

纪念《声学技术》创刊 30 周年特邀文章——

# 人工耳蜗植入者音乐感知研究

冯海泓, 原 猛, 陈友元

(中国科学院声学研究所东海研究站, 上海 200032)

**摘要:** 目前, 大部分人工耳蜗使用者在安静情况下的言语感知能力基本达到较高水平, 然而其音乐感知能力尚未达到理想水平, 音乐感知研究已成为目前人工耳蜗领域主要关注的研究方向之一。主要探讨人工耳蜗使用者音乐感知研究的研究方法、材料, 并分析影响其音乐感知能力的主要因素。对音乐的四大要素即音高、时长、响度及音色等的研究进展进行讨论, 从信息的传输和接收间的关系分析影响人工耳蜗植入者音乐感知的因素。对现有的主、客观测试以及调查问卷等研究手段进行总结。综合分析了目前音乐感知研究的进展, 并对未来的研究发展方向进行展望。

**关键词:** 人工耳蜗; 音乐感知; 听觉补偿

**中图分类号:** TB556

**文献标识码:** A

**文章编号:** 1000-3630(2012)-01-0053-08

**DOI 编码:** 10.3969/j.issn1000-3630.2012.01.008

## Review of music perception of cochlear implant users

FENG Hai-hong, YUAN Meng, CHEN You-yuan

(Shanghai Acoustics Laboratory, Chinese Academy of Sciences, Shanghai 200032, China)

**Abstract:** Nowadays, many cochlea implant (CI) users could achieve a relatively high speech perception in quite environment. However, their music perception is far away from optimization. Music perception has been a potential research direction. This paper focuses on the investigation of the research methodology, test materials for cochlear implant (CI) music perception, as well as analyzing the main factors that affect the music perception of CI users. The research progress on the four music factors, i.e. pitch, duration, loudness and timbre are discussed based on the relationship between the information transmission and reception, which affects the perception of music for CI users. The research methods for evaluating the CI music perception are summarized. At last, the future research direction on improving the CI music perception is discussed.

**Key words:** cochlear implant; music perception; auditory prosthesis

## 0 引言

在过去 30 年里, 人工耳蜗(cochlear implant)技术取得了突飞猛进的发展, 已有超过十万名重度到极重度耳聋患者因此而重获声音, 大多数人工耳蜗(CI)使用者均可获得良好的言语可懂度。人工耳蜗这一高精密电子产品的临床意义无疑是巨大的。然而, 该设备也存在一些缺点。由于人工耳蜗的最初设计理念是为了令使用者获得安静环境下的良好言语可懂度, 其噪声环境下的言语可懂度以及对音乐信号的传送效果并不十分理想。而对于人工耳蜗使用者来说, 音乐是除语音外最重要的声刺激, 但大多数人都抱怨使用人工耳蜗听到的音乐并不动

听。“我能听懂说话声, 但无法听清音乐”是许多人工耳蜗使用者对音乐聆听效果的普遍评价。为了更好地了解和探知导致人工耳蜗使用者音乐感知能力低的原因, 对目前的人工耳蜗技术、音乐感知研究情况以及提高人工耳蜗使用者音乐感知的方法进行了介绍, 并对未来的人工耳蜗技术发展方向进行展望。

总的来说, 音乐同言语声有许多相似性<sup>[1,2]</sup>。从声学角度来讲, 音乐声和言语声都是在相对较长的时间内通过频率的变化来传递某种信息。音乐同语言一样以沟通和情感表达为目的; 另一方面, 语言也同音乐一样具有节奏和韵律成分。在所有的民族和时代中, 两者都具有相同的复杂内部结构规则。因此, 只有那些音乐所独有的特征才能使人们区分出音乐这一独特的声音类型。

音乐欣赏需要对四个基本因素进行感知, 它们包括时长(duration), 音高(pitch), 响度(loudness), 以及音色(timbre)<sup>[3]</sup>。作为音乐基本要素之一, 节奏(rhythm)反映了音乐信号的时间特征, 通常的时长

收稿日期: 2011-11-01; 修回日期: 2012-01-18

基金项目: 国家自然科学基金项目(11104316)、上海市自然科学基金项目(11ZR1446000)资助。

作者简介: 冯海泓(1966—), 男, 博士生导师, 教授, 研究员, 研究方向为听觉补偿技术与水下声学。

通讯作者: 原猛, E-mail: ym@mail.ioa.ac.cn

为几秒到几分钟。准确的节奏感知需要很好地把握声音的时间传播特性。从许多方面来说,节奏是音乐最基本的特征,是一种时间特性的表达。音高,是另一个音乐基本要素,它与声音的频率有关,在音乐范围中以音符表示。虽然音高的差异是一种绝对频差,但这种音高差具有更大的意义,在一段音乐中,可使人们感受到由一系列音高所组成的旋律(melody)。和弦(harmony)是相邻音高间的联系,是人们区分重叠声质量的音乐属性,例如,一个大调三和弦或小调三和弦可使其他音乐特征(如对位音)升高。对于音色的理解,美国声学学会 1960 年给出的定义是“(音色)是一种令听者能够在响度、音高相同的情况下分辨出两个相似声音的听觉感知属性”<sup>[4]</sup>。它包含了声音中不与音高或响度直接相关的特征,通常在乐器辨识测试中被评估的量就是音色。音乐感知主要包含模式的感知,是节奏、音高、响度、音色的一种变化<sup>[5]</sup>。而音高的序列或模式形成了音乐中的旋律与和声,时间长度或模式的变化是节奏的基础。虽然音乐还有许多其他方面的因素需要更深入的理解,从本质上讲,音乐是节奏、音高、和弦的多重组合。而且,虽然从解析和心理声学角度来讲,音乐可解构成基本因素的组合,但对于聆听者来说,音乐是作为一个不可分割的整体而进行感知和表达的,而非是这些因素的集合。正是由于这种组合和相互作用,才使得音乐产生了其独特的听觉感受,并供人们聆听和欣赏。也正是由于其复杂性和独特性,使得对音乐感知的研究变得难以深入。对于人工耳蜗使用者来说,这些因素的有效表达和接收是他们能够获得音乐感知的前提。

