

不在声音里，亦不在你我之间：音乐作为身体间的“共调”事件

作者 高于涵 · SDE 学员 · 约 1.3 万字 · 发表于2026年7月31日 · 加固版

摘要

当代音乐本体论从“作品”向“事件”的转向，已将音乐从声音客体移到互动过程。然而，其核心模型——演奏者发出邀请，听者以身体偏移回应——暗中保留了一个未经审查的先验预设：音乐事件的参与者，首先是已经分立好的、以“发送”和“接收”为基本动作的两个主体。本文从一组让这一预设失效的边界现象出发，指出这一局限，并逼出一个更底层的新概念：共调。共调不是两个主体之间的“接通”，而是所有在场身体在一个先于意识的主体间张力场中，经由声音媒介实时相互调节彼此的神经生理状态，共同生成一个共享的节奏-张力-情感形态的动力学过程。本文在界定此概念后，给出其发生与中断的操作性判准；继而将其与现象学“身体间性”传统进行正面对话，厘清二者“同出而异路”的关系；并推衍其对音乐教育、表演哲学及录音体验的理论后果。最后，正面回应来自神经科学和文化相对主义的质疑，并诚实陈述本文的证伪条件。

关键词：音乐本体论；事件哲学；身体间性；共调；接应

一、引言：当我们同时哼出那一段旋律

在人类追问音乐本质的两千多年里，有一个问题几乎从未被正面提出。不是“音乐是什么”——对此的回答早已汗牛充栋：音乐是数的和谐，是情感的语言，是乐音的运动形式，是有意图地组织起来的声音。而是：“在音乐发生的那一刻，我们是谁？”

“我们”这个字，看上去像是一个修辞上的顺手捡起。演奏者在弹，听者在听——他们是两个人，不是一个“我们”。你弹你的，我听我的，在语法上毫无毛病的两句话，精确地標示出了这件事在绝大多数理论语言中被默认了的位置：音乐，是两个已经分立好的主体之间发生的一件事。从现象学到神经美学到音乐表演研究——也包括近年来最重要的“事件论”转向——它们共同的、未被充分主题化的地基，正是这个“先有分立，后有接通”的默认证。

本文的工作，就是把这个地基挖出来，追问一句：它真的是地基，还是只是事后被安上去的脚架？

1.1 被两个日常现象逼出的追问

让我们从两个几乎每个人都可能经历过的日常场景出发。

第一个场景。你在洗澡，水声哗哗，你张口哼出了一段旋律。你并不是在“唱给谁听”——门关着，房里没有别人。你甚至没有意识到自己在哼。你只是在那个热水冲刷背脊的当下，任由一股说不清的推力，从喉咙深处揉出一段声音来。

第二个场景。你和一位老友并肩坐在沙发上，电视关着，初秋的暮色从窗帘的缝隙里沁进来。沉默了一阵之后，你们忽然同时哼出了同一段旋律。哼出来之后，你们同时愣住了，然后相视而笑。那一刻，你们都被“我们竟然同时哼出了它”这件事本身吓到了。

在第一个场景里，“发送”与“接收”全无用武之地。在第二个场景里，那段旋律是谁“递出”、又是谁“接住”的？没有先后，没有主从。在它被你们意识到之前，它已经同时在你们两个身体里升了起来。那个“我们”，在你和我尚未来得及分清楚的瞬间，就已经在了。

1.2 从现象到概念：排除更简单的替代解释

这两个场景不能直接等同于“因此存在一个先于你我之分的共调场”。在抵达那个概念之前，必须先排除对它更经济的替代解释。

两个场景共同指向一个特殊的结构：在音乐被意识到之前，它已在身体里发生；而当它被意识捕捉到时，“谁先发起”这个问题根本找不到答案。在洗澡的例子中，那段旋律是被身体在那个当下的整体张力状态逼出来的，而非由意识决定。在两人哼唱的例子中，结构更尖锐：那段旋律，是从两人共享的同一片暮色、同一组沙发带来的身体微倾、同一个房间空气流动的缓慢节律所共同编织出的共享张力轮廓中，同时涌现出来的。

对此最直接的替代解释是“偶然同步”。但这一解释无法说明当事人“被吓了一跳”的反应。偶然同步引发的是“哦，巧了”，而不是“这怎么可能”。后者意味着当事人察觉到了一种比偶然更深的确定性——“它本来就该在那一刻从我们两个身体里一起出来”。“应该”一词，在事后追述的“被吓到”中，是确凿的。

更精密的解释来自镜像神经元理论：听者的运动前皮层在无意识中模拟了对方可能做的动作预备，当预测被命中的那一刻，听者便同步启动。这解释了“我如何知道你将要做什么”，却仍然无法解释“被吓到”中的那个“我们”。如果“我”全程保持在“我正在模拟你”的认知框架里，“你的动作准确命中了我的预测”只会带来满足，而非震惊。震惊，是分类框架失效的信号。被同时哼唱吓到，是“你和我”这个分类框架在那一瞬间暂时失效的信号。

因此，这两个场景所揭露的，是存在于两个身体之间的一种比个体意识更深的动力学结构——它在“你”和“我”尚未分立之前就已经在运作。这个结构没有现成的名字，但它既然反复出现在那些让“发送-接收”模型失效的边际地带，它就应该被命名、被界定、被检验。这就是“共调”被逼出来的路径。它不是一个更漂亮的名字，而是一步不走出去不行的概念推进。

1.3 已有的“事件论”推进到了哪一步，又停在何处

过去二十多年里，音乐本体论的讨论从“作品”转向了“事件”。这一转向最具代表性的工作是斯默（Small, 1998）提出的“musicking”概念：音乐不是物，而是一项活动，是所有人在一个音乐场域中做的一切事情的总称。这项工作的功绩不可低估，但它停留在了场域描述的层面：它告诉我们谁在做什么，却没有追问，这个场域的核心是怎样运作的——两个人“一起”做音乐时，那个让“一起”从一个漂亮概括变成真实事件的微观引擎，究竟是什么？

