

声刺激重放方式对优先效应回声阈值的影响*

曹 腾 余光正 饶 丹[†]

(华南理工大学 物理与光电学院声学研究所 广州 510641)

2024 年 7 月 25 日收到

2024 年 9 月 17 日定稿

摘要 针对优先效应研究中三种常见的声刺激重放方式对回声阈值的影响,开展了心理声学实验研究,同时测量了三种刺激条件下(基准、建立和破坏)四种类型刺激信号(噪声脉冲、语音、音乐和粉红噪声)的回声阈值。结果表明,耳机虚拟听觉空间法与扬声器重放法在回声阈值上没有显著差异,而它们的回声阈值与耳机双耳时间差法都存在显著差异,主要体现在长时刺激上。可能的原因是不同重放方式的空间定位因素差异。此外,随着刺激条件的变化(从基准、建立到破坏),三种刺激重放方式的动态回声阈值变化规律一致,即噪声脉冲刺激的回声阈值先增加后降低,三种长时刺激的回声阈值均未出现显著变化。

关键词 优先效应, 回声阈值, 声刺激重放方式, 定位因素

PACS: 43.66, 43.71

DOI: 10.12395/0371-0025.2024220

CSTR: 32049.14.11-2065.2024220

Effect of sound stimuli presentation method on echo threshold of precedence effect

CAO Teng YU Guangzheng RAO Dan[†]

(Acoustic Lab., School of Physics and Optoelectronics, South China University of Technology Guangzhou 510641)

Received Jul. 25, 2024

Revised Sept. 17, 2024

Abstract A psychoacoustic experiment is conducted to investigate the effect of three common sound stimuli presentation methods on echo threshold (ET) of precedence effect. The ETs of four types of stimulus (noise burst, speech, music, pink noise) under three stimuli conditions (Baseline, Buildup, Breakdown) are measured. The results show that the ETs of headphone-based virtual auditory space (VAS) presentation method and loudspeaker-based presentation method have no significant difference, while ETs of them are all significantly different from those of the interaural time difference (ITD) method, especially for long-duration stimuli. The possible reason is the differences in localization cues between various stimulus presentation methods. In addition, with the change of stimuli conditions (from Baseline, Buildup to Breakdown), the dynamic changes of ETs under three stimuli presentation methods are the same, that is, the ETs of the noise burst increase first and then decrease, and the ETs of the three long-duration stimuli do not show significant changes.

Keywords Precedence effect, Echo threshold, Sound stimuli presentation method, Localization cues

引言

在日常听音环境中,除了声源发出的直达声之外,还存在来自墙壁、天花板等界面的反射声。然而,即使在这样复杂的声环境中,人类依然能够较为准确地进行声源定位。该现象与一种被称作优先效

应 (precedence effect)^[1-2] 的听觉效应密切相关,即听觉系统主要依靠第一个到达人耳的声音(直达声)来进行声源定位,而对后到达的反射声起抑制作用。该效应对理解人类听觉定位机理有着重要的科学意义,同时在声重放系统、建筑声学设计以及声信号处理等方面也有一定的应用价值^[3-4]。

在优先效应研究中,为了易于找出相关规律,研

* 国家自然科学基金项目 (12074129) 和广东省自然科学基金项目 (2024A1515011446) 资助

[†] 通讯作者: 饶丹, phdrao@scut.edu.cn

究者们一般只考虑两个声源(即直达声源和单个反射声源)的情况,并通过一种被称作“领先声-滞后声”的范式(Lead-Lag paradigm)^[5]来实现。研究表明,对于短脉冲信号,当领先声-滞后声延时较小时(约 1~5 ms),听者一般只感知到来自领先声一侧的单个融合的听觉事件,此时优先效应起作用,领先声(直达声)信息主导着声源定位;而随着延时进一步增加并超过某一阈值时,听者逐渐感知到两个空间分离的听觉事件。这种情形下的滞后声一般被称作回声,而该阈值也被称作回声阈值(ET)^[5],即听者从感知到单个融合听觉事件过渡到两个空间分离听觉事件的临界延时,该阈值作为优先效应的上限延时阈值,反映了优先效应的鲁棒性。

优先效应研究中有三种不同刺激条件的回声阈值测量^[6-8]:最常见的是播放单对领先声-滞后声刺激的基准条件,其次是多次连续重复播放刺激对的建立条件^[9],以及在连续多次播放刺激对后,突然改变刺激的某些特征(如交换领先声和滞后声的空间位置)的破坏条件^[9]。通常情况下,建立条件的回声阈值比基准条件的大,而破坏条件的回声阈值比建立条件的小。

影响回声阈值的因素主要包括:阈值的定义和标准、刺激特性、刺激重放方式等,而本文主要探究刺激重放方式的影响。在现有文献中,声刺激主要通过扬声器或耳机进行重放。其中,早期大多数研究采用扬声器进行重放(以下统称为扬声器重放法)^[1,6,8,10-11],它是在消声室中利用两个扬声器重放刺激来分别模拟直达声和单个反射声。随着音频信号处理和声重放技术的发展,耳机重放凭着其能够独立控制定位因素和易于操作的优点逐渐被广泛采用。耳机重放又分为双耳时间差(ITD)法、双耳声级差(ILD)法和虚拟听觉空间(VAS)法,其中,ITD法^[12-14]和ILD法^[15]是通过独立控制定位因素ITD或ILD实现领先声(或滞后声)的偏侧化,而VAS法^[16-18]则是利用头相关传输函数(HRTF)^[19]对信号进行滤波实现声源的虚拟。相较而言,以往基于耳机重放的优先效应研究多数采用ITD法,因为这种重放方式的实现较为方便,并且也能获得良好的优先效应效果,但事实上,这种重放方式与耳机VAS法或扬声器重放法存在定位因素的差异,而这种差异是否会影响阈值结果,暂未有文献进行报道。