## 1 人工耳蜗技术

人工耳蜗(如图 1 所示)的核心即通过对听觉神经的直接电刺激从而使感应神经性耳聋患者获得有效的听觉感知。虽然目前已出现了许多种不同设计的人工耳蜗,但其主要特征还是相同的。所有人工耳蜗设备均通过麦克风采集声音信号,通常麦克风的摆放位置都类似于传统的耳背式助听器一样戴在使用者的耳廓上。由麦克风输入的声音信号通过模数转换变为电信号,并传送至人工耳蜗言语处理器单元。言语处理器是人工耳蜗体外机的核心部分,负责对输入信号进行处理,主要工作是从声音信号中提取所需的特征量并将编码后的特征量转化成神经电刺激信号,从而使人工耳蜗使用者获得听觉感知。对于声音处理的方法(即言语编码策略)

以及电路的设计比较灵活,因此对于不同公司的不同产品,其声音处理方法也不尽相同,从而导致电极输出的电刺激脉冲信号模式也不尽相同<sup>[6]</sup>。下面将对人工耳蜗植入体、体外声音处理器作一介绍。

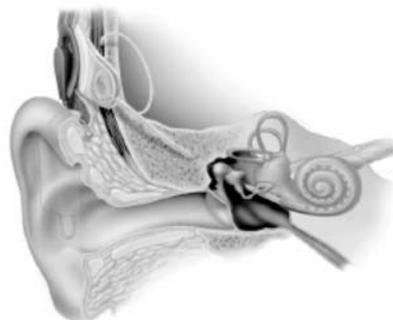


图 1 人工耳蜗系统示意图

Fig.1 Schematic of cochlear implant system

### 1.1 人工耳蜗体外机

人工耳蜗体外机由含有 1 到 2 个麦克风的语音处理器、传输线圈以及相关连接线组成。语音处理器将麦克风信号转换成电极的电刺激模式,刺激参数由处理器根据每位病人的电听力图(MAP)进行编程修改。该转换过程包含许多步骤,特定的步骤根据处理器及所用的言语处理策略不同而变化。所得的信息经编码成为射频信号,并最终通过线圈耦合的方式传送到人工耳蜗的体内接收器。

### 1.2 人工耳蜗植入体

人工耳蜗声音处理器输出为含有用于产生植入体电刺激的数字编码信号,该信号通常通过射频方式由分别置于皮肤内外的两个耦合线圈传输入体内植入体部分,同时,植入体的能量也由耦合线圈提供。植入体中的集成电路将信号解调从而得到所需刺激的电极序号及电刺激信号的幅度等信息。大多数现有刺激器的输出信号为双相对称脉冲电流序列。

在正常的耳蜗中,声音频率与最大听神经刺激位置间存在着——对应关系<sup>[7]</sup>。高频声音会在蜗底部位(base of cochlea)产生较大的神经活动,而低频声音会在蜗顶部位(apex of cochlea)产生较大的神经活动。这种音质分布关系不仅仅适用于耳蜗绒毛细胞,对于听觉神经的细胞体和树突也同样适用。这就意味着,即使对于只有极少量甚至没有耳蜗绒毛细胞存活的重度听觉神经性耳聋而言,人工耳蜗仍然可以利用多电极序列的结构来实现音质分布关系,从而激发残留听神经来产生听觉。

通常来讲,人工耳蜗的电极阵列包含一组独立的电极,该电极序列通过手术从圆窗附近植入到患

者耳蜗内,并沿蜗轴方向深入耳蜗。电极阵列中接近蜗顶的电极刺激将产生低频声,而接近蜗底位置的电极刺激将产生高频声。通常电极可被配置成三种模式:单极模式,双极模式,以及共地模式。在单极模式下,设置一个蜗外或靠近耳蜗的电极,该电极通常具有较大的接触面积,其它电极分布在蜗内。该较大接触面的电极的作用是作为其它蜗内电极的电流返回路径。通常单电极人工耳蜗系统采用单极模式进行电刺激。在采用单极模式的多电极耳蜗系统中,在手术时特别注意应使蜗内电极尽量紧靠蜗轴缠绕,从而保证每个电极产生的电刺激只激发听觉神经的一小片独立区域,尽量减小电极间的互扰。

多电极植入体的电刺激空间分离性可通过采用双极刺激模式得到改善。在此刺激模式下,两个相邻的蜗内电极被配置成一个电流的输出端和一个电流的接收端,这样电极的电刺激电流路径可通过最短的距离到达接收端,从而减低电极间的互扰。

