

调谐的角落化：音乐在被“听即校正”制度化之后被逼到了哪一层

作者 高于涵 · SDE 学员 · 约 2.5 万字 · 发表于2026年7月31日 · 加固版

摘要

本文提出并论证一个不可被既有音乐学概念还原的核心判断：音乐，在它尚未被任何一种评价体系捕获的原初层面，是两个人（或更多人）之间一次活的“共在调谐”——一种动态的、微秒级的、身体间的相互调整，它旨在共享一个正在发生的当下，而非准确再现一个既定对象。然而，一部音乐文明史，可被重述为“听即校正”不断制度化、从而系统性地压缩调谐发生条件的历史：从口传时代基于记忆存活的软约束，到记谱法与标准音高将“原作”固化、将每一次表演降格为“副本”的本体论翻转，再到现代评价体系与流媒体算法将“可测量的精准”确立为音乐唯一合法性来源。这套制度的核心操作，不是在排除“错误”，而是在每一个声音发生的瞬间，植入一套“证明它值得存在”的评价程序——它没有杀死调谐，而是将它逼到了特定的时间（深夜）、特定的关系（母婴）、特定的心理状态（“不在乎会不会弹错”）之下才能短暂苏醒。在此视角下，本文重新透视琴房中被内化为身体知觉的校正、民歌声响中被判定为“冗余”而被切除的调谐残片、职业音乐家自我体验中校正装置偶然停机时才浮现的调谐间隙，以及母亲哼唱这一最低现场为何是音乐发生最后的、不可被彻底优化的根据地。本文不否定精确与高效的文明价值，而是将它们从本体论独断还原为工具：服务于调谐，而非替代调谐。结语给出本文的证伪条件——只要一次不携带任何校正意图的共享听，仍然在专业音乐教育中被当作一种严肃的、值得追求的能力来培养，本文的叙事即告破产；同时指出，调谐的重新打开不是简单地“放下校正”即可回归的本源，而是在已被校正深度格式化的耳朵上进行的一次艰难的、需要制度支撑的知觉重构。

关键词：共在调谐；听即校正；音乐发生学；作品与场景；本体论翻转；音乐评价体系

一、导论：在任何一个评价发生之前，音乐是什么

在所有关于“音乐是什么”的回答中，有一个发生现场始终被形形色色的理论绕了过去。这个现场不是一个定义，而是一个极其朴素的、每天都在发生却几乎不被任何音乐学文本郑重对待的场景——但它不是一个先验的“本质原点”，而是本文的发生学视角所照亮的、此前未被当作理论出发点的第一现场。

两个人在同一个空间里。一个人发出了一串声音——哼了一段调子，或者在琴上随手按了几个音。另一个人听见了。在听见的那一秒里，两个人的身体同时发生

了某种几乎不可被测量的微调：发出声音的人，他在声音还没结束的时候，已经从对方肩膀的微动、呼吸的稍滞、眼神的某个闪动里，接到了一种无法被语言命名的回传信号。这个回传信号让他接下来发出的第二个音和第一个音稍稍不同——可能更轻一点，更慢一点，或者忽然转向了一个他自己此前并没有预设的方向。而聆听的那个人，他的身体也不是被动的接收装置：他在声音落在耳膜上的同一刹那，已经自发地、前意识地参与到了这个声音的行进之中——他可能微微屏住了呼吸，可能在某个不自觉的瞬间咬紧了下颌，也可能在某一刹那感到一种想要叹气或是想要笑出来的冲动，只是被他压住了。

如果一定要给这个现场一个名字，本文称之为一次共在调谐。这里“共在”一词，在下文全部论证中特指二人正在共同度过一段不可彼此替代的生命时间——你在这段时间里，我也在这段时间里；如果你不在了，这整段时间的质地就会不一样。“调谐”则指向他们之间的信号往来不是信息的编码与解码，而是一种持续的、微秒级的、不可事先写定脚本的身体间调整——发出者根据接收者的在场状态实时修正声音的走向，接收者根据声音的质地实时修正自己身体的在场状态。这个循环，完成一次，就是一次音乐的发生。

请注意这里的关键：在这个现场里，还没有人问“他唱得准不准”、“这首曲子对不对”、“这段声音好不好”。这些问题在共在调谐发生的那个刹那，根本就不在场。在场的只有一件事：两个人，正在通过声音，共同调节着彼此身体的状态，让这一刻在两具身体之间被共享，而不是被某一方单独地“表达”再由另一方单独地“欣赏”。共享与表达的区别，是整个音乐本体论翻转的轴心。表达有对象：指向一首已有的作品，或者一种已有的情感类别。共享没有对象——或者说，共享的对象就是这个正在进行中的、共同被活过的当下。

但文明在其漫长的制度化进程中，建立起一套截然不同的听觉框架，并最终将其确立为用以评判所有音乐事件的唯一合法尺度：听即校正。这四字概括的，不是某一种具体的听觉技巧，而是一种已经内化到几乎所有受过音乐教育的人身体里的知觉习性：在听到任何一个声音的那一刻，耳朵就会自发地启动一个比对程序——这个音的位置，与规定的位置之间，偏差了多少？这个节奏的节点，与节拍器规定

的节点之间，误差是几毫秒？这个动作的执行效率，是否已经逼近了该生理条件下所能达到的最小冗余？

听即校正一旦启动，共同调谐便立即退场。因为调谐的发生，有一个不可被绕过的前提：两个人都在听“正在发生的这个声音”——而不是在听“这个声音是否符合同一位置上的那个标准范本”。调谐的耳朵，是朝向此刻的、朝向对方的、朝向一个还在流动中、尚未被定性的声音事件。校正的耳朵，是朝向谱面的、朝向标准的、朝向一个在被听到之前就已经被写定、被判定好了的“正确版本”。两种耳朵不可能同时在场，正如一块指南针不可能同时指向北极和南极。

然而，在展开全文之前，必须做一笔关键的自我约束，以防止本文的论证滑入一种“现成本质等着被发现”的路子式的两极对立。本文不是在讲一个“调谐被校正杀死”的故事。校正从来没有、也不可能完全消灭调谐。本文所追溯的，是两者之间动态的此消彼长、相互塑形、在特定制度条件下失衡的过程。校正的耳朵不是从外部闯入、杀死调谐耳朵的入侵者，而是在特定条件——口传记忆的不稳定性催生了对记谱的生存需求，标准化的协作效率催生了对精确性的制度奖励——下，被一步步生成为“默认耳朵”的。在这个过程中，调谐本身也在变：它并没有被彻底逐出人类的听觉经验，而是被逼到了特定的时间（深夜）、特定的关系（母亲与幼儿）、特定的心理状态（“不在乎会不会弹错”）之下才能短暂苏醒。被逼到角落，与被杀死，是两个完全不同的发生学判断。本文的全部论证，都是对“被逼到角落”这个发生学格局的展开——而对于这个格局本身，本文的立场不是哀悼，而是诊断。并不是说曾经有一个完整的、健康的音乐文明，后来被校正制度夺走了一块。调谐从未完整地占有过全部——它从一开始就与记忆筛选、技能传递、社会协调中的各种软校正共生。变化不是“失去”，而是两者之间的空间比例发生了不可逆的倾斜，校正从共生的手段之一，膨胀为裁判一切发生之合法性的独断尺度。

有一种深具影响的想法认为，音乐中最珍贵的东西恰是那些不能被记谱的微妙质地——“韵味”“气息”“手感”——而标准化精确网格是它的反面。这种看法在民歌“原真性”话语（如部分民族音乐学对田野录音“未被污染”状态的怀念）以及某些反技术主义的音乐哲学论述中反复出现。本文的立场与此不同，而且更尖

锐：问题不在于网格的存在，而在于网格的边界被偷偷置换成了音乐本身的边界。那个调谐的现场，从来就没有离开过人类的生活——母亲还是会在深夜对幼儿不成调地哼唱，朋友们还是会在酒后跑调地齐唱老歌，独处的人还是会无意间在琴上摸出几个自己也不知道名字的音。但所有这些现场，在“听即校正”已被制度化为音乐唯一合法性来源的时代，都在发生的那一刻就被一层无形的评价薄膜包裹起来——“我唱得是不是很难听？”“我弹得这是什么东西？”“这算不算音乐？”这层薄膜不是一个从外部罩上去的罩子，而是耳朵被训练成“自动质检员”之后，主体对自己发出的每一个声音所做的前意识审查——这正是校正被内化为知觉习性的发生学形态。调谐还在发生，但它发生的刹那，就失去了被严肃对待的资格。本文要做的，就是把这个资格还给它。

二、从记忆的存活到本体的颠倒：校正如何一步步占据了“默认耳朵”的位置

在展开历史叙事之前，必须做一个方法论声明。本章将音乐文明史分为三个段落：口传时代的软校正、记谱法的本体论翻转、现代评价体系的制度化完成。这不是对实际历史进程的精确描述。真实的历史更可能是一个同时并存的、不均匀的、在不同传统中差速推进的过程——口传时代不是没有校正（回忆偏差导致的自然淘汰本身就是一种软校正），现代也不是没有调谐（本文自己就举了一堆现代条件里残存的调谐现场）。三段式的清晰，是以牺牲发生的混杂性为代价的。它的价值不在于忠实地复现历史，而在于把那条“校正从生存策略逐步升格为本体论独断”的发生学线索，从历史的混杂中抽绎出来。作为理想型，它不惧被反例打脸——一个在口传时代就存在的精确记谱尝试，或者一个在当代学院里刻意保留调谐空间的教学实践，都不会推翻这条线索，只会为它补充必要的复杂性。但本文必须首先承认：它使用的是理想型，而非编年史。

耳朵不是中性的。每一双耳朵在它“听到音乐”之前，已经被一种不可见的框架格式化了。本章追溯这个格式化过程的源头与关键翻转点——校正的耳朵是怎样从一种不得已的生存策略，慢慢变成规定“什么是音乐”的默认立法者的。

在口耳相传的漫长岁月里，校正的逻辑已经嵌入每一次传唱，但它不是作为审美标准，而是作为存活条件。那时没有谱子。一首歌要活下去，必须被记住，被另一个人能大致复现。人类的听觉记忆远不如视觉记忆牢固——一首歌在两次传唱之间偏移太多，就再也无法被识别为同一首歌，它就会从集体记忆里滑脱，永远消失。于是，每一次传唱都默认携带了一个筛选机制：那些偏差没有大到触发“断裂”的旋律骨架，留下来；偏差过大导致整首歌散架的，出局。请注意这个机制的残酷与温和并存。残酷之处在于，它无时无刻不在淘汰——绝大多数人类发出过的声音，根本没有被接住第二遍。温和之处在于，它不要求绝对精确，只要求偏差不越界。只要一首歌还能被下一棒的人认出来，那么每一棒唱出的那个唯一的、带着个人偏移的版本，都是“对”的。因为没有一份作为终极参照的“原作”可供校对。在那个世界，不存在“错误”——只有“被接住了”和“没被接住”。

