arXiv:2303.02523v2 [eess.AS] 3 May 2023

大众广泛采用辅助聆听技术的要求

*T*homas B. Kaufmann Mehdi Foroogozar JulieLiss Visar Berisha

美国亚利桑那州立大学言语与听力
科学系，亚利桑那州坦佩

摘要

辅助聆听系统（ALS）可大幅提高语音可懂度并减少聆听努力。几乎可以肯定，不仅听力损失者，几乎所有人都能从辅助聆听系统在几乎任何聆听场景中所提供的信噪比提升中获益。然而，除了重度至极重度听力损失者，其他人几乎不会使用辅助聆听系统。迄今为止，尚未对这种低采用率的原因进行系统调查。

作者推测，辅助聆听技术普及率低的原因包括：(1) 无法使用个人拥有的接收设备；(2) 缺乏高保真立体声；(3) 接收设备无法提供无堵塞聆听体验；(4) 对齐延迟造成的失真；以及 (5) 无法自动连接到可用的辅助聆听音频信号。

我们针对这些问题分别提出解决方案，旨在为辅助聆听技术在公众中的广泛普及铺平道路。

索引词——辅助聆听、听力辅助技术、对齐延迟、可懂度、聆听努力

1. 引言

拥有超过 100 个座位的大型礼堂和演出场所通常配备扩声系统，为观众中的每个座位提供足够的声压级，使听者能够清晰地听见音乐，并相对轻松地理解语音 [1]。然而，在大型礼堂、教室、剧院或体育场等嘈杂且混响的空间中，语音理解会变得极具挑战性，尤其是对于存在听力损失的人群，包括助听器使用者和人工耳蜗使用者 [2, 3]。为了改善此类环境下的聆听理解，辅助聆听系统（辅助聆听系统）应运而生。虽然辅助聆听系统经常被听力损失者使用，但公众通常不会使用它们。

辅助聆听系统（ALS）可在场馆的音响系统与听众佩戴的接收设备之间建立无线连接，绕过声学空间。研究表明，辅助聆听系统（ALS）能够将听众的信噪比提高多达 15 至 20 dB，效果远胜于降噪算法和定向麦克风 [4, 5]。符合许多国际无障碍法规的可用音频传输方式包括射频（RF）、红外线（IR）和磁感应 [6]。射频（RF）和红外系统要求每位听众向场馆借用接收设备，而音频感应环路系统（AFILS）则允许配备电话线圈的助听器和人工耳蜗用户直接连接场馆的音频广播 [6, 7]。2022年7月发布的蓝牙低功耗音频标准及其“Auracast”公共广播功能近期问世，有望实现与感应环路系统相同的直连助听器功能。Auracast 还将兼容人工耳蜗和个人聆听设备，例如头戴式耳机或耳塞式耳机。

2. 问题

辅助聆听系统提供的更高信噪比，很可能对几乎所有人在几乎任何聆听环境中都有益。然而，实际上，辅助聆听系统通常只供听力损失达到重度至极重度的人士使用。大众未能广泛采用这类系统的原因尚未得到充分了解；迄今为止，也没有人对此进行过系统性调查，因为听力辅助技术领域长期以来经费不足、研究匮乏。

听力损失与多种身体和心理健康风险相关，包括社会隔离、抑郁症、生活质量降低以及痴呆风险显著升高 [8, 9, 10]。平均而言，从符合助听器适用资格到实际开始使用助听器，个体需要等待 9 年；而尽早使用助听器可能有助于减轻听力损失带来的一些风险 [11, 12]。研究表明，通过助听设备中的电话线圈使用辅助聆听技术的人，对设备整体的满意度更高 [13]。因此，听力损失者和整个社会都将受益于辅助聆听技术的广泛普及

。

我们推测，以下五个原因可能会阻碍辅助聆听技术在公众中的广泛采用，并针对每个原因提出潜在的解决方案。

2.1. 个人拥有的接收设备

听者可能更愿意使用个人拥有的接收设备，而不是向场馆借用接收设备。这种偏好可能源于便利性、舒适度和卫生方面的考虑。由于射频系统和红外系统使用的频段各不相同，听者拥有兼容接收设备的情况可能并不多见。然而，音频感应环路系统采用全球通用的标准向听者传输音频信号。因此，目前市面上绝大多数助听器型号和所有人工耳蜗都兼容这项技术。Auracast 也将让听者能够使用个人拥有的接收设备收听。