在第三种电极配置即共地模式中,采用一个蜗内电极作为刺激电极,选取其他部分或全部的蜗内电极作为电流返回路径。该种配置模式的效果被认为介于单极刺激模式和双极刺激模式之间。

### 1.3 人工耳蜗言语处理策略

人工耳蜗言语处理策略对于人工耳蜗系统是十分重要的,它决定了何种声音将被植入者感知。刺激参数的变化可导致输入声音的电刺激表达的变化。通常,这些刺激参数包括采样率、刺激率、被激励电极、刺激顺序、刺激脉冲类型、刺激电流强度等。频谱信息通过改变电极间的刺激电流强度反映,时间信息通过每个电极的电流强度随时间变化来反映。由于不同言语处理策略所使用的参数有所不同,统一声音输入经不同言语处理策略处理后得到的电刺激形式也不尽相同。下面对当前几个主流的言语处理策略做一简要介绍。

同过去相比,当前使用的言语处理策略均为滤波器组策略,将输入声音进行分频带处理,每个频带对应一个电极。滤波器组的滤波器数量和带宽除根据不同策略有所不同外,对于使用同一个策略的不同人工耳蜗植入者来说也是不同的。

目前最广泛使用的言语处理策略主要有三种:连续间隔采样(Continuous Interleaved Sampling, CIS), SPEAK, 及 Advanced Combination Encoder (ACE)。这些策略旨在将声音信号中主要的频谱特征表达达到电极输出,但并未考虑所表达的频谱特征与言语间的关系。更重要的一点是,在这些策略的

使用中,通常会设定一个与输入信号无关的固定刺激率。

图 2 所示是一个典型的 6 通道 CIS 处理策略流程图。输入信号经过一组带通滤波器得到其短时谱。对每个滤波器输出的带通信号的包络进行估计,这些包络信号以特定的时间间隔进行采样,并将其幅度转换为适当的刺激电流值。在植入体部分,电脉冲信号通过相应电极输出并对听神经进行刺激。在大多数 CIS 策略的应用中,通常的刺激率为每通道每秒几千个脉冲。许多商用设备制造商和独立研究人员都对 CIS 策略进行了改变,但 CIS 的核心思想并未变化,从而也说明 CIS 是现今言语处理策略的基础。SPEAK 和 ACE 策略也是以 CIS 为基础,但通常拥有更多的电极数量。

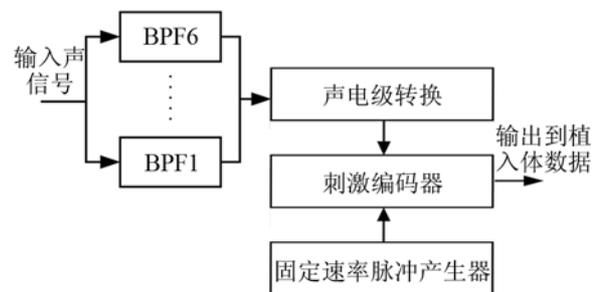


图 2 6 通道 CIS 声音处理策略流程图  
Fig.2 Flow diagram of 6 channel CIS processing strategy

图 3 是澳大利亚 Cochlear 公司的 Nucleus 系列人工耳蜗产品中所采用的 ACE 处理策略的流程图。与 CIS 策略类似,输入信号经带通滤波器组进行分频处理,并得到频带输出信号的包络,比较各频带包络的幅值并选择其中最大的几个作为输出,通常所选的最大幅值频带数为 6~10 个。这些幅值经转换为相应的刺激电流强度并从相应电极输出。

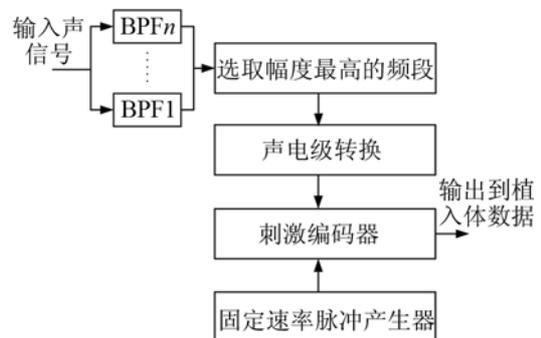


图 3 ACE 策略流程图  
Fig.3 Flow diagram of ACE strategy

整体而言,ACE 策略的刺激脉冲速率比 SPEAK 策略高许多。SPEAK 策略采用每通道 250 Hz 的脉冲速率。ACE 策略的总脉冲速率至少为 14.4kHz,但从功能上来说,SPEAK 和 ACE 策略是相近的。

## 2 人工耳蜗使用者的音乐感知

音乐感知对于人工耳蜗植入者来说是非常具有挑战性的。人工耳蜗中言语处理策略的主要功能是提取信号的瞬时包络信息,去除了瞬时精细结构信息。保留包络信息基本可以满足人工耳蜗植入者的言语识别,但瞬时精细结构信息对于音乐感知是至关重要的。因此,人工耳蜗植入者所感知的音乐的音质和细节信息都会严重下降,从而导致对音高感知能力的下降。事实上,人工耳蜗植入者经常将音乐描述成令人不快的、机械的、难于理解的声音<sup>[8-10]</sup>。因此,提高患者的音乐感知将有助于提高他们的生活质量和对声音的感知能力。另一方面,有研究表明,对于正常听力者(Normal Hearing, NH)来说,音乐体验可使其听觉功能得到改善,从而导致他们在更复杂的听环境下能够更易获得言语感知。这一发现预示了音乐体验对于嘈杂环境下的语音理解有积极的帮助。但如前所说,目前的人工耳蜗无法将声音中的瞬时精细结构信息及丰富的频率信息传递到听神经,从而导致无法获得良好的音乐感知。

