其分布也比较均匀,32 例受试者中,识别阈的信噪比得分最低为 -3.1 dB,最高达到 21 dB,识别阈的均值为 8.7 dB SNR。

图 1a~e 可使我们用肉眼对上述五项测试的结果是否符合正态分布有一个大致的判断,而通过 SPSS 18.0 计算出的上述数据分布的偏度系数(表 2)则更能提示数据的正态性与否。偏度系数的绝对值越接近零,表示数据越接近正态分布。表 1 中显示的是各项测试根据分布情况计算出的偏度系数。安静条件下扬扬格词、BKB 语句和 MHINT 语句识别率的偏度系数明显高于其他,识别率结果存在明显偏移,即普遍接近最大值。

### 2.2 安静条件下各测试识别率同嘈杂噪声下 HOPE 语句识别阈的比较

图 2~5 显示的是利用安静条件下四种测试(MHINT 语句、中文 BKB 语句、扬扬格词表、单音节词表)的识别率得分作为横坐标所对应嘈杂噪声下 HOPE 语句识别阈作为纵坐标绘制的散点图。在测试结果均值处添加辅助线,可见扬扬格词测试、中文 BKB 语句识别率和 MHINT 语句识别率

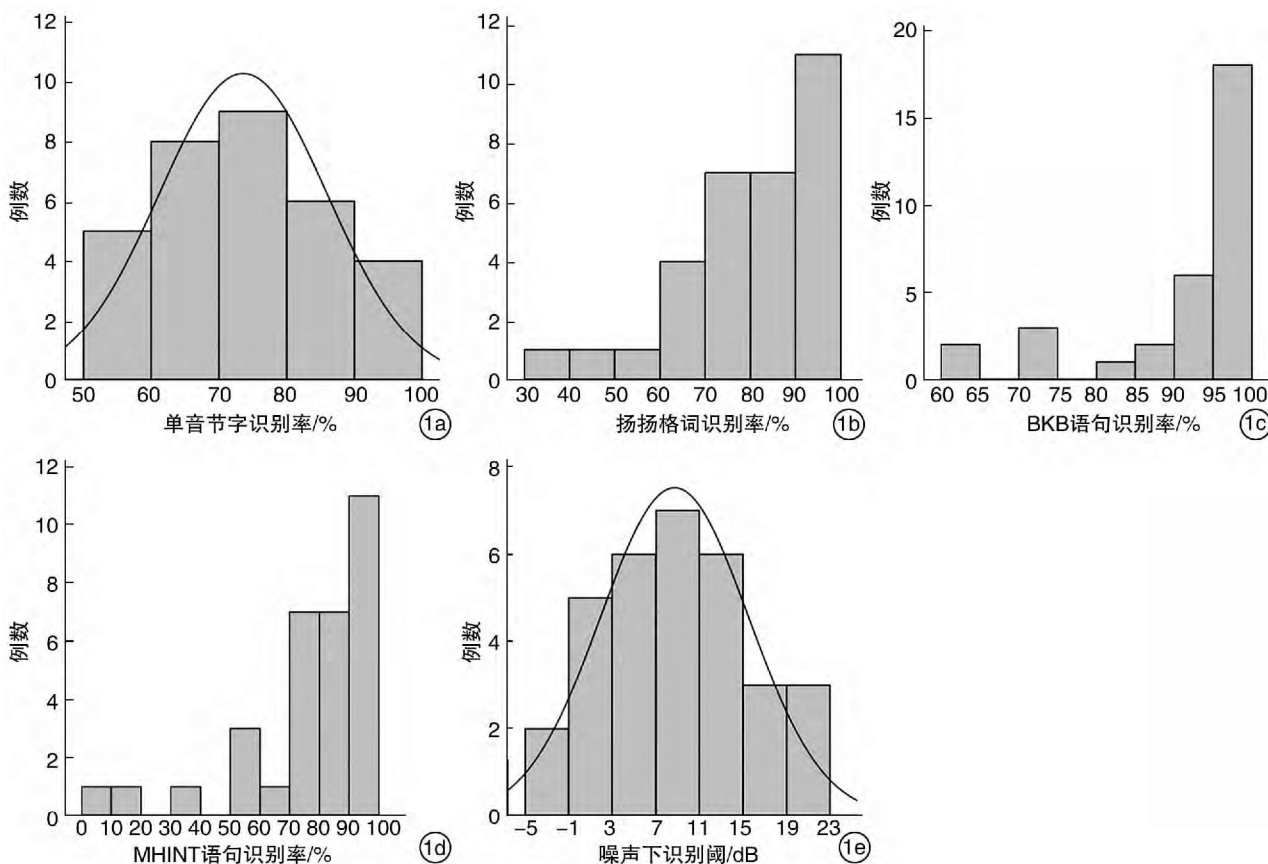


图 1a~e 各项测试结果的直方图

表 2 各项测试的得分分布情况

测试项目	均值	识别率 85%~<100%的人数	识别率=100%的人数	偏度系数
单音节	73.6%	6	0	0.144
扬扬格	80.1%	17	1	-1.14
BKB 语句	92.1%	26	14	-1.54
HINT 语句	77.4%	14	4	-1.53
噪声下识别阈	8.7 dB			0.16

得分点大部分集中于纵轴均值以上, 分别占总人数的 56.2% (18 例)、71.9% (23 例) 和 56.2% (18 例), 但是在横坐标上的分布较平均。

单音节词表测试中, 无一例受试者的识别率达到 100%, 且无论在纵坐标还是横坐标散点均没有显著聚集。散点图中添加分别以单音节字识别率均值和噪声下识别阈均值绘制辅助线, 可以将所有散点分成 4 个象限, 左上象限代表在两项测试中表现均较好的点, 而右下象限代表在两项测试中表现均不佳的情况, 这两种情况代表这两项测试的结果对应良好。左上和右下象限的点分别为 12 和 11, 占总人数的 71.9%。

