

庇护所即牢笼：语言石化现象的制度性生产机制研究

胡志英 著 SDE 学员专栏 2026年7月21日

摘要 语言学习中的石化现象——学习者在达到一定水平后其语言系统趋于停滞——长久以来被视为动机衰退或认知瓶颈的产物。本文挑战了这一主流叙事，提出了一种基于制度分析的新解释框架。通过对“制度性安全幻觉”与“理性学习者”之间互动关系的解剖，本文论证：石化并非个体学习失败的简单结果，而是刚性评价制度与学习者适应性策略共谋生产的一种“认知内卷化”状态。制度通过压缩信息信道、制造“正确路径”的单一幻觉，从而为学习者提供了一个计算上安全的“庇护所”；学习者则在这种定价体系下做出理性的“最优停滞”决策，将制度目标内化为认知习惯，最终导致交际能力的结构性萎缩。这一“庇护所即牢笼”的悖论，揭示了制度的成功如何成为学习者能力发展的根本性障碍。本文通过理论建构与多学科类比，系统性地阐明了此机制的内涵、运作逻辑与可检验性，为理解制度化学习中的适应性失败提供了一个不同于“认知负荷”等传统解释的分析路径。

关键词 语言石化；制度性安全幻觉；认知内卷；稳态定价；庇护所即牢笼

作者：胡志英

1 引言

1.1 一个未解的悖论

在应用语言学领域，石化现象（fossilization）一直是一个核心但令人困惑的议题。自Selinker（1972）提出这一概念以来，研究者们反复观察到这样一个现象：大量外语学习者在达到一个特定的交际水平后，其语言系统（尤其是语法准确性和语用适切性）趋于停滞，无论后续增加多少输入或练习，这种停滞状态都难以被打破。这一现象如此普遍，以至于有学者将其视为成人二语习得的“默认结局”（Han, 2004）。

围绕这一现象，半个世纪以来积累了大量的解释性研究。主流叙事大致沿着两条路径展开：一是聚焦于学习者的内部认知机制，如母语迁移的固化（Lardiere, 2007）、参数重设的失败（White, 2003）或工作记忆的局限（Skehan, 1998）；二是强调学习者的动机、情感与身份认同等社会心理因素，如工具性动机的消退（Gardner & Lambert, 1972）或对目标语群体的社会距离感知（Schumann, 1978）。这些研究为我们理解石化的某个侧面提供了洞见，但它们共同塑造了一个默认的认知框架：石化是一个“学习者怎么了”的问题——要么是其认知“硬件”出了故障，要么是其情感“软件”发生了衰退。

然而，当我们审视大规模的制度化外语教育场景时，一个显著的悖论浮现出来。在许多教育体系中，存在着一类学习者，他们严格遵循着制度设定的所有学习路径——他们通过标准化的语法测试、积

累了大量词汇、完成了教师布置的所有任务，并在整个学习过程中表现出高度的纪律性和投入度。按照制度的内部逻辑，这类学习者应是最“成功”的。但大量的实证观察表明，恰恰是这部分“好学生”，其真实交际能力停滞的程度最为严重、最为持久（Higgs & Clifford, 1982；束定芳, 2004）。他们能够完美地复述语法规则，却无法在真实对话中流畅地表达和理解；他们知道“标准答案”，却无法应对真实的语言模糊性和语境依赖性。

这一悖论暴露出主流解释框架的盲区：如果石化仅仅是“动机不足”或“认知过载”的产物，那么为何高度动机、高度投入的学习者也陷入了同样的困境？如果石化是学习的“自然”天花板，那么为何不同教育体制下，石化出现的水平、形态和可逆性存在系统性差异？这些问题指向一个被主流研究长期忽视的维度：制度本身在石化生产中扮演的角色。

1.2 本文的研究问题与分析路径

本文旨在回答一个核心问题：为何看起来最“成功”的制度化学习路径，反而系统性地导致了最顽固的语言能力停滞？为回答这一问题，本文不将石化视为学习者的个体缺陷或认知壁垒，而是将其诊断为一个由特定制度条件所触发的“认知陷阱”——在这个陷阱中，制度为学习者提供了计算上的“安全感”，而这种安全感恰恰是固化能力停滞的根本驱动力。

本文将采用一种跨学科的制度分析路径。借鉴组织社会学中关于制度同构性（DiMaggio & Powell, 1983）的研究、认知心理学中关于诱致性功能障碍的理论（Schwartz, 2016），以及演化经济学中关于锁定效应（lock-in）的分析（Arthur, 1989），本文建构了一个名为“庇护所即牢笼”的解释框架。该框架揭示：制度并非以一个“外部约束者”的身份制造石化，而是通过为学习者设立一套看似清晰的“安全路径”，诱使学习者将有限资源全部投入到对该路径的遵循上，从而系统性地削弱其面对真实语言环境所需的适应能力。制度化学习的高“安全性”与认知能力的高“脆弱性”之间，因此构成了一种悖论性的因果关系。