2.2. 高保真立体声

听者可能更愿意收听立体声，而非单声道；尤其是使用耳塞式耳机、头戴式耳机、助听器或人工耳蜗收听立体声，能够带来完整的空间化听觉体验，Starkey 的“Virtual Barber Shop”体验便展示了这一点，而且在网上很容易找到。听者也可能更偏好高保真听觉体验，但由于音频带宽较低、失真、干扰、延迟和丢包等问题，各种辅助聆听技术并不总能提供这种体验。

目前，任何符合法规要求的辅助聆听技术都无法提供高保真立体声。不过，Auracast 将支持高保真立体声。

2.3. 不遮挡聆听体验

辅助聆听系统主要用于有许多其他听者在场的群体场合。听者在聆听所播放的音频信号时，可能希望保持情境感知，并能够与附近的人交谈。这就需要不受遮挡的聆听体验，而场馆自有的头戴式耳机接收设备通常无法提供这种体验。

虽然听者可以自备开放式头戴耳机或骨传导耳机，但这种情况可能并不常见。佩戴助听器的人可能使用开放式助听器，从而获得无堵塞聆听体验。对于任何佩戴助听器或人工耳蜗的听者，都可以通过一种程序实现类似无堵塞聆听体验的效果：该程序将辅助聆听音频信号与设备麦克风接收到的声学音频信号混合。这类似于任何现代头戴式耳机和耳塞式耳机所提供的“通透模式”。例如，该模式允许用户聆听从智能手机传输的音频信号，同时将设备外部麦克风接收到的信号馈送至输出换能器，从而营造出无堵塞聆听体验的错觉。只有一只助听设备、另一只耳朵健康的人，也可以获得部分无堵塞聆听体验。

Auracast 将能够利用兼容接收设备的通透模式，提供不受遮挡的聆听体验和高保真立体声。

2.4. 对齐延迟造成的失真

通过场馆辅助聆听系统传输的音频信号通常以光速（299,792,458 m/s）传播。场馆扬声器系统产生的声学信号以声速传播；在 20 °C 的干燥空气中，声速约为 343 m/s，即 1.125 ft/ms。因此，听者几乎瞬间就能听到辅助聆听音频信号，而听到声学音频信号时则会有所延迟。延迟的大小取决于听者与场馆前方扬声器之间的距离；这些扬声器通常安装在与舞台前沿对齐的位置。例如，如果听者坐在距离舞台约 200 ft / 60 m 的地方，扬声器发出的声学信号会在其产生约 180 ms 后到达听者的耳朵。如果听者感知到的只有声学信号，这种延迟并不重要。然而，如果听者同时感知到辅助聆听音频信号和声学信号，且其中一个信号相对于另一个信号有所延迟，听者就会感知到失真信号。

对于大约 5 ms 或更短的极短延迟时间，由于两路信号之间的梳状滤波效应，这种失真会被感知为声音的“染色” [14]。对于长于 5 ms 但短于 30 ms 的延迟时间，这种失真往往会被感知为混响，而超过 30 ms 的延迟时间则往往会被感知为回声 [15, 16]。此类失真会降低语音可懂度并增加听者的聆听负担 [17]。大量实验发现，大多数听者认为所感知到的回声效应达到无法忍受程度的阈值约为 30 ms，而听力损失者往往比具有正常听力的个体能够容忍稍长的延迟时间 [18]。近期研究还表明，无论是否存在听力损失，听者通常都更偏好较短的延迟时间 [17]。

在辅助聆听场景中，对于如上所述听觉体验未受阻隔的人而言，感知到的失真效应最为明显，例如使用开放式助听器的结果。假设在当今市场上，超过 50% 的助听器采用开放式设计 [19]。鉴于在任何给定场馆中，绝大多数辅助聆听系统用户都能感知到这种延迟

效应，且最大可容忍对齐延迟约为 30 毫秒，因此，舞台边缘与距舞台最远座位之间的距离超过 35 英尺 / 10 米的任何场馆，都应配备能够考虑并补偿这种延迟效应的辅助聆听解决方案。

目前可用的辅助聆听技术都无法准确计算并补偿对齐延迟效应。如果希望获得最佳聆听体验，任何类型的现有或未来辅助聆听技术都必须具备对齐延迟补偿功能。下文将讨论相关解决方案。