本文所属的脉络，正是对这个缺口的直接回应。在这个方向上的先行工作中，“接应”概念被提出，用以指认演奏者与听者之间“递出邀请—身体偏移—再回应”的微观循环。从musicking到接应，迈进的重要一步在于：它不再只是

说“音乐是大家一起做的事”，而是试图把那个让“一起”真正发生的微观引擎指认出来。

但这一步有没有停在一个新的、尚未被审查的地基上？本文认为：有。接应模型描述的过程，其起点是两个已经分立好的角色——一个递出，一个接住。二者的分立，在接应的描述中被当作给定的事实，没有被追问过“这个分立本身是如何发生的”。在“你递出第一个音”之前的那一瞬，我的身体早已随着你吸气的微小肋间肌收缩，提前做了一次微弱的肩部下沉——这个下沉，不是在我“决定”听你之后才发生的。它在我知道“我”在听“你”之前，就已经发生在了“我们”的身体复合体里。而你的第一个音，也不是从“你”开始的——它是在那个已被你我共同的呼吸节奏、体温交互、空间张力充分预调节之后的身体里，被逼出来的第一个动作。

由此可以逼出一个比“接应”更底层的概念。它不再把你和我预先分开，再去描述“我们如何接通”。它直接从那个先于你-我之分的、被身体间的前意识调节所构成的共享张力场出发，把音乐事件的发生，理解为这个场自身的自我调节和自我显露。这个概念，本文命名为共调。

1.4 本文的核心判断：从“接应”到“共调”

共调的操作性定义如下：在音乐事件的现场，所有在场者的身体，在一个先于个体意图的身体间张力场中，实时地、相互地、不可逆地调节彼此的神经生理状态，并在此过程中共同生成一个为这个唯一的当下所独有的、共享的节奏-张力-情感形态。这个调节过程，本身就是音乐事件的本体。

这一概念需要当场澄清几点。第一，共调不是接应的对立面或替代者，更不是它的浪漫化“前奏”。接应是共调在角色已分化条件下的一个特例，但二者的关系不是简单地把时间线往前推一点。在高密度共调状态下，接应所依赖的那条“递出—偏移—再回应”的线性因果链，会被共调的涌现整体性所吸收：每一次身体偏移和声音应答，不再可以被切割成“你先我后”的微循环，而是同时从那个共享的张力形态中涌现出来的并行产物。第二，共调不预设意识的参与。第三，共调是实时

的——它不被事后追认，必须在音与音之间的毫秒级时间窗口内实际发生，否则那一刻便中断了。

在接下来的论证中，本文将从四个层面将这个概念推进为可操作的判准。第二章从边际现象出发，逼出“共调”的不可还原性。第三章正面展开共调的微观结构，给出其发生与中断的判准，并以一次钢琴教学现场演示其判别效力，同时给出共调与接应的可操作区分。第四章将共调与现象学的身体间性传统进行正面对话，厘清二者的关系，并在此基础上推衍其对音乐教育、表演哲学及录音体验的后果。第五章正面回应来自神经科学和文化相对主义的质疑，并诚实陈述本文的证伪条件。

在进入论证之前，最后一点需要交代：本文追问的是“音乐事件如何可能”的最低条件，而不是“音乐应该是什么”的规范性主张。它从那些让既有概念变得不够用的边缘现象出发，去追溯那个被默认地基所遮蔽的更底层。这不是对已有理论的否定，而是在它们铺好的路上继续往前走一步——松动一个至今未被松动的先验预设：在音乐事件的开端，你我就已经是分开的。

二、逼出“共调”：一组让“你我之分”失效的边际现象

一个新概念是否成立，取决于它是否能解释一些既有概念解释不了的现象。对于共调而言，这类现象的共同特征是：在音乐事件发生的当下，“发送者”与“接收者”的区分不再成立。

2.1 母子交互发声：谁在模仿谁？

在发育心理学中被充分观察的母婴交互发声现象，是共调的第一个硬证据。斯特恩（Stern, 1985）在其经典研究《婴儿的人际世界》中描述了“活力情感”（vitality affects）的跨模态同步：母亲与三个月大的婴儿之间的声音交互，并不存在一个清晰的“谁先开始”的线性链条。“咕咕”与“唧呀”声之间，在音高轮廓、节奏模式和强度渐变上存在一一对应的互调关系，且发生得极快——远在有意识的模仿决定之前。Trevarthen（1979）提出的“原初主体间性”概念表

明，人类婴儿从出生起就具有直接感知并响应他者身体节奏的能力，无需以认知上分离的“我”和“你”为前提。近年的多脑扫描研究也提示，这种互调的脑间同步常落在200毫秒以内，甚至趋近于零延迟的随机分布，意味着两个神经系统在一开始就被耦合成了一个临时的功能复合体。

接应的语言怎么描述这个场景？勉强可以说，母亲“递出”了一个声音邀请，婴儿用咿呀声“接住”了它。但这个描述的真实性有多大？在母亲发出第一声“咕”之前，婴儿早已用前声音的面部表情发出了信号，而母亲的身体已经接住了它。所谓“母亲先发出咕声”，是对一个首尾相衔的环做出的武断剪切——你把它剪开，然后指着被剪断的那一头说“这是起点”。共调的叙事则无需剪开这个环：声音互答，是这个“母亲-婴儿”身体复合体的自我调节运动在听觉域中的显露。这个叙事更经济，也更贴近实证发现。

2.2 被自己吓到的齐唱：“我”消失了

本文引言中描述的二人同时哼唱的场景，其理论分量正在于事后“被吓到”的反应。民族音乐学的田野记录也提供佐证：在西非集体鼓乐中，老练的鼓手们反复报告，某个节奏型会“自己冒出来”，没有人听到谁先开始它，但当它冒出时，每个人都在同时打它，并且立刻认出它是对的。