记谱法的出现，终结了这个只有“接住”没有“错误”的世界。这是音乐史上一次深不见底的本体论翻转。当一段旋律可以被记写为乐谱，它就从一个只存在于活人记忆与喉咙之间的时间性事件，变成了一件可以被翻阅、被保存、被运输的“物”。而谱子一旦成为物，它就在无声无息中完成了一次致命的身份升格：它从对一次发生的追记，变成了这同一首歌后来所有发生的“原作”。演奏不再是一次发生的本身；它变成了一次“还原”——还原得越逼近谱子，越是对；越偏离，越是错。那个在口传时代每一个歌者都可以依法留下一笔偏移、让歌本身在反复传唱里微妙生长的生态系统，被一份不变的纸页永久地锁死了。“歌”的本体，从那个在所有传唱者的身体之间流动的、活的、变形的共同事件，被偷换成了一页安静的、等距的、沉默的符号。

这个本体论翻转最隐蔽、最深远的一笔后果，不是有了一种叫乐谱的技术，而是随之催生了一双全新的耳朵：校正的耳朵。一个有谱子的演唱者，在唱出每一个音时，不再是单纯地“发出这个音”——他在发出的同时，已经在用自己的耳朵对正在发出的音进行实时监听，并与脑海中那个谱子位置上的“正确音”进行比对。误差一旦出现，他会在同一刹那启动一个身体微调：喉头收紧一点，气流加急一点，手指略微挪高一点。这个监听-比对-修正的闭环，不是他刻意“运用”的；它

在他拿到谱子的第一天就被安装进了他的知觉系统，成为他“听自己唱歌”时最基础的运作模式。

这个模式本身极其高效。它是整个西方多声部音乐能在几百年里积累起那么恢弘的复调与交响建筑的技术基底——没有校正的耳朵，一百个人怎么可能在同一个和弦上同时发出误差不超过几音分的声音？但高效的另一面，是一个容易被忽略的知觉代价：当监听-比对-修正成为“听”的主要功能时，“听”就不再是对一个正在发生的、朝向我的声音事件的全身心暴露；它变成了一份质量控制工作。那个正在发出声音的人，他的耳朵在执行质检，而不是在沟通。他听到的，不再是“这个声音在这个空间里、在这群人中正在生起什么样的涟漪”；他听到的，是“刚才那个音高了还是低了”“后面的那个技术难点快到了，我这次能过去吗”。这就是校正对调谐的第一次、也是最根本的一次收编：它把听的朝向，从“此刻正在发生什么”扭转到了“此刻正在发生的事，离那个应该发生的范本，还有多远”。

接下来的一切发展——标准音高体系把“音准”从相对关系变成绝对频率，节拍器把“时间”从音乐的呼吸变成等距刻度，钢琴制造工艺把音色从制琴师手工的不可复现偏差变成工厂标准的统一输出——都在持续加固同一个本体论格局：音乐，就是一个可以被一个精确网格全面捕获的、物化的对象；而好的演奏，就是对这个对象的尽量零误差的向心逼近。在十八世纪到二十世纪的古典音乐建制化进程中，这套格局完成了它的自我封闭。入学考试用音准和节奏作为最硬的淘汰筛，乐团招聘用拉幕盲听以确保只听物理精度，考级制度把从一级到十级铺成一条以错误率持续下降为唯一上升路径的阶梯。在这个阶梯的最顶层，站着一个能在最难的曲目上把误差压到几乎不被耳朵察觉的演奏者。他已经是这个精确体系所能给出的最高褒奖了。

然而，有一个问题，可以从这个闭环的内部撬开一条细缝。当这个演奏者走下舞台，回到自己的公寓，脱掉演出服，坐在沙发上，忽然随便在琴上弹了几个音——也许是一段他自己也说不清来源的旋律碎片，也许只是反复按同一个和弦、每次让手指的力度略微不同、听着那个和弦像心跳一样缓慢地变亮又变暗。在这一刻，他在做什么？他在校正吗？没有。没有谱子可供他校对。他在调谐——不是与

另一个具体的人，而是与他自己的耳朵、与自己被整晚演出高度动员而尚未平复的身体、与这间空房间里正在弥漫开来的那层安静，在调谐。这个场景，不会被写进任何演奏家传记，不会出现在任何音乐学分析里。但它是他最接近“音乐为什么从一开始会在他身体里发生”的那个原点的一刻。而这位演奏家被体制训练得如此彻底，他很可能在弹完几个音之后对自己轻声说出一句话：“我在瞎弹什么”——然后迅速把手从琴键上收回来，起身去倒一杯水。那个刚刚打开了一道缝的调谐现场，就这样被他自己体内那个已经无法关闭的校正审查员，重新关上了。

三、共在调谐——一个新概念的不可还原性及其与既有理论的切割

在继续深入琴房和民歌现场之前，必须先确立本文的核心分析工具。前两章的积淀已经足以逼出这个概念：它的种子，早已埋在口传时代每一次有偏移却被接住的传唱里，埋在记谱法翻转本体论时那个被无声替换的“歌”的本体上。现在是时候给它一个严格的发生学定义，并用它与一系列既有概念完成逐一的切割，以确立它的不可还原性。

共在调谐，指的是：在两个或多个人之间，通过声音的发出与接收，实时地、前意识地、不可事先写定脚本地进行的一种双向身体调整过程。它的目的，不是去准确再现一个既定的声音对象（作品、乐谱、标准），而是让“处在同一段不可切分的时间中”这个事实，作为一次可以被身体感知到的共同体验而发生。调谐的质料，是音高、音色、节奏、力度这些可以被记谱记下的参数，但调谐本身却不能化约为这些参数的任何组合——正如同一个微笑不能被化约为面部肌肉的位移量。调谐发生的时候，校正可以同时在场（比如一位专业歌手的音准控制越好，她越能在调谐中携带更精微的偏离），但调谐的发生不依赖于校正。母亲和婴儿的调谐，不需要任何一方具备音准知识。

这个概念能不能被任何一个已有的、在音乐学或相关学科中已被广泛接受的概念所1:1替换？本文必须在这里做最诚实的自我审问。

它不是默会知识。 迈克尔·波兰尼的默会知识指向的是个体在实践中所掌握的、难以言传的技能认知——比如一个提琴手知道如何在换弓的瞬间用肉眼看不见的手指微调来保持音色匀净。这仍然是在校正逻辑内部的认知：它解释的是演奏者为了更精准地达到某个既定声音目标，所运用的无法被完全外显化的窍门。共在调谐不朝向任何既定目标——它朝向的是正在共享当下的另一具身体的状态。

它不是“互动”。互动是一个过于宽泛的社会学概念，可以指任何两个主体之间的信号往来，包括商业谈判、课堂问答、路怒症司机之间的鸣笛对骂。共在调谐是互动的一个极窄的子类，它的限定条件相当苛刻：必须是通过声音进行的（不是文字、不是眼神）、必须是即时的（不是一方的演奏和另一方延迟的鼓掌）、必须是在事件发生的过程中持续地双向修改的（不是一方的独白和另一方的沉默接收）。把共在调谐叫作互动，就像把“接吻”叫作“口腔接触”——没有说错，但切掉了一切让这个概念非要存在不可的剩余物。更关键的一笔区分在于：互动语言学与对话分析所研究的那种微观时序调整——话轮之间的无缝衔接、听者“嗯”的节点与说着呼吸的同步——服务于“让交际继续下去”的功能性任务。共在调谐的调整，其目的不在于完成一个外在的交际目标，而在于“让此刻被共同度过”这一本体论承诺本身。前者是功能性的，后者是共在性的。共在调谐切开的，正是这条线。

它不是现象学意义上的“主体间性”。舒茨以降的社会学现象学传统，对“面对面情境”中意义的共同建构有极其精细的分析——两个主体如何在相互朝向的意识行为中，共同建构起一个共享的意义世界。这一传统正确地指出了“共同在场”在发生学上的优先性。但它所描述的“相互朝向”，仍然是发生在意识层面的——我意识到你正在意识到我，我根据这一意识调整我的表达。共在调谐与此的区分在于：它发生在比“意识”更底一层的前意识身体层面。在母亲对幼儿哼出第一个不成调的音之前，她没有先“意识到”幼儿的在场状态，然后“决定”调整哼声；她的身体——喉头松紧、气息深浅、音高走向——已经在接到幼儿身体传回的信号（扭动的频率、贴在胸口的重量、呼吸的深浅）的同一刹那，自发地发生了修正。这个修正在被意识捕捉到之前，就已经完成了。现象学的互动描述，始终需要

一个“意识指向”作为互动的发动机；共在调谐把这台发动机再往下卸了一层——卸到了身体之间的直接调频，那个调频不必经过意识的门禁。

它不是共情。共情是心理学概念，指向个体对另一个人情绪状态的认知性理解或情感性共鸣。一个人可以完全共情另一个人——他能精确地理解对方此刻的悲伤——但他坐在观众席上，他的身体没有参与到那个悲伤被声音调谐的过程中去。共在调谐更接近共情的身体发生学前提：在“我理解你的感受”这件事能作为心理事件成立之前，两个人的身体必须先声音的共在里，完成一次彼此的调频。

它不是行动中反映。唐纳德·舍恩在《反映的实践者》中描述的专业实践者在行动中对自身行动的实时监测与调整，仍然是以从业者自身为中心的认知框架：从业者通过与问题情境的对话来修正自己的策略。共在调谐不是从业者单方的自我监测——它是一个双向的循环：我发出声音，我从你的在场状态中接收到一个回传信号，我基于这个回传修改我的下一个声音，你接收到这个被修改的声音，你的在场状态再次发生改变，回传信号再次更新，如此无休止地循环。两个人在共同、而不是交替地调节着一个正在发生的当下。舍恩的模型里有对话，有修正，但没有共在。

