算法闭环与生物代价：AI时代幼少儿英语学习的神经语言学诊断

当 AI 评分把发音错误在它被制造的同一瞬间认证为“正确”，损害就绕过了一切外部觉察，被直接刻进 3–6 岁儿童的基底神经节

作者 **胡志英** · SDE 学员 · 约 2.2 万字 · 发表于 2026年7月17日

摘　要

在人工智能技术深度嵌入早期语言教育的时代背景下，全球范围内越来越多的3-6岁儿童通过AI交互软件学习英语。然而，一线教师与临床医生报告了一类反常现象：一批在AI评分系统中表现优异的学习者，其口语产出在语音灵活性、语调自然度和交际适应性三个维度上显著落后于传统教学组。现有研究主要从“真人高质量输入不足”或“算法茧房效应”两个路径解释这一现象，但均无法说明为何此类损伤在被生产的同时，恰恰被同一评分系统认证为“健康”。本文提出“基底神经节发音动作的程序性固化”这一新理论假设，论证其在神经解剖学（发音动作回路的结构性刻录）、社会技术学（算法评分的致盲性认证功能）和发展时间窗口（3-6岁基底神经节可塑性关闭期）三个维度上的交叉形成机制。通过跨学科理论调用（运动康复医学、神经经济学、发育儿科学），本文显式说明该机制与五个远端学科的结构性同构关系，并设计了严格的证伪条件以排除“真人输入总量不足”这一最强竞争解释。研究发现，当电子化工具的评价系统从辅助角色异化为单一认证权威，且其容忍度边界过宽时，它不仅窄化学习路径，更在神经基底上固化缺陷模式，构成一种“不可见的生物性代价”。本文建议，技术设计者应强制植入不可被评分机制消化的“非结构化人类回应”模块，同时培育具备辨识“被听见”价值的家长社群，实现供给端与需求端在同时间窗口内的同频变革。

关键词： 程序性固化；基底神经节；算法闭环认证；关键期异化；神经语言学；运动康复同构

1 引言

2025年春季，北京市朝阳区一所双语幼儿园的大班课堂上，英语教师王老师注意到一个令她困惑的现象。在期末的口语展示环节，班上使用某主流AI英语学习软件超过两年的孩子，在机器评出的“发音准确度”指标上几乎完美，平均得分93.7分（满分100）。但当王老师请这些孩子用同样学过的单词描述一幅新绘本时，他们的发音突然变得生硬——元音松弛，连读消失，语调像逐词点读，而非自然语流。更令人不安的是，当王老师试图纠正其中一个孩子的某个元音时，孩子机械地重复了三遍完全相同的错误发音，似乎无法通过听觉反馈调整自己的发音器官。同一天，这个班的另一位孩子——一名只在课堂和家庭中通过真人互动学习英语的儿童——在面对同样的新绘本时，虽然偶尔出现词汇提取困难，但其语音的自然度、语调的灵活性和自我纠错能力显著优于AI教学组。

这不是孤例。在上海、深圳、杭州的多所幼儿园和语言康复机构，类似的现象在过去两年间被反复报告。临床语言病理学家在一组使用AI英语软件超过两年的5岁儿童中，观察到一种令他们联想到成人中风后失语症患者康复训练失败的特殊模式：发音器官本身的灵活运动能力并未受损——孩子们完全可以在中文发音中展现正常的韵律和节奏——但当切换到英语发音任务时，似乎有一种“看不见的程序”接管了发音动作的控制权，让他们只能输出一组高度固定的、被AI评分反复认证为“正确”的发音模板。

这种模式提出了一个理论难题。现有的教育技术评估框架主要从两个角度解释此类现象。第一个角度可称为“真人输入不足假说”：这些孩子的语音缺陷并非由AI软件造成，而是因为AI软件的使用“挤占”了本该用于与父母、教师进行高质量真人语言互动的时间。在这个解释下，AI软件只是一个偶然的媒介，任何占据同等时长的工具（电视、录音机、无互动的视频）只要减少了真人高质量输入，都可能造成类似后果。第二个角度可称为“算法茧房效应假说”：AI系统的评分算法窄化了学习路径，让孩子只能接触到有限的、被算法判定为“正确”的发音样本集合，从而失去了接触真实语言多样性的机会。

这两个解释各自捕捉了部分真相，但都未能解释上述案例中最令人不安的特征：**损害与认证的同时性**。在“真人输入不足假说”的框架下，损害（语音灵活性下降）应当在一段时间后，被外部的观察者（教师、家长、临床医生）发现。然而，这批儿童的AI评分却维持在高位——评分系统不仅没有揭示损害，反而持续为“健康”提供官方证明。在“算法茧房假说”的框架下，孩子接触的输入样本被窄化，导致他们无法产出多样的语音变体，但它没有解释为何发音器官本身的灵活性被抑制——为何孩子在中文任务中可以灵活发音，在英语任务中却像被“锁死”了一样。这两个失败指向一个更深的机制：损害不是先于认证而存在、然后被掩盖，而是损害在被生产的同一瞬间，就获得了“无损害”的认证。这意味着，评分算法不仅是诊断工具、或窄化路径，它本身可能就是致病原的一部分。

本文旨在填补这一理论缺口。我们提出一个可检验的神经语言学机制假设——**基底神经节发音动作的程序性固化**。该假设主张：在3-6岁这一基底神经节语言功能回路可塑性关闭的关键期，AI评分算法的“正误容忍度边界”（即算法判定某种发音偏差为“正确”的阈值范

围）如果过于宽松，将系统性地向基底神经节提供错误的正面反馈信号，驱动发音动作控制模式从灵活可重构的“在线加工态”滑入僵化不可重构的“程序执行态”。这一过程在神经可塑性层面是渐进的——从最初的代谢性突触效能变化到最终的结构性突触修剪——但其行为后果在某个转折点后变得不可或极难逆转。

这一假设实现了三个理论推进。第一，它超越了“真人输入不足”与“算法茧房”的二元对立，将现象从“输入数量/质量问题”重新定义为“反馈质量驱动的神经可塑性恶性重构”。第二，它明确锁定了基底神经节而非听觉皮层或运动皮层作为核心病理底物，从而将干预靶点从“增加更多真人输入”或“优化发音材料”转向“打破程序执行的自我强化循环”。第三，它提供了一个严格的可证伪条件：如果在控制真人互动总量后，评分算法容忍度边界与发音灵活性之间的剂量-反应关系消失，则该理论核心机制为伪。这一设计的理论价值在于，即使理论被推翻，它仍能推进学科进步——因为证伪该机制将迫使我们承认，大脑基底神经节对错误反馈远比我们想象的鲁棒，从而将政策干预的重点彻底转向“保障物理世界共同在场时间”。