在优先效应研究中,虽然研究者们会根据不同的实验目的采用不同的声刺激重放方式,然而很少去比较和讨论不同重放方式下优先效应结果的差异。从一些研究结果来看,这种差异是存在的,但依

然缺乏统一的定论,特别是对回声阈值的影响方面。例如,Aoki等^[20]比较了同一批受试者(被试)在扬声器重放法、耳机ITD法和耳机ILD法下优先效应的起始(Onset)优势差异,结果表明这三种重放方式下信号的起始在听觉事件整体感知方面都具有相似的主导作用。Seeber和Hafter的初步研究表明^[21],对于同一批被试,在经过HRTF个体选择任务(即通过一些标准让被试从包含多对非个性化HRTF的数据库中选择一对)后的耳机VAS法和扬声器重放法在优先效应方面差异较小。此外,关于重放方式对回声阈值的影响,一些文献综述了前人的工作并指出^[5,22],基于耳机ITD法测出的回声阈值要比扬声器重放法更小,其中耳机ITD法的回声阈值范围在1~2 ms左右^[12],而扬声器重放法的回声阈值范围在5~10 ms甚至更长^[2]。然而,值得注意的是,这些观点并非基于同一批被试的测量数据得出。Krumbholz等^[23]以及Brown等^[24]系统比较了耳机ITD法和ILD法在优先效应不同刺激条件下的回声阈值差异,其结果表明,对于短脉冲信号(时长分别为120 μ s和123 μ s),ITD法在三种刺激条件下的回声阈值都比ILD法的更高,并且在ILD法中,破坏条件的回声阈值较建立条件出现了显著降低,而ITD法则未出现类似的阈值变化过程。这两项研究的结果表明,不同刺激条件下优先效应的鲁棒性与定位因素有关。除此之外,文献^[25]初步探讨了耳机VAS法和耳机ITD法在基准条件下的回声阈值差异,结果表明,这两种重放方式的定位因素差异是导致长时信号的回声阈值存在差异的主要原因。

综上所述,在优先效应研究中,系统地比较不同重放方式的回声阈值差异的研究较少,尤其是同时采用并比较三种常用重放方式(耳机ITD法、扬声器重放法、耳机VAS法)的差异研究,暂时未见文献报道。这使得不同重放方式对回声阈值的影响仍不明确。此外,多数研究都只采用短时脉冲信号,缺乏采用更能反映真实情形的长时信号,如语音、音乐、噪声等。

本文通过心理声学实验,在同一实验条件和同一批被试条件下,比较分析优先效应研究中常用的三种重放方式(耳机ITD法、耳机VAS法和扬声器重放法)分别在三种刺激条件(基准、建立和破坏)下的回声阈值差异。研究结果为优先效应研究中如何选择合适的重放方式以及实际应用提供了参考依据,同时对动态过程的初步探究也有助于加深对优先效应作用机制的理解。

1 实验方法

1.1 被试

本实验共招募了 12 名被试, 年龄范围在 20~29 岁之间, 经听力筛查判断均为正常听力者。同时, 本文涉及真人被试的各项实验内容均已通过华南理工大学附属第二医院伦理委员会的伦理审查。

1.2 实验信号

以往优先效应研究大多采用短脉冲信号(如噪声脉冲、猝发声), 相比之下, 更能反映真实情况的长时信号, 如语音、音乐和噪声等则较少被研究。此外, 刺激信号的声学特性(包括信号时长、频谱宽度以及瞬态特性)也会对回声阈值产生影响。基于以上考虑, 本实验采用了声学特性存在差异且具有一定代表性的刺激信号类型, 包括 1 种短脉冲信号和 3 种长时信号: 1 ms 的噪声脉冲、2.3 s 的普通话男声片段“我们都在学习电脑打字”作为语音信号、3 s 的短号声片段作为音乐信号、2 s 的全频带粉红噪声信号。其中, 噪声脉冲和粉红噪声由 MATLAB 生成, 音乐信号选自阿基米德项目录制 CD 的第 34 首作品 *Cadance from Concert* 的 0.5~3.5 s 片段^[26]。所有信号的采样率均为 44.1 kHz。同时, 本实验采用领先声-滞后声(Lead-Lag)刺激范式测量回声阈值, 并采用相对于听者头中心的顺时针球坐标系, 即以听者头中心为坐标原点, 当声源位于水平面上时, 方位角 $\theta = 0^\circ, 90^\circ, 180^\circ, 270^\circ$ 分别对应正前、正右、正后和正左方。本研究中采用的两个声源(领先声和滞后声)均位于水平面上。

本实验涉及多种刺激重放方式和刺激条件, 具体的实验信号生成过程如下:

首先, 生成不同刺激条件下的信号。基准条件下刺激只由 1 对领先声-滞后声(即测试对)组成, 而建立和破坏条件(两种动态条件)下刺激由领先声-滞后声序列(即条件对 + 测试对)组成, 测试对为整个序列中的最后 1 对领先声-滞后声, 所有条件下被试只需要对测试对进行回声判别。由于在以往的优先效应动态条件研究中, 研究者们对短时信号和长时信号分别采取了不同的刺激序列处理方式, 因此, 本实验参照文献[7,17]的做法, 对噪声脉冲和三种长时信号也做了不同的处理。其中, 对于噪声脉冲, 条件对总共包含 9 对领先声-滞后声, 相邻 2 对领先声-滞后声之间的延时间隔为 250 ms, 条件对中最后 1 对

领先声-滞后声与测试对之间的延时间隔为 750 ms, 而对于三种长时信号, 条件对只包含 2 对领先声-滞后声, 由于信号本身做了淡入淡出的处理, 因此条件对相邻两对刺激之间没有设置额外的延时, 条件对播放结束后, 间隔 750 ms 后播放测试对。在建立条件中, 测试对的领先声-滞后声与条件对中的每 1 对都完全相同, 而在破坏条件中, 测试对的领先声和滞后声进行空间位置的交换, 即测试对中领先声和滞后声的空间位置和条件对中相反。