它不是克里斯托弗·斯莫尔的“做音乐”。斯莫尔的“Musicking”概念，是将音乐重新理解为一个动词——一个所有参与者在其中都扮演角色的社会性事件。这在方向上与本文最为接近。但斯莫尔的“做音乐”仍然以“音乐”为最终归属：他的工作，是把“音乐”从对象概念扩展为活动概念。而共在调谐的中心，不是去重新定义音乐，而是去指出：有一个比“做音乐”更底层的发生——两个人的身体通过声音彼此调谐——这才是“音乐”这个词在历史上被发明出来用以追认的那个最初的事件。当母亲在深夜的踱步中哼出那不成调的声音时，她没有在“做音乐”；她只是在和她的孩子一起，用声音共度这一小段对谁都无法重来的时间。这件更底层的事，后来产生了那些被命名为音乐的东西，但它的发生并不需要“音乐”这个范畴的许可。

它不是约翰·布莱金的“人的音乐性”。布莱金在《人的音乐性》中以其对南非文达人音乐生活的田野研究，论证了一个核心命题：音乐是人的一种先于文化的

生物学能力，它在所有文化中都首先表现为一种身体间的、参与的、整合性的活动。这与本文的“共在调谐”有深层共振——它们共享同一个基本直觉：音乐的原初发生地，是两具以上的身体共同在场的那个声音事件。但布莱金的论证落点，在于证明这种身体间的共同参与本身就是音乐——它是一种普遍的人类生物能力，不应被西方“作品中心论”的音乐定义所排斥。本文的推进在于指出：即便是这种身体间的共同参与，也有一个更底层的、尚未被分解的发生学前提——那个参与之所以能发生，是因为两具身体在任一音符被“组织成音乐”之前，已经通过声音彼此调谐了在场的状态。布莱金给了参与一个生物学基底——人有与生俱来的音乐性；本文给这个基底补上了它的发生学前序——这种音乐性要能在具体的两具身体之间被激活，需要先完成一次共在调谐。二者不是对立的，而是追问的层面不同：布莱金回答的是“人为什么能音乐”，本文回答的是“一次具体的音乐事件，在最开始的那一刻，是什么让它发生起来的”。

它不是托马斯·图里诺的“参与性音乐”。图里诺在《音乐作为社会生活》中区分了“参与性音乐”和“呈现性音乐”——前者强调所有在场者的共同参与，不区分表演者与观众；后者则以表演者对观众的呈现为核心。这个区分在描述层面与本文的“调谐”与“校正”有显著的重叠：参与性音乐的现场，的确比呈现性音乐更依赖在场者之间的实时互动；而呈现性音乐则天然地向校正逻辑倾斜——因为它预设了一个需要被准确呈现给听众的音乐对象。然而，本文的“共在调谐”与图里诺的“参与性音乐”之间存在一个质的分界。图里诺的区分的单位是音乐事件的社会组织形式——一场印尼甘美兰的集体合奏是参与性的，一场西方交响音乐会是以呈现为主的。而本文的区分的单位是在任何一个音乐事件内部，两具身体之间是否发生着那个微秒级的双向调谐。一场图里诺意义上的“参与性音乐”完全可能充斥着校正——一个甘美兰的新手在整个合奏中都在拼命校对自己的节奏是否与锣的击打时刻对齐，他的耳朵在校正，不在调谐。反过来，一场图里诺意义上的“呈现性音乐”——一场看似单向的独奏会——在某个极其珍贵的瞬间，演奏者与台下某一排某个完全陌生的听众之间，却可能发生一次无声的、以呼吸和身体微动为媒介的实时调谐。共在调谐因此不是一个音乐类型的标签，而是一个在所有音乐类型中都可能发生、但被校正制度系统性压缩了的发生层面。图里诺的参与性/呈现性区

分，切割的是音乐事件的社会学形态；共在调谐切割的，是比社会学形态更底层的身体间动力学。

它不是“夹带”。“夹带”（entrainment）在节奏研究中特指两个或多个有节奏的行为在时间上的相互锁定——心跳与鼓点、步伐与号子、两个提琴手的弓速。它描述了调谐的一种可见的声学后果（同步），但它本身并不区分“为准确再现作品的同步”和“为共享当下的同步”。一支管弦乐队在指挥棒下严格对齐拍点，是夹带；两个朋友一边走路一边唱歌、步伐与歌声在无意识的微调里渐渐靠拢，也是夹带。夹带捕获了调谐的声学足迹——两个有节奏的事件在时间轴上趋于一致——却没有捕获调谐的发生学朝向：那个朝向对方在场状态的意向性。管弦乐队的齐奏，乐手们可以完全不关注彼此的身体在场状态——只盯着指挥棒——而完成高精度的夹带；但母亲对幼儿的哼唱，没有指挥棒，没有节拍器，她唯一追踪的信号，是幼儿身体的在场状态（呼吸、松紧、扭动）。夹带是可测量的声学同步现象，调谐是携带着朝向的、不可被声学参数穷尽的身体间修正。前者可以通过技术训练被稳定地复制，后者的发生每一次都需要两具身体在当下重新接通。

它不是亲职声学调适（IDS）研究中的“母亲语”。发展心理学与声学领域已有大量研究证实：母亲在与婴儿互动时，会自发地调整声学参数——更慢的节奏、更高的音高、更夸张的轮廓、更多的重复——并且这种调整是实时地、根据婴儿的反应动态修正的。这与“共在调谐”描述的现象在表面上有高度重合：即时的、双向的、声学的、以共享当下为目的的身体调整。然而，这恰恰是本文必须做的最关键的一笔切割。IDS研究自始至终站在校正的立场上看母亲的声音：研究者用声学仪器测量母亲语的音高与成人语相比“高了多少”，语速“慢了多少”，频段分布“集中在哪个区域”，然后将这些参数偏离定义为“母亲语”的声学特征。这套研究范式的默认操作，是把母亲的声学调整当成一种可以测量的、可以从其他声学输出的背景中“被识别出来”的客观声学现象来研究。而共在调谐要指出的恰恰是：母亲在哼唱的当下，她从不关心——也完全不知道——自己的音高比某个“标准值”高了多少。她不是在“执行”一套可以参数化的声学策略。她的耳朵和身体，被一根更深、更原始的线索牵引着：婴儿的呼吸有没有变缓？身体有没有更沉

地靠到她的胸口？那种刚喝完奶后的、让怀抱变重了一点的彻底松弛，有没有到来？IDS研究所测量的那些声学参数——音高轮廓的夸张度、元音的延长率、短语之间的停顿时长——在发生学上是调谐过程的物理残留，是副产品，不是调谐本身。IDS研究者测量这些偏离值并建立统计模型，恰恰证明他们已经事先占据了一个校正的位置：他们把母亲与婴儿之间的双向调谐降维为“一种具有特定声学特征的语音输入”，然后研究这种输入与婴儿语言发展之间的因果关联。这个研究不但没有捕获调谐，反而在操作层面完成了对调谐的一次学术收编——把它变成了校正逻辑可以处理、可以测量、可以发表论文的经验对象。共在调谐因此切开了IDS完全无法触及的一个层面：调谐的朝向不是声学参数，而是对方的在场状态。这个层面，不论IDS研究把声学测量精度提升到多高，都无法触达。

它也不是阿尔弗雷德·舒茨的“共同内在时间流”。舒茨在《共同创造音乐》中论证，音乐现场的本质是两个主体共享同一个“内在时间流”——不共享外在的钟表时间，而是共享一种伯格森意义上的、在声音的前后相继中被共同经历的时间。这个现象学概念已经在音乐哲学中占据了本文试图占据的那个位置——音乐的根本不是在物理时间里组织音高，而是制造一段被共享的内在时间。如果共在调谐只是把这个现象学命题用更身体化的语汇重新说了一遍，它就不是一个新概念。二者之间的区别在于：舒茨的“共享内在时间流”仍然以感知同一段时间流为中心，它是一个认知-现象学的命题——两个人可以在闭上眼睛的沉默中，通过音乐的绵延“共同经历”一段时间。但共在调谐要求的不仅是共享时间，更是双向的身体修正。如果舒茨的两个人关掉了彼此身体状态的实时回传——他们听不见、看不见对方的呼吸、微动、肌肉张力的实时变化——他们仍然可以共享一段内在时间流（例如，他们分别戴上耳机听同一段录音），但他们无法发生调谐。调谐的中断，不来自时间流的中断，而来自身体回传信号的关闭。舒茨给了音乐一个现象学的第一现场——两段内在时间在声音中交叠。本文给他的现场补上一个更底层的发生学前提：在为同一段时间所流贯之前，两个人的身体必须先通过声音完成一次基本的调频。这个调频，不是认知层面的“我感知到你也在这段时间里”，而是身体层面的——我的呼吸斜率因为你的在场而改变，我的下一个音的力度因为你肩膀刚才的微

微一沉而做出调整。舒茨的两个人，共享着一段内在时间；本文的两个人，在共享时间之前，先互相用身体校准了彼此的在场状态。

因为这个概念不是任何既有概念的子集或变体，它才构成了一次真正的、不可还原的命名——它切开的，是一个此前被所有那些精确的概念网络共同漏掉的发生学层面：在那个层面里，音乐尚未成为“作品”，甚至尚未成为“活动”；它只是一次两个身体之间的共同调谐——而这件事，是一切音乐的生物-人际基准。

3.1 调谐需要校正吗？——对一桩互生关系的辨析

概念确立之后，在将其推入制度分析之前，必须正面回应一个本文所能预见的最强反驳。这个反驳不来自某种理论立场，而来自一个几乎每一位有经验的演奏者或声乐教师都会追问的直觉。

调谐的能力本身，是不是需要校正来培养？

反驳者可以沿着本文自己的案例序列进行追问。母亲抱着幼儿在深夜踱步哼唱——这位母亲之所以能让自己的哼声携带一种让孩子安静下来的质地，是因为她在此前无数次与孩子的互动中，已经用她的身体“学会”了怎样控制气息的深浅、怎样让喉头的松紧刚好对应孩子的扭动频率。这种“学会”，是不是一种隐性的、未经命名的、在反复试错中被身体吸收的校正过程？那位山歌歌手之所以能在第三个音上反复复现那枚滑音——而不是在某一次唱时滑到了别的音上、破坏了旋律的可辨认性——是因为他的嗓子已经在几十年的田野与灶台边，被无数次“滑到这里刚好、再多一点就破了”的边界经验训练过。这种对嗓子边界的精确控制，是不是某种形式的校正？那个在琴房漫游的孩子，之所以能按出一个让他自己感到兴奋的奇异和弦，是因为他在此之前的几个月里已经积累了一定量的手指独立性与琴键位置的空间记忆——这些能力的获得，难道没有依赖最基础的手指规范训练（恰是某种软校正）？

更进一步，反驳者可以越过案例层面，直接攻击本文最核心的那个论断——“母亲和婴儿的调谐，不需要任何一方具备音准知识”。婴儿的发声能力有显著的生物学局限：在最初几个月里，他只能发出几种元音化的咿呀声，他的喉部肌

