
计算机辅助的中文言语测听平台的建立

郝昕¹ 黄高扬² 冀飞¹ 贾珈² 陈艾婷¹ 赵阳¹ 蔡莲红² 韩东一¹

¹解放军总医院耳鼻咽喉-头颈外科

²清华大学计算机技术科学系

言语测听和纯音测听一样，是听力学最基本、最重要的测试手段^[1]。它既可在临床诊断中与纯音听力图相互印证，提供患者的言语听敏度信息，又能够评估患者在言语察觉、分辨、识别和理解等方面的能力。

言语测听的普及取决于言语测听材料和言语测听方法两方面的进展，二者是内容和形式的关系。测听材料的发展不断拓展着言语测听的应用范畴，从这项基本的心理物理测试中挖掘出更多的有用信息；测试技术的改进则提高着言语测听的效率和可靠性，使其提供的信息更加细化和准确。两者相互依存，共同推进着现代言语测听技术的发展。而数字化信息技术的介入^[2]已成为未来发展的方向。

目前英文言语测听，无论在测听材料还是测听方法方面，都已发展得比较成熟^[3]。常用的测听材料包括：测试言语识别率的 CID W-2、NU-6、CNC 等单音节词表和 BKB、CUNY 等短句表，以及一些噪声下语句测听，如 SIN、Quick SIN、HINT 等。近年来国外也开发了一些计算机辅助的言语测听系统，但尚未在临床中广泛使用。

中文言语测听的普及长期受到“测听材料匮乏”和“测听方法不规范”两个瓶颈的制约，有很大的改进空间^[4]。解放军总医院耳鼻咽喉科研究所近年来一直致力于发展标准化的普通话（单音节、扬扬格词、短句及噪声下语句）言语测听材料，出版了国内第一张言语测听 CD^[5]，经大量的全国各方言区的多中心临床验证^[6]，已能满足临床上对测试信度、效度和实用性的要求。但全国各地普遍存在着“言语测听方法不规范”的问题。(1)言语测听手法掌握不当；(2)对影响测试结果可靠性的诸多因素（如声学校准、学习记忆效应、受试者应答方式等）认识不足；(3)对言语测听结果的临床意义存在误读；(4)测听人员仍然使用“一张纸、一支笔”的记录方式，需手工计算识别率，数据不易汇总分析。国内虽也有准计算机化言语测听的零星研究^[7,8,9]报道，但未能普及。

本研究将基于解放军总医院业已完成的一系列标准化的普通话言语测听材料，结合临床常规听力计的外接音源输入及言语测听模块，研制开发出一套计算机辅助的中文言语测听平台^[10]，以 Microsoft Visual C++编写完成。该测听软件集受试者信息管理、声学校准、语音播放、测听流程自动化、测听报表生成、数据分析管理等功能为一体。可以辅助进行单音节识别率、扬扬格词识别率及识别阈、安静下语句识别率及识别阈、噪声下的语句识别率及识别阈等测试，在一定程度上克服了上述“测听方法不规范”的弊端，为临床应用及科研工作提供了一个实用而有效的计算机辅助工具，也为中文言语测听的推广提供了切实有效的途径。

基金项目：“十一五”国家科技支撑计划(2007BAI18B12, 2008BAI50B01, 2008BAI50B08)；解放军总医院苗圃基金(08MP-11)