近几年,大多数已发表的针对人工耳蜗植入者的音乐感知的研究都是基于一个基本假设,即音乐能够用含有一定数量的基本特征所组成的声音序列来表征。这些基本特征主要包括节奏、旋律、音色。和声、响度等声音属性也是构成音乐的因素。

除了这些客观的声音特征外,各异现象对于音乐聆听体验也是十分重要的,主要包括主观的音质、情绪等。对于音乐聆听的主观评测也是音乐感知研究的重要组成部分。

### 2.1 节奏的感知

人工耳蜗传送声音时保持了相对较高的瞬时特征。因此,人工耳蜗植入者的节奏感知也由于时间特征的保留而保持得较好。也就是说,音乐信号中的瞬时模式通常很容易被人工耳蜗植入者分辨。在所有音乐元素中,最早有关植入者感知的研究表明,节奏信息是人工耳蜗植入者最易感知的元素,这一结论已被一系列研究证实<sup>[5,11-13]</sup>。

需要指出的是,瞬时特征并不只是对节奏感知有用,对于音高感知和旋律感知也一样有用。人工耳蜗植入者感知瞬时特征的能力已经有多位专家探讨。Kong 测试了植入者的节拍分辨、节奏识别以及节奏对旋律感知的作用<sup>[14]</sup>。此研究的结论是植入者的节拍分辨同正常听力者相近,节奏可能会对旋律感知提供重要线索。当去掉节奏轮廓信息时,

旋律识别将变得非常差。对于节奏识别来说,作者的结论是耳蜗植入者通常无法分辨节奏模式间的区别。这些结论反映了一个事实,即音乐节奏信息并不仅在瞬时特征上不同,在有限时间段内所包含的音符数量也不同。在最近的一篇有关节奏模式再现的研究中,耳蜗植入者能够准确地再现含有相同数量音符或时间间隔,但每个音符的时间分布不同的节奏模式<sup>[15]</sup>。对于基本的节奏分辨来说,大量数据表明耳蜗植入者的分辨能力同正常听力者相近。但普遍认为更复杂的节奏分辨能力测试将有助于揭示植入者与正常听力者的差别。最重要的是,植入者听到的多数真实的音乐含有许多乐器共同演奏产生的多层次的节奏信息,对于这样复杂的音乐信号,植入者很难从众多同时存在的乐器中捕捉到乐曲的节奏信息。

### 2.2 旋律感知

旋律感知受一个人对旋律的熟悉程度以及对构成旋律的各种元素的感知程度的影响。旋律的基本元素包括旋律轮廓、绝对或相对时长,以及节奏。其他特征诸如歌词、特定的歌手或乐器等均对感知有所影响。

#### 2.2.1 曲调辨识

人们对于旋律识别的一个重要因素即为曲调的辨识。曲调辨识能力取决于很多方面,诸如对曲调的熟悉程度,是否可以在听到它的时候就说出其名字。这种能力取决于很多因素,如音乐素养、音乐训练,所处的社会文化环境,对曲调和其名字的对记忆能力等。曲调的识别同样也受音乐播放的环境条件影响。例如,生日歌是西方音乐传统中人们最熟悉的旋律<sup>[16,17]</sup>,即使在节奏、音符错误的条件下,几乎所有人都能轻易地识别出这首旋律。因此,准确感知音调和瞬时模式等声音的基本因素并不总是旋律识别的前提条件。

如前所述,大多数植入者的节奏感知能力与正常听力者相近。这一观察预示着植入者可能更容易识别具有规整节奏模式的旋律。Schulz 等在 1994 年通过对 8 名使用单通道人工耳蜗的植入者进行的研究已经证实了该假设<sup>[18]</sup>。在该研究中,四种儿童乐曲中常见的曲调经不同的音乐编排方式进行处理。实验结果表明,正常听力者可以获得接近 100% 正确的曲调识别率,而植入者只能获得很低的识别率(大概 50%左右)。在对植入者数据根据节奏是否规整进行分类分析时发现,具有规整节奏的曲调识别率较反之高 15%。在另一个使用 49 名多通道电子耳蜗植入者及 18 名正常听力者进行 12 个熟悉的

曲调的辨识研究也有相似的结果。结果显示,植入者的旋律识别率大概为 19%,而正常听力者的旋律识别率约为 83%。有规整节奏的旋律较无规整节奏的旋律识别率高 12%。使用不同言语处理策略得到的旋律识别率无显著差别。Kong 于 2004 年的研究对节奏信息对人工耳蜗植入者的旋律感知的重要性做了进一步的支持<sup>[14]</sup>。在他们的实验中,6 名多道电子耳蜗植入者依照指示分辨听到的 12 种熟悉歌曲。这些歌曲被处理成含有节奏或不含有节奏。另外的 6 名正常听力者也参与了测试,结果显示,无论是否含有节奏信息,正常听力者均可获得接近完美的旋律感知。对于植入者来说,当有节奏信息时的平均旋律感知为 63%,对于无节奏信息的情况下的旋律识别率仅仅接近随即猜测水平。2004 年,Looi 等使用 15 名 Nucleus 植入者对 10 种熟知的没有歌词和伴唱信息的旋律进行辨识测试<sup>[19]</sup>。这些旋律包含正常的节奏内容以及音高信息。6 名受试者使用 SPEAK 言语处理策略,其余受试者使用 ACE 策略。整体结果表明,植入者仅能识别出 50%的曲调,而正常听力者可识别 100%的曲调。研究分析表明,对于植入者而言,节奏和音高信息对于其旋律感知是有贡献的。