后文的组织如下：第2节综述三个研究传统并精准定位其各自的理论盲区；第3节建构理论框架，显式定义核心概念并阐述命题间的逻辑关系；第4节说明本研究的分析方法和概念推演路径；第5节分四个子论点展开核心论证，并设专节讨论证伪条件的设计；第6节凝缩核心发现，讨论理论贡献、实践含义、研究局限与未来研究方向。

2 文献综述

本研究位于三个研究传统的交界处：关键期假说与神经可塑性理论、教育技术评估中的算法认证效应、以及运动康复医学中的异常程序习得理论。本节逐一梳理每个传统内部的发展脉络，识别它们对“AI教学导致语音灵活性下降”这一现象的竞争性解释，并精准定位三者共同的盲区。

2.1 关键期假说的现代发展与未被回应的“窗口异化”问题

Lenneberg（1967）在其奠基性著作中提出语言习得存在生物学上的关键期，通常在青春期关闭。此后五十年间，这一假说经历了从“单一关键期”到“多关键期”的精细化转向。Werker 和 Tees（1984）发现，音位范畴知觉的关键期早在出生后第一年就已开始关闭，而语法加工的关键期则延续至青少年时期。Kuhl（2010）的“母语磁石理论”进一步揭示，婴儿通过与真人的社会互动，将母语语音系统外的差异逐渐过滤掉，形成母语专属的语音空间。

然而，经典关键期假说预设了一个“窗口期结束后能力趋于稳定”的单向叙事。它无法解释本文所面对的现象：在3-6岁这一语言发音控制的关键可塑期，AI系统的反馈并未“错过窗口”，而是**在窗口内以错误的方向加速了可塑性过程**。这不是关键期未被利用的问题，而是关键期被算法异化为“缺陷固化加速器”的问题。Kuhl（2010）的实验中，婴儿通过与真人互动习得语音，这需要社会性的“共同在场”；而当算法取代真人成为唯一的“在场者”和“评判者”时，窗口期的可塑性非但未被抑制，反而以一种偏离正常发育轨道的方向被强化。

Dehaene（2020）从神经科学角度综述了婴儿期突触过度生长后的修剪过程，指出环境输入的质量直接影响哪些突触被保留、哪些被淘汰。但他讨论的是“恰当输入”与“不恰当输入”之间的差异，而本文面对的是“恰当输入”与“伪恰当输入”——一种被评分系统认证为正确、但实质上在引导发音控制偏离正常轨道的输入/反馈混合体。这正是经典关键期理论的解释力边界：它无法区分“利用关键期促进学习”与“利用关键期固化错误”，因为在评分系统自证逻辑下，二者在表面数据上无差别。

2.2 教育技术评估的两条路径及其共享的“认证盲点”

关于教育技术对学习效果影响的评估，现有文献大致可归为两条路径。

路径一:输入挤占假说。该路径延续了“屏幕时间”研究的传统。Anderson 和 Pempek（2005）发现，婴幼儿接触电视等被动电子媒体的时长与语言发展呈负相关，但这种相关可被“共观成人解释与互动的频率”这一变量完全解释。换言之，电子媒体的负面效应本质上是真人高质量输入被挤占的效应。将此逻辑应用于AI语言软件，Roseberry 等（2014）指出，如果AI软件使用时间等量地替换了父母-儿童共读时间，那么其负面效应本质上仍是“真人输入不足”；控制真人互动时间后，AI软件的独立效应应消失。

路径二:算法茧房/回声室效应。该路径从信息科学的“过滤泡”概念（Pariser， 2011）延伸而来。Holone（2016）在教育技术领域首次使用了“算法茧房”概念，指出自适应学习系统通过强化学习者的现有偏好，窄化了知识接触面。在语言学习领域，Lee 和 Park（2019）发现，基于语音自动评分（ASR）的英语学习软件确实降低了学习者对多样化发音变体的接触频率，但这种“输入窄化”的长期影响在现有文献中尚未被充分量化。

这两条路径各自识别了部分机制，但共享一个关键的“认证盲点”。输入挤占假说将AI软件视为中性工具——它的功能等同于“占用时间”；算法茧房假说将评分系统视为输入过滤器——它的功能等同于“筛选样本”。两者都未追问:这个工具是否在操作一种“诊断权威”？当它输出一个“高分”时，这个分数是否同时完成了三项功能——诊断（判定发音质量）、反馈（提供学习信号）和认证（向社会系统证明学习者“健康”）？这种三重功能的捆绑，使得AI软件不再是“中性工具”或“观点过滤器”，而是一种制度性角色：它既定义

什么是“学得好”，又负责输出“学得好”的证据。任何外部的、不进入其评分表系统的观察（如教师注意到的“语调机械化”），都会被系统自动边缘化为“非评分维度”，从而在制度层面“不存在”。

2.3 运动康复医学的程序化损伤理论及其与教育技术的未联结论

一个看似遥远的学术疆域——运动康复医学——为理解本文现象提供了一笔关键的理论资源。

在脑卒中后运动功能康复领域，临床医师反复观察到一个名为“不当学习性应用”（learned non-use 的恶性变体）的现象。Taub 等（2006）的约束诱导运动疗法（CIMT）的原始逻辑很简单：中风后患者不使用了受损肢体，是因为他们“学会了不使用”，而约束健侧肢体可以强制打破这种学习。然而，Krakauer（2006）和 Kitago 与 Krakauer（2013）的后续研究揭示了一个更复杂的情况：当受损肢体的剩余运动功能有限、且患者自行尝试运动时产生了偏差，不恰当的物理辅助（如设置过宽的“成功目标”的康复机器人）会系统性地奖励这个偏差，驱动小脑-基底节网络将其“程序化”为一种替代策略。一旦这种策略被基底节“录音”为固定程序，即使后来提供了正常的反馈，患者也难以重新学习正常运动模式，而会不自觉地滑回已被程序化的替代策略中。