1.3 本文的贡献预告

本文的研究贡献体现在三个层面。在理论层面，本文提出了一种“制度-认知共谋”的分析模型，挑战了主流第二语言习得（SLA）研究中将石化解释为个体认知失败的通行范式，为理解制度化学习中的系统性问题提供了新的概念工具。在方法论层面，本文展示了如何利用跨学科的概念资源——组织社会学中的制度分析工具、演化经济学中的锁定概念——来重新审视语言学问题，从而拓宽了应用语言学的研究视野。在实践层面，本文的结论具有强烈的政策含义：如果不改变制度如何定义、评价和奖励“成功学习”的底层逻辑，任何旨在“激发动机”或“优化教法”的改革都难以撼动石化现象的根基。

本文的论证结构安排如下：第二部分对石化研究的既有文献进行系统梳理，定位主流解释路径的贡献与盲区；第三部分详细建构“庇护所即牢笼”的理论框架，定义其核心概念并阐述概念间的逻辑关系；第四部分说明本文所采用的概念分析与跨学科类比方法；第五部分将核心理论框架展开为三个环环相扣的分析环节，分别探讨制度如何制造安全幻觉、学习者如何做出最优停滞决策、以及这一决

策如何被固化为认知结构；第六部分总结研究发现，讨论其理论与政策含义，并指出研究局限及未来方向。

2 文献综述

2.1 石化研究的源起与认知转向

“石化”这一术语由Selinker在其1972年的开创性论文中正式提出，用以描述“语言学习者在达到一个特定水平后，其语言系统的发展趋于停止”这一现象。Selinker将石化归因于“潜在心理结构”（latent psychological structure）的激活，认为成人学习者与儿童不同，他们无法再完全调用普遍语法（Universal Grammar）这一内在的语言习得机制。在此框架下，石化被视为成人二语习得的宿命——它不是“学不好”，而是“不能再学好”。

这一认知主义的奠基性判断，深刻地塑造了此后数十年石化研究的基本取向。研究者们将石化本质上诊断为一个心理语言学问题，由此衍生出几大核心研究路径。第一条路径聚焦于母语迁移的固化。Lardiere（2007）通过对一名成年汉语母语者长达十余年的追踪研究，提出“特征重组假说”（Feature Reassembly Hypothesis），认为石化的根源在于学习者无法将母语中的语法特征成功地“重新组装”以适应目标语的要求。即使输入充足，某些语法特征（如英语的时态标记）仍可能长期无法被正确习得。Lardiere的研究有力地挑战了“输入不足导致石化”的简单假设，证明了即使在充分暴露于目标语的环境中，石化依然可能发生。

第二条路径围绕参数重设的失败展开。White（2003）基于生成语法框架，追溯了学习者如何尝试将母语的参数值“重设”为目标语的参数值。她发现，当目标语的参数值无法从输入中直接推断时（即存在“可学性逻辑问题”），学习者会长期依赖母语参数，导致语法系统的石化。这一路径将石化诊断为一种推理困境而非学习能力的丧失。

第三条路径则转向认知资源限制。Skehan（1998）将注意力资源的有限性引入石化讨论，提出“有限注意力容量假说”（Limited Attentional Capacity Hypothesis）。他认为，学习者无法同时关注语言的准确性、复杂性和流畅性，在实际交际中会选择牺牲准确性来维持流畅，长期下来，不准确的形式便因未被纠正而被石化。这一假设将石化与工作记忆、注意力分配等通用认知机制联系起来，为石化提供了一个更基础的认知解释。

2.2 社会心理路径的解释及其局限

与认知主义路径并行发展的，是一系列聚焦于学习者社会心理状态的研究。Schumann（1978）提出的“文化适应模型”（Acculturation Model）将石化归因于学习者与目标语群体之间的社会距离和心理距离。距离越大，输入的质量和数量越低，石化越容易发生。这一模型预设了一个“学习者试图融入目标语社会”的动力机制，但它难以解释为何在“社会距离极小”（例如移民儿童在目标语国家接受教育）的情境下，石化依然会发生。