肉控制精度极低，他的呼吸与发声之间的协调才刚开始建立。这种局限本身，就构成了对调谐深度的约束。母亲当然可以在这种约束下完成最基本的调谐——用声音让孩子感到自己的在场。但更复杂的调谐形态——比如两个爵士乐手之间那种一句接一句、彼此摸透对方下一句要往哪个和声拐的实时即兴——显然要求参与者具备相当程度的声音调控能力。而这种能力，至少在大多数音乐传统中，不可能完全在“纯调谐”内部自然生长；它需要某种形式的、带有规范性反馈的重复练习——即带有校正成分的训练。

本文的回应分三步。

第一步，承认反驳的合理部分。调谐与校正之间，从一开始就不是截然分离的两个领域。在每一次有效的调谐中，都渗透着参与者在过往经历中积累起来的声音调控能力——而那些能力的获得路径，确实在大多数场合下包含了某种程度的校正训练（显性的或隐性的、制度化的或经验性的）。本文从未主张存在一个“完全没有校正的人类调谐黄金时代”。口传时代的长老在教年轻猎手唱狩猎歌时，他会纠正他——“不对，这一句声音要往上扬，往下沉了就惊不到鹿”——这就是校正，虽然它不依赖于乐谱。校正作为一种基本的教与学机制，和人类声音文化本身一样古老。本文的批评对象，从来不是校正作为手段的存在。

第二步，精准定位问题所在。问题不出在校正本身，而出在一件极其容易发生的事上：校正从手段，偷偷升格为目的。一个钢琴老师让一个六岁的孩子反复练同一条音阶，起初是为了“让四指和五指的力量更均匀”——这是手段。但到了考级前三个月，这条音阶的反复练习变成了什么？变成了“在一次就成功的模拟考中拿到优秀”。那个曾经的工具性目标（均匀的手指力量是为了将来能弹出更自由的音乐）被悄然替换为自指性的目标（弹出均匀的音阶本身就成了被评价、被奖励的成绩）。手段升格为目的的那一刻，校正在制度层面就完成了对调谐的僭越：它不再服务于一个超出它之外的、朝向共享当下的发生的可能性；它把自己变成了那个发生的全部意义。本文的诊断因此不是“校正应该被调谐取代”，而是“一个健康的音乐文明，需要制度性地防止校正僭越为自指的目标——以确保手段始终在为目的服务”。

第三步，给出一个正面的制度设想作为论证的落脚点。如果本文的批评只到“手段僭越为目的”这个诊断为止，它仍然只是一篇批评文章——告诉读者哪里做错了，却没有提供任何可以被实践的、哪怕极其微小的替代方案。以下正面案例，不是某个已经大规模实施的“成功模式”，而是一个在现有的校正制度内部、靠某位教师的个人判断力所撑开的微小空间——它的价值恰恰在于，它不需要推翻整个制度才能发生。

一位在音乐学院附中任教的钢琴老师，她的班上有一群正在准备期末考试的十三四岁的学生。期末考试曲目已经布置下去——每人一套练习曲、一套复调、一首奏鸣曲的一个乐章、一首自选乐曲。按照常规的教学节奏，考前最后几节课应该用来反复打磨细节：这里踏板换早了半拍，那里左手的伴奏盖过了右手的旋律，诸如此类。但她在倒数第三节课上做了一个不在教学大纲里的决定：她把两个学生叫到同一架钢琴前，让他们四手联弹一首莫扎特的奏鸣曲慢乐章——不是考试曲目，不规定速度，不要求风格准确，只要求一条：“你们两个人，谁也不许管自己弹得对不对。你们只做一件事：听对方的手。听他的手走到哪里，你的手就跟到哪里；你慢了他等你，你快了他追你；弹错了键也不要紧，弹错了就错了，用下一个音去回答它。”学生们愣了一下。这样弹琴，是从第一天进琴房就被禁止的——乱弹琴。

他们开始弹。前两小节是灾难：两个人各自按照自己背下的版本在走，谁也没有在听对方。老师没有喊停。到第三小节，一个学生忽然在某个本来该弹三拍的长音上多停了半拍——因为他听到另一个人的手在那个瞬间犹豫了一瞬，他不知道那种犹豫是什么，但他的手指自动等了一等。那半拍的等待，是他们人生中第一次发生共在调谐。接下来的一分多钟里，他们仍然弹错了无数个音，忘记了无数处本该做的力度标记，但那首曲子第一次——在他们之前几个星期反复校准、反复自我审查的练习中从未出现过的那种第一次——像一根活的东西，在他们两个人的手之间，被共同地、小心翼翼地托着行进。弹完之后，两个人都没有说话。老师也没有点评技术。她停了几秒，说了一句：“这遍你们自己留着。回去以后如果还想这样弹，就再来找我。”

这当然只是一个极其脆弱的案例。它发生在期末考试的间隙，能发生的条件相当苛刻：一位有足够的制度自信去“浪费”一节课的老师，两个已经具备足够技术基础因此能有余力去听对方的孩子，一所还没有把每一节课的教学产出量化到分钟的学校。这些条件在任何一处被收紧——更刚性的考纲、更密集的课时考核、更焦虑的家长——这个案例就不会发生。但这恰恰是它的发生学价值所在：它证明了一件被本文全部论证所指向的事——调谐可以在校正制度内部发生，但它在当前条件下只能在特定的、脆弱的、靠个别人的判断力才撑得开的缝隙里发生。这不是一堵密不透风的墙被推翻了；这是一扇能被推开的门被找到了。而找到这扇门，比站在墙前面重复“墙是错的”更有实践意义。

四、校正如何从技术升格为制度性独断——以钢琴考级为例

第二章结尾留下了一条尚未焊接的裂缝：记谱法翻转了音乐的本体论——谱子从追记变为原作，演奏从发生变为还原——但“校正的耳朵”是怎样从一个为多声部协作服务的技术工具，变成“每一个学音乐的孩子从第一节课就被要求长出来的身体器官”的？这里缺了一整个制度传递机制。本章以钢琴考级为例，补上这个发生学环节。

钢琴考级制度在全球范围内的扩散——从十九世纪末英国三一学院和皇家音乐学院的本地考级，到二十世纪下半叶覆盖东亚、北美、澳洲数百万琴童的庞大产业——是校正从技术升格为制度性独断的最关键的传递装置之一。它不是一套中立的评估工具。它在三个制度节点上，一步一步地把“精确还原乐谱”从一种手段推升为音乐学习的全部可见意义。

第一个节点：曲目体系的建立如何制造了“唯一可见的评价维度”。任何一套考级大纲，核心是一份分级曲目单。每一级有三到四首规定曲目——练习曲、复调作品、奏鸣曲乐章、自选乐曲（近年部分考级体系中自选曲目已被取消，全部改为规定曲目）。每首曲目下面，标注着考评标准：音准、节奏、速度、力度记号、指法、触键、踏板、风格。这八到十项标准，乍看是全面的——它们覆盖了从技术到表现的所有层面。但仔细看：每一项都可以被精确地描述为“偏离谱面标记的程

度”。音准——音符是否弹对；节奏——时值是否精确；力度——谱面上的f和p是否被遵守；指法——是否使用规定指法。所谓“音乐表现力”，在考级评分表上，被操作化为“是否做出了谱面上的力度对比和速度变化”——这不是在衡量表现力，这是在衡量学生是否精确地执行了谱面上关于“如何表现”的全部指令。一个孩子在这个评价体系里学琴十年，他会内化一件极深的事：一首曲子弹得对不对，就是它离谱子有多近。而他永远不会在这十年里遇到任何一次被制度严肃对待的提问：“你刚才弹的那一遍，有没有哪一个瞬间，你感觉不是你在控制手指，而是手指在带着你走？”

第二个节点：音乐师范教育的标准化如何让“校正的耳朵”成为教师的默认听觉装置。一个考级体系的运转，需要数以万计能按统一标准打分的考官。这些考官的耳朵，是在音乐学院和师范院校的音乐教育专业里被批量训练出来的。以中国为例，各省师范院校音乐教育专业的课程体系，从1980年代以来逐渐趋同：视唱练耳课从大一上到大二，每周四个课时，训练的核心能力是“听准”——听出音高、音程、和弦、节奏的精确数值，并将听出的内容以标准记谱法写下来；钢琴必修课以拜厄、车尔尼、小奏鸣曲为主轴，评价标准是“完成度”——错音率、节奏准确率、速度达标率；教学法课程所传授的“如何教”，其核心内容是“如何帮助学生更快地消灭错误”。四年下来，一个师范毕业生被训练出的不只是一双能听出细微音高偏差的耳朵，更是一种不可撼动的知觉信念：音乐教学中“教”的本质，就是帮助学生缩小他弹出来的声音与谱面规定的声音之间的差距。这个信念如此基本，以至于它不再被当作一种需要被论证的立场——它就是“正常的音乐教育”本身。而当一个持有这种信念的毕业生成为考级评委，她坐在考场里，听到一个考生在某小节忽然把速度放慢了一点点——不是因为技术跟不上，而是因为那个和声让他想在那里多停一下——她的评分表上会写下什么？“节奏不稳定。”

第三个节点：闭环的自我验证。当考级体系、师范教育、社会期待这三道工序合流之后，一个完美的自我验证闭环就完成了。老师用校正的耳朵在音乐师范院校里被训练出来——她进入社会后，用这双耳朵教学生——学生用这双耳朵学——学生被送到她当年的老师（如今是考级评委）面前用同一双耳朵打分——家长看到考

级证书上的“优秀”，掏钱续费——社会各界的共识被不断加固：“学音乐，就是把曲子弹对。”这个闭环最隐蔽的力量在于：它的每一个参与者，都不是在被迫执行一套外在的规训。他们在自己的感知层面，真诚地认为“这双耳朵就是用来听音乐的全部”。当那位老师在课堂上对学生的慢板说“节奏不稳定”时，她不是在违心地执行一套标准；她是真的——在她的知觉系统所能触及的全部范围里——听不到别的。校正已经从一种评价方式，变成了她身体里那双耳朵唯一能识别的知觉语法。

在这个闭环里，调谐不是被禁止了——没有人说“不准调谐”。调谐是被不可见化了：没有任何一个制度节点，会为“这个学生刚才在某个瞬间，和另一个学生的身体产生了实时调谐”这件事，提供任何可以命名、可以记录、可以传递的评价语言。当一件事无法被任何制度认可，它在制度的意义上，就不存在。