通讯作者：韩东一，hdv301@263.net；蔡莲红，CLH-dcs@tsinghua.edu.cn

一 软件结构

言语测听软件具有友好的输入输出界面，提供的功能主要包括数据管理、言语测听流程管理两部分。言语测听软件主要功能的流程如图 1 所示。

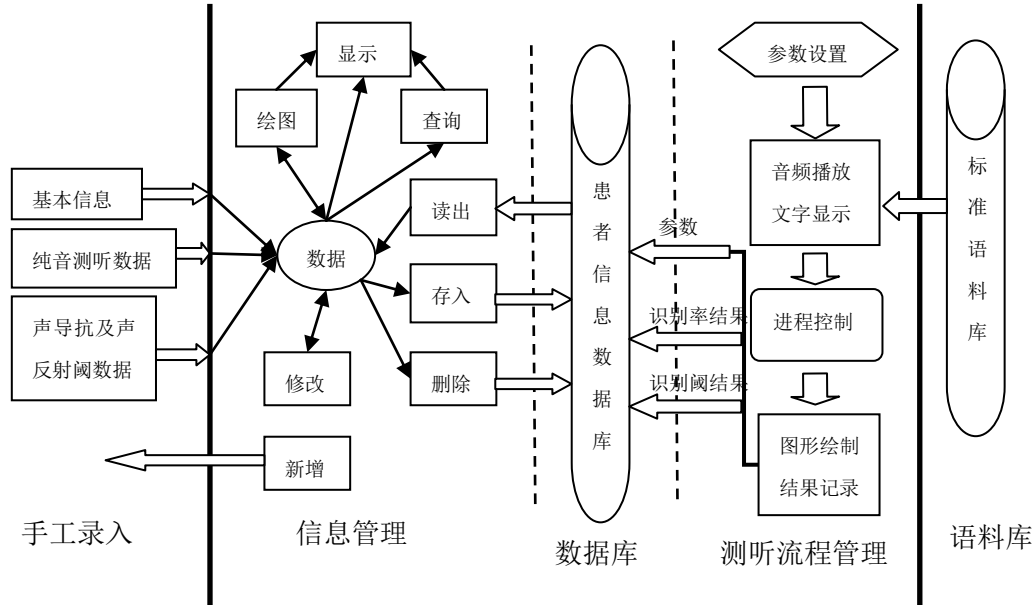


图 1 言语测听软件主要功能流程图

1 数据管理

“数据管理”功能负责管理患者基本信息、常规听力学（纯音及声导抗测听）数据、言语测听条件及结果，具有增、删、查、改、导入\导出、报表生成的功能。

(1) 患者基本信息

患者的基本信息包括姓名、年龄、性别、母语方音、常住地、普通话水平、耳科诊断、助听装置使用状态等。

(2) 常规听力学数据

包括患者双耳的纯音（气、骨导）听力图、鼓室图及声反射阈。纯音听力图的录入可采用数值录入或在图上直接标画特定符号两种方式，软件自动计算出患者左右耳在 0.5、1、2kHz 三个频率点的平均听阈(3FA)和在 0.5、1、2、4kHz 四个频率点的平均听阈(4FA)。

(3) 言语测听条件及结果

同步记录言语测听所采用的换能器种类（TDH 系列耳机、插入式耳机、扬声器及扬声器的入射方位角）、侧别（左、右、双耳）、测听内容及结果（识别率、识别阈）等。可绘制同一条件下受试者的识别率随言语声强变化的函数(Performance - Intensity, P-I)曲线，并生成图形文件保存或导出。

(4) 数据管理

对患者的基本信息可实现新增、删除、修改、检索等功能。其中的检索功能，可按姓名、年龄、性别、测听日期、病种、听损程度等多个检索条件组合进行检索，极大地方便了日后的统计分析和科研工作。测听结果可以直接打印成临床报表，或转成文本文件或 Excel 文件保存。对于在多台电脑上使用该软件所生成的数据文件，可以方便地实现导入\导出，从而有利于同一家医院门诊、病房等多个测听场所的数据汇总，也为在多家医院间开展多中心临床研究提供了方便。

2 言语测听流程管理

“言语测听流程管理”功能则负责对各种测听流程的控制。如图 1 所示，计算机播放的测听材料以文字方式显示在电脑屏幕上，其音频信号被送至听力计的外接信号输入端，由听力计控制言语信号的给声强度，并经由其内置功放输出到相应的换能器（TDH 系列耳机、ER-3A 插入式耳机、扬声器或骨振器等）。听力师借助听力计的监听耳机及对讲系统，实时监控音频文件的播放，并以人机交互方式记录患者的应答、操控测听流程的进行。

(1) 声学定标

既然计算机播放的言语信号需经由听力计来控制给声强度并以听力级(dB HL)为单位表述，这就涉及到听力计的声学校准及外接言语信号的定标问题。在英文中虽均以 calibration 表述，但实则涉及的是两个层面的问题。

听力计属国家强制检定设备，每年须校准一次，主要完成对各类换能器的纯音、窄带噪声及言语信号的零级校准^[11]。比如，对于 TDH 系列耳机而言，其言语识别基准阈级（即言语信号的听力零级）为 20 ± 2 dB SPL。应由专业的校准工程师调整听力计内部的输出控制电路，使得听力计表盘显示为 70 dB HL 时，耳机输出的言语信号强度为 90dB SPL。

而对外接信号的定标，则是对来自计算机声卡的言语信号电平做适度微调，使之符合听力计内置功放的输入电平要求。但言语信号是在时域与频域上都不断变化的非稳态信号，因而国家标准《GB/T 17696 声学 测听方法 第三部分：语言测听》^[12]规定，言语测听录音成品中应包括定标言语听力计的信号音，应不短于 60s，可以是 1kHz 为中心的 1/3 倍频程窄带噪声，或带宽不少于 1/3 倍频程的 1kHz 调频啁音，调制信号应为重复率在 4~20 范围内的正弦或三角波。定标信号音的强度应与该测听材料中的全部测试表的平均声级一致。因此本软件启动之后，需先播放 60 秒 1kHz 调频啁音。此过程中同步调整听力计的输入电平微调，使 VU 表的示数指到 0dB（外接信号电平过低时，也可为-10dB），言语测听软件会要求听力师录入这个 VU 表基值。实际测试时将听力计的表盘示数加上此 VU 表基值，便是人耳实际听到的言语强度。只有完成外接言语信号的定标，屏幕上的言语测听模块才由灰转黑，允许进行正式测听。