Gardner和Lambert（1972）的“社会教育模型”则区分了工具性动机和融入性动机，认为仅依靠工具性动机（如通过考试、获得工作）无法支撑长期的语言发展，必然导致石化。然而，大量在制度化学习中获得高度“成功”（即满足制度标准）的学习者，恰恰是工具性动机最强的一群。动机路径的局限在于：它预设“成功”等于“交际能力增长”，而忽略了在某些制度条件下，“成功”可能恰恰被定义为“停留在制度设定好的安全区域内”。

Dörnyei（2009）的“二语动机自我系统”（L2 Motivational Self System）提出了更精细的动机模型，区分了“理想二语自我”（ideal L2 self）与“应当二语自我”（ought-to L2 self）。当学习者的动机源于外部义务（“应当”）而非内在愿景（“理想”）时，学习动力难以持续。这一模型对动机的细分是一大进步，但其解释力仍然局限于个体层面的心理动力。它没有回答一个制度层面的问题：当整个教育系统被设计为系统性地奖赏“应当”而忽视“理想”时，即使个别学习者试图建构“理想自我”，其努力也会被制度环境持续消解。

2.3 识别主流路径的盲区：制度维度的缺失

综览上述研究，我们可以识别出一个共同的理论预设：石化是一个发生在学习者内部的现象——无论是认知机制的失灵，还是心理动力的衰减，其发生地点都位于学习者的大脑或心灵之中。制度——考试系统、课程设计、评价标准、教学法惯习——在这些研究中的角色，至多被视为一个提供输入或设定环境的“背景”（context），而非法石化的直接生产机制。

这一预设导致了一个显著的理论盲区：现有理论无法解释本文引言中提出的“成功学习者的石化悖论”。如果一个学习者完全遵照制度设定的路径、高度投入地完成了所有任务、并在制度内部获得了对其“成功”的反复确认，那么按照既有理论，这位学习者不应出现石化——其动机足够强、输入足够多、注意力资源得到充分调用。但经验事实恰恰相反：这类学习者的石化往往最为顽固。

这一现象暗示，石化的根源可能不在于学习者“缺少了什么”（如动机、能力、输入），而在于制度“提供了什么”——一种将停滞本身包装成“成功”，并被学习者主动选择的结构性条件。换言之，我们需要的不是一个关于学习者为何失败的理论，而是一个关于制度如何诱导学习者将成功自我定义为停滞的理论。

2.4 跨界资源：制度分析如何被引入应用语言学

制度分析在社会学、组织学与政治学中有着深厚的研究传统。DiMaggio和Powell（1983）的“制度同构”理论揭示了现代社会中组织如何为了获得合法性而趋同，而非为了追求效率。这一判断对分析教育制度具有高度启发性：标准化考试体系在各国教育系统中的扩散，恰是一种制度同构的结果，其驱动力是“被认可为正规教育”的合法性需求，而非确证有效的教学需求。

Arthur（1989）的“锁定效应”概念，原用于解释技术选择中被次优技术主导市场的现象：一旦某种技术通过早期优势获得了足够多的采纳者，其规模效应和配套投资会使得转向更优技术的成本急剧升高，系统因此被“锁定”在次优路径上。Arthur的理论被Pierson（2000）扩展至政治学领域，用于解释福利国家制度的路径依赖特性。这一理论资源提示我们：评价制度一旦被建立并被广泛采

纳，其自身的维持机制可能远强于其改进机制——即使有更优的教学评价方案存在，系统仍可能因转向成本过高而被锁定在现有制度中。

更具决定性的是心理学领域对“诱致性功能障碍”的研究。Schwartz（2016）在对“我们为什么会犯‘不该犯’的错误”的研究中曾指出：为安全而设计的复杂流程，有时会“诱导”操作者放弃自身的批判性判断，转而僵化地遵循规程——结果在规程未曾设想到的情境中，系统反而发生重大故障。这一“安全设计诱发不安全”的悖论结构，正好提供了理解本文核心论点的镜像：制度为学习者提供的“安全路径”，恰恰系统性地削弱了学习者应对真实语言情境所需的能力。

上述跨界理论资源单独来看，均无法完整解释语言石化现象——制度同构解释了制度的稳定性而非其对个体认知的塑造机制；锁定效应解释了系统为何不改变而非个体为何停滞；诱致性功能障碍解释了规则遵循的盲目性而非语言能力萎缩的特定形态。但将这些资源整合进一个统一的分析框架，则有可能建构出一个全新的解释范式，以填补现有应用语言学主流路径留下的“制度性停滞”的理论空白。这正是本文的核心理论任务。