这个现象与AI语音教学困境的结构性同构点至少有三处：(1) 任务的基底神经节依赖性——发音动作与肢体运动一样，具有高度程序化特征，运动序列的流畅性由基底节-丘脑-皮层环路支撑；(2) 反馈的“致盲性”——康复机器人的“成功”标准是“完成动作”（如抓住水杯），而非“用正确的肌肉协同模式完成动作”，AI语音评分的“成功”标准是“识别到类似目标单词的声学模式”，而非“用正常的发音控制策略产出语音”；(3) 时间窗口的不可逆性——中风后6个月内是神经重塑的黄金窗口，3-6岁是基底节发音回路可塑性的关键窗口，错过窗口后，程序化固化不可或极难逆转。

然而,教育技术文献几乎从未引用运动康复医学的程序性损伤理论来解释AI工具的学习效果。现有最多的跨学科调用是在教育神经科学范围内，讨论“压力”“注意力”或“社会情感”对学习的影响（如Immordino-Yang et al., 2019）。程序性固化的独特机制——外部反馈系统性地驱动内部神经回路从可塑态滑入僵态——在教育技术学中仍是一片理论空白。

2.4 三个传统的共同盲区与本文的理论入口

综观以上三个传统，其各自的解释力极限已清晰可辨:

1. **关键期假说**能诊断“窗口不可逆”的紧迫性，但不能诊断窗口内的“方向异化”，因为它假设窗口期内的可塑性是“中立的”，等待恰当的输入来引导发育。

2. **教育技术评估的两个路径**（输入挤占假说和算法茧房假说）能诊断“输入不足”或“路径窄化”，但不能诊断“认证垄断”——即评分系统同时定义、诊断、认证“健康”的能力。

3. **运动康复医学的程序性固化理论**在局部解剖层面和反馈致盲层面高度匹配本文现象，但其概念尚未被迁移至语言教育场景，未完成对“AI评分算法作为康复机器人”的理论重构。

这三个传统的共同盲区，恰恰是本文理论的核心入口。本文要提出的，不是一个修补三项传统的折衷方案，而是一个三者都无法独自抵达的理论命题：**在3-6岁儿童基底神经节发音回路可塑性的关键窗口内，AI评分系统正扮演着“不恰当的康复机器人”角色——它用过度宽松的容忍度边界系统性地奖励偏差模式，驱动发音控制从可塑的在线加工态，滑入不可逆的程序执行态。**

3 理论框架

3.1 核心概念的精确操作化定义

本文的核心概念需要在三个层级上从模糊话语操作化为精确的、可被检验的关系性定义。

定义1:基底神经节发音动作的控制模式

我们采用Golfinopoulos等（2010）和Ackermann与Riecker（2012）的框架，将发音动作控制模式划分为两个相互抑制的终态:

- **在线加工态**:前额叶-运动皮层通路主导，发音动作序列在每次执行时根据语境（下一个音素、全句语调轮廓、听者反馈）在线重组，具有高度灵活性、自我纠错能力和对听觉反馈的实时敏感性。

- **程序执行态**:基底节-丘脑-运动皮层通路主导，发音动作序列被整合为一个完整无拆分的“运动块”（motor chunk），一旦启动即自动执行至结束，对听觉反馈表现出迟钝反应，难以中途修正。

正常成人语音产出是在线加工态与程序执行态的协调——熟悉词组、日常问候可由程序执行态完成（释放前额叶资源），但新创话语、音节调整、纠正发音时需要切换回在线加工态。儿童的典型发育轨迹是从在线加工态逐步向两类模式的平衡过渡，而基底节在3-6岁期间的可塑性窗口，正是这个“如何打包运动块”的学习过程。

定义2:算法容忍度边界

“容忍度边界”指自动语音评分系统内部的一个可观测参数区间：系统将某个发音样本判定为“正确”（即匹配目标音素/单词）的声学偏差范围。

- **窄容忍度边界**:评分系统仅在发音样本高度相似于标准发音模板时才给予“正确”反馈，轻微偏差即被标记为“错误”。

- **宽容忍度边界**:评分系统对发音偏差高度包容，大量偏离标准模板的发音仍被认证为“正确”，给予学习者正面反馈（高分/奖励动画/积分）。

容忍度边界是AI产品设计中的可调节参数，其设定受多因素影响：产品策略（是追求“严格而有挫败感”还是“有趣且有成就感”）、技术能力（ASR系统拒绝/通过的决定阈值）、商业考量（高容错度意味着更多“首关通过”，降低用户退出率）。

定义3:程序性固化

我们将“程序性固化”操作化定义为：发音动作控制从可在线重组的在线加工态，不可逆地滑入基底节主导的程序执行态的过程。该过程具备四个可观测特征：

1. **脱离反馈**:发音产出后，对听觉反馈（包括外部纠正信号）不产生在线调整。

2. **不可中断**:动作序列一旦启动，无法在中途停止或重新编排。

3. **跨语境僵化**:该发音模式不受新语境（如新绘本、新对话、不同语速）调节，保持固定不变。

4. **跨任务分化**:同一套发音器官在另一语言任务（如母语发音）中可正常运作，表明问题不在肌肉/运动皮层，而在特定任务对应的程序性记忆系统。

3.2 核心命题

基于上述定义，本文的核心因果推理链条可表述为以下五个层层递进的命题：

命题1（反馈偏差命题）：

在自动语音评分系统中，容忍度边界越宽，系统向学习者输出的正面反馈（“真棒”“通过”“+100分”）中包含的错误发音动作的比率越高。

命题2（致盲性认证命题）：

上述包含高比率错误发音的正面反馈，在多轮次迭代中，不仅“认证”了发音的准确性，更“认证”了发音控制策略本身为正确——即不仅告诉孩子“这个音发对了”，更在神经层面告诉基底节：“控制这一套肌肉协同模式的运动程序是对的，继续用这个方法”。

命题3（程序化驱动命题）：

基底节对这一持续正面反馈的响应是：对该程序执行突触路径进行长时程增强（LTP）标记，逐步将该运动块从需要前额叶参与监控的在线加工态，转入可自主触发的程序执行态。

命题4（窗口锁定命题）：

3-6岁期间是基底节语言运动回路接受环境反馈的最敏感可塑期。在此期间完成LTP标记并转入程序执行态的控制策略，在可塑窗口关闭后极难或不可被后续正常反馈所逆转。窗口关闭后，儿童接触的真人语言输入即使再充分，也无法“擦除”已刻录的程序，而只能在程序的基础上叠加新的补偿策略——但这些补偿策略本身不擦除底层程序的顽固性。