接应的语言在此彻底失效。失效的原因不在描述不精确，而在于接应的底层预设——先有分立，而后接通——在这种类型的发生中根本不存在对应的经验内容。这里没有任何一方“邀请”了对方。那段旋律在两个身体里升起来的时刻，是同一个时刻。它不是“被接通”的，它是“被共调”出来的。在长期共享聆听记忆与身体协调经验后，两人的自主神经系统形成了高度灵敏的共振结构。在某个特定的环境触发下——一次不自主的深呼吸、窗外光线的变化——这个共振结构在低于各自意识阈值的时刻，自动产生了一个共享的神经生理激发模式，并同时外化为可听的声音。它不是“我”发出被你听到的；它是“我们”这个暂时形成的身体间复合体的自我调节运动，在那个可听的瞬间，把它自己展示给了“我们”。

2.3 洗澡间里的忘我哼唱：“你”还未出生

独处时的哼唱，常被解释为“自我倾听”——一个内部分裂的“我”在为自己演奏。这在逻辑上并无漏洞，但在理论上却是冗余的。身体本身就是一个高度自耦合的振动系统：发声器官与听觉、呼吸与心率、肋间肌与喉部，在没有任何他者介入时，已构成一个微型的共调回路。当你在洗澡时哼出那段旋律，不是“你在给自己演奏”，而是你的身体在借助声音实时调节它自己当下的整体张力状态。独处的音乐事件，只是这个复合体的自我调节在听觉域中最鲜明的呈现。当另一个身体进入场域后，这个回路被扩大和复杂化了，但其基本的本体论结构并未改变——它从来就不是“两个分立主体之间的沟通”，而是一个动态的“身体间复合体”的自我调节。

2.4 从边际回到中心：共调是常态，角色分化是特例

这些边际现象共同指向一个结论：在“你我”分立之前，已存在一个前意识的、身体间的共调过程。音乐事件的发生，不是分立后的主体努力去“接通”，而是这个共调场在运作中借由声音把自己推到可共享感知的听觉表面。既有事件论证——从musicking到接应——全部以已完成角色分化的表演场景为默认原型来建立。从本体论上看，这种可见性恰恰是遮蔽：它把那个在分化之前就已运作的共调场，推到了理论的盲区。

现在可以更完整地表述本文的核心判断：音乐事件的本体论位置，不在声音里，不在你或我之中，也不在你我之间的那条“连线”上——而在于那个先于你我之分、将我们一并裹挟其中并借由我们的身体持续自我调节的“共调场”之中。接应是共调在功能已分化条件下的特例，不是音乐事件的最小必要结构。共调才是那个最小必要结构。

三、共调的微观结构：它如何运作、如何可见

如果一个概念不能在真实的音乐场景中被操作性地识别，它就只是漂亮的文字。本节的任务是将共调从一个本体论概念推进为可判别的过程。

3.1 共调的发生判准

在一个有 ≥ 2 个身体共处同一物理空间的当下，以下四个条件同时满足，即可判定共调正在发生：

(a) 双向自发对称。参与者的身体状态（呼吸率、心率变异性、姿势肌张力、微表情流）出现非随机的互趋同现象，且这种互趋同不可还原为“一方模仿另一方”的单向因果链。

(b) 声学驱动而非意识驱动。上述身体互趋同时序上以声音事件（音头、乐句节点、休止、和声转折）为触发锚点，而非以参与者的意识决策。事后询问时，参与者无法报告自己“决定在那一刻改变呼吸”或“主动调整姿势”。

(c) 高精度时间锁定。身体互趋同的启动时间窗口，在音乐结构节点的 ± 300 毫秒以内发生。这一窗口的设定基于人际同步的实证研究——如Lindenberger等人（2009）对吉他二重奏中脑间同步的经典工作。超过此窗口的身体变化，更可能是一般性的环境适应或注意力漂移，而非音乐事件驱动的共调。

(d) 不可逆路径偏移。此判准是本文最具原创性的理论假说，需稍详展开。

在共调状态下，任何一个参与者自发产生一个微弱的偏离——比如比合奏伙伴稍早的换气、力度上极微的减弱——至少会被另一个参与者的下一音乐动作以一种事后可被观测到的方式所“接续”：对方的下一动作在客观上沿着那个偏离打开的新方向推进了，而非将其抹平回原轨。这个“接续”不一定成功；偏离也可能无人接，共调在那一刻中断。

此处需要正面处理一连串尖锐追问：偏离为何是“不可逆”的？本文提出一个明确标注为有待检验的机制推测：当偏离发生，其声学痕迹（力度、时间、频谱上的微差）已在空气中的声场里留下了特定的物理印记；接续者在下一个音上的动作，其运动预备已在亚意识层面将这个偏离纳入了神经-肌肉系统的参考副本

(efference copy) 之中，因此他接续的动作是从已被更新的声场状态出发的，而非从原轨出发的。这个推定的机制可以解释为什么偏离不能被“抹平”——因为运动预备自身已经被偏离开了。

“接续”需要多高的成功率，才算是判准满足？本文设定一个宽泛但可操作的最低窗口：在一个持续的音乐段落（不少于十个乐句节点）中，若被第三方观察者编码为“发生了偏离”的事件里，接续的发生率显著高于随机接续基线（随机基线定义为：在音乐结构上等价的节点上，接续者做出沿偏离方向推进的概率不高于50%），则可判定判准(d)满足。

接续的物理标志是什么？本文给出一个可被客观观测的候选指标：在偏离发生后的紧接下一个音头上，该音头的力度矢量（在三维声场或MIDI力度坐标系中）在方向上更靠近偏离所指向的“新方向”，而非更靠近该节点在乐谱或该演奏者惯常执行中的平均力度值；同时，该音头的准备动作（如指尖下落前的悬停时间）与此演奏者在无偏离条件控制下的平均准备时间之间存在可测量的差异（缩短或延长，依偏离性质而不等）。重要的是，这个差异并非事先计划好的：当演奏者事后被问及时，他们不应能准确复现这个刻意做任何“决定”的过程。若能，则是接应而非共调。这些指标不要求仪器级别的精度即可被有经验的音乐家观察者以现场记录、事后录像微分析的方式判定，它们因此是第三方可操作的。