然后, 对上述三种刺激条件的信号分别进行耳机 VAS 法、耳机 ITD 法和扬声器重放法的信号处理。

对于耳机 VAS 法, 虽然以往研究都分别采用过个性化 HRTF 和非个性化 HRTF, 但这些研究都未将它与其他重放方式进行系统的比较, 因此, 为了方便实验探究以及考虑到实际 VAS 应用中采用非个性化 HRTF 的普遍性, 本文的耳机 VAS 法只采用非个性化 HRTF。首先, 基于 KEMAR 人工头模型, 在 Mesh2HRTF^[27] 中采用边界元法^[28] 计算了水平面上 1.5 m 的 HRTF, 然后利用傅里叶逆变换, 计算得到相应的头相关脉冲响应 HRIR (采样率为 44.1 kHz, 采样点数为 512), 最后将相应方位角的 HRIR 与刺激进行卷积得到耳机中的虚拟声信号。其中, 将方位角 $\theta = 45^\circ$ 的左右耳 HRIR 与刺激卷积得到滞后声双耳信号, 同理, 方位角 $\theta = 315^\circ$ 的为领先声信号, 将二者整合得到最终的双耳信号, 图 1(a) 为噪声脉冲刺激在耳机 VAS 法下的双耳信号示例图。

对于耳机 ITD 法, 先对耳机 VAS 法采用的 HRIR 数据进行 1500 Hz 的低通滤波, 再利用双耳互相关法^[19] 计算相应的 ITD 值。其中, 45° 方位角(对应滞后声)的 ITD 值为 $+436 \mu\text{s}$, 315° 方位角(对应领先声)的 ITD 值为 $-436 \mu\text{s}$ 。最后, 利用相应 ITD 值对刺激进行处理并整合得到最终的双耳信号, 图 1(b) 为噪声脉冲刺激在耳机 ITD 法下的双耳信号示例图。

对于扬声器重放法, 直接将信号馈给一对扬声器, 并通过调节扬声器通路延时模拟领先声和滞后声。图 1(c) 为噪声脉冲刺激在扬声器重放法下的双耳接收到的信号示例图。

1.3 测试方法与流程

所有实验均采用二下一上自适应阶梯法^[8,17] 和二项迫选法进行测量, 并采用双轨道程序。在每一个试次(trial)中, 领先声-滞后声延时从两条自适应轨道的其中一条随机获得, 其中, 一条轨道的初始延时设置为 1 ms, 以便所有被试都能感知到侧向的单个融合听觉事件, 另一条轨道的初始延时设置为一

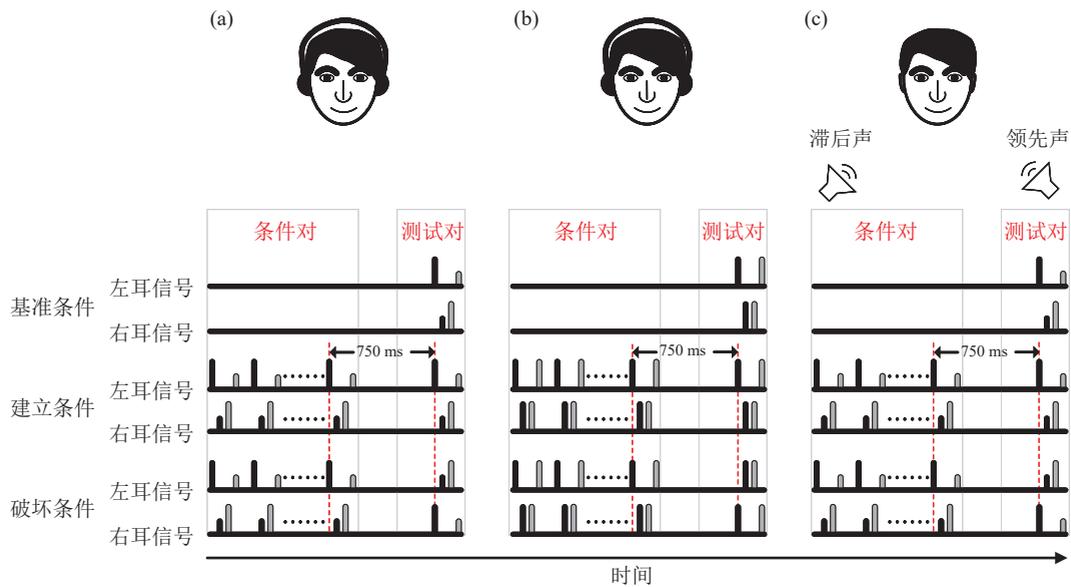


图1 基于三种刺激重放方式下三种刺激条件的噪声脉冲双耳信号示例(基准条件只包含测试对, 建立和破坏条件由条件对+测试对组成, 黑色和灰色条柱分别表示领先声和滞后声信号) (a) 耳机 VAS 法; (b) 耳机 ITD 法; (c) 扬声器重放法

个更大的值(约 30~90 ms, 取决于不同被试和不同条件), 以便所有被试都能够清晰地感知到两个空间分离的听觉事件。两条自适应轨道都遵循二下一上的准则, 基于心理测量函数值为 70.7% 的占比感知两个空间分离的听觉事件。当被试连续两次给予正反馈(即感知到两个空间分离的听觉事件)时, 延时会降低, 当被试给予一次负反馈(即未感知到两个空间分离的声像)时, 延时会增加。延时变化的规则采用对数步长类型^[24], 一开始采用 0.2 的步长因子(新延时 = 旧延时 $\times 10^{0.2}$), 当每条轨道的反转次数达到 4 次之后, 步长因子降为 0.05(新延时 = 旧延时 $\times 10^{0.05}$), 并直到整个程序运行结束。每条自适应轨道的终止条件为反转次数达到 8 次或试次数达到 100 次, 而每条轨道的回声阈值为最后 4 个反转点对应延时的均值, 最终的回声阈值取两条轨道阈值的平均值。