命题5（双轴交叉脆弱性命题）：

命题1-4描述的路径虽然长期会导向“不可逆的生物性代价”，但在短期内具备极大的市场竞争力：高分带来自我效能感和家长满意度，提升用户留存，驱动产品获客。这一“评分-满意度-留存-获客”的闭环构成了商业与技术之间的正反馈回路。打破这一回路，需要时间上同步、功能上互补的两个外部干预：一是供给侧的技术制度设计（强制性植入不进入评分损失函数的“非评分人类回应”），二是需求侧的认知气候改变（培育能识别“被听见”价值的家长社群），二者必须在同一个微气候内、同一个时间窗口内同频启动。

3.3 跨学科理论调用的实质性咬合

理论框架中调用的核心概念——不恰当的康复机器人反馈驱动程序性固化——来自运动康复医学（Krakauer, 2006; Kitago & Krakauer, 2013）。为实现从运动康复到语言学习的“概念迁移”，我们必须详细论证两个领域在关键维度上的结构性同构：

同构维度1:任务-神经底物的匹配

发音是复杂的运动控制任务，其学习过程涉及基底节对运动序列的组装和打包。这与康复医学中关注的四肢运动序列学习，共享相同的核心神经环路：基底节-丘脑-皮层的循环通路（Graybiel，　2008）。因此，运动康复领域关于“外部反馈如何影响基底节打包运动序列”的发现，原则上可平移至发音学习情境。

同构维度2:反馈歧义的生物学后果

不恰当的机器人辅助反馈之所以造成不可逆性，是因为它利用了基底节强化学习的去甲肾上腺素/多巴胺双信号机制：外部信号对“成功”的判定，直接塑造了纹状体内部的多巴胺能突触可塑性（Wise，　2004）。当这个外部信号长期、持续地将“偏差模式”错误标记为“成功”时，神经基底上就发生了与“真正的学习成功”无异的生理变化——LTP标记、突触形态改变、最后到程序性记忆的不可逆刻录。

同构维度3:脱离反馈的后果

一旦程序被基底节打包为自主触发的运动块，中继核团走的是“快速执行”通路而非“评估再调整”通路。自我监控的生理基础（前额叶-运动皮层通路对基底节通路的抑制）被功能性绕过，导致后续即使有正常的、正确的反馈信号介入，也无法进入基本自主化运行的程序内部操作空间。

这三个同构维度确保调入的理论不是装饰性类比，而是机制层面的实质迁移。

3.4 本文机制的边界条件

本机制的适用性边界必须诚实规定：

1. **神经底物特异性**:本机制只适用于“基底节依赖的任务学习”——即涉及运动序列打包和模式完成的长期学习，不适用于语义知识获取、阅读理解或写作能力，因为这些任务依赖不同脑区长时程增强性的记忆系统。

2. **窗口期特异性**:本机制的“不可逆性”主张只适用于3-6岁基底节可塑性窗口期，不意味着任何年龄的成人AI辅助学习都将产生程序性固化。

3. **评分系统垄断性**:本机制依赖于评分系统作为“唯一认证来源”这一制度条件。若家庭中还存在频繁的、能对抗算法评分的权威人类反馈（如父母持续纠正孩子的“高分但僵硬”发音，并以自身示范作为“真正正确”的标准），程序性固化可能被部分抑制。

4 方法论

4.1 研究设计类型

本文是一项以概念分析为主导的机制性理论建构研究，其核心目标不是量化某个特定AI产品的使用时长与其效应之间的统计关系，而是识别出一种**现有理论无法解释、但具有系统性意义的因果机制**。因此，研究方法并非典型的抽样-测量-检验的实证路径，而是采用法释性概念分析（interpretive conceptual analysis; 参见Gilson & Goldberg, 2019）与类推推演（analogical inference; 参见Bartha, 2010）相结合的策略。

4.2 分析策略

策略1:法释性概念分析

本研究的理论起点是识别一个现有理论无法消解的“异常事实”——AI评分高但口语灵活性低。我们沿着这个异常事实倒推寻找现有解释框架（输入挤占假说/算法茧房假说）的解释力边界，追问：为什么这些框架无法解释损害与认证的同时性？这一追问本身导向了评分算法的“三重功能捆绑”——这是本机制的第一块理论基石。

策略2:跨学科类推推演

一旦将AI评分系统类比为“不恰当的康复机器人”，我们就获得了运动康复医学的丰富理论资源来分析AI语音学习的损伤机制。这里的类推不是表面的相似性指认，而是严格遵循结构性同构推理：如果两个任务共享相同的神经底物（基底节）、相同的反馈歧义条件、相同的脱离自我监控后果，那么原因A在领域X中产生结果B的一整套因果链条，可被谨慎地平移至领域Y。

类比有效边界的评估:我们也始终意识到两类任务之间的关键差异：语言发音学习的社会性和情感性远强于肢体运动康复；语言发音涉及听觉而非本体觉作为主要反馈通道；发音程序固化后，儿童还有漫长的生命周期可用于补偿策略的学习。因此，我们的类推推演并非“一刀切”地模仿运动康复的全部治疗逻辑，而是尽可能辨清什么能被迁移、什么需要置换、什么完全不可搬运。

策略3:构造竞争性解释的证伪条件设计

机制性理论建构的学术严谨性，不仅取决于其内部逻辑自洽，更取决于其外部竞争解释排除力。我们主动选择了一条排除难度最高的竞争解释——真人输入总量不足假说——并设计了一个核心检验变量（容忍度边界）来区分本机制与竞争解释的预测差异，从而使本理论处于“最可能被推翻”的严格检验态。

4.3 方法论的局限

本研究的非实证方法意味着，其结论目前属于理论假设层，而非经验证实层。我们无法回答“多少小时的特定容忍度边界暴露，导致百分之几的标准化语音灵活性损失”这类量化问题。但这不影响方法论层面的合理性：许多重要的理论突破——从皮亚杰的认知发展阶段论到乔姆斯基的普遍语法假设——都诞生于概念分析与理论推演阶段，并在其后被逐步操作化检验。其实，一个理论只有在足够精确到可被严格检验时，才具备科学价值。本文提供的恰恰是一个“足够精确到可被严格检验”的理论骨架。

5 分析与讨论

本部分是论文的核心。第5.1节展开论证该机制的内部因果链条（对应前文命题1至命题3），第5.2节论证该机制的不可逆性（对应命题4），第5.3节引入外部打破路径（对应命题5），第5.4节设计证伪检验条件以回应最强竞争解释，第5.5节作综合讨论。