3.2 共调的中断判准

以下三种情况任一出现，即判定共调过程中断：

(a) 自我监控侵入。任一参与者开始以“我发出的声音是否合乎某个外在标准”为依据来调整自己的音乐动作。其可观测标志包括力度的非音乐性收窄、节奏微延迟、以及音色中高频泛音的突然衰减——后者与面部及喉部肌肉的防御性收紧有关，这在关于表演焦虑的实证研究（如Yoshie et al., 2009）中已有清晰记录。

(b) 注意力逃逸。任一参与者的注意重心从声音事件本身滑向与此刻声音无关的内部思绪，可被事后报告为“我刚才走神了”。

(c) 截断式中断。外部环境因素强行切断正在进行的声学驱动链条，且中断后参与者无法在紧接的下一个乐句中恢复此前已在运行的共同张力形态。

需要说明，中断不等于失败。共调总是在“接通→中断→再接通→再中断”的循环里运行的。真正的音乐事件不是从头到尾完美同步的神话，而是那两个身体在一次次中断后仍一次又一次重新找到彼此的动态过程。

3.3 一次教学现场的演示：共调如何被自我监控短路

为使上述判准直观可感，本节借一次钢琴教学现场进行演示。一位有近二十年教学经验的老师，在琴房里为学生示范一首肖邦前奏曲。学生坐在老师右侧约一米处。

第一次演奏（教学示范版）。老师弹出第一个和弦。力度饱满，节奏精确。但与此同时，她全部的注意力都集中在“这个音对不对？学生有没有看清楚我的指法？”这一自我监控指令上。她的手指运动变异性显著收窄，颈部肌肉微紧，呼吸被缩短以便在弹的同时预备口头解说。声音上，力度的起伏被压缩在极小的安全区间内，节奏近乎节拍器般规整，音色中缺乏活生的泛音变化——声学分析可预计其在触键后的高频泛音衰减速度显著较快。学生在此全程中无自发同步的身体互趋同——呼吸稳定，肩部静止，面部停留在礼貌性关注。事后他至多评价“弹得很好”。这是一次精确的声音传输，但共调的四条判准全部缺席。被“示范”这一社会角色引入的自我监控，在声学驱动触及身体之前就把它短路了。

第二次演奏（空琴房版）。学生离去。暮色渗进琴房。老师在某个瞬间不自觉地再次把同一段前奏曲弹了一遍。这一次，她的腕关节在触键前的准备窗口中变得更柔软，动作不再由预置的运动程序严格前端驱动。在某个乐句本应渐强的地方，她却做了一个极轻微的渐弱——因为她自己的身体在那一片暮色里，无意中选择了那个更向内收的姿势。从声学上看，这是在规定为渐强的节点上发生的方向反转，且反转发生在触键后极短时间内，这意味着它不是意识层的从容选择，而是身体对那个共享环境的直接响应。

如果此时有另一个身体在这个空间里——比如走廊上驻足的学生——把耳朵放在这个动作流的附近，他的身体将自动被卷入。听者的身体会在那个意外的渐弱处同步产生肩部微沉，这个微沉在房间的微弱声场变化中被演奏者无意识地感知，进

而在她的下一个触键中被“接住”：下一个和弦，会比她惯常的力度又轻了极微的一点。这次的轻，不是她决定的——它是共调回路在那个瞬间借她的手做的一次自我调节。判准(a)(b)(c)(d)全部满足。

这同一个老师、同一首曲子、同一架琴的两次演奏，其间的差异在声学上可测量，在体验上可判别为“有没有共同发生过的厚度”。这个案例的论证分量不在于生动，而在于它所描述的行为差异可以用两个控制条件操作性地分离。本文以此演示：共调不是这些差异的诗化解释——它就是这些差异本身所共同构成的那个整体的、不可切割的过程。

3.4 共调与接应的可操作区分

本文提出的最根本辨别维度是：共调不以动作-反应的时间延迟为其特征标记，而是以那个延迟的坍塌为特征标记。在接应所描述的状态中，听者的偏移与演奏者的再回应之间存在一个可识别的时间间隔（200–400毫秒量级）。而在共调高度激活的状态中，这个延迟会坍塌——它不再作为可被独立测量的因果间隔而存在。听者的微倾和演奏者的下一个音在时间上的先后关系，进入±100毫秒以内的零附近随机分布；在此精度上，人类无法建立“A引起B”的时间知觉，因果箭头消失。此时发生的，不再可以被描述为“你偏移，然后我接住”，而只能被描述为“两者同时从同一个正在升起的共享张力形态中被逼出来”。共调发生时，不是两个人在“快速地接应”；而是接应的因果链本身，被共调的涌现整体性所吸收。

四、共调与身体间性：同出而异路

至此，本文已经正面建构了共调的概念骨架。但在把它推衍到教育、表演和聆听之前，必须完成这次哲学对话——这个对话的缺席，是本文初稿最严重的理论疏漏。

4.1 现象学早已在那里了

本文从头至尾攻击的核心预设——在音乐事件的开端，你我就已分开——恰恰是现象学传统，尤其是梅洛-庞蒂，整个哲学生涯都在对抗的东西。梅洛-庞蒂在

《知觉现象学》和《可见的与不可见的》中反复论述的“身体间性”（intercorporeality），明确指出：婴儿在尚未形成“自我-他者”区分之前，就已经通过身体图式直接感知并回应他者的身体意向。在存在论层面上，“我”与“他者”从来就没有真正被分隔开过。胡塞尔在《笛卡尔沉思》中关于“原初的结对”（primordial pairing）的论述——自我与他者的身体在先于意识区分的被动综合层面上，就已结对为一个互属的共同体——也在处理同构的问题。

既然本文在批评一个已经被现象学深刻批评过的先验预设，那本文与现象学之间是什么关系？共调，是一种新的发现，还只是把梅洛-庞蒂的“身体间性”换了一套生理学术语又讲了一遍？