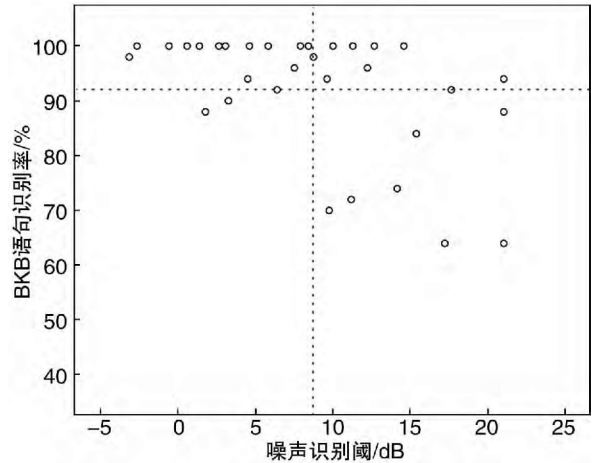


图 4 BKB 语句识别率与噪声识别阈比较的散点图

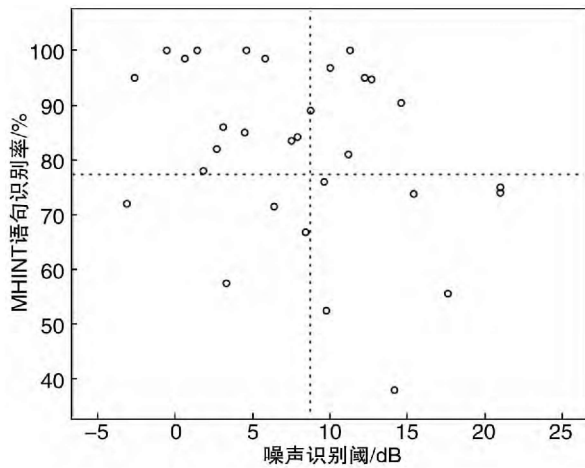


图 2 MHINT 语句识别率与噪声识别阈比较的散点图

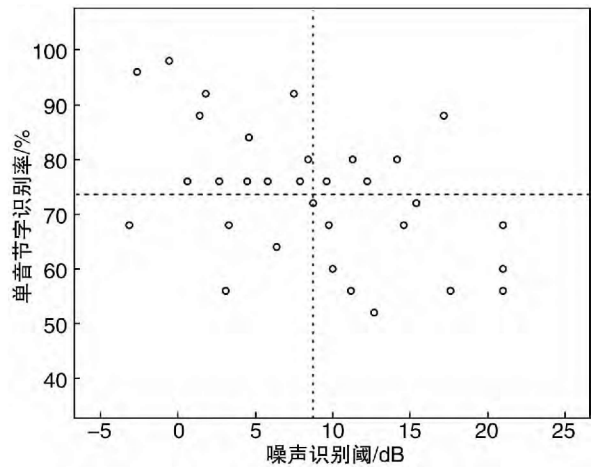


图 5 单音节字识别率与噪声识别阈比较的散点图

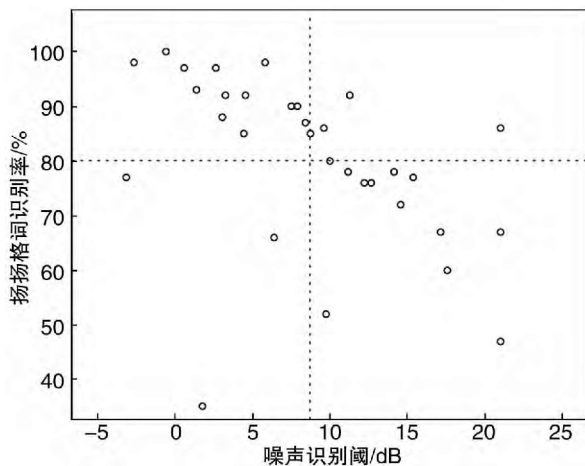


图 3 扬扬格词识别率与噪声识别阈比较的散点图

安静条件下 MHINT 语句识别率、中文 BKB 语句识别率、单音节字识别率、扬扬格词识别率以及嘈杂噪声下 HOPE 语句识别阈的结果的相关性可以通过对以上各项测试进行相关性分析明确。经 Pearson 相关性分析可知, 安静条件下所有测试项的识别率结果都与噪声下 BKB 语句的识别阈结果显著相关(表 3)。

### 3 讨论

#### 3.1 不同言语测试材料结果的分析

通过实验得到的直方图分布及偏度系数计算可见, 安静条件下的扬扬格词识别率测试、中文 BKB 语句识别率测试、MHINT 语句识别率测试都

表 3 安静条件下识别率与噪声下识别阈的相关性分析