这个过渡节一旦焊上，本文的发生学线索就从“跳跃的锐利”变成了“推进的不可撼动”：记谱法翻转了本体论（第二章）→考级与师范教育把翻转后的本体论固化为唯一合法的知觉语法（本章）→这套知觉语法在琴房里被内化为演奏者的身体习性（第五章）→它在民歌现场把调谐残片识别并切除为冗余（第六章）→它在流媒体时代被算法推到极致（第七章）。

五、琴房中的身体：校正如何被内化为知觉习性

前一章理清了校正如何通过考级与师范教育，从技术升格为制度性独断。本章进到更深的肌理层：当校正不再是外在的评判，而是被内化为演奏者身体知觉的一部分时，它如何重新塑造了演奏者对他自己、对他发出的每一个声音的体验方式。

琴房是校正制度最精微的末端毛细血管。一个孩子被抱到钢琴前，他伸出食指，在黑键和白键上随意地按。他听到一个低沉的声音，像远处闷雷的尾巴；他又按了一个清脆的高音，像一颗小石子落进井里。第二个音和第一个音之间没有节拍，没有调性，没有任何一本教材能追认的旋律逻辑。他只是在听：这个声音出现之后，下一个声音可以是什么？他的手腕上下浮沉，身体不时倾斜，用肩膀、肘

部、呼吸来配合手指对琴键重量的试探。这一刻，他不在“学音乐”——他在用自己的整个身体，测绘一片他从未进入过的声音大陆。他每按下一个键，都是在向这片大陆发出一句询问，而每一次询问得到的回答——那个声音在空气中的衰减方式、那个音和前一个音之间在耳朵里留下的“还想再听”或“想停下来”的张力——都决定了他下一次落指的方向。

这个场景在规范的音乐教育话语里有一个专门的名字：“乱弹琴”。这三个字本身已经是判决，而且它同时触犯了校正逻辑的两条核心禁令。禁令一，它没有构成任何一首“对”的曲子——它不是拜厄某一条，不是小奏鸣曲某一章，不是任何能被对照谱子来逐音纠正的对象。它因此无法被收编进那条“从弹不对到弹对”的进步叙事。它作为声音事件，在“是否正在学音乐”的评价网格里，没有合法的身份。禁令二，它浪费了时间。一个孩子要弹到能演奏一首奏鸣曲，需要几千小时的有效训练。这十分钟的漫游，不仅没有为那几千小时贡献任何可计量的产出，还可能“养成了坏习惯”——手腕那么歪，以后跑动怎么提速？

琴房里的成年人在这两条禁令的同时作用下，做出了那个全世界每天发生几百万次的动作：他把孩子的身子扶正，把手腕放平，把教材翻到这次该练的那一页，把手指摆放到正确的起始键位上。“我们开始练吧。”他说这句话时没有恶意。他爱这个孩子，希望他学成。但他的这个动作，在发生学层面的实质是：他用手腕的角度、教材的页码、起始键位的规范，把孩子刚才在声音大陆上的整个探险，一次性判定为不值得发生的冗余行为。

这里必须做一个关键的方法论声明。本章对琴房场景的叙述——包括上述“漫游的孩子”及其后续——不是对某一特定个体童年经历的经验报告，而是一种理想型叙事。它的功能，是将校正制度内化为身体知觉这一发生学机制，以一种可被直观把握的场景形态呈现出来。本文无意断言每一个琴童都经历了完全相同的轨迹；本文所展示的，是校正制度在其逻辑极限处所趋近的那个典型形态——当考级体系的评价维度、师范教育的听觉训练、家长的社会期待完整合流之后，一个进入这个体系的孩子，他所面对的那套“把漫游识别为冗余”的知觉语法，其运作方式是高度结构化的。理想型的价值不在于每一个细节都能被经验证实，而在于它把制度逻

辑的纯粹形态从经验的芜杂中析出——正如韦伯的新教伦理并不宣称每一个清教徒都那样计算自己的恩典状态，而是把一种在历史中实际起作用的精神气质推到了它的逻辑终点。本文使用理想型叙事出于同样的方法论考虑。

这个理想型叙事所揭示的，并不是一个孤立的儿童教育学问题。它是“听即校正”作为内化了的知觉习性，如何在身体层面运行的精确模型。一个经过长期训练的学习者，他会自发地在演奏过程中分裂成两个人：一个是正在动作的手指与呼吸，另一个是飘在头顶上方的、隐形的评审员。这个评审员实时监控着音准、节奏、触键均匀度、声部平衡，并在每一个参数偏离预设的刹那，发出一个无声的警报。久而久之，他不再需要这个评审员“刻意”发出警报——他的整个身体已经被重新校准：他的手腕在即将塌下去的前一帧会自动收紧，他的呼吸在即将进入困难段落时会自动调整到备战模式，他的耳朵会自动过滤掉那些与“弹准”无关的声音质地。他用数年的训练，把自己锻造成了一台高精度的自我校正装置。

这台装置，在行业里有一个被高度赞誉的名字：“技术过硬”。而本文要指出一件被这一赞誉掩盖的事：这台装置在完成自我锻造的同时，也完成了一次对自我体验深刻的重组。重组的结果是，演奏者已经很难再问自己一个最简单的问题：我为什么要弹这首曲子？不是他不想问，而是他的身体和耳朵已经不再为这个问题提供任何可感知的线索。当他演奏中觉得“哪里不对”时，他的第一反应——几乎是生理反射——是去查技术参数：刚才那个换把是不是慢了？琶音的颗粒是不是不够匀？左手的颤音频率是不是掉下来了？他极难将自己的不满意追溯到更深的源头：那个乐句，他想要的究竟是什么？为什么在那个小节之后，他突然感到整首曲子失去了方向？这些问题指向的是一个没有被校正坐标系事先网格化的、开放的、不指向可被精确诊断的病因的领域。在这个领域面前，他会感到一种无法被技术训练消化的焦虑，于是他迅速退回到技术上。技术至少是可控的。手腕角度微调两毫米，呼吸提前半拍准备，那个音就准了。准了之后，那个焦虑就被暂时压住。每一次这样做，他都在用自己的身体重复一个发生学的选择：放弃调谐的开放性，换取校正的确定性。

古典音乐评价体系里，“技术分”和“音乐表现分”在纸面上是分开的。但在任何一位资深评委的实际感知中，二者从未真正分离。一个音色黯淡、律动不稳、力度层次模糊的演奏，极难被评为“音乐表现深刻”——不是评委不想，而是在这种物理条件下，演奏者意图传达的任何“音乐内涵”，根本没有足够的技术资源送达听者的耳朵。反过来，一个技术上几乎无懈可击的演奏，即使在音乐理解上毫无独到之处，也极容易拿到一个体面的总分。那份技术本身的“完成的满足感”，在评委的经验里，会自发地转化为“音乐上也是成立的”这种判断。这不是任何一个个体的错误，而是校正逻辑彻底内化之后的一个必然的感知后果：当所有人都被训练成一架精密的校正装置时，装置们自然会认可另一台装置的高性能，并把这种性能命名为“好”。

最后用一个极端的身体案例来为本章压卷。一位旅美的顶尖钢琴家，在一次访谈里被问到一个问题：“你弹了一辈子琴，你有没有过在台上，完全忘了自己在弹什么曲子、忘了手指、忘了观众，只觉得自己像是一根空心的管子、整个宇宙的风从你身上吹过去的那种时刻？”他想了一会儿，很诚实地回答：“有过，但是很少，而且每一次都只在弹我不在乎会不会弹错的曲子的时候。”弹我不在乎会不会弹错——这七个字，说穿了全部。那个被校正装置长期占用的身体，只有在校正装置短暂地、偶然地停机的间隙，才第一次知道自己仍然能成为一根空管，被不属于自己的风吹透。这个间隙，就是调谐的现场。而这位钢琴家余生能够回到那个现场的次数，已经屈指可数。这不是因为他缺乏天赋，而是因为他的天赋，大半辈子都在被用来喂养那台永远开着的校正录音机。

（此处引述系本文作者基于多位职业演奏家在公开访谈中的自述所做的复合性重构。多位在职业生涯中达到顶尖技术水准的演奏者——包括但不限于几位在二十一世纪初接受过《纽约客》《卫报》深度访谈的独奏家——都曾以不同方式描述过类似的体验：最深层的音乐沉浸，往往发生在“不再担心技术失误”的罕见时刻。本文将这种多源证词凝缩为一个代表性场景，以揭示校正机制停机时调谐才得以短暂浮现的发生学条件。此类复合性重构的价值不在于精确引证某一句话，而在于它捕捉了一种跨越个体、跨越访谈语境的重复出现的身体经验。）

六、民歌现场的调谐残片如何被判定为冗余：制度中介的微观分析

前面几章在制度和身体的层面，分析了校正如何完成对调谐发生条件的系统压缩。本章必须把论证下降到具体的声音事件上，展示那些在校正制度的缝隙里仍然存活的调谐残片，以及它们在遭遇制度中介时如何被识别、被判定、被切除——并且，这种“切除”本身是一个发生学事件，而非一个“原本完整的有机体被外力割掉一块”的静态损伤。

第一个现场发生在民歌比赛的排练厅。一位五十多岁的民间歌手，从小跟父亲在村里唱一种山歌调子。那调子有一个固定的旋律骨架，但他每次唱到某一句的第三个音时，都会在音尾加上一个极其微弱的滑音——声音在往下掉的半路上突然被喉头轻轻兜了一下，像一片叶子被一阵看不见的气流托住，颤了一颤，然后才落到那个本该直接落到的音上。他自己说不太清楚为什么要在这里滑那一下。如果不滑，嗓子完全可以直接落下去——他做得到。但他在那个地方，就是“不舒服”。仿佛不滑那一下，身体里有什么东西没出来。

后来在一次闲聊里他提起，很小的时候听父亲在灶台边唱这首调子，那天灶口返烟，满屋子的烟呛得父亲在那一句的第三个音上咳了半声——嗓子一滞，音就滑了一下，又多颤了一拍才收住。这滑音，是父亲的嗓子在那个下雨的下午与灶口的湿柴之间一次无意的合作。当时正在喂鸡的母亲听见了，从灶台另一头笑着说了句：“刚才那个样子好听。”父亲记住了那个颤。后来他每次唱到那里，都故意滑那一下。在他自己也不知道的深层意识里，那一滑已经不再是技术，而是对那个下雨的下午、对那个被烟呛住的嗓子、对那个从灶台另一头传过来的笑声——的一次身体的纪念。