(2) 乱序播放与同步计分

以往使用言语测听 CD 或其它制成品，听力师选定测听表号并在预先印制好的计分表上记录患者的应答。各测听表及表内各个测听项之间的排列顺序都是固定的，对于短期内

进行过多次测听的受试者而言，就有可能已经记住了部分言语内容。为了克服这一记忆效应的影响，本软件不允许同一受试者重复使用同一张测听表，并对每张表内各测听项以随机顺序播放，相应的文字内容同步显示在电脑屏幕上。受试者复述之后，听力师直接在文字下方相应的判断框点击“正\误”，每表测听完毕就会自动计算并显示言语测听结果。

(3) 练习表的应用

为了克服练习效应^[13]的影响，每类测试之前都应要受试者进行充分的练习，以熟悉测听环境及流程。为此软件对每类测听都提供了数量不等的练习表。

二 测试内容

言语测听主要涉及言语识别阈、言语识别率两大类。言语识别阈是指受试者能够正确识别当前播放言语材料的 50%时所对应的最低声级。而言语识别率是指用百分数表示的受试者正确重复测试项数占总测试项数目的比例。两项测试各有其特点，测试结果既可与纯音测听的结果相对照并用于诊断，又可用于评估患者的听敏度及治疗、康复的疗效。

1 言语识别率

言语识别率的测试内容包括单音节、扬扬格词表、安静下句子、噪声下句子四种类型，包括 22 张单音节表（每表 25 个音节）、5 张扬扬格词表（每表 40 词）、12 张安静条件下语句表（每表 10 句 50 个关键词）和 28 张嘈杂噪声下语句表（每表 9 句 50 个关键词）。测试前，需要选择测试表号，设置计分单位，填写给声强度等信息。软件控制言语识别率的测试流程。可在多个强度下重复进行言语识别率测听（已使用过的表号已被屏蔽，不可重复使用），所得结果可手动绘制成 P-I 函数曲线。

单音节、扬扬格词与语句的识别率测试流程基本一样。不同的是：①各类测试的计分单位可视临床或研究工作的需要而改变。对于单音节测试项，既可以整个音节计分，也可将音节拆分成声母、韵母和声调三部分来计分；对于扬扬格词测试项，既可以双字词作为基本计分单位，也可拆分成两个单字计分；对于语句测试项，既可以以整句计分，也可以句中关键词最为计分单位。②语句识别率测试可在安静和噪声背景下进行，所选噪声为 4 人同时谈话而混叠产生的嘈杂语噪声(Babble)。此时需同时记录语句的强度与噪声的强度。③单音节与扬扬格词测试时，测试项的播放是每间隔 4 秒自动切换的。听力师在这有限的时间内若出现误操作，软件还允许在播放下一个测试项时进行更正。而语句测听则采用手动播放的方式，允许听力师有充分的时间——点击各关键词下方的判断框。

2 言语识别阈

言语识别阈的测试内容包括扬扬格词表、安静下句子、噪声下句子三种类型，包括 4 张阶梯下降式（每播送 5 个词，强度就自动下降 5dB）的扬扬格词表、6 张安静条件下语句表（每表 20 句 100 个关键词）和 14 张嘈杂噪声下语句表（每表 18 句 100 个关键词）。测试前，需要选择测试表号，填写初始给声强度（或信噪比）等信息。软件分别以 ASHA 美国听力言语语言协会 1988 年推荐的言语识别阈测听指南^[14]和 Plomp & Mimpfen 提出的强

度自适应的方式，快速测得扬扬格词与语句的言语识别阈。

(1) 扬扬格词识别阈

扬扬格词是最为常见的言语识别阈测听材料。软件中扬扬格词识别阈测听，依照 ASHA(1988)推荐的言语识别阈测听指南进行。首先依据纯音平均听阈(3FA)摸索一个初始给声强度，确保在该强度下 5 个扬扬格词均能正确识别。如果不能，软件则会提示“增大初始给声强度”。接着软件会自动控制播放强度阶梯式下降：每播放 5 个词，强度自行降低 5dB。因而听力师不必再操心强度控制，只需判定患者的应答正确与否。当患者在某一强度下全部 5 个扬扬格词均不能正确识别，测试终止。软件会计数患者在整个阶梯下降过程听对的数目，以初始给声强度减去该数目，再加上一个 2.5dB 的校正因子，即为该患者的言语识别阈。

