

计算机辅助汉语言语测听软件的研究与实现

黄高扬^{1,2,3} 贾珈^{1,2,3} 蔡莲红^{1,2,3} 郝昕⁴

¹ 普适计算教育部重点实验室

² 清华信息科学与技术国家实验室(筹)

³ 清华大学 计算机科学技术系 100084

⁴ 解放军总医院耳鼻咽喉头颈外科 北京 100853

摘要: 本文给出了一种基于计算机辅助的汉语言语测听软件。软件实现了患者信息管理及测听自动化,利用本软件,既可以辅助进行纯音测听,也可以辅助进行言语测听,包括测试单音节识别率,扬扬格词识别率及识别阈,安静下句子识别率及识别阈,噪声下句子识别率及识别阈等,为言语测听的临床实验及研究工作提供了实用而有效的计算机辅助工具。

关键词: 言语测听, 听力图, 识别率, 识别阈, PI 曲线

1. 引言¹

听觉在人们日常交流中占有重要地位,听力检测则能反映出听力障碍者的残疾程度和社会交往能力。听力学发展早期,利用纯音测听可以简单地检测出听力损伤的情况。然而纯音测听却不能正确反应日常言语交流能力。言语测听则是了解受试者言语感知能力的有效的工具^[1]。因此言语测听受到言语声学、听力学、临床医学的广泛关注。

言语测听使用大量的语料,通过比较不同条件(如不同声强、有无噪声)下的测试结果,可达到诊断听力疾患,评估听敏度的目的,并且在评价听觉中枢、语言中枢的功能,选择干预方案和评价康复效果等方面发挥着不可替代的作用^[2]。

言语测听涉及到的因素较多,测试过程繁琐。手控测试难以实施测试条件的设置与调整,如随机调整语料的播放顺序,难以保证测试的准确度,难以实现测试结果的查询与比对等。因此,设计一个计算机辅助的言语测听系统是很有必要的。

本文实现了一个集受试者信息管理,测听流程自动化等功能为一体的言语测听软件。软件基于 Microsoft Visual C++编写完成。计算机将提供语音播放、患者信息管理、结果记录等功能,可以辅助进行纯音测听、单音节识别率、扬扬格词识别率及识别阈、安静下句子识别率及识别阈、噪声下句子识别率及识别阈等测试,为言语测听的临床应用及研究工作提供一个实用而有效的计算机辅助工具,方便管理及研究。

2. 言语测听系统设计

为了实现言语测听的自动化,提高测听的准确性和工作效率,本文设计了一个计算机辅助言语测听平台 CA-MSAS (Computer Assisted Mandarin Speech Audiometry System)。该平台是由言语测听软件、测听语料及听力计组成,测听过程在隔音效果良好的测听室内进行。下面分别介绍各部分的工作原理。

2.1 计算机言语测听软件

言语测听软件提供的功能主要包括信息管理、言语测听流程管理两部分。信息管理负责管理患者信息、听力数据、测试条件及结果,具有增删查改的功能。而言语测听流程管理负

资助项目:中国自然科学基金(60805008,90820304),国家基础研究计划(973 计划)(No. 2006CB303101)和国家高技术研究发展计划(863 计划)(No. 2007AA01Z198)