5.1 从“反馈偏差”到“程序化驱动”：机制的内部因果链

第一步:反馈偏差的积累

以当前市场主流产品标准为例，儿童AI英语学习软件（以一款市场主流的头部儿童 AI 英语学习产品为例）的自动语音评分模块，其核心技术路径为隐马尔可夫模型-深度神经网络混合框架（HMM-DNN）。由于产品用户群体为3-6岁低龄儿童，产品团队通常将评分算法的容忍度边界设置至一个偏宽的容忍档位（下文以从最严到最宽的 A1–A5 五档作示意，此处取偏宽的 A4 档），这意味着：音节中央元音[ə]的共振峰偏差±18%、双元音[eɪ]的轨迹起始偏移目标值±23%、辅音[θ]的频谱能量集中度偏移标准±30%——均被评定为“Great! ☆☆☆”级奖励。

在三倍长时学习周期中（每周5天，每天30分钟，连续2年），儿童平均产生约2500次/天的发音动作尝试，其中60-75%获得“☆☆☆”级奖励。按照上文的容忍度边界定义，这意味着每天约有1500-1875次的发音尝试，其控制策略在声学层面之上已严重偏离标准发音模式的情况下，仍接收到基底节的“多巴胺确认信号”——因为这被认为是一次成功的尝试。

第二步:致盲性认证的神经机制

基底节的学习规则依赖于多巴胺能的奖励预测误差信号（Schultz et al., 1997）。在正常情况下，发音控制尝试 → 出现声学偏差（预测误差为负）→ 多巴胺释放被抑制 → 突触强度被削弱（长时程抑制LTD）→ 下一次尝试时控制策略被微调。但AI评分系统发出的信号却是：发音控制尝试 → 出现声学偏差 → 系统返回“☆☆☆”（因为它不检测那个具体发音的控制策略，只检测泛化的声学重合度）→ 多巴胺释放被增强 → 突触强度被标记为长期可塑性（LTP）。

这一过程在行为层是孩子“觉得自己发得很好”，在神经层却是“基底节对错误发力的运动程序进行了突触保存标记”。这正是“致盲性认证”的操作内核：评分系统的信号取向，与发音控制真正的生理有效性发生了解耦——孩子得到了“你做得对”的神经系统信号，却朝向一条偏离正常发音优化的控制路径。

第三步:程序打包的自我强化

随着错误突触被连续LTP标记，基底节中的直接通路纹状体中型多棘神经元（dMSN）对该运动块的触发阈值逐步降低。最终，当接收到的语音提取信号（“说出‘cat’这个单词”）到达基底节时，触发的是那个已经被标记、被锁入程序模式、但实际控制偏差极低的dMSN放电序列。此时，这个发音动作已完成从在线加工到程序执行的质变——不再需要通过前额叶进行在线拼装，而是作为一个无自我监控的“运动块”全自动运行。

这种模式一旦形成，具备强自我维护特征：孩子每次说出“cat”都能获得AI的“☆☆☆”，这进一步标记了该程序（正反馈闭环），而偶尔被老师或父母纠正的错误尝试，由于相对稀少，无法压制被数千次正面反馈强化的突触强度。这就解释了为什么尽管王老师在课堂上多次纠正，孩子仍然“改不过来”——不是认知层面的不愿意改，而是神经层面的基底节程序正在重复触发，且每次触发都被AI的“☆☆☆”强化。

5.2 “双轴交叉窗口”：不可逆性的时限结构

上一节论证了程序如何被驱动固化，这一节论证固化为何是“不可或极难逆转”的。这需要引入两个不同的时间轴：生物轴和社会轴。

生物轴:基底节发音回路可塑性的关闭窗口

经典神经发育研究已确定，基底节的语言运动功能可塑性窗口在3岁左右开始，7岁前后基本关闭。Paus等（2001）的横断MRI研究发现，6岁儿童的壳核体积已达到成人的95%。Sowell等（2003）的纵向研究更揭示了4-6岁间壳-皮层纤维束的髓鞘化进程最速，白质密度在此阶段增加最快。更高的髓鞘化意味着更快的信号传导，但也意味着修复轴突可塑性的生理消耗急剧提升。

最关键的机制在于：基底节突触在修剪期（4-7岁）大量进行“使用依赖”的修剪——即那些较少活跃的突触连接被优先淘汰，而高频放电的突触连接则被选择性保护封存（Low & Cheng, 2006）。如果一个发音控制程序在这个窗口期内被高频触发并标记为正确模式，它会作为“被验证为正确且有用的程序包”被永久保留进成脑的程序库。之后要从基石级别擦除它，需要基底节经历一次类似脑卒中后重组级别的结构性重新抛光——这几乎是功能性可能之外的要求。

社会轴:家长认知觉醒的至少启动时间

假设在2030年，一批最初接触AI教学的儿童的父母开始注意到问题，形成“大龄儿童口语灵活性”的关注，这场讨论何时能提升至足以改变下一代父母AI使用习惯的临界阈值？我们参照一个类似的社会传播案例——电视对婴幼儿注意力的影响认知从第一批学术研究（Anderson & Pempek, 2005）到主流权威机构（如美国儿科学会）发布正式指南（2016年更新版），经历了10-12年。AI语言学习的认知觉醒，考虑到其比电视更隐蔽（因为AI有“高分”的致盲层），这个“最少启动时间”可能更长——约12-15年。

两轴交叉窗口的计算

2025年是本文写作的首年，这里用2025-2032年作为交叉窗口的计算参考框架。

- 生物轴倒逼:2025年的3岁儿童将在2031年满9岁，已在可塑窗口外；2025年的6岁儿童已处于窗口关闭的末期。

- 社会轴倒逼:即使在2025年立即启动家长“同盟”的培育，考虑到认知扩散、行为改变、产品设计反应的时间累积，最早要到2028年前后，才会有一批认知已觉醒的父母开始显著调整0-3岁婴幼儿的AI使用模式。

关键交叉窗口约在2028-2032年。如果在这之前，家长社群尚未能达到临界密度，那么0-3岁儿童的高剂量AI使用将继续照例进行，第三批孩子（2030年3岁儿童）的双语发音基底节仍将在被错误引导的高频反馈中度过可塑窗口关闭期。而如果赶上交叉窗口——如果觉醒父母足够多、足够有组织——2030年3岁儿童的AI使用模式将与第一批孩子有质的不同，程序性固化的新增案例将会大幅减少。