耳机重放测试在经过声学处理的安静房间内, 扬声器重放测试在全消声室(有效空间尺寸为 13 m \times 6.5 m \times 5.5 m)中进行。在耳机重放测试中, 被试采用坐姿正对电脑屏幕, 通过鼠标进行自主测听和反馈, 双耳信号由计算机播放, 并通过 Roland Rubix 44 声卡和 TOPPING A50s 耳机放大器输出到入耳式耳机(Etymotic Research ER 2)进行重放。在扬声器重放测试中, 被试采用坐姿面向前方, 通过手持键盘进行自主测听和反馈, 同时让被试尽量保持头部不动, 扬声器信号由计算机进行播放, 并通过 Roland Rubix 44 声卡馈给两个扬声器(Genelec 8010A)进行重放, 其中, 45°方位角的扬声器重放滞后声刺

激, 315°方位角的扬声器重放领先声刺激, 两扬声器高度均为 1.2 m, 与听者的耳道口基本在同一水平面, 两扬声器距离被试头中心均为 1.5 m, 噪声脉冲刺激在被试头中心位置的最大声压级约为 55 dB(A), 粉红噪声刺激在被试头中心位置的等效声压级约为 59 dB(A), 语音和音乐刺激与其基本一致。所有实验的参数设置界面、测试界面和计算均通过 MATLAB 程序进行。

根据刺激重放方式和刺激条件, 整个实验总共分为 9 个模块(3 种刺激重放方式 \times 3 种刺激条件)。每个模块都需要进行 4 种类型刺激的回声阈值测量, 因此, 对于每一位被试, 总共需要测量 36 个回声阈值(9 个模块 \times 4 种刺激类型)。每个模块的实验时长约 30~40 min, 分多次完成, 在实验过程中被试随时可以休息。所有被试被要求判断是否感知到两个空间分离的听觉事件, 并且尽量忽略音色、响度、谐波等变化, 主要以空间感知作为判别标准。在正式实验前, 所有被试都经过详细的指导和训练, 以确保他们熟悉实验流程和回声测听标准。

2 实验结果分析与讨论

对所有被试数据进行了重测信度检验^[29-30], 结果表明数据的一致性较好。本文对实验结果的分析主要从两个方面进行。首先, 对基准条件(优先效应研究中最为常用的刺激条件)下的实验数据采用双因素重复测量方差分析, 结合关于定位因素差异的分析^[25], 综合讨论这三种刺激重放方式的回声阈值

差异。其次,结合两个动态条件的实验结果,采用三因素重复测量方差分析,综合分析三种刺激重放方式对不同刺激条件下回声阈值的影响和回声阈值变化规律的影响,其中,对变化规律的影响是这个部分的分析重点。

2.1 基准条件的结果对比及分析

图2为基准条件下三种刺激重放方式的回声阈值平均结果。首先,从图中可以发现,在三种刺激重放方式下,三种长时刺激(语音、音乐和粉红噪声)的平均回声阈值都比噪声脉冲明显偏大。此外,对于不同重放方式的阈值差异,结果表明,其与刺激的类型密切相关,对于噪声脉冲,三种重放方式的平均回声阈值之间没有出现明显的差异,并且其回声阈值范围都在10 ms左右或10 ms以下,而对于三种长时刺激,不同重放方式的回声阈值则表现出较为明显的差异,具体而言,耳机ITD法的平均回声阈值比耳机VAS法和扬声器重放法都明显偏大,而耳机VAS法的平均回声阈值和扬声器重放法相近,这种差异在音乐和粉红噪声刺激中尤为明显。值得注意的是,对于回声阈值的离散程度,三种长时刺激的阈值离散程度都比噪声脉冲偏大,尤其是音乐和粉红噪声,而对于所有刺激信号,三种重放方式的阈值离散程度基本一致。

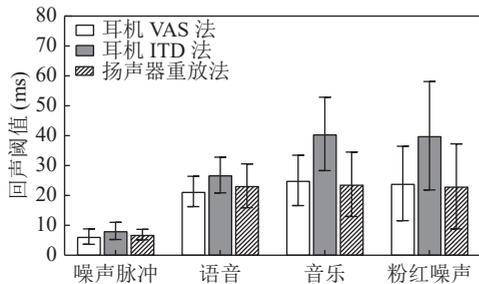


图2 基准条件下三种刺激重放方式的回声阈值平均结果 (12名被试,误差条代表均值的标准差)

对上述实验结果采用双因素重复测量方差分析(3种刺激重放方式 \times 4种刺激类型),因变量为回声阈值。首先,夏皮洛-威尔克正态性检验(Shapiro-Wilk test)结果显示,12组数据均满足正态性检验($p > 0.05$)。此外,Mauchly球形度检验结果显示,部分数据违背球形度检验(刺激类型: $p = 0.022 < 0.05$;刺激重放方式 \times 刺激类型: $p = 0.018 < 0.05$),这部分数据在后续的主体内效应检验结果中采用Greenhouse-Geisser校正项。主体内效应检验结果显示:刺激重放方式对回声阈值有显著的影响 [$F(2, 22) = 22.909, p < 0.001$];经过Greenhouse-Geisser校正后,刺激类型

对回声阈值有显著的影响 [$F(1.690, 18.594) = 37.350, p < 0.001$],且这两个因素对回声阈值的交互作用也是显著的 [$F(3.354, 36.891) = 7.754, p < 0.001$],说明三种重放方式的回声阈值差异与刺激类型密切相关。

对于不同类型刺激的回声阈值差异,简单效应分析结果显示:无论在何种刺激重放方式下,三种长时刺激的平均回声阈值都比噪声脉冲显著偏高(所有 $p < 0.01$)。这与以往分别采用不同重放方式的基准条件研究结论一致^[10,14,16,21],说明听者对噪声脉冲刺激和长时刺激的回声阈值感知存在较大差别。部分原因在于,噪声脉冲刺激由于时长较短而不会导致领先声和滞后声在时域上存在重叠,而长时刺激则会出现一定程度的时域重叠,从而进一步掩蔽回声。