负责各种测试流程的控制。

言语测听软件主要功能的流程如下图1所示。

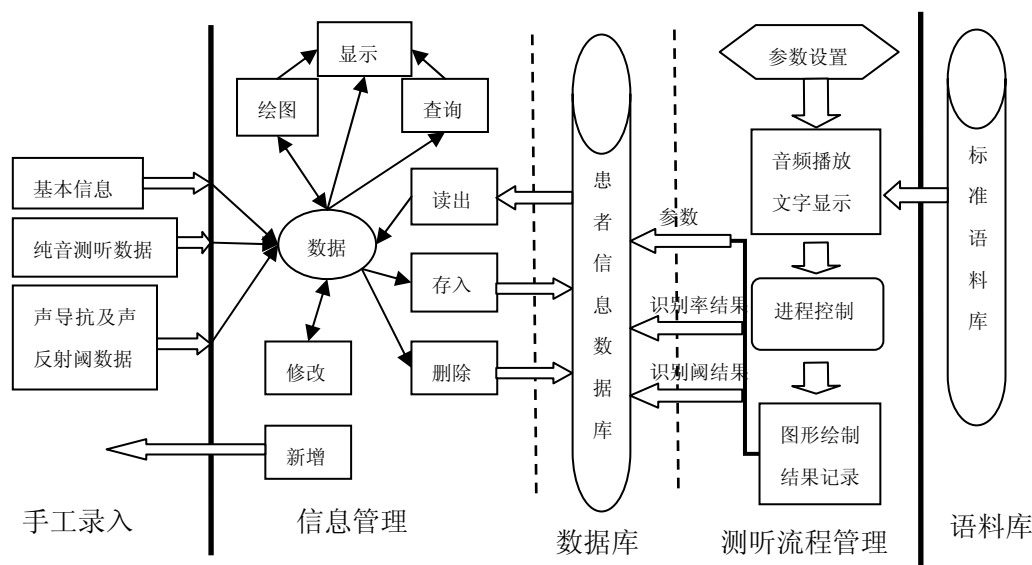


图1 言语测听软件主要功能流程图

2. 2 言语测听语料库

本软件所使用的语料是由北京 301 医院耳鼻喉科设计制作的，包括 22 张单音节表，9 张扬扬格词表，28 张噪声下句子表，12 张安静下句子表以及若干张练习用表。该测试语料既考虑了语音平衡^[3]、又避免了生僻的字词，并进行了语料等价性的测试^[4]。对于录制的语音还进行了响度平衡处理。

2. 3 听力计

听力计与计算机相连，将音频信号进行功率放大，并输出到扬声设备（耳机、扬声器或骨振器等）。听力计可调节、计算、显示音频信号的强度，因此在测试过程中，测试员可以通过听力计，调整、监视音频信号的强度。在纯音测听中，听力计直接生成纯音信号。

2. 4 校准

实际测试时，只需要将计算机的音频输出接口与听力计相连，就完成了测听的硬件准备工作。为了消除计算机声卡输出电平的差异，在测试之前计算机先输出一段校准音，调节听力计校准表盘的示数，使其指到 0 或-10（听力计的校准表盘只有这两个值是精确的），并由言语测听软件记录下这个校准值。在测试中，将听力计左右声道的声强示数加上此校准值，便是人耳实际听到的声音强度。

3. 信息管理

患者的基本信息、纯音听力数据、声导抗及声反射阈数据需要用户手工录入，而言语测听的信息则是由言语测听软件自动记录的。每个患者的所有信息都以文件的格式存放在数据文件夹中，这些文件构成了言语测听软件的简单数据库。各种数据按不同的类型与测试时间分别存放，在软件中可以随时调看，方便了追踪治疗及康复效果。

4. 言语测听

言语测听常见的测试指标是言语识别阈、言语识别率。言语识别阈是指受试者能够正确理解当前播放言语材料的 50%时所使用的最低声级。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，测试结果既可直接用

于诊断，又可与纯音测听的结果相对照，诊断听力损失，评估患者听敏度。

4. 1 纯音测听

纯音测听与言语测听相对应，是一项基础的听力检查项目^[5]。在纯音测听中，首先固定纯音信号的频率，通过变化信号强度和受试者的反应，测得受试者的听阈^[6]。依次测试受试者在几个特定频率的纯音信号下的听阈，将测试结果绘制在以频率为横轴、以强度为纵轴的直角坐标系中。该图也是判断患者听力损伤情况的重要依据。

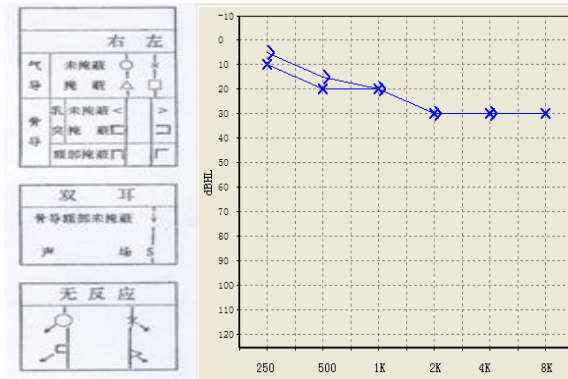


图2 言语测听软件绘制的纯音听力图（左耳）

纯音测听的结果对言语测听具有指导意义。在扬扬格词识别阈测试开始前，需要测试员手工输入一个初始强度，该强度通常会取纯音气导平均听阈（500，1000，2000HZ 三个频率下听阈的平均值）^[7]。