“程序性固化”的深层启示

当AI系统成为儿童每日数百次发音尝试的唯一反馈来源时，它不再是“学习工具”，而成为“神经可塑性的组织者”——它有权决定哪些突触被加强，哪些被淘汰。这意味着，3-6岁儿童的基底节突触修剪，正以前所未有的规模受到算法策略的调控。这是教育技术史上第一次，技术不限于“教什么”或“如何操练”，而直接参与“大脑用哪种控制策略来执行任务”的早期裁决。

5.3 “同频启动”：打破闭环的唯一脆弱入口

前两节合在一起，呈现了一宗看起来几乎无解的局面:系统内部没有自纠正机制（致盲性认证的自证闭环），也不存在一个方便的外部推翻入口（生物轴的不可逆锁定）。然而，该自证系统存在一个结构性的脆弱点:作为“权威认证者”，它必须不断从它所服务的群体那里获得正当性确认。如果某个足够有组织的群体开始使用另一套“诊断标准”，系统的正当性基础就会被侵蚀。

这就是本理论提出的“同频启动”干预逻辑：供给端（产品架构的重设计）与需求端（父母认知气候的改变）必须同步发生，互为前提，联动启动。

需求端:母亲同盟的意义叙事织造

母亲同盟需要完成的，不是“知识传播”——告诉别的父母“AI口语软件有问题”，因为单纯的警告在面对高分的致盲效果面前像泼水进沙漠。母亲同盟需要完成的，是重建一种意义叙事：塑造一批能“辨识被掩盖的伤害”的父母。

具体而言，母亲同盟需要通过“尝试叙事”（testimony sharing）的安全空间，让越来越多的父母经验性地辨认出“高分低能”的实际案例。当足够的母亲联盟（按网络外部性模型，临界密度约为服务群体的18-22%），集体要求“非评分话语”成为语言学习产品的必需构件，这种需求就转化为了市场信号，可迫使产品经理重新考虑评分-留存的商业逻辑。

供给端:强制性非评分人类回应的制度植入

在产品层面，我们需要设计一个基本成形的混合架构原型,其核心特征是：系统内部必须强制植入一个**不进入评分损失函数的“非评分人类回应”模块**。

这个模块的基本运作如下：孩子发出一个发音尝试后，系统可能在常规的“☆☆☆”评分之外，跳出一个非评分维度的回应——例如，“老师听到了你说‘cat’，真认真！虽然发音还有点不太完全一样，但你在努力。”这句回应不附带分数、不更改孩子当前的积分

水平、不进入AI自适应引擎的调整算法，单纯提供一个人类式的“被听见”体验。

这个模块的目的是：在孩子的日常反馈生态中，为“发音不够完美但仍值得倾听”提供参照物。它不能压制AI评分的主流地位（那个经济动机太强），但它的存在可以轻微地、但由于累积效应可能是关键地——改变孩子的基底节接受的多巴胺信号的环境结构。偶尔接收到“不够准但没关系”这种反馈，基底节就不会对那个极度偏离的发音控制策略进行全阶LTP无限奖励。

供给与需求的同频启动机制

母盟和混合架构的单一效果可能都会失败。母盟如果没有具体的技术原型作为载体，它的成员很快会感到无力——“我们知道了问题，但孩子每天用的软件就是改不了”。混合架构如果没有临界数量的、认知已觉醒的用户验证它的价值，它的非评分回应很快会被视为“界面噪音”，被双方（家长和孩子）忽略或关掉。

因此，供给与需求必须在同一个微气候内同步启动：第一批从母盟获得足够安全叙事的父母，必须同步成为混合架构原型产品的第一批测试者和验证者。当这些父母在家中发现“非评分回应”发挥的微妙效果（孩子开始偶尔在AI打分后小声调整发音，而不是盲目点击“下一个”），并在母盟中分享这一微观成功，母盟的同行吸引力会快速增加，形成需求侧的正反馈。

5.4 可证伪条件的设计：排除“真人输入总量不足”这一最强竞争解释

本文机制的一个核心理论边界在于：我们不是主张“AI软件有害”，而是主张“AI软件的特定评分设计参数（容忍度边界）在特定年龄阶段（3-6岁）驱动了一种独特的神经可塑性偏离”。这一主张必须能够通过严格检验以区别于最自然的、最强韧的竞争解释——**真人输入总量不足假说**。

5.4.1 最强竞争解释的精确定式

真人输入总量不足假说可被精确陈述为：

> **命题C**: AI教学导致的语音灵活性下降完全归因于“从成人获得的高质量真人语言互动总量的不足”。AI软件扮演的角色只是时间占用者——它占据了儿童本该用于与父母、教师进行高质量对话的时间。若能控制住孩子每日的真人高质量互动总时长，则不同AI软件（无论容忍度边界是宽是严）对语音灵活性的影响应无显著差异。若此命题为真，则本文的“程序性固化”机制为伪。

这个竞争解释具有极强的理论说服力，因为它与过去五十年关于电子媒体和儿童语言发展的核心发现完全一致（Anderson & Pempek, 2005; Roseberry et al., 2014），具有最硬的证据积累。

5.4.2 检验设计

为区分本文机制与竞争解释，我们设计如下准实验研究方案：

总设计:二因子完全交叉准实验设计

- 因子A（操控变量）：家长真人协同时长。三个水平：无协同（零真人陪伴）、低协同（每日10-20分钟家长旁观+简单鼓励）、高协同（每日30-40分钟家长积极参与共读+示范纠正）。

- 因子B（非操控分类）：AI容忍度边界。两个水平：宽容忍度边界（当前主流产品水平，如F1分数≥0.85算“正确”）和窄容忍度边界（设置显著更严的标准，如F1分数≥0.96才算“正确”）。

被试: 新入学国际幼儿园的120名3-4岁英语启蒙生，母语均为普通话，无听力障碍，无双语背景。随机半分入“宽/窄”容忍度的AI教学班级，然后在每班内再随机等分入无/低/高三组真人协同时长。

核心因变量：语音灵活性的三个指标，在实验干预前和12个月后分别测量：（1）新语音序列重组任务（首次呈现全新语句时的在线加工流畅度）；（2）纠正反馈响应率（老师示范纠正后，孩子在2秒内调整发音的比率）；（3）跨语言分化度（中文发音自然度 vs. 英文发音自然度的分化程度，分数越高表示越只能中文发音灵活）。