对于不同刺激重放方式的回声阈值差异,简单效应分析结果显示:对于噪声脉冲,三种重放方式的平均回声阈值之间没有显著差异。值得注意的是,文献[5,22]曾提出短脉冲刺激在耳机ITD法的回声阈值较扬声器重放法更小的观点,但在本实验中并未得到相似的结论。事实上,这两篇文献中的观点并非基于相同条件和同一批被试的数据得出,而是通过对比前人不同研究工作得到的,其结论的合理性还有待进一步考证。相比之下,本研究中基于同一批被试和相同条件得到的结论更为合理。此外,在本研究中,对于语音刺激,扬声器重放法的回声阈值和两种耳机重放方式都不存在显著差异,只有耳机VAS法和耳机ITD法的回声阈值存在显著差异($p = 0.001$);对于音乐和粉红噪声,只有扬声器重放法和耳机VAS法的回声阈值不存在显著差异,而它们的回声阈值和耳机ITD法均存在显著差异(所有 $p < 0.01$)。总体而言,对于四种刺激信号,耳机VAS法和扬声器重放法这两种重放方式的回声阈值之间都没有显著差异。

不同重放方式的本质差异在于定位因素的不同,因而导致空间听觉感知存在差异。文献[25]的结果表明,这些差异会引起回声阈值的显著差异。具体到本实验中采用的重放方式,耳机ITD法只控制单一定位因素ITD实现声源偏侧化,其ILD值为零。因此,耳机ITD法存在定位因素ITD和ILD的不一致性^[24],例如,对于领先声信号,ITD值为 $-436 \mu\text{s}$ (对应方位角为 315°),ILD值为0 dB(对应方位角为 0°)。而耳机VAS法和扬声器重放法的定位因素ITD和ILD在空间位置上是对应一致的。耳机ITD法中定位因素的不一致性导致两种空间听觉感知的变化:领先声-滞后声空间分离感知程度以及单声源听觉事件宽度。一方面,由于控制非零ITD或

ILD 都能使听觉事件发生偏侧, 因此相较于同时控制多种定位因素的耳机 VAS 法和扬声器重放法, 只控制单一定位因素的耳机 ITD 法中听觉事件的偏侧程度相对更小, 导致其对应的领先声和滞后声听觉事件的空间分离程度更小, 进而更难感知到两个空间分离的听觉事件; 另一方面, 基于瑞利双工理论^[2], 低频成分的定位主要由 ITD 主导, 高频成分主要由 ILD 主导, 因此, 对于宽频带信号, 耳机 ITD 法会使单声源信号的不同频率成分偏侧到不同的位置 (低频成分偏侧程度较大, 高频成分偏侧程度较小), 从而导致听觉事件出现展宽。而根据信号检测理论^[31], 当两个听觉事件的空间距离固定时, 单声源 (领先声或滞后声) 听觉事件的宽度 (对应于在听觉空间的分布范围) 越大, 越难区分两个听觉事件。

综上所述, 在基准条件下, 长时刺激在耳机 ITD 法下的回声阈值基本都要比耳机 VAS 法和扬声器重放法显著偏大, 主要原因在于, 只有耳机 ITD 法存在定位因素的不一致性, 这种不一致性造成两种空间感知变化 (更小的领先声-滞后声空间分离感知程度和更大的单声源听觉事件宽度), 进而更难感知到两个空间分离的听觉事件。

2.2 不同刺激条件的结果对比及分析

结合基准条件和两个动态条件的实验结果, 综

合比较三种重放方式在三种刺激条件下的回声阈值差异。图 3 为四种类型刺激在三种刺激条件下的回声阈值平均结果。

根据图 3 可以发现, 首先, 对于不同类型刺激的回声阈值差异, 两个动态条件 (建立和破坏) 下三种长时刺激的平均回声阈值都比噪声脉冲刺激更高, 这与 2.1 节基准条件的结果一致, 即便噪声脉冲刺激在建立条件出现了阈值的增加。其次, 在两个动态条件下, 噪声脉冲在三个重放方式的平均回声阈值是相近的, 而对于三种长时刺激, 尤其是音乐和粉红噪声, 耳机 VAS 法和扬声器重放法之间的平均回声阈值差异很小, 而它们都比耳机 ITD 法明显偏低。这与基准条件的结果也是一致的。

从图 3 还可以发现, 随着刺激条件的变化 (从基准、建立到破坏), 噪声脉冲的平均回声阈值都是先增加后降低, 且阈值变化的程度相一致, 其中耳机 VAS 法的阈值增加和降低分别为 8.29 ms 和 6.77 ms, 耳机 ITD 法的阈值增加和降低数值分别为 9.85 ms 和 5.25 ms, 扬声器重放法的阈值变化分别为 8.50 ms 和 7.94 ms; 对于三种长时刺激, 三种重放方式下均未观察到动态回声阈值的明显变化。这表明, 三种重放方式对动态回声阈值变化规律的影响是一致的。