4. 2 言语识别率测试

言语识别率的测试，是一种通过受试者对不同强度的标准的言语样本的听辨，评估受试者言语识别总体情况的方法。所使用的言语样本，可以是单音节字、扬扬格词或句子^[8]。

从单音节字到扬扬格词，再到语句，越来越贴近实际生活中的言语。这三类测试，通过对言语过程的不断深入的探查，不仅考察了听没听到，听没听对的感知上的问题，还考察了人们对言语的分辨与理解这些后天形成的能力。

汉语音节是由声韵母相拼形成，具有不同的声调。虽然扬扬格二音节词是日常生活中常见的规范词语，但前后字的关联与受试者的常识可能影响测试结果。句子虽然更好地代表了日常生活的言语，但受试者的方言基础、社会文化背景，可能影响测试结果。因此除了语料的精心设计外，还要仔细考虑测试条件、计分单位和测试流程^[9]。

4. 2. 1 言语识别率测试流程

识别率的测试流程如下图3所示。

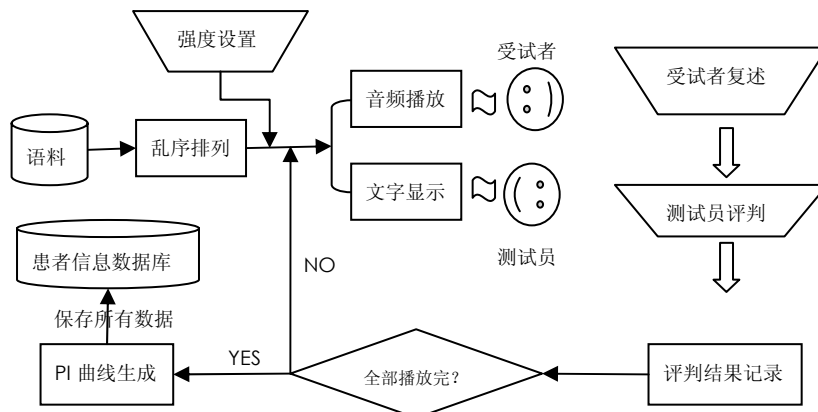


图3 识别率测试流程图

计算机负责控制言语识别率的测试。流程中，包括言语样本选择、乱序、音频播放、显

示文字、接收测试员输入、PI 曲线生成、测试数据保存等关键步骤，均由计算机自动完成。由于在一次测试中，没有声强的变化，所以，计算机将以固定的强度播放所有的音频文件。而如果下一次测试需要改变声强大小，则需要测试员手工调整听力计输出端的表盘示数来实现声强的改变。

可选择的言语样本包括单音节、扬扬格词表、安静下句子、噪声下句子四种类型，每类都包括若干张表。测试前，需要选择测试表号，设置计分单位，填写音频强度等信息。

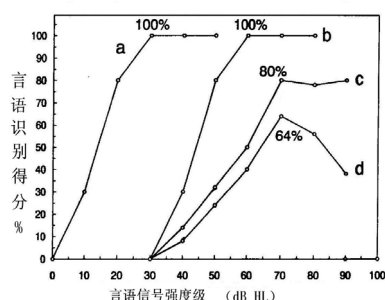
乱序播放的目的，是为了避免了重复测试时因受试者的记忆效应而对结果造成的影响。计算机利用一个数组存放音频文件的地址，对此数组进行随机排序，再顺序播放数组中的音频，即实现了对音频文件的乱序播放。

测试时，计算机播放音频文件给受试者，同时将音频文件的文字内容将显示在测试员观看的屏幕上。在受试者复述之后，测试员可通过点击文字下相应的判断框记录结果。所有的测试信息与结果都将在测试完成后录入数据库。

单音节字、扬扬格词与语句的识别率测试的流程基本一样。不同的是：①各类测试的计分单位不一样。语句测试时还需要忽视虚词的影响，因为虚词对人们理解语句的作用微乎其微。②语句测试分安静下句子与噪声下句子，在噪声下句子测试时，需要同时记录语句的强度与噪声的强度。

4. 2. 2 PI 曲线

一次测试结束，计算受试者复述正确的字或音位占总数的百分比（识别率）。根据受试者的得分，增大或减小言语信号强度，进行再测试。直到得到该受试者在多个不同的言语信号强度下的言语识别率。