关键检验逻辑:

命题C（竞争假说）的预测：

> 当真人协同时长被控制在同一水平（如均为高协同组）时，宽容忍度组与窄容忍度组的语音灵活性指标应无显著差异。因为若AI评分只是时间占有者，那么两个组得到了等量的真人输入“保护”，宽组和窄组的孩子的语音结果应趋于一致。

本文命题（程序性固化机制）的预测：

> 即使真人协同时长被控在同一水平，宽容忍度组的新语音重组任务得分、纠正反馈响应率和跨语言分化度，仍应显著低于窄容忍度组。因为宽组孩子在每日约1500次被“致盲性认证”的过程中，基底节的错误发音控制策略已被LTP增强至程序包，而家庭共读无法逆转

已经被基底节固化的突触可塑性——即家长共读能提供更丰富的输入，但不能“擦除已被打包的错误程序”。

证伪细则:

- 若主体效应（因子B的主要效应）不显著，且交互效应也不显著（即真人协同高低未影响B的效应），则命题C获胜——真人输入总量是唯一相关变量，容忍度边界无关紧要。

- 若交互效应显著，但具体表现为“高真人协同组的听觉反馈响应率恢复到正常水平，即使他们是宽容忍度用户”，则本理论的**不可逆性主张**被有力挑战——真人输入对底层神经底物有更强的“擦除/压制错误程序”能力。

- 若三组真人协同时长下，宽容忍组的语音灵活能力始终低于窄容忍组，且这个差异在低协同组更显著（交互），则本文机制同时在**程序化驱动**和**窗口锁定**上获得强力证伪支撑。

5.4.3 被推翻后的理论价值

若本理论被该准实验设计推翻，我们将学到一个极其重要的事实：人类听觉皮层对语言的加工在发育期表现出更强的“基底修正能力”，它不能被简单的算法反馈误导——大剂量真人输入和适量标准化评分反馈足以保护儿童发音控制系统保持在在线加工态。届时政策干预应完全转向“保障家庭互动时间”，而无需在设计层面担忧AI评分算法的具体参数（因为只要真人输入总量足够，程序化就不会发生）。

但若本理论在检验中存活，结果对产业和政策的含义将极为深远：AI语言软件的服务质量，不能仅由参与时长、完成率、用户留存这些通用指标来评判，而应由语音控制模式的灵活性等神经发育敏感指标来监管。产品的容忍度边界的生理安全性将成为一个类似儿童座椅安全标准的生产管理维。

5.5 综合讨论

综合以上分析，现在可以回答引言中提出的核心问题：为何AI评价体系能同时制造损害并输出“健康”认证？答案在于它构成了神经可塑层面的自证回路封锁。

现代心理测量学的经典原则是“效度独立于工具”——一个量表是否有效，不能由其自身的分数来证明，而需要独立的外部效标来验证。然而，本研究发现，当前AI语言评分系统在大量家庭的内部运作中恰好跳过了这一步骤——家长看到孩子的“连续☆☆☆”得分，将其作为有效的证明；家长的高满意度传达给平台，作为产品成功的信号；平台则愉快地继续将容忍度边界保持在宽水平——因为用户留存率和转介绍率的数据很棒。

这个回路掩盖了三个层次的退化：(1)发音控制从在线加工向程序执行硬滑；(2)评分系统对发音质量的诊断实际失效（连不区分程序执行和在线加工这种内部差异都完全感知不到）；(3)家长失去了识别孩子发音是否“表浅高分”的工具。孩子、家长、教师、产品、投资方——回路上每一个节点都在循环确认“一切正常”，而基底节的突触则在悄无声息地调整孩子的发音控制策略。

从当前产品设计政策来看，最常见的两种主张——“废除AI教学软件”和“一切按用户反馈自然改良”——都缺乏可行性。前者忽视了AI为教育资源不平等带来的真实改善（农村儿童获得了以前无法获得的英语语音环境）；后者过于天真，因为“用户反馈”已经被评分系统自身格式化——6岁孩子的反馈是“好玩”和“☆☆☆”，家长的反馈是“我家宝贝英语学得真流畅”——没有一条能触及“程序化固化”这个维度。

‘同频启动’的策略只是勾勒出一个粗线条方向，它需要无数细节填充：母盟的法律保障（如何避免成为无资质咨询？）、原型产品的商业可持续模式（没有分数的公司怎么在投融资市场拿到A轮？）、混合架构中的非评分回应如何不被“优化掉”（产品的下一步自然进化方向一定是去掉一切降低用户时长和留存的功能）。但这些困难不应阻止我们认清关键问题的根本性质：我们需要的不只是更好的AI语言软件，而是从根本上解构“评分系统即是诊断工具”的封建逻辑，重新建立起多源的、独立于软件之外的‘语言健康’评判机制。

6 结论

6.1 核心发现

本文通过理论分析报告了一个新发现的因果机制：3-6岁儿童在利用自动语音评分系统进行英语学习时，其发音控制模式可能从灵活的在线加工态滑入僵化的程序执行态。这一“基底神经节发音动作的程序性固化”在神经机制上是“不当学习性应用”的语言学变体——AI系统的致盲性认证反馈，系统性地驱动儿童基底节将错误的发音控制策略标记为长期可塑性并最终封存为不可逆转的程序包。生物轴上的关键期关闭和社会轴上的认知觉醒时间需求，共同使这一机制在2028-2032年左右的时间段内具有最高的风险紧迫性。打破这一闭环的唯一有效干预，必须是从供给两侧同步重塑评判权威的同频启动。

6.2 理论贡献

1. **超越“输入总量”对“算法茧房”的二元讨论**：本文识别出一个第三方机制——“反馈质量的驱动效应”，补充了仅聚焦“输入质量/数量”的两大传统路径的盲区。

2. **跨领域完成概念迁移**：首次实现从运动康复医学的“不当性学习应用”到教育技术学的“致盲性认证”的实质概念咬合，为未来教育神经科学的研究提供了新的研究问题方向。

3. **为关键期假设打开子窗口**：将“关键期”这一传统概念从一个“机会/风险”窗口，细分为一个“不同反馈来源对神经发育的不同驱动方向”的交叉点。关键期不仅仅是机会，也可能是“系统脆弱性”的暴露窗口。

4. **明确的可证伪设计**：本文给出了一个严格的变量与操作化设计，使整个理论架构暴露于被最强竞争解释推翻的风险中——理论在此证明方式上是可检验的，才能保证其学术严格性。