相关系数( <i>r</i> )	单音节字识别率	扬扬格词识别率	中文 BKB 语句识别率	MHINT 语句识别率
噪声下 HOPE 语句识别阈	-0.482 ( <i>P</i> <0.01)	-0.453 ( <i>P</i> <0.01)	-0.497 ( <i>P</i> <0.01)	-0.491 ( <i>P</i> <0.01)

不同程度受到天花板效应的影响,而安静条件下单音节字识别率测试同噪声下 HOPE 语句识别阈测试基本不受天花板效应干扰。

嘈杂噪声下语句识别阈的获得采取的是“自适应”的方法,即根据患者自身的能力调整信噪比,最终得到的是患者“信噪比损失”。而在单音节测试中,单音节字表能很好地降低词的熟悉程度和人的记忆效应带来的干扰,因此,安静条件下单音节字识别率测试同嘈杂噪声下 HOPE 语句识别阈测试收到天花板效应的影响较小。

扬扬格词为双重读词,所含信息高于单个音节的字,受试者能够一定程度地根据日常经验判断,但所含信息少于短句,因此得分普遍高于单音节字但又低于中文 BKB 语句识别率。

中文 BKB 短句最初开发的目的是用于对听力损失的儿童进行语句识别测试,选取的语料来源是北京地区 4~5 岁儿童的日常口语,成人患者在测试中很容易利用听到的片段推测出上下文,故安静条件下识别率测试得分较高,受到天花板效应影响最明显。