4.2 同出而异路：一个被忽视的区分

回答是：同出而异路。

“同出”：共调与身体间性，确实共享同一个哲学出发点——二者都拒绝把“已经分立的主体”作为分析人与人之间经验的原初单位。现象学将此视为存在论事实（我们从未真正分立过），共调将此视为操作起点（在音乐事件中，分立是后发的、已被发生出来的，而非先在的）。在这一点上，本文从现象学中受益，并坦率承认。

“异路”：梅洛-庞蒂的身体间性描述的是一种存在论层面的、持续的背景性互属状态——它是“我”与“他者”在知觉世界中无时无刻不在共享的、沉默的、不可分割的基底。但共调描述的是一种有判准的、可发生也可中断的、毫秒尺度的实时动力学事件。它不是一个恒定的背景条件，而是一个在时间中每时每刻都可能接通也可能断开的动态过程——一个可以被自我监控切断（§ 3.2a）、被注意力逃逸悬置（§ 3.2b）、被环境干扰截断（§ 3.2c）的发生。梅洛-庞蒂的“身体间性”不会在某一刻“中断”——它作为存在论事实，始终在那里。但共调会。这个区分，在现象学的框架中尚不具备，也不可能具备——因为现象学的任务不是为身体间互属状态设定毫秒级的、可被第三方观测的中断判准。

因此，本文对现象学传统的推进——而非否定或重复——正在于此：它将现象学对“你我之分”先验预设的批判，从一个存在论的背景条件，推进为一个在时间中有生灭、有判准、可中断、可人为关闭和重启的动力学事件。理论收益是双向的：共调因这一对话而明晰了自己的哲学位置，不再是“精致的综合性命名”；现象学的身体间性也因这一推进，获得了一个通向微观时间尺度、可被经验研究操作化的大门。

4.3 由此而来的几个关键后果

这一区分一旦奠定，立刻收紧了几处在初稿中松动的论述。

首先，本文不再——也不能——再将物理共场中的双向共调称为“完整形态”。不存在一个先在的“完整共调本质”，然后录音是它的“残缺版”。物理共场和录音体验，是在不同的身体-技术约束下被生成的两种不同的共调形态。物理共场同时包含向外接住与向内生长两个回路；录音则是一个去除了实时互调回路后、因偏移无处释放而被迫逼出独立的向内生长回路的另一种结构形态。二者之间不存在“完整-残缺”的等级关系，只存在结构组件的不同配置。这一修订，使本文对录音的本体论定性不再滑回“把结论当起点”的老路的“本质先于形态”的叙事。

其次，共调并不绑定“物理共同在场”这个媒介条件为形而上学本质。它只绑定于“两个身体之间的实时互调回路的完整性”。物理共场是这个回路在当前技术条件下最可靠的保障，但它不是这个回路本身。如果有朝一日，某种未来技术能在异地身体之间完整重建同一密度的双向实时互调回路——届时，本文对“物理共场”的特殊主张可以安全地放弃；但对“共调回路是音乐事件的本体”这个核心主张的放弃，绝不依赖于技术乐观主义的证伪，而只依赖于另一种证明：存在这样一个事件，其中共调的四条判准全部不满足，但参与者仍持续地、跨情境地报告“这就是一次完整的音乐事件”。那才是本文真正会认错的唯一情况。

五、推衍：如果本体是共调，教育、表演与聆听意味着什么

以下推衍严格从本体论位置出发：它只问“如果本体如此，那么实践的结构性后果是什么”，而不以规范性语言替代结构性分析。

5.1 音乐教育：在“标准化”的微观矛盾中，共调如何被一步一步截断——又被逼成新形态

在主流音乐教育体系中，核心目标可概括为“准确的再现”：学生被训练去准确再现乐谱上的音高、节奏和表情标记。这个目标的直接操作后果，是在每一个音的决策点上，学生被训练去参照一个外在于此刻的标准——乐谱、老师的示范、评分表——而非参照自己身体当下的真实冲动，以及与现场另一身体间的实时互调状态。

但本文在此不想只讲一个“原初完好—被败坏了”的丧失故事。发生学要求追问得更彻底：这个以标准化为核心的教育体系，其自身是如何在特定历史张力中被发生出来的？它的“标准”与“评分表”作为调节装置，解决了什么问题，又制造了什么新矛盾？

现代音乐教育体系的标准化进程，在相当程度上是应对“音乐民主化”需求的一个历史产物。当音乐教育从师徒制的个体传承转向班级授课和大规模评估时，它必须解决一个此前不存在的“如何公平地、可比较地评估一大批不同身体条件的学生的音乐表现”这一共同体内部的政治-伦理矛盾。“准确再现”作为一个调节装置，在此被发生出来，不是因为它更接近音乐的本质，而是因为它回应了一个特定的历史张力：大规模教育的公平性诉求与音乐经验不可通约性之间的矛盾。评分表以及它所框定的那五项标准（音准、节奏、音色、风格、完整性），是这个矛盾被暂时结算出的一个临时稳态。

而它一旦被建立为一个稳态，就开始以它自己的逻辑反向塑造身体实践，制造它自己的新矛盾。当一个六岁的孩子第一次坐在钢琴前，用手指去触碰琴键，感觉键下沉的反作用力从指尖传回手腕——他在自发地与这件乐器建立共调。然后老师

握着他的手，示范了“正确的”手型。这个“正确手型”立刻切断了他那个在键盘上以指尖皮肤的微变触键角度来感受键阻力的自发性调谐回路。他的手掌肌肉因维持被指定的姿势而进入恒定的轻微紧张，这个紧张状态让他再也无法分辨出D键和E键之间那些极微细的、只能通过柔软的指关节才能感知到的阻力差异。这是他身体里第一个微小的矛盾：自发的探询冲动与一个外在于此刻的姿势标准之间的冲突。这个冲突被结算出的结果，是前者被抑制。

此后若干年，这种截断以无数面目反复出现。他因自发延长一个休止符而被纠正“这里数拍子”；他因在练琴时随手改了谱子中两个音而内心涌起“真好”的感觉，但那个感觉迅速被他内化的评分表自我诊断为“不务正业”。这些日复一日的微观矛盾，逐步在他身体里建立起一个持续的紧张：朝向此刻的互调冲动，与朝向一个不在场的标准的自我监控，构成了他音乐身体内部的一对常在张力。