将受试者多次测试的结果表示在以识别率为纵轴，声强为横轴的坐标系中，就得到了不同声强下的得分点，将这些得分点顺序连接，就得到了受试者的 PI 曲线

言语测听软件将自动绘制 PI 曲线。对于听力正常者，经测试可分别生成单音节字、扬扬格词、语句的 PI 曲线。将受试者的 PI 曲线与听力正常者相应的 PI 曲线相比较，可以诊断受试者的听力损失情况与发病原因（见下图4）。

图4：不同类型听力损失患者的PI曲线。

- a：正常听力者PI曲线 b：传导性听力损失患者，PI曲线向右平移约30dB；
c：感音性听力损失患者，PI曲线在70dBHL以后达到峰值约80%；
d：蜗后听力损失患者，PI曲线的峰值约64%；且在峰值之后出现明显回跌。

4. 3 言语识别阈测试

言语识别阈测试将听力水平具体化为一个阈值。与识别率相对应的话，识别阈将对应 PI 曲线上得分为 50%左右时的声强。言语测听软件集成了扬扬格词、安静下句子及噪声下句子识别阈测试，三者各有不同。

言语识别阈的测试样本从扬扬格词到安静下句子，再到噪声下句子，同样越来越贴近真实的说话场景。言语识别阈考察的是受试者在多大的信号强度下能听懂言语，而对汉语而言，最小的语义单位是词，只听对了一个双音节词中的一个字，是无法理解这个词的意思的，也就是听不懂。因此，在识别阈测试时，计分单位都是固定的以词为单位。在句子测试时，需要事先人为切分好关键词。同一受试者的识别阈值与识别率值相结合，可以验证纯音测听结果准确性及真实性，还可对受试者听力损失程度、性质作出判断^[10]。

4. 3. 1 扬扬格词识别阈

其主要流程如下图5 所示

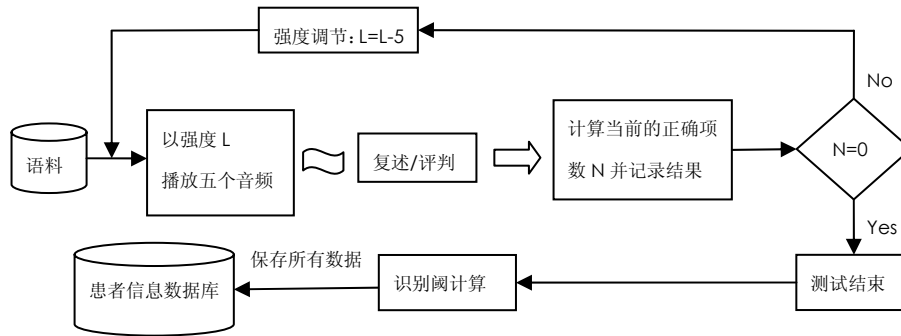


图5 扬扬格识别阈测试流程图

音频播放，文字显示，结果记录，强度调节，识别阈计算等关键步骤由计算机自动完成。

测试前，需要测试员输入一个初始强度，以此为基准，播放五个音频。播放完毕后，若受试者没有完全听错，则降低声强，继续播放后面五个测试项，直到受试者完全听错或听不见，结束测试。与识别率测试不同，识别阈测试中降低声强这一步骤是由计算机通过对音频文件的处理而自动实现的。计算机与听力计相连，将听力计对输入的增益值固定，当输入信号的强度变化时，输出信号的强度也将线性地变化。

临床上常将纯音气导平均听阈与扬扬格词的识别阈相比较，当二者相差在 12dB 以内时，可视为一致。而当超过该值，则可提示有听神经，听处理不良等病症的可能^[7,10]。

4. 3. 2 安静下句子识别阈

安静下句子识别阈测试流程与扬扬格词类似，不同的是：每播放一个句子都需要改变一次强度，而且变化幅度是与每句的关键词数相关的。测试时，计分单位都是关键词。

开始时，先播放第一句，让受试者熟悉测试环境及条件，类似于热身。该句用于调节起始强度：若受试者完全听对，则强度下降 2dB，播下一测试项；若未完全听对，则强度上升 2dB，继续播放该句，直到完全听对。热身阶段后，再播放后四个测试项，强度变化规则跟热身时一样。到第五句播放完时，强度的变化程度将视每句关键词的总数以及受试者复述正确的关键词个数而定。为了直观地显示测试过程中的强度变化情况，言语测听软件将在每一句语料播放的同时绘制强度变化折线图，其横轴为播放次数，纵轴为强度（见下图 6）