为了验证刺激重放方式、刺激条件和刺激类型对回声阈值是否存在显著影响, 利用 SPSS26.0 统计

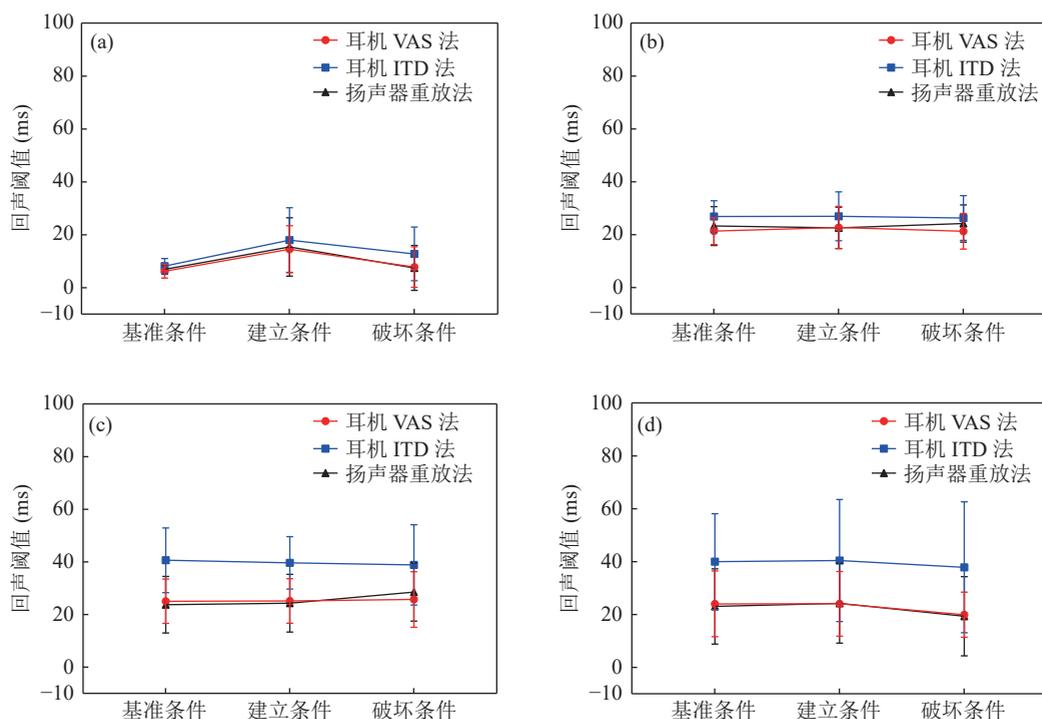


图 3 四种类型刺激在不同刺激条件下的回声阈值平均结果 (12 名被试, 误差条代表均值的标准差) (a) 噪声脉冲; (b) 语音; (c) 音乐; (d) 粉红噪声

分析软件对上述 36 组实验结果进行三因素重复测量方差分析 (3 种刺激重放方式 \times 3 种刺激条件 \times 4 种刺激类型), 因变量为回声阈值。首先, 夏皮洛-威尔克正态性检验 (Shapiro-Wilk test) 结果显示, 其中 29 组数据均满足正态性检验 ($p > 0.05$), 而剩下的 7 组数据都满足偏度小于 3 且峰度小于 8 的条件, 因此也都近似服从正态分布^[32]。此外, Mauchly 球形度检验结果显示, 部分数据违背球形度检验 (刺激类型: $p = 0.003 < 0.05$; 刺激重放方式*刺激类型: $p = 0.022 < 0.05$; 刺激条件*刺激类型: $p = 0.027 < 0.05$; 刺激重放方式*刺激条件*刺激类型: $p < 0.05$), 这部分数据在后续的主体内效应检验结果中采用 Greenhouse-Geisser 校正项。

主体内效应检验结果显示, 刺激重放方式对回声阈值有显著的影响 [$F(2, 22) = 19.501, p < 0.001$]; 刺激条件对回声阈值没有显著影响 [$F(2, 22) = 2.287, p = 0.125$]; 经过 Greenhouse-Geisser 校正后, 刺激类型对回声阈值有显著影响 [$F(1.412, 15.533) = 32.787, p < 0.001$], 刺激重放方式和刺激类型对回声阈值有显著的交互作用 [$F(2.884, 31.724) = 8.966, p < 0.001$]; 刺激条件和刺激类型对回声阈值有显著的交互作用 [$F(3.541, 38.956) = 6.471, p = 0.001$]。此外, 其他交互作用均不显著。

进一步地, 采用 Bonferroni 校正的两两比较的事后比较, 结果显示: 耳机 ITD 法和扬声器重放法在回声阈值上有显著差异 ($p = 0.002$), 耳机 VAS 法和扬声器重放法在回声阈值上没有显著差异 ($p = 1.000$), 而耳机 ITD 法和耳机 VAS 法在回声阈值上有显著差异 ($p = 0.002$)。这一结果综合表明, 耳机 ITD 法和扬声器重放法对回声阈值的影响在一定程度上存在较大差异, 耳机 ITD 法这种只控制单一定位因素的重放方式并不能有效地模拟自由场的优先效应情形, Brown 也曾提出这一观点^[33]。相比之下, 耳机 VAS 法更能有效地模拟自由场的优先效应情形。值得注意的是, 虽然本研究中耳机 VAS 法采用的是非个性化的 HRTF, 但其回声阈值与包含个性化 HRTF 的扬声器重放法的结果并没有显著差异。当然, HRTF 的个性化是否会对回声阈值造成显著影响, 仍然需要更详尽的研究去证实。

此外, 刺激类型和刺激条件对回声阈值有显著的交互作用。进一步的简单效应分析表明, 对于噪声脉冲, 建立条件的回声阈值比基准条件显著偏高 ($p = 0.002$), 破坏条件的回声阈值比建立条件显著偏低 ($p = 0.001$), 基准条件和破坏条件的回声阈值之间没有显著差异 ($p = 0.211$), 对于三种长时刺激, 三种

刺激条件的回声阈值两两之间均不存在显著差异。这一结果表明, 随着刺激条件的变化, 噪声脉冲刺激和三种长时刺激的动态回声阈值变化规律是存在显著差异的, 其中噪声脉冲的回声阈值先显著增加后显著降低, 三种长时刺激的回声阈值均未出现显著变化, 而这些阈值的变化趋势在三种刺激重放方式中是一致的。