但这并不意味着“共调能力被毁灭了”。那个被二十年专业化训练重塑的身体，也许不再能产生“洗澡时随口哼出”的那种未经训练的自发性旋律，但它是否在另一些不为外人所知的微观时刻——比如在极度精密的指触控制中，当某个音色被推到临界边缘的那个瞬间——经历了另一种全然不同的、未经训练的身体绝不可能进入的共调体验？那种体验，发生在演奏者与他正在操控的那个发声体之间，在毫秒尺度的动作-声学反馈回路上，其互调密度和精度或许远超任何日常的二人合奏。因此，本文对音乐教育体系的结构分析，不是“教育毁掉了共调”，而是：这个体系在运转的每一分钟，都在以标准化手段切断身体自发的共调回路，但与此同时，也在它自己的严苛条件下，于极个别、极边缘的时刻，逼出了另一种高度特异的共调形态——一种只属于高精度控制的、在极窄的动态窗口中发生的、不为未经训练的身体所知的共调。

这个分析不指向“回归初心”的怀旧方案。它指出的结构性问题不是“应该回归本真”，而是：这个体系如果要在客观上成为“音乐事件的发生场所”，它需要调整的不是“理念”或“口号”，而是它如何在每一次纠正动作中，给那个自发的身体探询冲动留下一丝不被立即判为“错误”的缝隙。

5.2 表演哲学：不是“送”出去，而是“调”进去

在“作品-表演者-听众”的三分框架中，表演被定义为将音乐信息“传达”给听众。共调视角完全翻转这幅图景：表演者的第一个音不是“把她内部已经完成了的音乐送出去”，而是“把此刻她体内那股还未成形的张力抛出去，作为一个试探”。这个试探被听者的身体接住并产生微弱的偏移后，表演者的下一个动作就不再是她自己单独可以做出来的事——它是共调场借她的手做出的一次从未被独自做过的动作。

对表演训练的结构性含义在于：一个演奏者要让演出有可能进入这种状态，她所需要的不只是“更深入地理解乐谱”，而是必须有能力在上台的那一瞬间，有意地松开她对“将要发生什么”的预置控制，把一部分控制权交给那个正在实时生成的高密度共调场。她必须把自己的身体训练成一个足够敏锐的声学共调工具——不仅要能读出她自己身体的冲动，还要能读出那些从几百具其他身体中涌向她的、无声但有力量的偏移。这意味着一套不同于“演奏技术精确化”的训练内容。在此，本文给出一个具体的、可被表演者拿去实验的微观训练建议：在合奏排练中，有意识地在某一个你本来已完全确定力度的乐句节点上，延迟到触键前约200毫秒时才做出力度决定——把这200毫秒交给你的身体，去读对方此刻的呼吸与肌肉微张力，让它在这个窗口里推你往某个你未预置的方向走半步。不需要你做对，只需要你不阻止你的身体已经在做的微调。

5.3 录音体验：共调的形态重构

此处必须回应的一个逻辑后果是：如果共调要求在身体间的实时互调中发生不可逆的路径偏移，那么录音——这种听者的身体独自在场、而演奏者的身体动作早已被封存完毕的听觉体验——在本体论上处于什么位置？

本文的回答是：录音是共调在技术条件下的一种独立的形态重构。在录音的聆听中，听者的身体仍然发生真实的共调——呼吸会被乐句的衰减推着变浅，肩膀会在休止处微沉。共调回路在他自己的身体内部完整运转了一次。但它缺失了回路的一半：那个微沉的肩，没有被另一个身体接住。演奏者的下一个音早已封存在介质

里，不会为任何听者的呼吸而偏移。这意味着，录音中的共调，是有去无回的。你的偏移无处可去，于是它不会像在现场事件中那样被化解、接住、消耗；它就悬在你的身体里。而正是这个“悬着”的状态，打开了录音体验中一个在现场事件中极难出现的新回路：因为你发出的偏移无人来接，它便在你日后的呼吸、步伐、沉默里反复回响、自我发酵，并最终与你的其他经验化合，长成某种完全属于你自己的东西。

现场事件的完整实时互调结构，迫使偏移在毫秒尺度内被回应和消耗，因而较难被个体内化为长期心理结构；录音则让那个无处可去的偏移被迫进入身体的长期记忆回路，在反复重放中被一层层重新编码、重新赋义。这不是共调的“降解”，是共调在技术条件下被逼出的一种全新的存在形态——它关闭了向外接住的回路，却打开了一个向内生长的回路。本文不哀悼现场，而是在命名一个被技术生出来的、全新的音乐事件形态。

六、回应最有力的质疑

6.1 神经科学的质疑：“共调不过是已知生理机制的现象学重述”

神经科学家可以指出：你描述的身体间互趋同、呼吸-心率耦合、前意识运动模拟——所有这些都是已知的生理机制。共调是不是只是把这些机制用现象学语言重装了一遍？

这个质疑正确地指出了共调所依赖的生理层，但它把三个不同层次的解释混为一谈。第一层是生理机制层：镜像神经元、自主神经耦合——它们回答“共调发生时身体里哪些系统在运作？”这一层必要但不完整。第二层是事件判准层：本文的判准正是在生理机制之上叠加额外的筛选器，将“音乐事件中的身体间共调”从“一般性人际同步”中识别出来。两个心率曲线可以高度相关，但这种相关既可以来自双向互调，也可以来自两人都在跟随同一个外部节拍器。人际同步这个概念本身不包含“这个同步如何产生”这一层信息的区分，而共调恰恰做了这个区分。第三层是本体论位置层：神经科学能展示两枚大脑的耦合形态，但它无法在自