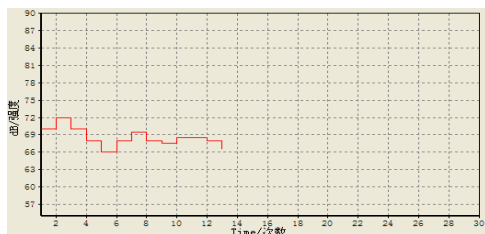


图6 语句识别阈测试时的强度变化折线图

结束条件一是全部的句子播放完毕，条件二是当强度变化幅度较小，定量来说就是一系列变化值的标准差小于一个阈值，便可认为强度已收敛到阈值，可以结束测试。手工实时计算标准差难度很大，但在计算机上则可简单实现。

4. 3. 3 噪声下句子识别阈

噪声下句子的背景噪声是嘈杂的言语，与安静下句子相比，更贴近真实生活中的说话场景，而且在评估助听器及人工耳蜗的效果方面也有着重要意义^[11]。其测试流程类似于安静下句子识别阈测试。不同是，噪声下句子测试时，变化的是信噪比，语句强度是固定的。

5. 总结

计算机辅助言语测听软件大大降低了测听过程的复杂性，提高了测试结果的精确性、可读性，并为每一位受试者建立了详细的听力数据库，为追踪治疗效果，了解康复状况提供了便捷的工具体，是从事汉语言语测听事业的研究人员及医务人员的有效的辅助工具，也为今后进一步的研究工作打下了基础。

参考文献

- [1] 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.
- [2] Hall JW., Mueller HG., Speech Audiometry, In Hall JW. & Mueller HG. Ed, Audiologist's Desk Reference, San Diego: Singular Publishing Group, 1997:115-174.
- [3] 张华, 王硕, 王靓, 等, 普通话言语测听双音节词表的编辑与初步等价性评估, 中化耳鼻咽喉头颈外科杂志, 2006, 第 41 卷, 第 6 期: 425~427.
- [4] 王靓, 张华, 王硕, 等, 普通话单音节词言语测听材料的等价性分析, 中国耳鼻咽喉头颈外科, 2006, 第 13 卷, 第 6 期, pp397-401.
- [5] 张勉, 李兴启, 重视纯音听阈测试及结果的临床应用, 听力学及言语疾病杂志, 2006, 第 14 卷第 3 期: 163~164.
- [6] 孟曦, 郑军, 周婉荣, 刘千, 纯音气导听阈不同测试手法比较, 临床耳鼻咽喉杂志, 1987 年第 1 卷第 3 期: 142~143
- [7] 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
- [8] 郝昕, 顾瑞, 冀飞. 发展言语识别率测试材料的理论框架. 听力学及言语疾病杂志, 2006, 14(6): 401-404.
- [9] 郝昕, 冀飞, 陈艾婷整理, 怎样发展标准化的中文言语测听材料, 中华耳科学杂志 2008, 6(1), 13~16.
- [10] 李蓓, 郑虹, 言语识别阈测试, 川北医学院学报 2007 年第 22 卷第 3 期, 284~286.
- [11] 区建国, 金昊, 许由, 等, 双侧人工耳蜗植入者在噪声环境下的言语辨别能力 中华耳鼻咽喉科杂志, 2001, 36(6), 433~435

The Research and Realization of Computer Assisted

Mandarin Speech Audiometry System

Huang Gaoyang^{1,2,3} Jia Jia^{12,3} Cai Lianhong^{1,2,3} Xi Xin⁴

¹Key Laboratory of Pervasive Computing, Ministry of Education;

²Tsinghua National Laboratory for Information Science and Technology ;

³ Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China

⁴ Department of Otolaryngology, Head and Neck Surgery Institute, Chinese PLA General Hospital, Beijing 100853, China

Author: E-mail: hgy04@126.com

Keyword: speech audiometry, audiogram, Speech Recognition Score, Speech Recognition Threshold, PI-curve

Abstract: This article introduces a software for computer-assisted Mandarin speech audiometry, allowing for automation of patients' information management and speech audiometry. The software is apt to assist in not only pure tone audiometry, but also speech audiometry, in respects that includes not only Speech Recognition Score(SRS) of monosyllable、spondee word、sentence in noise and sentence in quiet, but also Speech Recognition Threshold(SRT) of spondee word、sentence in noise and sentence in quiet. It is a practical and effective computer-aided tool for clinical trials and research in the field of speech audiometry.