2.5. 自动连接

听者可能希望其辅助聆听接收设备能够自动连接到可用的辅助聆听音频流。对于大部分清醒时间都佩戴助听设备、且经常转换聆听环境的听者而言，这种需求可能尤为强烈。例如，这些听者可能使用助听器、人工耳蜗或具有通透模式的耳机。

目前，连接射频或红外系统需要听者向场馆借用接收设备。连接感应环路系统则需要用户按下助听设备或遥控器上的按钮，或者在智能手机上选择感应环路程序。

使用 Auracast 时，个人可能会接触到多个可用的广播音频流，因此会产生选择连接哪个音频流的困惑，听众需要像连接公共 Wi-Fi 网络一样，手动连接到某个音频流。建议的 Auracast 音频流连接方式包括：在智能手机或远程控制设备上选择音频流、将智能手机碰触近场通信（NFC）设备，或扫描二维码。这些方式都无法提供与连接感应环路系统同等的无障碍体验。对于因任何原因而无法操作智能手机或远程控制设备的人来说，这可能会构成障碍。

目前可用的辅助聆听技术都无法自动连接到辅助聆听音频信号。下文将讨论一种拟议的解决方案。

3. 解决方案

目前可用的辅助聆听技术存在各种不足，可能会阻碍其被公众大规模采用，也可能妨碍助听器得到更广泛的应用。在上述讨论的技术中，Auracast 最有希望成为一种全球通用标准，能够解决上文讨论的所有问题。然而，迄今为止，对补偿对齐延迟的要求以及自动连接问题尚未得到解决。我们提出了针对这两个问题的潜在解决方案。

3.1. 对齐延迟造成的失真

由于声学音频信号的传播速度比辅助聆听音频信号慢，因此需要补偿由此产生的对齐延迟。对此，可以采用多种解决方法。

3.1.1. 多重广播

一种可行的解决方案是为场馆配备多个辅助聆听发射器。每个发射器都必须发送带有额外对齐延迟的音频信号，并且需要根据特定观众区域调整对齐延迟。例如，对于一个从舞台边缘到最远座位的最大距离为 200 英尺 / 60 米的场馆，如拥有 2,000 多个座位的大型表演艺术场馆，需要 3 个辅助聆听发射器，以确保每个座位的对齐延迟都低于 ±30 毫秒。随着场馆规模扩大和／或对最大可容忍延迟时间的要求更加严格，所需发射器的数量也会增加。

目前，复杂的感应环路系统安装采用的就是这种方法。多个感应环路放大器由数字信号处理器馈送，处理器会为安装了天线导线阵列的各个观众区域添加适当的对齐延迟。不过，为各观众区域提供的对齐延迟只是近似值，并未根据每位听者的位置精确调整。由于感应环路天线采用近场天线设计，听者只需改变所处的物理位置即可“选择”所需信号。然而，对于射频和红外系统，听者或场馆工作人员必须根据听者的座位位置选择所需频道。例如，在设有站席的场馆以及户外音乐活动中，这种解决方案还会带来额外的挑战。

当前的蓝牙低功耗音频规范定义了一个名为“呈现延迟”的参数，由发射器（广播源）通过扩展广播数据包进行广播，同时还会广播描述实际音频流参数的元数据，包括编解码器设置和信道信息。呈现延迟参数的主要用途是同步多个接收设备（广播接收端），但也可用于为不同观众区域提供适当的对齐延迟。

然而，要让广播接收端实现呈现延迟，广播接收端必须能够缓存音频数据，从接收到数据起一直到将其播放给听众为止。当前版本的蓝牙低功耗音频规范并未要求支持超过 40 ms 的呈现延迟。因此，许多兼容 Auracast 的接收设备可能无法

提供足够的缓存时间，以便在大型场馆中使接收到的辅助聆听音频信号与声学音频信号实现时间同步。

此外，在现行规范下使用 Auracast 配置多个广播流，可能会导致广播大量相同的音频流，从而受到空中传输时间和频谱占用的限制。在蓝牙低功耗音频中，音频平面与控制平面相互分离，这意味着广播音频流仅包含原始音频数据包，而命令与控制参数则通过广播数据包单独传输。因此，可以考虑仅提供一个广播音频流，同时提供多个扩展广播数据包流；每个数据包流包含一个唯一的呈现延迟值，并且都指向同一个广播音频流。然而，现行蓝牙低功耗音频规范不允许采用这种配置。现行蓝牙低功耗音频规范也未提供可针对每位听众单独配置本地对齐延迟的参数。