身的概念体系中回答“这个耦合本身，是音乐事件的一个附带症状，还是它的本体？”这个回答需要哲学本体论来给出。本文的工作就在这第三层上。

进攻性预判：如果人际同步理论声称已充分解释音乐事件中的所有身体间现象，它应当预测：在控制生理同步水平的前提下，音乐事件经验与非音乐社交经验之间，不存在可被受试者独立报告的本质差异。共调模型则做出相反的预测：即便保持同等生理同步水平，声学驱动的同步与视觉/触觉驱动的同步，在受试者的事件体验报告上会有系统性的显著差异——前者会被更多地报告为“我们一起在做一件事”，后者会被更多地报告为“我们在互相配合”。这个差异源自判准层，而非生理机制层。此预测原则上可被执行良好设计的实验检验。

6.2 文化相对主义的质疑：“你的模型不适用于集体音乐传统”

来自民族音乐学的质疑是：你用来逼出“共调”的场景，听起来全是从极度个人化的、内省式的西方音乐聆听传统中抽象出来的。把它拿到西非的集体鼓乐或巴厘岛的加麦兰等大型集体音乐实践中，它那套精致的二体互调节描述会显得太窄。

本文对此的回应分三段。第一段：接受。共调作为概念，确实是从西方语境下被充分内在化了的聆听体验中逼出来的。它当前的论述框架，是从一个身体面对另一个身体——即二体互动的极限情形——出发建立的。对于大型集体音乐实践中同时涉及数十甚至上百个身体的多体共调场，本文不声称在此已给出充分的一般化描述，这需要另一个专门的工作。

第二段：但是，把二体互动作为最简模型来建立核心机制，是理论工作的标准第一步。经典力学从无摩擦质点在真空中的运动开始，不是在否认世界上有摩擦力和空气阻力，而是在给出那个最简的运动方程，然后才能写出受多个力同时作用的炮弹运动方程。同样，共调的二体模型，是在为音乐本体论写它的动力学方程的最简形式。它写完以后，多体模型的推广才有一个可以出发的基点。

第三段——进攻性预判。如果文化相对主义者声称“音乐事件的本质在不同文化中完全异质、不可能有跨文化的最小结构”，那么共调模型会做出以下可检验的预测：在任何一种被其文化共同体内部承认为“音乐事件”的集体实践中，当人为

地在声学节点上引入一个超过±500毫秒的延迟（即切断共调判准c所要求的高精度声学驱动），共同体成员报告的事件体验会系统性地降低——他们会说“今天好像不太对”“没有通”“没有热起来”；而当你切断的只是视觉通道（蒙上眼睛但保留声音的实时互调），事件体验的降低程度则显著低于切断声学通道的降低程度。如果这个预测在跨文化实证检验中被支持，那么它意味着：声学驱动的实时互调精度，而非视觉或文化的象征结构，才是音乐事件跨文化成立的最小生理-本体论条件。这不等于“共调是所有文化都意识到的概念”——文化完全不需要有这个概念。这只等于：不管你用什么文化语言来描述它，当你把声学互调的时间精度切掉以后，音乐事件在任何一个文化中都会在同一地方垮掉。

七、结论：音乐作为不能被占有的发生

现在可以给出一个不同于以往的回答。音乐不在声音里——这是二十年前 musicking 已经说清楚的事。音乐也不在你我之间那条“送一接”的连线上——这是本文在 musicking 和接应之后试图继续往前推的一步。音乐在那里，在那个比你我之分更早、在第一个音送出之前就已经在默默运作的、将我们两个身体一并裹挟其中、并借由我们的呼吸和微颤来调节它自身的共调场里。你我并非先分开、再用音乐来重新连上——而是音乐在发生的那一瞬间，暂时地解除了“你和我”这个二分，让我们成为一个更大身体的两个器官。

这个判断走到这里，必须守住自己的诚实边界。第一，本文的共调模型是从二体互动的极限情形出发建立的最简结构，对于大型集体音乐实践中的多体共调场，尚未给出专门的一般化论证。第二，本文对录音体验的分析将其定位为共调的一种独立形态重构——它关闭了向外接住的回路，打开了向内生长的回路。但这一定位是否适用于某些带有实时弹幕互动或虚拟现实反馈等高度交互化的数字音乐实践，本文未展开讨论；那些场景中的身体间关系可能同时激活多条回路，形成一种混合形态，需另行研究。

最后，一个诚实的判断，必须给出那个能让它认错的条件。什么样的事实一旦坐实，本文的核心判断就需要被放弃？不是在技术实现了异地全感官实时双向往返

的那一天——那一来，放弃的只是本文对“物理共场”的保守坚持，而非核心主张本身。本文真正认错的条件只有一个：在未来，存在这样一个被参与者持续、跨情境地报告为“这就是一次完整的音乐事件”的事件，并且在其中，本文在第3.1节所设定的共调全部四条判准都被严格控制和测量证明了其缺席。如果这样一个事件被确凿地检验出来——且其结果可被独立团队重复——那么本文对“共调是音乐事件的最小必要结构”这一核心主张，就必须被放弃。在那一天到来之前，本文所做的工作，就是对“音乐一直在哪里发生”这件事，做出的一次尽可能诚实的记录。

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）：舒茨、沟通性音乐性与节奏夹带

正文第四节已与梅洛-庞蒂的身体间性做过一次正面对勘，得出"同出而异路"。但还有三家离本文更近，正文未及。其中第一家尤其要紧——它就躺在本组第四篇《调谐的角落化》的参考文献里，本篇却一次也没有碰它。

（一）与舒茨「相互调谐关系」的切割——本篇最重的一处空门

阿尔弗雷德·舒茨在《共同制作音乐》（1951）中论证：一切音乐的社会关系奠基于一种**相互调谐关系**（mutual tuning-in relationship）——演奏者与听者

在内在时间流中"共同变老",这种共同流动先于并奠定了一切符号性的沟通。这比梅洛-庞蒂离本文更近:舒茨也讲"共同",也讲时间,而且讲的正是音乐。

分离线有两条。**其一是单位**。舒茨的共同流动仍以两条已经构成的意识流为单位,"共同"指的是两条流的并行;本文主张分立本身是事后的,不是并行的两条,而是一个尚未分化的张力场先在。**其二是有没有中断刻度**。相互调谐在舒茨那里是关系的奠基条件,它没有毫秒尺度的接通/中断判据;本文给了四条发生判准与三条中断判准,并主张音乐事件正是在"接通—中断—再接通"的循环里运行的。