后来，这位歌手被推荐去参加省里的民歌比赛。排练时，声乐指导对他建议：“老师傅，您这一句的气息在这个地方有点偏，喉咙没完全锁稳。去掉这个小滑音，线条会更干净，整体表现力会更好。”这个建议不是凭空而来的。它背后是一整套在当代中国声乐教育中被制度化的审美范畴——“干净”。要理解这个“干净”怎样从一个描述性的形容词变成一条规范性的尺，需要稍微追溯它在几部影响

广泛的声乐教材与教学体系中的确立过程。从1980年代起，随着各大音乐学院声乐系对“科学发声法”的系统引进，意大利美声学派关于“连贯”（legato）与“统一音色”的主张，经由沈湘、周小燕等宗师级教师的教学体系，被转化为中国声乐教育的一套可操作标准：气息必须均匀托住每一个音，声区过渡不能有可辨识的“坎”，音与音之间的过渡必须平滑无痕。同一时期，中央电视台青年歌手大奖赛从1984年起，将“音准”“节奏”“音色统一”“气息连贯”作为评委打分的主要技术指标。连续二十余年的电视播出，让这些标准从专业圈内部的行话，变成了几亿观众耳熟能详的“好嗓子”的定义。在这种审美范式中，任何一个不被旋律骨架所要求的、在音尾多出来的、让声音线条出现轻微波动的东西——比如喉头半松半紧时自然带出的一枚滑音——都会被识别为“不干净”。

指导是有道理的：去掉它，声音离那根由完美气息托住的理想弧线，确实更近了一步。歌手听了。他改掉了那个滑音。比赛时唱得非常干净，评委给了高分。

但他后来对一位采风的音乐研究者说过一句话。那次比赛之后几个月，有一天傍晚他一个人在院子里坐着，嘴里不自觉地哼起了那首调子。哼到那一句第三个音的时候——他原以为已经被自己改掉的那一滑——嗓子没有任何预兆地自己滑了下去。他说那一下他停住了，愣了好几秒。“没掉，”他说，“那个东西没掉。它在里头。我不知道什么时候它又会出来。”但在台上，在全国比赛的舞台上，他再也没有唱过那个滑音。调谐的残片没有在他身体里被消灭——但它被他自己的身体在“这一次要唱对”的场合下，主动地压了回去。

这里需要做一个关键的发生学澄清。本文说这个滑音被“切除”，不是在暗示有一个“完整的、本真的”山歌版本曾经在那里稳稳地站着，后来被一刀砍掉了一块。这个滑音，从一开始就不是这首山歌的“构成要素”——它是那个下雨的下午、那口灶的湿柴、母亲那句笑、父亲那半声咳嗽，在声学上的一个残余痕迹。在它被声乐指导识别为“冗余”之前，它根本不是一个“被切除物”——它只是歌手身体的一部分，是他在每一次唱到那里时不自觉地复现某种身体记忆的方式。正是“干净”这套评价网格的介入，才第一次把它建构成了一个可以被剥离、被指

认、被建议移除的“多余的东西”。切除是一个发生学事件——它制造了自己的操作对象，而非在一个预成的身体上施行手术。

（此案例系本文作者基于对中国多个省份民歌传承人口述材料的综合阅读所做的复合性重构，不系于某一个具体的可指名歌手，但它所呈现的“滑音作为身体记忆”的发生学结构——以及歌手在制度介入后亲历的“身体记忆未被消灭但被主动压制”的经验——是跨个案的真实存在。本文采用复合性重构的理由是：此类口述材料所捕捉的身体经验在发生学上具有结构性的相似性，将它们凝缩为一个代表性场景，能够更清晰地展示该结构在制度压力下的运行方式。作者清楚这样做在经验可核查性上付出的代价，并接受这一代价作为理论建构阶段的方法论取舍。）

这不是孤例。在每一个被纳入现代评价体系的民间音乐传统中，同样的识别与切除手术都在以“科学”“规范”“提升表现力”的名义静悄悄地执行。民族音乐学自查尔斯·西格以来，已有大量文献批评西方五线谱系统对非西方音乐声呐的裁剪——微分音、不确定音高、滑音技巧，在记谱过程中被系统性地“规整”为最接近的半音或全音。这条批评线索是重要的，但它关注的焦点是记谱裁剪了什么音。本文要指出的与此不同，而且更进一层：即使那些音本身被精确地记录下来，记谱也无法保留一个更根本的东西——那个滑音在歌手身体里所携带着的、朝向某一次具体共在调谐的关系性记忆。记谱不仅裁剪了音的形态，更裁剪了音所承载的共在关系。这不是声学精度的损失，而是发生了学层面的遮蔽。

第二个现场更加朴素，也更加拒绝被记谱或录音所捕获：一位母亲在深夜抱着半岁大的幼儿，轻轻踱步，鼻子里哼出一种不成调、没有固定音高、节奏松松垮垮地跟着孩子身体扭动的声音。没有谱子，没有标准，没有正确。从任何校正的视角听，这十分钟的哼唱在每一个参数上都落到评分表的底部——音准无从测量，节奏无稳定节拍，旋律无任何可辨认的结构，冗余极高。但她没有停。她多哼的那几遍，从功能上，是浪费。从共在上，是唯一有效的信号：正是因为它们是多余的，没有在“哄睡”这个任务完成的那一刻立刻停止，它们才能让孩子在迷迷糊糊之中感知到——此刻不是两个齿轮在完成一对功能性的咬合；此刻是两个人，在共同度过一段生命时间。

这就是调谐的不可优化性。共在的全部质地，恰恰在于那些“可以不做却仍然做了”的声音。它们是一个人对另一个人发出的、比任何语言都更原始的承诺：我在这里，我的身体在和你一起消耗生命时间。而这个承诺，是音乐能够在一代又一代人身上重新发生所需要的最低的、不可再还原的地层。校正逻辑当然不会禁止母亲哼唱。但它每天在做的事，是把每个学音乐的孩子的耳朵，训练成听不到母亲哼唱里那个共在信号的耳朵——乃至这个孩子长大后，他会用同一套评分逻辑去听自己孩子的“乱唱”，然后温和地纠正他：“来，爸爸教你唱，这句应该这样唱。”

七、从作品到场景：调谐之发生条件被制度压缩的路径

概念确立、制度机制焊接、身体与案例展开之后，该从结构上回答本研究的核心问题：调谐的发生条件，是怎样在整个文明的进程中，一步一步被校正制度系统性地压缩的。

先从一对区分开始。任何一段可以被我们称为“音乐”的人类声音事件，同时可以在两个截然不同的存在论层面上存在。在第一层面，它是一件作品：被固定的音高与时值排列，被记谱或录制，可以脱离其产生时的那个特定的在场，而被独立地传递、学习、评价、相互比较。在第二层面，它曾经是——或者在某一些珍贵的情况下，它仍然是——一个场景（scene of co-presence）：在某一次具体的、不可复现的、发生在特定活人之间的共同在场中，由至少两具身体共同调节出来的那个声音事件。“场景”在全文中均在此严格意义上使用，指声音事件发生的那个不可复现的、活人之间的共同在场，与“作品”构成一对本体论概念，而非日常语义中的“舞台场景”或“场景布置”。作品的存在论基础，是校正——它需要被准确地写定、被准确地识读、被准确地还原。场景的存在论基础，是调谐——在共同在场里的那种实时双向修正，不需要任何固定的参照，它的唯一合法性来源是“这一切正在发生”。

这二者并不必然冲突。人类最伟大的音乐作品，从格里高利圣咏到巴赫受难曲，从贝多芬晚期四重奏到爵士即兴，它们最初全都是从无数次的场景里——从修

道院回廊里两个人的一唱一答，从教堂里独唱者与管风琴之间的博弈，从四重奏排练厅里四个人为一句乐句的速度激烈争论又最终一起滑进去的那个时刻，从小酒馆里萨克斯手朝钢琴手丢过去的一句乐句引子——被一次次地凝练出来的。作品是从场景里长出来的；它被写定之后，又成为新的场景的引子。

但制度化的现代音乐教育，恰恰把这个“长出来”的动态关系，扭成了一条单向的不归路：场景被定义为通向作品的训练台阶，作品被定义为台阶顶端唯一的产出物。一个学生被要求排练一首奏鸣曲，不是被要求“在这个奏鸣曲的框架内，与你一起排练的同伴之间发生出一段活的共同调谐”。他被要求的是：在期末考试之前，把这首奏鸣曲弹到足以让评审打出体面分数的程度。他的排练搭子对他的全部价值，在于对方的伴奏是否稳、是否准、是否能让他安心地把技术难点磨过去。如果这位搭子在排练中某一段忽然慢了半拍——不是技术失控，而是他被那一段的和声突然触动了，他想在那个延长记号里多沉溺一秒——练完之后，两位合作者之间会发生一场对话，而这场对话的内容，几乎一定是：“你刚才那一句节奏不太稳，咱们对着节拍器再合一下。”而不是：“你刚才慢那一下，我好像突然明白了这一段为什么要这么写。”前者是校正的回馈，后者是调谐的延续。在前者被高度制度化的环境里，后者越来越没有合法的发声空间。

这里必须在“技术”与“制度”之间做出一个关键区分，以防止论证滑向技术决定论。记谱法与录音技术本身并不必然劫持调谐。一份乐谱，完全可以成为调谐的工具——两个歌手拿着同一份谱子，但在每一次对唱中根据对方当天的嗓音状态、呼吸深度、情绪起伏，实时调整自己的力度与速度，让谱面上的同一行音符在每一次场景里都产生不同的共享当下。录音技术亦然：一份录音，完全可以成为一次新的调谐的媒介——这一点将在下文专门展开。真正将调谐的发生条件系统性地压缩的，不是这些技术本身，而是制度——评价体系、教育体制、考级机制——将技术所固化的“作品”确立为音乐的唯一合法形态，将“场景”降格为通向作品的训练台阶或可被裁减的冗余。制度，而非技术，才是本文诊断的那个真正的机制。乐谱没有罪，有罪的是那个把偏离乐谱一寸都判为“错误”的评价法庭。