### 3.2 不同测试材料之间相关性的分析

虽然通过相关性分析能够得出嘈杂噪声下 HOPE 语句识别阈与其他所有的安静条件下测试的识别率结果均具有相关性,但是并不能认为这些安静条件下得出的识别率结果可以对噪声下言语识别阈做出预测。本研究中共有 4 例患者在安静条件下 MHINT 语句测试中获得识别率的满分,但是这些患者噪声下 HOPE 语句识别阈的信噪比结果分别为-0.6 dB,1.4 dB,4.6 dB 和 11.3 dB,最高和最低的信噪比损失之间的差距高达 11.6 dB。结果表明,虽然受试者安静条件下 MHINT 识别率均为高分,但并不代表他们的言语能力处在同样的水平;14 人的 MHINT 识别率达到 85%,占总人数的 43.8%,但他们的噪声下识别率分布在-2.6 dB~14.6 dB,这个信噪比的分布范围已经分别接近了最低和最高的信噪比。同样这一问题也适用于安静条件下中文 BKB 语句和扬扬格词测试的结果。

正是由于目前安静条件下扬扬格词识别率、中文 BKB 语句识别率和 MHINT 语句识别率都或多或少受到天花板效应的影响,因此这些识别率测试的得分,特别是高分段的得分,难以客观评价人工耳蜗使用者,尤其是已经多年使用的成熟植入者的言语能力。相对的,安静条件下单音节字识别率、

嘈杂噪声下 HOPE 语句识别阈测试受到天花板效应的干扰较小,结果更可靠。对于高水平的植入者,此两者结合有助于判断患者某项测试的得分究竟是患者自身言语能力问题还是受到了疲劳、设备故障等外在因素的影响,结合测试结果辅助听力师进行人工耳蜗调机。

### 3.3 当前临床言语测试材料应用情况

对植入者的言语识别能力的评估是检验人工耳蜗植入效果的最直接和最根本的手段<sup>[5]</sup>。早在 1996 年,美国听力学会(AAA)、美国耳鼻咽喉头颈外科学会(AAO-HNS)和人工耳蜗制造商代表组成的委员会推出了一套用于评估人工耳蜗术后康复效果的开放式言语测试材料,称为最小言语测试集(minimum speech test battery, MSTB)<sup>[6]</sup>,此项测试集包括辅音-元音-辅音(consonant-nucleus-consonant, CNC)测试和选自 HINT<sup>[7]</sup>的部分语句材料,均为开放式测试,前者评估植入者的单音节识别能力,后者评估植入者在安静和(或)噪声条件下的语句识别能力。这是成人耳蜗植入成效评价的最基本内容。

由于人工耳蜗技术的进步,植入效果也不断改善, Gifford 等<sup>[8]</sup>就曾提出,高达 71% 的成人耳蜗植入者在安静条件下 HINT 语句识别率可以达到 85% 甚至更高,测试呈现出一定的天花板效应,同时还指出 AzBio 句表比 HINT 语句能更好地评估患者的言语识别能力。因此 2011 年 6 月该委员会重新修订了 MSTB,除继续保留 CNC 词表外,建议采用更高难度的 AzBio 句表<sup>[9]</sup>和 BKB-SIN 噪声下言语测试,以分别评估人工耳蜗成人植入者在安静及噪声环境下的表现。

参照英文言语测听的发展轨迹,中文的言语测听材料经历 60 年的发展,也建立了包括单音节字、扬扬格词、安静及噪声下短句的成人普通话言语测听的完整体系<sup>[10]</sup>,并已应用于语后聋人工耳蜗植入者的言语感知能力评估。张华等<sup>[11]</sup>开发了 10 张 50 字的音位平衡表,其中 7 张具有良好的表间等价性。郗昕等<sup>[12]</sup>创编了 22 张彼此等价的 25 字单音节短表,并出版了商品化的单音节语音 CD。李剑挥等<sup>[13]</sup>编辑并验证了 6 组每组 40 词的双音节测听词表。有关中文语句测听材料,黄丽娜等研发了普通话版 HINT(MHINT),不过为实现同英文版本的兼容性,她将汉语语句加长到 10 字<sup>[14]</sup>。郗昕等则开发了一套噪声下的希冀(HOPE)短句集,每句

6~8个字,采用4人交谈混叠后的嘈杂噪声以模拟日常环境中的嘈杂场景<sup>[15]</sup>。该测试短句也可以在安静条件下播放,由于其难度等同于英文BKB语句,故安静条件下播放时又称为中文BKB句表。