针对本实验中展现的动态回声阈值的变化规律, 进行相关分析和讨论。

首先, 噪声脉冲刺激的结果表明, 当噪声脉冲重复播放后, 大多数被试都经历了较为明显的优先效应建立现象, 逐渐感知不到独立的回声, 同时, 在突然交换领先声和滞后声的空间位置后, 这种建立现象被破坏, 被试再次感知到独立的回声。这一阈值的变化趋势与以往分别采用耳机 VAS 法^[16]和扬声器重放法^[6-7]的结果一致, 而与以往采用耳机 ITD 法^[23-24]的结果存在一些差异。在 Krumbholz 等^[23]以及 Brown 等^[24]的研究中, 基于耳机 ITD 法下破坏条件的平均回声阈值较建立条件并未出现显著的降低, 而本文中则出现了显著降低, 可能的原因在于实验条件的不同, 例如: 在 Krumbholz 等的研究中, 耳机 ITD 法采用了两种具有不同偏侧程度的 ITD 值 ($\pm 300 \mu\text{s}$ 和 $\pm 600 \mu\text{s}$), 结果发现, 只有 ITD 值为 $\pm 600 \mu\text{s}$ 时出现了破坏条件阈值显著降低的现象。而本实验中采用的 ITD 值为 $\pm 436 \mu\text{s}$, 处于其实验采用的两个数值之间, 因此在破坏条件下也可能出现较为明显的回声阈值下降。

其次, 对于三种长时刺激, 三种重放方式下动态回声阈值均未出现显著变化。Djelani 等^[16]在研究中提到, 回声抑制主要取决于刺激的时长。他们的实验观察到, 语音或噪声刺激的回声阈值一直增加直到相应的刺激时长达到 250 ms, 此时可能已经开始出现优先效应的建立过程。而在本实验中, 采用的三种长时刺激都显著大于 250 ms, 因而其优先效应建立过程在单对刺激播放时 (基准条件) 就已经发生了, 此时听者已经获得了关于声源足够多的信息, 刺激的重复播放 (建立条件) 对其不再有明显的影响。此外, 本实验中长时刺激在破坏条件下也未出现阈值的显著变化。一些研究认为^[6,16], 当领先声-滞后声的空间位置发生突然变化时, 听者对回声的感知变化是暂时性的, 而这种暂时性的感知变化可以视为优先效应破坏过程和重新建立过程的一种指示。图 4 为回声阈值随时间变化而出现的动态变化过程, 在刺激的重放过程中, 回声阈值随时间动态变化, 建立过程回声阈值会显著增加, 建立完成后阈值

将保持稳定,一旦刺激特性的突然变化(如交换领先声和滞后声的空间位置),回声阈值将显著降低。但是由于优先效应的建立和破坏过程经历的时间非常短,这就导致破坏过程发生后,短时间内又出现了重新建立并迅速完成建立过程(如文献[16]提到的 250 ms),如图 4 中从 A 经过 B 再到 C 的过程,这一过程相对来说较为短暂。而本实验要求被试需从整体上对长时刺激进行回声感知判别,这一感知和决策过程需要一定的时间积累,如图 4 中阴影区域。因此,对于长时刺激,被试对破坏条件下(交换领先声-滞后声空间位置)回声感知的判别很大程度上依然发生在建立完成时,从而使得回声阈值未降低。

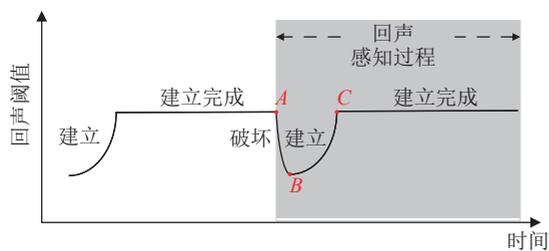


图 4 回声阈值的动态变化过程(阴影区域表示破坏条件发生后听者对长时刺激的回声阈值感知过程)

此外,关于刺激重放方式对动态回声阈值变化规律的影响,实验结果显示,三种重放方式的动态回声阈值变化规律是一致的。其中,噪声脉冲刺激的回声阈值都是先增加后降低的过程,三种长时刺激均未出现显著的阈值变化,而对于噪声脉冲刺激,本研究的动态回声阈值变化规律(先增加后降低)与以往基于不同重放方式的研究结果基本一致。综合而言,本研究的结果一定程度上说明,不同重放方式所存在的定位因素差异(如 ITD、ILD)和空间感知差异并不影响优先效应的回声阈值变化规律,这也是本实验发现的主要物理机理。实际上,虽然耳机 VAS 法和扬声器重放法重现了比耳机 ITD 法更多的定位因素,但是正如文献[25]以及本文 2.1 节所提到的,这些重放方式的空间定位因素差异造成的听觉感知差异,如领先声-滞后声空间分离感知程度、单声源听觉事件宽度,主要体现在长时刺激上,而长时刺激的时长较长,回声阈值主要由单对领先声-滞后声决定,从而导致刺激的重复播放和动态变化对其回声阈值不产生显著的影响。

3 结论

本文主要基于经典的领先声-滞后声实验范式

(即在没有额外反射影响的自由空间环境中),研究了三种刺激重放方式(耳机 ITD 法、耳机 VAS 法和扬声器重放法)对三种刺激条件(基准、建立和破坏条件)下优先效应回声阈值的影响差异,实验结果主要适用于自由空间。主要结论如下:

(1) 基于耳机 VAS 法和扬声器重放法的回声阈值不存在显著差异,而这两种重放方式的回声阈值都与耳机 ITD 法存在显著差异,主要体现在长时刺激上。原因与耳机 ITD 法的定位因素不一致性相关,这些结果表明,在进行优先效应回声阈值的测量研究中,耳机 VAS 法比耳机 ITD 法更适合模拟自由场的情形,并且一定程度上能够代替扬声器重放法。

(2) 三种刺激重放方式的动态回声阈值变化规律是一致的。随着刺激条件的变化,噪声脉冲刺激的回声阈值变化都是先增加后降低,而三种长时刺激的回声阈值均未出现显著的变化。这表明,三种重放方式的定位因素差异对动态回声阈值的变化规律没有显著影响。