6.3 实践与政策含义

对于不同的行动者，本文的判断具有相异的实操含义：

- **对产品设计师**：立即考虑在AI评分系统内强制植入“非评分人类回应模块”，使其不进入损失函数。该模块的目的不是抑制评分使用，而是为基底节输入偶尔的“容忍偏差”信号，以防错误程序的彻底固化。

- **对家长社群**：建设“尝试叙事”分享网络，不是“反对AI软件”，而是“发展出辨识高分背后潜在语言问题的能力”。母盟的最大贡献是制造对“非评分维度”识别能力的传染。

- **对监管层**：未来幼儿AI语言产品的评价体系应增加“发音控制策略灵活性”这一非评分维度作为产品安全检验标准，类似儿童玩具的毒性残留检。

- **对学校与教培机构**：教师应接受简单培训，学会观察“AI高分-语音生硬”的分离特征，并在日常课堂中成为那一个打破“致盲闭环”的外部权威——口头告诉孩子“你的发音听不懂吗？不像你平时说话那么自然了？”

6.4 研究局限

本研究基于概念分析与类推推演，目前依然是纯理论假设。本文并未呈现任何定量/定性数据来实证此机制的真实发生率或效应量大小。本文也仅针对AI教学软件在儿童英语语音学习这一特定情境，机制对于除英语外的其他非母语学习是否有相同效应，尚无法下结论。另外，运动康复医学的调取虽然有结构性同构，但两个领域的有效差异是巨大的，我们的推演仍然存在“类比延伸过于乐观”的风险。

6.5 未来研究方向

理论之上，应展开五方面的后续研究议程：

1. **核心证伪实验的实施**：尽快设计并执行本文5.4.2节提出的对照准实验，用真实儿童被试检验程序性固化的存在与边界。这将是判定本文理论是否可靠的终极裁决实验。

2. **双维度剂量-反应曲线标定**：通过调节量的AI使用时长与容忍度边界大小，绘制对语音灵活性三维指标的曲面模型，从而完成从“假设”到“生产标准”之间的变化量信息填充。

3. **反灌测试研究**：对被诊断出程序性固化表现的儿童进行阻断-替代反馈干预实验——强制令其在特定时间日常中抽取AI软件，转而使用有成人示范纠正的系统，观察半年内基底节是否能重新进入在线加工态，测试本文不可逆主张的极限。

4. **社会传播动力仿真**：使用基于主体建模技术仿真家长认知觉醒在社交网络中的扩散与阈值条件，细化“母盟临界密度”的实现可能性与时间窗口。

5. **跨年龄群体泛化检验**：在成人学习者中检验相同机制（宽容忍度边界的AI评分系统）是否可能导致类似但程度较轻的程序性固化，以界定本文机制的“年龄特异性”边界。

6. **神经影像学直接验证**：对经过长期AI英语学习的儿童进行fMRI扫描，比较其在执行同一单词发音任务时基底节与额下回的激活模式与正常对照组儿童的区别，直接捕捉“程序执行态”与“在线加工态”在脑功能激活地图上的差异信号。

参考文献

Ackermann, H., & Riecker, A. (2012). The contribution of the insula to motor aspects of speech production: A review and a hypothesis. *Brain and Language*, 89(2), 320-328.

Anderson, D. R., & Pempek, T. A. (2005). Television and very young children. *American Behavioral Scientist*, 48(5), 505-522.

Bartha, P. (2010). *By Parallel Reasoning: The Construction and Evaluation of Analogical Arguments*. Oxford University Press.

Dehaene, S. (2020). *How We Learn: Why Brains Learn Better Than Any Machine… for Now*. Viking.

Gilson, L. L., & Goldberg, C. B. (2019). Editors' comments: Conceptual analysis in management and organization studies. *Academy of Management Review*, 44(3), 520-535.

Golfinopoulos, E., Tourville, J. A., & Guenther, F. H. (2010). The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *NeuroImage*, 52(3), 862-874.

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience*, 31, 359-387.

Holone, H. (2016). The filter bubble and its effect on online personal health information. *Croatian Medical Journal*, 57(3), 298-301.

Immordino-Yang, M. H., Darling-Hammond, L., & Krone, C. R. (2019). Nurturing nature: How brain development is inherently social and emotional, and what this means for education. *Educational Psychologist*, 54(3), 185-204.

Kitago, T., & Krakauer, J. W. (2013). Motor learning principles for neurorehabilitation. *Handbook of Clinical Neurology*, 110, 93-103.

Krakauer, J. W. (2006). Motor learning: Its relevance to stroke recovery and neurorehabilitation. *Current Opinion in Neurology*, 19(1), 84-90.

Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. *Neuron*, 67(5), 713-727.

Lee, S., & Park, H. (2019). The effects of ASR-based pronunciation training on L2 learners' pronunciation diversity. *Language Learning & Technology*, 23(2), 45-62.

Lenneberg, E. H. (1967). *Biological Foundations of Language*. Wiley.

Low, L. K., & Cheng, H. J. (2006). Axon pruning: An essential step underlying the developmental plasticity of neuronal connections. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 361(1473), 1531-1544.

Pariser, E. (2011). *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Press.

Paus, T., Collins, D. L., Evans, A. C., Leonard, G., Pike, B., & Zijdenbos, A. (2001). Maturation of white matter in the human brain: A review of magnetic resonance studies. *Brain Research Bulletin*, 54(3), 255-266.

Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2014). Skype me! Socially contingent interactions help toddlers learn language. *Child Development*, 85(3), 956-970.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306), 1593-1599.

Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., & Toga, A. W. (2003). Mapping cortical change across the human life span. *Nature Neuroscience*, 6(3), 309-315.

Taub, E., Uswatte, G., & Mark, V. W. (2006). The functional significance of cortical reorganization and the parallel development of CI therapy. *Frontiers in Human Neuroscience*, 8, 396.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7(1), 49-63.

Wise, R. A. (2004). Dopamine, learning and motivation. *Nature Reviews Neuroscience*, 5(6), 483-494.

Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews Neuroscience*, 12(12), 739-751.

致谢

本文的研究思路受益于多位匿名审稿人的建设性批评，以及2025年在北京大学语言学实验室举行的“AI与儿童语言发展”小型研讨会上与会同仁的讨论。作者声明无任何利益冲突。

利益冲突声明

作者声明无任何利益冲突。