有学者认为应该将语句作为言语测试材料,因为句子比单、双音节词和无意义信号更接近于实际聆听情况,符合语言的自然规律<sup>[16]</sup>,但是作为言语测试材料,具有明确语义的扬扬格词和语句含有较大的冗余度和较强的学习效应<sup>[17]</sup>,如果需要满足患者长期的、不同阶段的测试需求,就应该有大量等价的、具有更高难度的语句测试表,这些等价的测试表可以保证同一患者在每次测试的结果之间都具有可比性。

随着国内人工耳蜗植入术的进展以及相关技术和康复水平的进步,尤其是对于语后聋患者及早期植入者,需要有相应配套难度的测试材料能够客观评估其言语发展情况,有利于正确指导调机和术后康复训练。因此,我们也有必要发展难度更高的言语测试材料,避免本文中已经出现和随着技术进步必将会出现的更多的安静条件下的天花板效应问题。

#### 参考文献

- [1] HOLMES D. Lasker Foundation honours cochlear-implant pioneers [J]. *Lancet*, 2013, 382: 926-926.
- [2] 亓贝尔,刘博,张宁,等.人工耳蜗植入者噪声环境下普通话言语测试方案的优化研究[J]. *听力学及言语疾病杂志*, 2011, 19(4): 305-310.
- [3] 李佳楠,郝昕,洪梦迪,等.语后聋患者多通道人工耳蜗植入后听觉效果分析[J]. *临床耳鼻咽喉头颈外科杂志*, 2010, 24(13): 580-582.
- [4] MACKERSIE C L. Tests of speech perception abilities. [J]. *Curr Opin Otolaryngol Head Neck Surg*, 2002, 10: 392-397.
- [5] ZENG F G, REBSCHER S, HARRISON W, et al. Cochlear implants: system design, integration, and evaluation[J]. *IEEE Rev Biomed Eng*, 2008, 1: 115-142.
- [6] NILSSON M, MCCA W V, SOLI S D. User manual: A minimal speech test battery for adult cochlear implant patients[M]. Los Angeles: House Ear Institute, 1996: 4-5.
- [7] NILSSON M, SOLI S D, SULLIVAN J. Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise [J]. *J Acoust Soc Am*, 1994, 95: 1085-1099.
- [8] GIFFORD R H, SHALLOP J K, PETERSON A. Speech recognition materials and ceiling effects: considerations for cochlear implant programs[J]. *Audiol Neurootol*, 2008, 13: 193-205.
- [9] SPAHR A J, DORMAN M F. Performance of subjects fit with the Advanced Bionics CII and Nucleus 3G cochlear implant devices [J]. *Arch Otolaryngol Head Neck Surg*, 2004, 130: 624-628.
- [10] 张宇晶,郝昕.成人人工耳蜗植入相关的中文言语识别评价体系的建立[J]. *听力学及言语疾病杂志*, 2012, 20(4): 387-389.
- [11] 张华,王靓,王硕,等.普通话言语测试单音节词表的编辑与初步等价性评估[J]. *中华耳鼻咽喉头颈外科杂志*, 2006, 41(5): 341-345.
- [12] 郝昕,冀飞,陈艾婷,等.汉语普通话单音节测试表的建立与评估[J]. *中华耳鼻咽喉头颈外科杂志*, 2010, 45(1): 7-13.
- [13] 李剑挥,郝昕,冀飞,等.一组汉语普通话双音节测试词表的编辑与等价性研究[J]. *中国基层医药*, 2010, 17(10): 1297-1298.
- [14] WONG L L, SOLI S D, LIU S, et al. Development of the Mandarin hearing in noise test(MHINT)[J]. *Ear Hear*, 2007, 28(2 Suppl): 70S-74S.
- [15] XI X, CHING T Y, JI F, et al. Development of a corpus of Mandarin sentences in babble with homogeneity optimized via psychometric evaluation[J]. *Int J Audiol*, 2012, 51: 399-404.
- [16] 郝昕,冀飞,陈艾婷,等.怎样发展标准化的中文言语测试材料[J]. *中华耳科学杂志*, 2008, 6(1): 13-16.
- [17] 赵阳,郝昕,冀飞,等.嘈杂语噪声下汉语语句测试中的学习效应[J]. *听力学及言语疾病杂志*, 2009, 17(2): 107-110.

(收稿日期:2014-10-23)