### 2.2.2 旋律模式识别

分辨两个不同的音高轮廓测试与旋律分辨相关,但通常更加困难,因为测试材料的听觉线索的减少。在一个典型的旋律模式识别测试中,听者被要求判别两个音高序列是否相同。构成每组音高序列的音符都有相同的节奏,但没有歌词信息。因此,这种分辨主要依赖于听者对音高变化模式的感知能力。然而,对于每个音符来说,该测试并不需要听者对其绝对或相对的音高进行精确的分辨。例如,对整体音高轮廓的感知(如音高的整体上升或下降)即可分辨出两个旋律模式的不同。音高轮廓(或旋律轮廓)测试的例子如图 4 所示。

测试方法:选择 9 选项-强迫选择(9-alternative forced-choice, 9AFC)方法进行测试。每次将呈现一段旋律,受试者需要选择听到的旋律。这段旋律由五个音符组成,如图 4 所示。测试步骤:选定音色和频率范围后,9 种轮廓、4 种音高差(1, 2, 4, 6 个半音),共 36 个组合将被随机给出。

2003 年,Leal 等对 29 名 Nucleus 植入者进行了旋律轮廓评估,这些植入者使用的言语处理策略为 ACE 或 SPEAK,测试包含 12 对音高序列[Leal2003]。有三分之二受试者的识别率在 90%以上。在一个相关的测试中,相同的受试者被要求

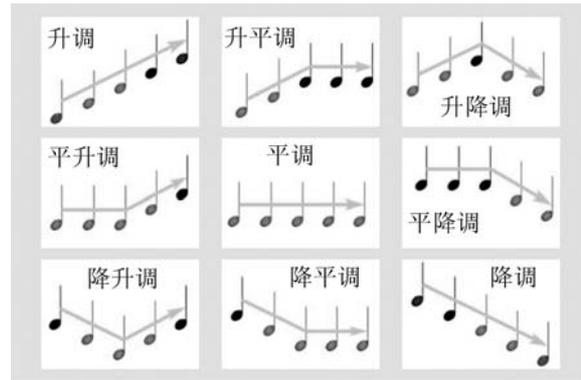


图 4 9 种轮廓示意图

Fig. 4 Nine melody contours presented in this study

描述每个序列的音高是否有变高或变低,并指出在序列中的音高变化位置。全部受试者的平均分辨率为 73%。然而,由于该测试并无正常听力者做对比测试,使得这些结果难于解释。

### 2.3 音高感知

大量研究表明,人工耳蜗植入者的音高感知能力较正常听力者差<sup>[5,18]</sup>。Gfeller 比较了植入者和正常听力者的高/低音高分辨能力。该测试旨在确定受试者所能分辨的音高差异的最小音节差(音高差阈)。测试结果表明 CI 和 NH 间存在显著差异。其中,NH 的差阈为 1.13 个半音,CI 的表现具有很大变化,平均差阈为 7.56 个半音,变化范围为 1 到 24 个半音。需要指出的是,在 Gfeller 的研究中使用了自适应步长方法来决定最小差阈。当正确回答的题数达到一个约定值时,自适应步长将变小;当错误的回答题数达到一个约定值时,自适应步长将增加。这种方法的前提假设是当步长增加时,受试者更易答对,相反,当步长减小时,受试者更易答错。然而,此假设并不总是成立,由于人工耳蜗植入者的音高排序(pitch ranking)能力具有很大的不一致性,他们音高排序的准确度受刺激信号的基频(F0)的影响很大。例如,即使步长不变,植入者的音高排序准确性也会因音符的 F0 的不同而发生从 100% 正确到 100% 错误的变化。因此,较大的步长并不代表更好的分辨率,反之亦然。当研究中使用相同的测试方法但选用不同的频率范围或在相同的频率范围内选用不同的参考基频时,所得到的结果或差阈很有可能会十分不同。

在 Sucher 的研究中,8 名植入者(Nucleus CI24 或 CI22 使用者)和 10 名 NH 受试者进行了音高排序测试。同预料的一样,CI 使用者的能力较 NH 受试者差很多。测试中使用了 6-半音和 1-半音作为音高差。尤其对于 CI 使用者来说,他们对 1-半音差的音高排序只能达到随机猜测水平。

最近,平利川等对美国 Advanced Bionics(AB)公司的 8 名 HiRes 90K 植入者和 8 名 NH 进行了音高差阈测试和音高变化方向分辨测试。测试中使用了 4 种不同音色和 3 个不同频率区间。这些植入者使用的言语处理策略为 HiRes120 或 HiRes-P。实验结果表明,同国外研究者的结论相似,无论使用 HiRes 或 HiRes120 策略,我国人工耳蜗植入者的音高差阈分辨能力为 1.8~10.7 个半音,音高变化方向分辨能力在 2.1~13.6 个半音,均较正常听力者差很多。分析表明,电刺激信号模式的复杂度直接影响音高分辨能力,具有清晰、规整的低阶谐波成分的声音对音高分辨十分重要。F0 相关的瞬时模式在 F0 比较低时对音高分辨有作用。