可裁决格:「同流而不共调」。取一批异地网络合奏者(往返延迟稳定在300–500 毫秒、无物理共场)。若他们持续地、跨情境地报告"我们一起走过了那段时间"式的共同时间体验——舒茨的判据满足——而共调的四条判准、尤其是判准(d)不可逆路径偏移系统性地不满足,则本文所命名的东西已被舒茨覆盖,共调只是它的一次生理学再描述,本篇应当退位。反之,若这批人报告的恰恰是"在一起,但那件事没有发生",共调才拿到独立的地位。这一格现在就可以做。

(二) 与「沟通性音乐性」的切割:禀赋,还是事件

马洛克与特雷瓦森提出的**沟通性音乐性**(communicative musicality)——婴儿与照护者互动中的脉冲、质地与叙事三要素,是人类天生的互动节律能力——比本文已引的特雷瓦森 1979 年"原初主体间性"更近。本文第二节用的正是母子交互发声的现场,因此必须划界。

分离线在**禀赋与事件不在一层**。沟通性音乐性说的是人身上有这样一套装置;共调说的是这套装置在此时此刻是否正在运转、有没有被自我监控切断。禀赋恒在,事件可断——本文全部的经验主张都落在"可断"这一侧。

可裁决:若两位沟通性音乐性测量得分都很高的成年人,在被明确告知"本次演奏将被打分"的条件下仍稳定满足共调四判准,则本文关于自我监控切断共调的机制失败,共调应并入禀赋论。

（三）与「节奏夹带」的切割——本篇最干净的一刀

夹带研究（entrainment，见克莱顿等人对合奏中相互适应的建模）把演奏者之间的相互调整处理为耦合振子的周期与相位锁定，其核心预测是**回到共同周期**：偏离会被相互修正、被抹平。

本文判准(d)的预测**正好相反**：偏离不被抹平，而是被接续、被沿着它打开的新方向推进，轨迹因此不可逆地改变。两条预测落在同一段录音上直接冲突，而且都可测。

可裁决：对合奏录音中的自发微偏离逐个编码，测紧接的下一乐句走向。夹带模型预测回归基线；本文预测显著偏向偏离方向。若回归基线占优，本文最原创的那一条——判准(d)——当场撤销。

附论二·可裁决预测

预测一（延迟坍塌的排他性）。在物理共场的合奏中，听者身体微偏与演奏者下一音之间的时间差分布，应在 ± 100 毫秒内近乎零对称；**同一对搭档**在单向监听条件下（听者不可被看见、其反应不可被听见），该分布应退回到 200–400 毫秒的单向延迟——也就是退回作者本栏前作所说的"接应"。若两种条件下分布无差异，则"共调不同于接应"这一区分失去经验支点。

预测二（自我监控的即时性）。对同一位演奏者引入评价性观察（告知本次将被打分），共调四判准的失效应发生在**第一个乐句之内**，而不是随时间逐渐衰减。若失效呈现为缓慢的疲劳曲线，则短路机制应改写为注意力资源模型，本文的"瞬时短路"表述过强。

预测三（不可逆的方向性）。即附论一（三）给出的编码检验：若第三方编码显示接续方向与偏离方向无关，判准(d)撤销，本文退回为一种关于同步的描述。

参考文献

（原稿自带的七条文献原样保留在前；末三条为本轮最近邻切割所补入。）

- Konvalinka, I., Xygalatas, D., Bulbulia, J., Schjødt, U., Jegindø, E. M., Wallot, S., ... & Roepstorff, A. (2014). Synchronized arousal between performers and related spectators in a fire-walking ritual. *Proceedings of the National Academy of Sciences*, 111(48), 17042–17047.
- Lindenberger, U., Li, S. C., Gruber, W., & Müller, V. (2009). Brains swinging in concert: Cortical phase synchronization while playing guitar. *BMC Neuroscience*, 10, 22.
- Merleau-Ponty, M. (1964). *Le visible et l'invisible*. Gallimard. (《可见的与不可见的》)
- Small, C. (1998). *Musicking: The Meanings of Performing and Listening*. Wesleyan University Press.
- Stern, D. N. (1985). *The Interpersonal World of the Infant: A View from Psychoanalysis and Developmental Psychology*. Basic Books.
- Trevarthen, C. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. In M. Bullowa (Ed.), *Before Speech: The Beginning of Interpersonal Communication* (pp. 321–347). Cambridge University Press.
- Yoshie, M., Kudo, K., Murakoshi, T., & Ohtsuki, T. (2009). Music performance anxiety in skilled pianists: Effects of social-evaluative performance situation on subjective, autonomic, and electromyographic reactions. *Experimental Brain Research*, 199(2), 117–126.
- Schutz, A. (1951). Making music together: A study in social relationship. *Social Research*, 18(1), 76–97. (本轮补入的最近对手)
- Malloch, S., & Trevarthen, C. (Eds.). (2009). *Communicative Musicality: Exploring the Basis of Human Companionship*. Oxford: Oxford University Press.
- Clayton, M., Sager, R., & Will, U. (2005). In time with the music: The concept of entrainment and its significance for ethnomusicology. *European Meetings in Ethnomusicology*, 11, 3–142.

编辑说明

本页据作者 07-30 投稿的独立文档版排印（同一批另有一个“一键包”版本，其论文①的标题位被生成过程的开场白占据，本页未采用该版）。编辑改动三处，均不涉论证：①清除正文一处“审稿人提出的追问”字样的过程语，改为“一连串尖锐追问”，被追问的内容与回应一字未删；②“发现学”这一学派内部的对置用词改为平实说法；③补入附论零、附论一、附论二与三条文献。作者的判准、案例与结论未动。

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。