在流媒体时代，场景的退场速度被猛地提上了一档。算法不需要矫正你的音准——它有Auto-Tune和后期混音，可以把任何声音修成物理上“完美”的成品。但它极端地放大了校正逻辑的另一面：它不是校正你的身体，而是校正你的听觉注意力。前奏超过十秒还没进高潮，是冗余——跳切掉。间奏第二段和第一段太像没有变化，是冗余——加音色钩子。整首歌超过三分钟完播率会跳水，是冗余——压缩到两分半。所有不能被“抓耳曲线”这个公式所吸收的声音纹理——沉默、犹豫、粗糙的和声、没有清晰功能走向的即兴片段——都被判定为浪费了宝贵的、可被计量的注意力窗口。歌曲的创作端，从写歌的那一刹那，就已经被这套反冗余逻辑预定格式化了。一首歌在被任何人听到之前，就已经先被“如何证明自己值得被听完”这个内置的校正算法审查过了一轮。而这恰恰是场景不能生长的环境：场景需要一种不确定的、不预先为“值得”担保的氛围——朋友在一起胡弹瞎唱，没有人问“这首歌好不好听”“这段和弦是不是太简单了”“我们唱得是不是太难听了”。在这些问题被迫进来之前，三个人已经在那十几分钟里共同渡过了一段无法计量的生命时间。而一个永远开着校正滤网的耳朵，会在第一时间就把这些声音挡在“不值得我们继续听”的门外。场景，就是这样被退场的。

然而，本文必须在此明确回应一个由这一叙事自然引出的质疑：如果调谐依赖于即时的共同在场，那么当代音乐的绝大部分形态——通过录音和流媒体传播的作品——是否都被排除了发生调谐的可能？两个人一起听同一首录音，在某一段钢琴独奏响起时不约而同地停下手中的事、交换了一个眼神、同时叹了口气——这不算一次共在调谐？

答案是：算，但它发生的条件比在同一个物理空间里即时生成声音要苛刻得多。当两个人共享一段录音时，调谐的发生不依赖于对声音本身的任何修改——录音是固定的，它不会因为听者的呼吸而改变力度或速度。调谐发生的空间，被压缩到了两个人对这段录音的共同反应之间的那条窄缝里：你们在同一刹那屏住了呼吸，在同一乐句结束后同时沉默了半秒，你的手在她的手背上轻轻按了一下而她在那同一刻把头靠在了你的肩上。这些不是“对音乐的反应”——它们是两个人通过共享一段已经固化的声音，而彼此调谐了自己身体的在场状态。调谐仍然发生了，

但它的声学质料不再是那个可以被实时双向修改的声音，而是那个不可修改的录音在两个人身体上激起的、在同一时间窗口内同步浮现的微反应。这要求更苛刻的时间同步性、更敏感的相互觉察、以及——最重要的——一段没有被“抓耳曲线”切成碎片的、足够长的、允许两个人慢慢沉进去的听觉时间。而流媒体算法的反冗余逻辑，恰好是在破坏这最后一个条件：它不让一首曲子有足够长的、不被切断的时间来在两个人之间织起共同的在场。调谐在录音时代没有被彻底消灭，但它的发生所依赖的听觉注意力连续体，正在被算法一刀一刀地切碎。

这个回应也恰好将本文与已有的几支音乐社会学与声音研究传统做出区分。蒂娅·德诺拉在《日常生活中的音乐》中深入论证了音乐如何成为社会行动的资源、如何在日常生活中建构主体性——这与本文的“调谐”同样关注音乐如何“做事情”而非“是什么”。但德诺拉的分析单位是个体如何运用音乐来调节情绪、组织注意力、建构自我身份；共在调谐的分析单位则是两具身体之间的实时双向修正。一个人戴着耳机在公共汽车上用一首歌来“让自己振作起来”——这在德诺拉的框架中是典型的“音乐作为自我技术”的案例，但它显然没有发生共在调谐，因为现场没有另一具正在被实时调谐的身体。同样，从皮埃尔·谢弗的“听觉对象”到晚近的声音研究转向，都在不同意义上拆解了“作品”作为音乐分析默认单位的霸权，把注意力从谱面引向了声音本身的物质性——它的频率、包络、空间定位。本文与这条线索共享同一个批判目标（作品的霸权），但本文的进路不同：本文不是在“作品”之外另立“声音”作为更真的分析对象，而是指出，即便是最精微的声音本体论分析，只要它还在分析“声音”这个对象的属性，它就仍然没有触到那个比声音更底层的发生学单元——两具身体之间通过声音进行的相互调谐。调谐不在声音的物理属性里，不在声音的社会使用里，而在两个在场者身体状态的那个永不休止的、前意识的互倒原料的循环里。

八、结语：在一间不评分的琴房里，音乐可否重新发生

本文提出了并论证了一个不可被既有音乐学概念1:1替换的新概念——共在调谐，并以此为核心，重置了一套理解音乐文明史的叙事框架。这套框架不从人类如

何征服转瞬即逝讲起，而从人类如何一步步将校正制度化、从而系统性地压缩调谐发生条件的过程讲起。这个叙事的关键节点，不是口传时代的调谐被记谱法一刀斩断，而是校正的逻辑如何从那以后持续膨胀——从乐谱到节拍器，从标准音高到考级体系，从电视大奖赛的评分标准到流媒体算法的抓耳曲线——每一次膨胀，都把调谐的发生条件向更窄的缝隙里逼退一步。但膨胀不是消灭。本文反复展示的那些调谐残片——灶口返烟的滑音被切除后歌手在院子里发现它没掉的那一刻、母亲踱步哼唱时的不可优化性、那位钢琴家在“不在乎弹错”时才回到的空管状态、两位学生在莫扎特慢乐章里错漏百出却第一次让曲子活起来的那一分多钟——它们都在共同证明同一件事：调谐从未被彻底消灭。它只是被逼到了角落里。而角落本身，就是校正制度最诚实的一份自供状——它暴露了这套制度永远无法纳入的剩余物。

在结语处，必须诚实面对整个论证的界限。本文的判断，在什么情况下会错？这个证伪条件是：如果一个学生只是把一首自己刚学会的奏鸣曲，原封不动地、按照谱面上所有记号的约束，毫无差错地弹给她的祖父听——她的演奏是高度校正的。但如果在那一个朴素的客厅里，祖父坐在旧沙发上，听到某个他年轻时也听过的旋律，眼眶忽然湿了；而孙女在眼角余光里瞥见那一丝湿痕的瞬间，她接下来的触键被一种她从未在这个曲子里用过的力度所改变——不是故意做的，是她的手指在接到那个回传信号之后，自己变的——那么，在这个现场里，精确地再现一首作品，与在二人之间发生一次活的共在调谐，并不需要互相排斥。它们简洁地抵达了彼此。于是，“校正制度化不可逆转地压缩了调谐”这个强势的历史叙事，就被推翻了。

这里也必须正面回应一个更深层的质疑。如果校正——记谱、节拍器、标准音高、考级体系——是音乐得以专业化、传承、并在技术上抵达如此高度的前提，那么“压缩调谐”是否是一种为了专业化而不得不承受的代价？换言之，本文是否在暗示，专业音乐教育在本质上就无法避免压缩调谐——因而它陷入了一个内在悖论？

本文的回应是：把“压缩调谐”接受为专业化不可免的代价，这恰恰是本文所批评的那种校正独断论的最后一道防线。它背后的预设是：调谐与校正之间，只有

一种零和关系——校正多一分，调谐就少一分。但本文第六章的四手联弹案例已经证明，这种零和关系不是必然的：在两个已经具备足够技术基础的孩子之间，当他们被给予明确许可去“不管对错只听对方”时，调谐不但没有被技术能力所阻碍，反而因为技术能力的支撑而获得了更灵敏、更精微的实现——那个在第三小节多等了半拍的感知，恰是以数年校正训练建立的手指独立性与听觉灵敏度为条件的。专业化并不必然以压缩调谐为代价；真正压缩调谐的，不是专业能力本身，而是那种把校正从手段推升为唯一目的的制度性操作——当考级曲目、师范教育、社会评价三者合流，把“精确还原”变成了音乐学习的全部可见意义时，才制造了“调谐被排挤到缝隙里”的那个格局。专业化没有悖论；有悖论的，是一种放弃了手段与目的之间制度性制衡的专业化。

本文的批评对象，因此不是校正本身，不是精确，不是乐谱，不是考级，不是任何一件有助人类更高效地共享声音的技术工具。本文批评的，是一种制度性的听觉独断论：它把那对在乐谱面前被训练出来的校正的耳朵，偷偷地、不经审判地，立为用以裁定一切声音事件是否有资格被称为“音乐”的唯一法官。在这个独断论的统治下，所有不能被校正评价体系所识别的声音发生方式，都被置于一个双输的绝境：要么你按照校正体系的要求改造自己，从而被吸纳为“合格的音乐”，但在这个过程中，你失去了自己曾经携带的那一粒不可替代的共同调谐的质地；要么你拒绝改造，于是你就只能在“非正式的”“民间的”“业余的”“不规范的音乐实践”这类边缘头衔下被温和地容忍，却永远不会被纳入“当我们严肃地讨论音乐时，我们要讨论什么”的那个中心。本文要做的，不是推翻这条边界，而是指出：边界本身是人画的。把边界画在这里，让它把那个母亲的哼唱、那个灶台边的滑音、那节被用来“浪费”的四手联弹课，画在“音乐”的外面——这是一个可以被重新纳入、也应该被重新纳入的文明选择。说“重新纳入”而不是“收回”，是要强调：这不是对某个失落本源的回归，而是一个需要制度性行动来完成的、对既有边界的重新划定——它需要新的教学大纲、新的评价语言、新的教师培训方案，需要一个文明有意识地为调谐留下一块合法的制度空间。

但必须在此处做一笔关键的收束：重新纳入这个选择，不是一件可以通过“放下校正的耳朵”这一姿态便一劳永逸完成的事。那位在倒数第三节课上让学生四手联弹的老师，她之所以能撑开那个缝隙，不是因为她的耳朵比同事更“天然”——恰恰相反，她自己在数十年的专业训练中，耳朵已经和所有同行一样，被深度格式化为校正的耳朵。她在听两个学生错漏百出的莫扎特慢乐章时，她的校正装置一直在报警：那里节奏塌了，那里和声残了，那里声部完全黏在一起了。只不过，她在那节课上做出的判断是——“我这次先不跟从这套警报”。这个“先不跟从”，不是一个自然耳朵的复位，而是一个已被校正深度格式化的耳朵在对自己实施的一次艰难的、高度自觉的抑制。它在制度层面的含义是：调谐的重新打开，需要的不是“回归天然”——已经没有天然可供回归了，每一双在现代音乐教育体系中浸泡过的耳朵，都被校正改写了知觉结构。它需要的是制度留出合法的时间与空间，让那些已经被校正格式化的耳朵去练习一种它们从未被正式教过的听觉方式——去听那个不被谱子担保的、正在发生的、只能靠彼此在场来导航的声音事件。那个已经被格式化的耳朵，能不能重新学会听这种声音，不是一个心理学问题，而是一个制度设计问题。如果评价体系不为这种“不可被计量的听”提供任何制度性激励——不为它加分，不为它列进教学大纲，不为它在学期考核里留出一格合法的位置——那么少数教师凭个人判断力撑开的缝隙，就永远是缝隙，不会变成门。