#### 2.4 音色感知

在音乐相关研究中,音色感知通常以乐器分辨测试作为研究方法。Gfeller 等对 28 名 Clarion CI 使用者及 41 名 NH 受试者对小号、单簧管、小提琴、及钢琴等四种乐器的分辨能力进行了测试<sup>[8]</sup>。结果表明,NH 受试者对乐器的分辨能力较 CI 使用者强很多。在另一项研究中,Gfeller 比较了 CI 和 NH 受试者对八种不同乐器的感知。NH 受试者的分辨率为 91%,CI 受试者的为 47%。他们还发现,如长笛、小提琴、钢琴等高频乐器声对 CI 受试者来说更嘈杂和难于感知,然而对于 NH 受试者并无此感受。言语识别同其音色感知间并无紧密的对应关系。Looi 比较了 15 名 CI 和 15 名 NH 受试者对单一音乐乐器的辨识能力<sup>[20]</sup>。实验表明,植入者的单乐器识别能力比多乐器识别能力好。Stainsby 对复杂声信号的频谱信息量进行了评估,并计算了声音的物理频谱和内在频谱间的关系(即实际频谱与听者感知的频谱信息量间的关系)<sup>[21]</sup>。4 名 NH、5 名 CI 和 3 名助听器使用者参与了测试。结果表明,听力受损者的实际与内在频谱相关性较正常听力者弱。对于 CI 使用者来说,虽然他们具有十分有限的音色分辨能力,经人工耳蜗传递的频谱信息量较正常听力者来说是很少的,这也是导致其无法准确辨识乐器的原因。总的来说,人工耳蜗所能提供给植入者的音色感知能力并不足以使其准确地辨识音乐乐器。

### 3 音乐感知调查问卷

音乐欣赏调查问卷经常作为听辨测试的辅助测试内容为研究人员所使用。国外常见的调查问卷主要包括爱荷华大学音乐背景和听赏水平问卷

(Iowa Musical Background and Appreciation Questionnaire, IMBAQ)以及成人人工耳蜗使用者音乐体验问卷。IMBAQ 由美国爱荷华大学 Gfeller 及同事编写,在人工耳蜗使用者的测试中得到应用<sup>[22]</sup>。问卷第一部分为了解受试者的音乐训练和活动参与情况,主要包括术前各个教育阶段(包括初中、高中和大学)的情况,教育结束后的活动,或者是否有过自学经历,以及人工耳蜗植入术后的情况。第二部分主要调查人工耳蜗植入效果稳定后植入者的音乐聆听体验和愉悦程度的自我评价,包括听音乐的频繁程度,是否获得愉快感受,不同来源和形式音乐的感受差异,音质评价,歌声和乐器的分辨能力等。成人人工耳蜗使用者音乐体验问卷表为 Mirza 及同事编写的自我评估工具<sup>[23]</sup>,问卷包含患者在耳聋之前和人工耳蜗植入后参与音乐活动和体验的各个方面,某些问题同 IMBAQ 相似,可作为 IMBAQ 的辅助调查问卷。

### 4 提高人工耳蜗植入者音乐感知的方法

提高音乐感知能力主要可通过改善现有人工耳蜗系统提高电极输出脉冲信号中与音乐感知相关特征的表达,也可在现有装置的基础上通过增加音乐训练的途径改善其音乐感知能力。

#### 4.1 改善声音处理器

本文中回顾的大部分研究表明,多数人工耳蜗植入者的音高和音色感知能力并不理想。现有人工耳蜗声音处理算法可通过几种可能的改进方法为植入者提供更多的音高相关信息。一种可能是改善用于声音信号短时谱分析的滤波器组设计。Geurts 等在 2004 年提出的改进思想是通过在包含基频信息的频率段增加滤波器的个数,从而提供更精细的频率分辨率,进而提高位置音高信息。另一方面,瞬时精度也能被提高。

另一种修改方法目的在于减少或消除植入电极间的相移,研究普遍认为相移是导致植入者无法从与基频相关的包络调制信息里提取有效信息的主要根源。增加幅度调制深度也可使音高信息变得更加明显<sup>[13]</sup>。

第三种修改方法是直接用基频信息控制电极刺激率。虽然该技术曾被用于 MPEAK 策略,而随着使用固定刺激率的声音处理策略(如 ACE、SPEAK 和 CIS)获得的言语理解度的提高,诸如 MPEAK 等早期策略已经不再使用,然而心理物理

学研究表明,电刺激速率的变化可产生对音高变化的感知,并且,音高与刺激速率间的关系也同正常听力者的音高与声调频率之间的关系相似。因此,似乎刺激速率是一个通过改善音高信息在刺激信号中的表达来改善音乐感知的合适手段。

在实际应用中,基频估计需做到实时处理,并能在很多实际的听音环境(如噪声或回声环境)下达到良好的估计效果。估计得到的基频应同时对所有活动电极的电刺激输出信号进行调制。已经证实此类算法改进可以使植入者至少在一定基频范围内获得更好的音乐音高感知。