3.1.2. 接收设备上的可变对齐延迟

确保辅助聆听音频信号与声学音频信号在时间上对齐的另一种解决方案，是采用带有“延迟补偿”功能的辅助聆听接收器。该功能会在接收器从发射器接收到辅助聆听音频信号后、将音频信号呈现给听众前，添加可变时长的对齐延迟。对齐延迟的时长需要根据每位听众与舞台或最远扬声器阵列之间的具体距离进行定制，以确保实现完全或近乎完全的时间对齐，从而为所有观众带来尽可能出色、毫无妥协的聆听体验。

3.1.3. 超声波发射器

第三种自动补偿实际对齐延迟的方案，可能是采用超声波载波的辅助聆听系统。超声波载波的传播速度与声速相同，就像声学音频信号一样。因此，辅助聆听音频信号和声学音频信号会同时到达听众的耳朵。目前市售的MEMS麦克风可以接收高达 100 kHz 甚至更高频率的超声波信号，带宽充足，能够传输高保真的立体声音频信号。许多助听器和人工耳蜗已经配备了此类麦克风。

也可以通过这种超声波载波传输数据。例如，可以利用这一功能将音频流的实时转录内容发送给听众，以便在智能手机或平板电脑上显示。

3.2. 自动连接

为满足自动连接至可用辅助聆听音频流的要求，需要作出一些假设。可以假设，听众可能会同时接收到多个广播流。还可以假设，当提供辅助聆听音频流时，至少有一个广播流与听众的声学环境相关。最后可以假设，与听众的声学环境相关性最高的广播流，就是听众希望连接的广播流。

根据这些假设，可通过以下方式满足自动连接的要求：(1) 扫描环境以查找可用的广播音频流；(2) 将每条音频流与听众的声学环境进行比较；(3) 自动连接到与听众的声学环境最相似的音频流。听众也可以通过按钮、语音控制、智能手机或遥控器等方式取消或覆盖此选择。

4. 讨论

要让辅助聆听技术得到广泛应用，所提供的聆听体验需要达到无听力损失的音响发烧友的要求，并带来无缝、流畅的用户体验。我们提出了多种解决方案，以弥补当前辅助聆听技术的不足。

为了确保在任何听众就座距离超过 35 英尺 / 10 米的场馆中，各类辅助聆听系统都能成功使用，辅助聆听系统需要配备一种方法，以补偿对齐延迟。这种延迟是由于辅助聆听音频信号的传播速度比场馆扬声器系统发出的声学音频信号慢而产生的。解决方法包括提供多个广播流、让接收设备能够添加可变时长的对齐延迟，或为辅助聆听音频信号使用超声波载波。

蓝牙 Auracast 有望让辅助聆听技术更加普及、易于使用且价格实惠，并可能因此提高助听器的使用率。然而，当前蓝牙低功耗音频规范中唯一可用于在时间上对齐辅助聆听音频信号与声学音频信号的方法，可能导致大型场馆受到空中传输时间和频谱占用方面的限制。蓝牙低功耗音频规范应进行修订，以便：(1) 允许多个扩展广播数据包流指向同一个广播音频流；(2) 要求广播接收端支持至少 500 ms 的呈现延迟值；(3) 提供一个参数，使远程控制设备（控制器／广播助手）能够向广播接收端传达“本地对齐延迟”参数；(4) 要求广播接收端提供自动连接到与听众声学环境高度相关的广播流的选项。

5. 参考文献

[1] Anna K. Nabelek 和 Amy M. Donahue，《礼堂中扩音系统的比较》，《美国声学学会期刊》，第 79 卷，第 6 期，第 2078–2082 页，1986年 6 月。

[2] Jin Sook Kim 和 Chun Hyeok Kim，《助听器辅助聆听设备与数字无线技术综述》，《韩国听力学杂志》，第 18 卷第 3 期，第 105 页，2014年。

[3] Michael F. Dorman 和 Rene H. Gifford，《人工耳蜗佩戴者在复杂聆听环境中的语音理解》，《言语、语言与听力研究杂志》，第 60 卷第 10 期，第 3019–3026 页，2017年10月。