(3) 长时刺激未出现动态回声阈值显著变化的原因与刺激的时长、优先效应动态过程的暂时性以及测听标准密切相关。

参 考 文 献

- Wallach H, Newman E B, Rosenzweig M R. A precedence effect in sound localization. *J. Acoust. Soc. Am.*, 1949; **21**(4S): 468
- Blauert J. Spatial hearing: The psychophysics of human sound localization. Revised edition. Cambridge, MA: MIT Press, 1997
- 葛鑫颖, 周沫, 王彦君, 等. 听觉优先效应在下丘的作用机制研究进展. *中华耳科学杂志*, 2022; **20**(6): 961-965
- Wühle T, Merchel S, Altinsoy M E. The precedence effect in scenarios with projected sound. *J. Audio Eng. Soc.*, 2019; **67**(3): 92-100
- Litovsky R Y, Colburn H S, Yost W A, et al. The precedence effect. *J. Acoust. Soc. Am.*, 1999; **106**(4): 1633-1654
- Clifton R K. Breakdown of echo suppression in the precedence effect. *J. Acoust. Soc. Am.*, 1987; **82**(5): 1834-1835
- Freyman R L, Clifton R K, Litovsky R Y. Dynamic processes in the precedence effect. *J. Acoust. Soc. Am.*, 1991; **90**(2): 874-884
- Yang X, Grantham D W. Echo suppression and discrimination suppression aspects of the precedence effect. *Percept. Psychophys.*, 1997; **59**: 1108-1117
- Blauert J, Braasch J. Acoustic communication: The precedence effect. Forum Acousticum, D. R. Institute, Budapest, 2005: 992-990
- Miller S D, Litovsky R Y, Kluender K R. Predicting echo thresholds from speech onset characteristics. *J. Acoust. Soc. Am.*, 2009; **125**(4): EL134-EL140
- Agaveva M Y, Al'tman Y A. Echo thresholds measured in the vertical and horizontal planes. *Hum. Physiol.*, 2008; **34**: 678-684
- Yost W A, Soderquist D R. The precedence effect: Revisited. *J. Acoust. Soc. Am.*, 1984; **76**(5): 1377-1383

- 13 Litovsky R Y, Shinn-Cunningham B G. Investigation of the relationship among three common measures of precedence: Fusion, localization dominance, and discrimination suppression. *J. Acoust. Soc. Am.*, 2001; **109**(1): 346–358
- 14 Braasch J, Savitala H, Blauert J. Modeling the precedence effect for percussive sounds with different attack transients. 19th International Congress on Acoustic, ICA, Madrid, Spain, 2007: 2–7
- 15 Dizon R M, Colburn H S. The influence of spectral, temporal, and interaural stimulus variations on the precedence effect. *J. Acoust. Soc. Am.*, 2006; **119**(5): 2947–2964
- 16 Djelani T, Blauert J. Investigations into the build-up and breakdown of the precedence effect. *Acta Acust. united Acust.*, 2001; **87**(2): 253–261
- 17 Robinson P W, Walther A, Faller C, *et al.* Echo thresholds for reflections from acoustically diffusive architectural surfaces. *J. Acoust. Soc. Am.*, 2013; **134**(4): 2755–2764
- 18 Wendt F, Höldrich R. Precedence effect for specular and diffuse reflections. *Acta Acust.*, 2021; **5**: 1
- 19 Xie B. Head-related transfer function and virtual auditory display. 2nd edition. USA: J. Ross Publishing, 2013
- 20 Aoki S, Houtgast T. Onset dominance in the precedence effect with loudspeaker reproduction in a free field. *J. Acoust. Soc. Jpn.*, 1994; **15**(3): 197–199
- 21 Seeber B U, Hafter E R. Failure of the precedence effect with a noise-band vocoder. *J. Acoust. Soc. Am.*, 2011; **129**(3): 1509–1521
- 22 Guzman S J. Buildup of the precedence effect: Fusion, masking, localization and lateralization of reflections. Doctoral dissertation, Chicago: Loyola University, 1999: 6–7
- 23 Krumbholz K, Nobbe A. Buildup and breakdown of echo suppression for stimuli presented over headphones—the effects of interaural time and level differences. *J. Acoust. Soc. Am.*, 2002; **112**(2): 654–663
- 24 Brown A D, Stecker G C. The precedence effect: Fusion and lateralization measures for headphone stimuli lateralized by interaural time and level differences. *J. Acoust. Soc. Am.*, 2013; **133**(5): 2883–2898
- 25 Cao T, Yu G, Rao D. Effect of headphone-based stimuli presentation method on echo threshold. *Appl. Acoust.*, 2024; **219**: 109942
- 26 Hansen V, Munch G. Making recordings for simulation tests in the Archimedes project. *J. Audio Engin. Soc.*, 1991; **39**(10): 768–774
- 27 Ziegelwanger H, Kreuzer W, Majdak P. Mesh2HRTF: Open-source software package for the numerical calculation of head-related transfer functions. 22nd International Congress on Sound and Vibration, IIAV, Florence, Italy, 2015: 388–395
- 28 Wu R, Yu G, So R H. Key anthropometric parameters of pinna correlate with individualized head-related transfer functions. INTER-NOISE and NOISE-CON Congress and Conference Proceedings, Institute of Noise Control Engineering, Hong Kong, China, 2017: 4023–4028
- 29 孟子厚. 音质主观评价的实验心理学方法. 北京: 国防工业出版社, 2008
- 30 金瑜. 心理测量. 第2版. 上海: 华东师范大学出版社, 2005
- 31 Moore B C. An introduction to the psychology of hearing. 6th Edition. Boston: Brill, 2013
- 32 Kline R B. Principles and practice of structural equation modeling. 4th edition. New York: Guilford Publications, 2015
- 33 Brown A D, Stecker G C, Tollin D J. The precedence effect in sound localization. *J. Assoc. Res. Otolaryngol.*, 2015; **16**: 1–28