#### 4.2 改善刺激方式

除在体外机的声音处理策略上进行改进外,新的电极技术也通过加入声音信号的精细结构信息来提高音高感知。精细结构信息包含声音信号中的快变信息,是对音高感知非常重要的声音特征。目前,研究人员通过新的电极序列设计增加了耳蜗的刺激位置,进而提高了对声音精细结构的表达,并将电极序列放置在距离听神经更近的位置对降低电极间的电流互扰也有积极作用<sup>[24-26]</sup>。

声电联合刺激也是一种改善植入者音乐感知的可能方法。在过去 20 年的研究中发现,在使用人工耳蜗的植入者中有许多都保留有一定的低频残余听力。低频区域被认为是对音高感知十分重要的区域,保留有低频残余听力意味着可以在使用人工耳蜗改善其中高频听力的同时,利用助听器对低频残余听力进行声刺激从而改善其对音高的感知能力<sup>[27,28]</sup>。对于这类病人,应使用短电极进行耳蜗植入,在手术期间,也应尽量避免对其耳蜗低频区域造成损伤,并在对侧耳或同侧耳配置助听器<sup>[29,30]</sup>。因此,通过这两种设备的联合使用,患者可以最大限度的获得声音感知。有研究表明,通过使用联合刺激可以使植入者在如熟悉旋律识别、乐器识别、音乐欣赏方面获得更高的水平<sup>[27,29,31]</sup>。Kong 等对 5 名在非植入耳配戴助听器的人工耳蜗植入者进行了旋律识别测试,比较了三种听音情况——只使用 CI,只使用助听器,和两者同时使用<sup>[27]</sup>。测试结果表明,只使用助听器和两者共同使用可获得相近的识别率,比只使用人工耳蜗的情况高 17%。该作者的观点是使用助听器可使植入者获得更好的低频段精细结构信息,从而增加了植入者从信号中提取 F0 信息的机会。虽然上述实验结果尚未体现出联合刺激在提高人工耳蜗植入者音乐感知能力方面的贡献,但从理论上来说,声音中的精细结构信息对于复杂声音的感知是肯定有帮助的,只是目前这

部分增加的低频信息对于人工耳蜗植入者的作用尚待发现。Reiss 等的研究报告也指出,这种联合刺激模式的效果会随着时间的增加而改变,预示着这项新技术在未来也将是研究的重点方向<sup>[32]</sup>。

#### 4.3 音乐训练

虽然现今的言语处理器并不能很理想地传输音高和音色信息,但有研究表明音乐感知可通过训练得到改善<sup>[9,33]</sup>。Gfeller 等开发了一套工具可供人工耳蜗植入者在家中多种音乐训练项目<sup>[34]</sup>,主要包括音高感知、音色识别、音色欣赏、旋律辨识、音乐风格欣赏等。11 名植入者完成了该训练(30 分钟一节,一周训练 4 次,共 48 节),对训练程序的整体满意度和训练收获通过调查问卷形式进行了评估,平均满意度为 3.45(满分 4 分)。Gfeller 分析了训练前后的旋律感知及欣赏变化。结果显示,训后旋律识别率较训前有提高,但无统计意义差别。但可以看到训后受试者在复杂歌曲测试中较训前可大幅提高对旋律的识别。就音乐聆听度来说,训后受试者的打分较训前明显提高<sup>[9]</sup>。Galvin 等开发了一套用于旋律轮廓辨识的自适应训练程序<sup>[35]</sup>。6 名人工耳蜗植入者参与了对训练程序的评估。结果显示,训练可使全部 6 名受试者的旋律轮廓识别率由 15.4% 提高到 45.4%。研究同时报告称,即使在受试者停止音乐训练一个月后,他们仍能保持很好的训练效果。

## 5 结论

总的来说,这些研究均说明即使目前的人工耳蜗系统还存在无法提供足够音高、音色信息的不足,训练是一个能够让人工耳蜗植入者获得音乐感知帮助的一种潜在手段。

#### 参 考 文 献

- [1] Bernstein L. The unanswered question[M]. Cambridge: Harvard University Press, 1976.
- [2] Patel A D. Language, music, syntax and the brain[J]. Nat Neurosci 2003, 6: 674-681.
- [3] Krumhansl C L, Iverson P. Perceptual interactions between musical pitch and timbre[J]. J Exp Psychol, 1992, 18: 739-751.
- [4] ASA. American standards acoustic terminology[M]. New York: Acoustical Society of America, 1960.
- [5] Gfeller K, Woodworth G, Robin D A, Witt S, Knutson J E. Perception of rhythmic and sequential pitch patterns by normally hearing adults and adult cochlear implant users[J]. Ear Hear, 1997, 18: 252-260.
- [6] Loizou P C. Mimicking the human ear[J]. IEEE Signal Processing Magazine, 1998, 115: 101-130.
- [7] Greenwood D D. A cochlear frequency-position function for sever-