最后，回到一个最朴素的现场门槛上，作为全书论证的最低检验。今天，一个初学钢琴的孩子，被要求去考级、去比赛、去证明自己那几年练习的每一分钟都是“有效的”。但如果她不想考级，她只是想在周五的傍晚，坐在琴前，随便按键，听那些声音在黄昏的光里慢慢暗下去；如果她弹完了一首自己偷偷学会的、从爷爷收音机里听到的老歌，然后回头看了她的老师一眼，而她的老师没有对她说“节奏有一点不稳，我们对着节拍器再过一遍”，而是轻轻说了一句：“你刚才弹的那个东西，能不能再弹一遍给我听听”——这个老师在这一秒里所做的，在本文的全部论证中，就是一件在制度层面上极微小的、却在发生学层面上颠覆性的事：她把她自己那双被数十年校正训练深度格式化的耳朵，暂时地、有意识地、艰难地放到了调谐的位置上。她让音乐，在还没有被审查过值不值得发生之前，发生了。

附论零·本组四篇的分工

本组四篇同出于一个原初问题——音乐的本质是什么——并且都朝同一个方向下刀：音乐不在声音里，而在活的身体之间。这带来一个真实的风险：**共调、接隙、共在调谐、接住之声**，可能只是同一件事被命名了四次。本节把分工写明，并给出这份分工在什么条件下应当被推翻。写在正文之后，是因为它不是本篇的论证，而是本篇与另外三篇之间的边界声明。

（一）四篇各占一层

①《不在声音里，亦不在你我之间》（共调）占本体层。它追问的是：在音乐事件的开端，“发出”与“接收”这个分工本身是否原初。它的主张最强，也最危险——分立是事后的。

②《不在声音里：音乐作为“接隙”事件》占回路层。它接受收发之分已经成立，转而追问一个完整的音乐事件最少需要几步：留隙、填入、回接，并给出这个回路何时算闭合。

④《调谐的角落化》占制度层。它不追问事件本身，而追问事件的发生条件如何被“听即校正”的制度化历史一步步压缩，以及被压缩到了哪几个角落。

③《守其所不知》（接住之声）占教育层。它追问在教师认知的尽头处——面对一个连教师自己都认不出的声音——凭什么还能接住，以及接不住时具体是怎么失手的。

四者在时间上首尾相接：①说明“我们”如何在声音里成形；②说明这个“我们”靠什么样的最小回路维持；④说明这个回路的发生条件如何被制度收窄；③说明在收窄到极限的教育现场里，凭什么还能把它重新打开。

（二）这份分工在什么条件下应被推翻

最危险的一格是①与②。若把“延迟是否坍塌”这一个变量（①第三节给出的±100 毫秒判据）从两篇中同时去掉，而①与②对同一段演奏仍然逐条做出相同的

预测，则①应当并入②，本组该是三篇而不是四篇。这一格本组自认最险，不回避。

④的承重结构可被抽掉。若④所描述的压缩效应，在没有作品概念、没有记谱、没有考级的口传传统里同样出现，则"制度史"就不是④的承重结构，④应改写为①②的一个推论，而不是独立一层。

③可能不成立独立一层。若③的每一次"接不住"都可以还原为②意义上的"回接失败"，或还原为作者本栏前作《接应》意义上的"通道未打开"，则③应被吸收。③在自己的附论一（三）里给出了这一格的判据。

四篇合起来只提出一个主张：音乐是两个身体之间一件会发生也会中断的事。四个命名如果不能各自守住上面那条被推翻的条件，它们就该合并。

附论一·最近邻切割（续）：作品概念、噪音与灵韵

正文第三节已把共在调谐与既有音乐学概念做过一轮切割，参考文献里也已列出舒茨、西格、布莱金、德诺拉、图里诺等真实条目。但本文第二节所讲的那件事——记谱与标准音高把"原作"固化、把每一次表演降格为副本——在音乐哲学里有一位公认的正主，而全文零次点名。

（一）与戈尔《作品的想象博物馆》的切割——本篇最重的一处空门

莉迪亚·戈尔论证：**作品概念**在 1800 年前后才取得规约性地位，此后演奏被重新定义为对作品的忠实再现，整个音乐实践据此被重编。本文第二节所说的"本体论翻转"与之几乎同题。不切开它，本文会被读作戈尔命题在中国音乐教育场景中的一次应用。

分离线：**戈尔处理的是概念史**——作品概念如何取得规约地位；**本文处理的是发生条件的压缩**——每一个声音发出的瞬间被植入一道"证明它值得存在"的评价程序。两者可以分离：评价程序的植入不必以作品概念为前提。

可裁决格：没有作品概念的口传传统里的严格正误规范。某些仪式音乐与师承体系中存在“错一个音即为不敬”的强规矩，却从未有过作品概念，也没有记谱与考级。戈尔预测：无作品概念处不应出现副本化；本文预测：只要“听即校正”被制度化，调谐照样被压缩，哪怕从未有过作品概念。若田野证据显示这类传统里调谐现场并未被压缩，本文应并入戈尔，“制度史”不再是本文的承重结构。

（二）与阿塔利《噪音》的切割：压缩的动因是复制还是评价

阿塔利把音乐史读作政治经济学的预言，其“重复”阶段——机械复制——解释了仪式性音乐的消退。这与本文的叙事共振，因此必须划界。

分离线在**动因**：阿塔利押在复制与商品化上，本文押在评价程序上。这两个变量在历史上高度相关，但可以分开测。

可裁决：比较两类场域——录音流通频繁而无考级体系的地区，与录音稀少而考级密集的地区——的调谐现场指标。阿塔利预测前者压缩更重，本文预测后者。这是一项现在就能做的比较研究，两条预测正面冲突。

（三）与本雅明「灵韵」的切割

灵韵讲的是作品的此时此地被复制技术剥夺。分离线：灵韵的丧失以“曾经有一个本真的原件”为前提；本文讲的是一个**从来就不是作品**、也从未有过“原件”形态的双身体事件，其发生条件被收窄。本文因此不承诺任何黄金时代——正文第七节已明确拒绝复古，这一点与灵韵叙事的怀旧结构相反。

附论二·可裁决预测

预测一（遮盖评价的即时效果）。对同一批有稳定被评价史的学琴者，给出一个可信的承诺：“本次演奏不做任何评价、不进入任何记录”。本文预测调谐现场指标——自发偏离的被接续率、事后“有没有一起发生过”的自评——应在**当次**就显著上升，而不需要长期养成。若需要数周才出现变化，则“角落化”应改写为习性理论，本文的即时性表述过强。

预测二（三个角落共享同一个变量）。本文主张深夜、母婴、“不在乎会不会弹错”这三个角落共享同一个变量：评价程序的暂时离线。推论：在这三个场景中测得的自我监控指标应显著低于同一个人在琴房中的水平，**且三者之间无显著差异**。若三者之间差异显著（例如母婴场景独低），则这三个“角落”不是同一个机制的三个位置，本文的统一叙事失败。

预测三（考级作为自然实验）。以考级为断点做前后设计。本文预测：考级前四到八周内，学琴者的自发偏离率与被接续率显著下降，**且考级结束后不自动回升**——因为评价程序已被内化为知觉习性（正文第五节）。若考后一个月内回到基线，则“内化为知觉习性”这一主张过强，应改为情境性压力反应。

参考文献

（原稿自带文献表与一段自我限权声明，原样保留在前；末三条为本轮最近邻切割所补入。）

（本文遵循一项严格的文献诚实原则：只列入作者确有把握其真实存在的著作。凡无法核实具体出处的研究——尤其在当代中国声乐教育制度与民歌传承领域——均不以虚构条目充数，而在正文注释中明确交代其作为“领域内广为人知的公开事实”或“基于多源公开材料的复合性重构”的性质。下文所列文献虽然未一一在正文中直接援引，但构成本文核心论证的学术对话背景。）

Blacking, John. 1973. *How Musical is Man?* Seattle: University of Washington Press.

DeNora, Tia. 2000. *Music in Everyday Life*. Cambridge: Cambridge University Press.

Foucault, Michel. 1975. *Surveiller et punir: Naissance de la prison*. Paris: Gallimard.

Polanyi, Michael. 1966. *The Tacit Dimension*. Garden City, NY: Doubleday.

Schaeffer, Pierre. 1966. *Traité des objets musicaux*. Paris: Éditions du Seuil.

Schön, Donald A. 1983. *The Reflective Practitioner: How Professionals Think in Action*. New York: Basic Books.

Schutz, Alfred. 1951. “Making Music Together: A Study in Social Relationship.” *Social Research* 18 (1): 76–97.

Seeger, Charles. 1958. “Prescriptive and Descriptive Music-Writing.” *The Musical Quarterly* 44 (2): 184–195.

Small, Christopher. 1998. *Musicking: The Meanings of Performing and Listening*. Middletown, CT: Wesleyan University Press.

Turino, Thomas. 2008. *Music as Social Life: The Politics of Participation*. Chicago: University of Chicago Press.

本文的证伪条件：如果一个专业的音乐教育机构，将“不携带任何校正意图的共享听”作为一门正式课程纳入其教学大纲，并以制度化的方式持续培养这种听觉能力，使其毕业生能在自己的教学与演奏中稳定地复现调谐现场，则本文关于“校正制度化系统性地压缩了调谐发生条件”的叙事即被证伪。本文作者诚恳期待这样的反证出现。

- Goehr, L. (1992). *The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music*. Oxford: Clarendon Press. (本轮补入的最近对手)
- Attali, J. (1977). *Bruits: Essai sur l'économie politique de la musique*. Paris: Presses Universitaires de France. (本轮补入)
- Benjamin, W. (1936). *Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit*. *Zeitschrift für Sozialforschung*, 5(1), 40–68. (本轮补入)

编辑说明

编辑改动两处，均不涉论证：①“发现学”这一学派内部的对置用词改为平实说法；②补入附论零、附论一、附论二与三条文献。原稿自带的文献表与那段自我限权声明——凡无法核实具体出处的材料，明确交代其为“公开事实”或“复合性重构”，不以虚构条目充数——原样保留，未作改动。

本文为 SDE 金点子发生器的思想拓展产出：是在特定方法引导下的思想实验与观点探索，用于拓展思路、激发思考，不构成正式学术研究成果或事实性结论；如需正式引用，请自行核验并遵守学术规范。