(2) 安静及噪声下的语句识别阈

传统的语句识别阈测听方法，需要根据语句识别的 P-I 曲线来推算，比较费时。Plomp & Mimpen (1979)提出了一种类似于 Békésy 测听的自适应测听方法，根据患者上一句正确应答率而相应增大或减小下一句播放的强度。随后 HINT 等噪声下语句测听材料也采用了这一方法，每次强度增减的步距为 2dB。参照此方法，本软件自行制订了强度（或信噪比）增减的步距规则，结合每句中关键词的正确识别率，步距分别设定为 0、0.5、1、1.5、2dB，使得语句的播放强度（或信噪比）总在患者的语句识别阈附近上下振荡。取每次上升、下降轨迹的中值点，即时统计测听过程中诸多中值点的均值和标准差。该均值即为患者的语句识别阈。当标准差小于 1dB 时，表明测听强度（或信噪比）已逐渐收敛于患者真实的语句识别阈，测听可提前终止。

三 意义

当今计算机技术已经深入听力学各个部分，声导抗测试、听觉诱发电位、耳声发射等检测技术，无不是借助于计算机技术而提高了工作效率和诊断准确度。但是作为听力学核心技术之一的言语测听多年来并没有多少实质性改变，始终作为听力计的一个附属功能存在，计算机技术给言语测听所带来的便利仅仅在近几年才得到部分听力设备厂家的重视。

中文言语测听近年来已取到长足进步^[16]，出现了一批分别适用于儿童和成人的单音节、扬扬格词及语句测听材料。但在推广过程中还是受到听力学专业人才匮乏、基础知识欠缺、外部设备繁复等问题的制约。因而只有开发集简易型、实用性和灵活性于一体的言语测听系统才便于在国内普及推广。本软件的开发，只是对计算机辅助的标准化言语测听系统的初步尝试，但已展现出良好的应用前景，相信会对推动国内听力学的发展具有积极意义。

参考文献

1. 韩东一, 杨伟炎. 普及言语测听, 提高耳科学诊疗水平. *中华耳科学杂志*, 2008, 6(1): 7-8.
2. Letz R. Continuing challenges for computer-based neuropsychological tests.

Neurotoxicology. 2003, 24(4-5): 479-89.

3. 郝昕. 言语测听的历史与现状. *中国听力语言康复科学杂志*. 2005, 2(1): 20-24.
4. 卜行宽, 倪道凤. 推进中文言语测听材料的标准化和临床应用. *中华耳科学杂志*, 2008, 6(1): 9-10.
5. 郝昕, 冀飞. 普通话言语测听 CD: 单音节识别率测试. 解放军卫生音像出版社, 北京, 2009.
6. 冀飞, 郝昕, 韩东一等. 汉语普通话单音节测听表的多中心复测信度研究. *中华耳鼻咽喉头颈外科杂志*. 2010, 45(3): 200-205.
7. 蔡宣猷, 张玉林, 黄银爱, 廖月红等. 计算机听力计专用广州话及普通话言语测听软件研制的初步报道. *中华耳鼻咽喉科杂志*, 1994, 29(1): 52-54.
8. 段家德. 一种实用的汉语言语听力测试系统及临床应用. *北京生物医学工程*. 2005, 24(1): 42-45.
9. 张华, 王硕, 陈静, 等. 普通话言语测听材料的智能化研究. *中华耳鼻咽喉头颈外科科杂志*, 2008, 43(6): 406-408.
10. 黄高扬. 计算机辅助汉语言语测听软件的研究与实现. 第 18 届全国多媒体学术会议 (NCMT), 2009.
11. 陈洪文, 于黎明. 听力计鉴定和声场校准. *中华耳科学杂志*, 2003, 1(4): 63-67.
12. 国家声学标准化委员会. 《GB/T 17696 声学 测听方法 第三部分: 语言测听》1999.
13. 赵阳, 郝昕, 冀飞, 等. 嘈杂语噪声下汉语语句测听中的学习效应. *听力学及言语疾病杂志*, 2009, 17: 107-111.
14. American Speech-Language-Hearing Association. Guidelines for Determining Threshold Level for Speech. *ASHA*, 1987, 29: 141-147.
15. Plomp, R, and Mimpen, A. Improving the reliability of testing the speech reception threshold for sentences, *Audiology*, 1979, 18: 43-52.
16. 郝昕. 中文言语测听材料的新进展. *中国眼耳鼻咽喉科杂志*, 2008, 8(6): 341-343.