- al species--29 years later[J]. *J Acoust Soc Am*, 1990, **87**: 2592-2605.
- [8] Gfeller K, Knutson J F, Woodworth G, Witt S, DeBus B. Timbral recognition and appraisal by adult cochlear implant users and normal-hearing adults[J]. *J Am Acad Audiol*, 1998, **9**: 1-19.
- [9] Gfeller K, Witt S, Stordahl J, Mehr M, Woodworth G. The effects of training on melody recognition and appraisal by adult cochlear implant recipients[J]. *J Acad Rehabil Audio*, 2000, **33**: 115-38.
- [10] Lassaletta L, Castro A, Bastarrica M, et al. Does music perception have an impact on quality of life following cochlear implantation[J]. *Acta Oto-Laryngologica*, 2007, **127**: 1-5.
- [11] Gfeller K, Lansing CR. Melodic, rhythmic, and timbral perception of adult cochlear implant users[J]. *J Speech Hear Res*, 1991, **34**: 916-20.
- [12] Leal M C, Shin Y J, Laborde M L, et al. Music perception in adult cochlear implant recipients[J]. *Acta Otolaryngol*, 2003, **123**: 826-835.
- [13] McDermott H J. Music perception with cochlear implants: a review[J]. *Trends Amplif*, 2004, **8**: 49-82.
- [14] Kong Y Y, Cruz R, Jones J A, Zeng F G. Music perception with temporal cues in acoustic and electric hearing[J]. *Ear Hear*, 2004, **25**: 173-185.
- [15] Kim I, Yang E, Donnelly P J, Limb C J. Preservation of rhythmic clocking in cochlear implant users: A study of isochronous versus anisochronous beat detection[J]. *Trends Amplif*, 2010, **14**: 164-169.
- [16] Gfeller K, Turner C, Mehr M, et al. Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults[J]. *Cochlear Implants International*, 2002, **3**: 29-53.
- [17] Looi V, Sucher C M, McDermott H J. Melodies familiar to the Australian population across a range of hearing abilities[J]. *Austr NZ J Audiol*, 2003, **25**: 75-83.
- [18] Schulz E, Kerber M. Music perception with the MED-EL implants. In: Hochmair-Desoyer H, Hochmair ES, editors. *Advances in cochlear implants*[M]. Vienna: Datenkonvertierung, Reproduktion und Druck, 1994, 326-32.
- [19] Looi V, McDermott H J, McKay C M, et al. Pitch discrimination and melody recognition by cochlear implant users[C]// *Proceedings of the VIII International Cochlear Implant Conference*, Indianapolis, Indiana. Elsevier, 2004, 197-200.
- [20] Looi V, McDermott H, McKay C, Hickson L. Music perception of cochlear implant users compared with that of hearing aid users[J]. *Ear Hear*, 2008, **29**: 421-434.
- [21] Stainsby T H. *The Perception of Musical Sounds with Cochlear Implants*[D]. Melbourne: The University of Melbourne, 2001.
- [22] Gfeller K, Witt S. *Iowa Musical Background and Appreciation Questionnaire (IMBAQ)*[M]. Iowa City: University of Iowa, 1998.
- [23] Mirza S, Douglas S A, Lindsey P, Hildreth T, Hawthorne M. (2003) *Appreciation of music in adult patients with cochlear implants: a patient questionnaire*[J]. *Cochlear Implants International*, **4**: 85-95.
- [24] Wilson B S, Lawson D T, Zerbi M, Finley C C. Recent developments with the CIS strategies[C]// *Hochmair-Desoyer of the Third International Cochlear Implant Conference*, 1993, 103-112.
- [25] Donaldson G S, Kreft H A. Place-pitch discrimination of single-versus dual-electrode stimuli by cochlear implant users[J]. *Journal of the Acoustical Society of America*, 2005, **118**(2): 623-626.
- [26] Koch D B, Downing M, Osberger M J, Litvak L. Using current steering to increase spectral resolution in CII and HiRes 90k users. *Ear and Hearing*, 2007, **28**(2): 38-41.
- [27] Kong Y Y, Stickney G S, Zeng F G. Speech and melody recognition in binaurally combined acoustic and electric hearing[J]. *J Acoust Soc Am*, 2005, **117**(Pt 1): 1351-61.
- [28] Tyler R S, Parkinson A J, Wilson B S, Witt S, Preece J P, Noble W. Patients utilizing a hearing aid and a cochlear implant: Speech perception and localization[J]. *Ear Hear*, 2002, **23**: 98-105.
- [29] Gantz B J, Turner C. Combining acoustic and electrical speech processing: Iowa Nucleus hybrid implant[J]. *Acta Otolaryngol (Stockh)*, 2004, **124**: 344-347.
- [30] Gantz B J, Turner C, Gfeller K E, Lowder M W. Preservation of hearing in cochlear implant surgery: Advantages of combined electrical and acoustical speech processing[J]. *Laryngoscope*, 2005, **115**: 796-802.
- [31] Gfeller K E, Olszewski C, Turner C, Gantz B, Oleson J. Music perception with cochlear implants and residual hearing[J]. *Audiol Neurootol*, 2006, **11**(Suppl): 12-15.
- [32] Reiss L A, Gantz B J, Turner C. Cochlear implant speech processor frequency allocations may influence pitch perception[J]. *Otol Neurotol*, 2008, **29**: 160-167.
- [33] Gfeller K, Christ A, Knutson I F, Witt S, Murray K T, Tyler R S. Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients[J]. *J Am Acad Audiol* 2000, **11**: 390-406.
- [34] Gfeller K. Aural rehabilitation of music listening for adult cochlear implant recipients: Addressing learner characteristics[J]. *Music Ther Perspect*, 2001, **19**: 88-95.
- [35] Galvin J J, Fu Q J, Nogaki G. Melodic contour identification by cochlear implant listeners[J]. *Ear Hear*, 2007, **28**: 302-319.