[4] D. B. Hawkins，《使用助听器和调频系统的轻度至中度听力受损儿童在噪声中的语音识别比较》，《言语与听力障碍杂志》，第 49 卷第 4 期，第 409–418 页，1984年。

[5] M. Samantha Lewis、Carl C. Crandell、Michael Valente 和 Jane Enrietto Horn，《噪声中的言语感知：定向麦克风与调频（FM）系统的比较》，《美国听力学会杂志》，第15卷第6期，第426–439页，2004年。

[6]Thomas Kaufmann、Juliette Sterkens 和 John M Woodgate，《助听环路：首选的辅助聆听技术》，《音频工程学会杂志》，第63卷第4期，第298–302页，2015年。

[7]Timothy H. Riehle、Robert J. Knuesel、Patrick A. Lichter 和 Dorin Panescu，《三轴电话线圈助听器：改善与公共感应环路的连接》，载于2015年第37届IEEE生物医学工程学会国际年会（EMBC）。2015年8月，第97–100页，IEEE。

[8]Roaa Zayed Atef、Bernhard Michalowsky、Anika Raedke、Moritz Platen、Wiebke Mohr、Franka M¨uhlichen、Jochen Ren´e Thyrian 和 Wolfgang Hoffmann，《助听器对认知障碍和痴呆患者认知衰退、抑郁症及生活质量进展的影响》，《阿尔茨海默病杂志》，2023年。

[9] Alison R. Huang、Kening Jiang、Frank R. Lin、Jen- nifer A. Deal 和 Nicholas S. Reed，《美国老年人的听力损失与痴呆患病率》，《美国医学会杂志》，第329卷第2期，第171–173页，2023年。

[10] Zhifei Wen、Sihan Peng、Liping Yang、Hongyan Wang、Xinqi Liao、Qingfang Liang 和 Xiangeng Zhang，《老年人社会隔离相关因素：系统综述与荟萃分析》，美国医疗主管协会期刊，第24卷第3期，第322–330.e6页，2023年。[11]Megan Knoetze、Vinaya Manchaiah、Bopane Moth- emela 和 De Wet Swanepoel，《影响成年人听力求助行为和助听器使用率的因素：过去十年的系统综述》，《听力趋势》，第27卷，第23312165231157255页，2023年。

[12] Annie N. Simpson、Lois J. Matthews、Christy Cassarly 和 Judy R. Dubno，《从符合助听器适用资格到采纳助听器的时间：一项纵向队列研究》，《耳与听力》，第40卷，第3期，第468–476页，2019。

[13] Erin M. Picou，《MarkeTrak 2022 调查中的助听器获益和满意度结果：功能特性与听力保健专业人员的重要性》，《听力研讨会》，第43卷，第4期，第301–316页，2022，出版商：Thieme Medical Publishers, Inc.

[14] Georg Stiefenhofer，《开放式助听器中的助听器延迟——正常听力者和听力受损者的音色-音高辨别》，《国际听力学杂志》，第 1–9 页，2022 年 3 月。

[15] Ruth Y. Litovsky、H. Steven Colburn、William A. Yost 和 Sandra J. Guzman，《优先效应》，《美国声学学会期刊》，第 106 卷，第 4 期，第 1633、8 页，1999 年。

[16] Lars Bramsløw，《开放式验配中的首选信号路径延迟和高通截止频率》，《国际听力学杂志》，第 49 卷，第 9 期，第 634–644 页，2010 年 9 月。

[17] Dina Lelic、Georg Stiefenhofer、Ellen Lundorff 和 Tobias Neher，《开放式助听器中的助听器延迟：正常听力和听力受损听者偏好的音质》，《JASA 快报》，第 2 卷，第 10 期，文章编号 104803，2022 年 10 月。

[18] Tobias Goehring、Josie L. Chapman、Stefan Bleeck 和 Jessica J. M. Monaghan*，《言语产生与感知的可耐受延迟：听力能力和助听器使用经验的影响》，《国际听力学杂志》，第 57 卷第 1 期，第 61–68 页，2018 年 1 月。

[19] Matthias Froehlich 和 Veronika Littmann，《关闭开放式验配：优化言语理解的有效方法》，《Hearing Review》，第 26 卷第 4 期，第 16–20 页，2019 年 4 月。