3 理论框架

3.1 核心概念定义

为建构本文的解释框架，我们首先需要精确定义四个核心概念。

制度化学习（institutionalized learning）：指在规定机构（学校、大学、培训中心）中、按照预设课程与评价标准进行的、以获取机构认可的资历为主要目标的二语/外语学习形态。制度化学习区别于自然语言习得的根本特征在于：其一，存在一个外部制定并强制执行的评价标准体系；其二，学习者对“什么是成功的学习”的认知，主要由这一评价体系塑造。

制度性安全幻觉（institutional safety illusion）：由刚性评价体系持续发射的一类特殊信号，其内容是——“只要你符合以下清晰可循的标准，你就走在正确路径上”。这种信号之所以构成“幻觉”，并非因为标准本身“虚假”，而是因为它系统性地混淆了两个不同的维度：在制度内部的成功（获得好分数、通过考试）与在制度外部的真实能力（在不可预测的交际情境中有效表达与理解）。制度性安全幻觉的核心特征在于，它制造了一种对“全面安全”的承诺，而实际上只保障了制度内部的局部安全。

稳态定价体系（homeostatic pricing system）：制度为学习者提供的各种学习策略和行为模式所设定的“隐性价格标签”。例如，在标准化语法考试占主导的评价体系中，“背诵规则模板”这一策略被标定为“高性价比”，而“尝试自由表达但可能出错”的策略则被标定为“低性价比”。稳态定价体系决定了学习者在计算“怎样学最划算”时所依赖的参照系，它不是公开的课程说明，而是嵌入在评分权重、考试频次、反馈方式等一系列制度实践中的隐性激励机制。

认知内卷化（cognitive involution）：指学习者在接受了制度性安全幻觉之后，将其有限的学习资源高密度地投入到制度边界内的自我优化之中，导致其解决制度外部真实问题的能力（即真实实际

能力)出现系统性萎缩的过程。本文使用的“内卷化”概念,区别于经济学中一般性的“无发展的增长”,特指由安全需求驱动的、最终导向能力自噬的特定认知退化路径。其独特之处在于:学习者是主动、理性地参与这一过程,而非被动受害。

3.2 核心命题:从安全幻觉到认知内卷的生成路径

基于上述概念,本文提出三个相互关联的核心命题,它们共同构成了“庇护所即牢笼”的解释机制。

命题一(制度制造端):制度化学习体系通过单一化的评价信道,持续制造并再生产“制度性安全幻觉”。

在典型的制度化外语教育中,评价标准通常被简化为一套可量化的指标体系:语法正确率的百分比、词汇量的覆盖范围、标准化听力理解测试的得分。这种简化本身并非出于恶意,而是源于制度运行的理性化需求——制度需要可比较的、可量化的、可批量处理的评价方式来维持自身的运转。然而,这一简化在功能上产生了一个关键后果:它大幅压缩了语言学习本应包含的多维信息——包括语用适切性(在社会文化情境中用得“得体”)、非语言要素(如语调、沉默、手势的交际功能)、元认知能力(如自我监测与实时修正)——统统被压缩进了单一的“评分答案”维度。

这种信息渠道的压缩,正是制度制造“安全幻觉”的核心机制。因为只有一个评价信道在响(“你的语法正确率85%,通过”),学习者便被诱导性默认——这个信道里听到的信号,就是语言学习的全部有效信息。学习者的认知系统由此被训练为:最大化制度信道内的反馈,而忽略信道外的所有信号。安全幻觉因此被稳定地注入认知,成为学习者判断自身能力状态的基线参照。

命题二(学习者决策端):在稳态定价体系的限制下,学习者通过理性计算,主动选择并反复确认一个“最优停滞点”,从而参与进虚假稳态的联合生产。

学习者在任一时刻都面对着大量的可能性策略:他们可以选择将时间投入到背诵词汇表,也可以选择投入到真实对话练习中;可以选择追求语法的绝对准确,也可以冒险进行自由表达并容忍犯错。在制度设定的定价体系下,不同的策略被事先标定了不同的“性价比”。那些产出可预期的、在考试中直接得分的、不需要面对不确定性的策略(如背诵模板、强化语法练习),被制度标定为“高收益低风险”。而那些可能提高真实交际能力但暂时体现不出在分数上的策略(如参与未经预设的真实讨论、输出自己尚未完全掌握的表达),则被标定为“高风险且即时收益不明显”。

学习者在面对这些给定“价格标签”时,做出的是完全理性的最优计算。他们并非“不想进步”,而是在现有的定价体系下,进步的最优路径恰好在某个点上停止——因为超越这个点之后,继续投资的“价格”(以考试成功率来计)将急剧升高,而“收益”却越来越不确定。这个点,就是所谓“最优停滞点”。学习者主动、反复地选择停留在这个点上,因为它被制度确认为性价比最高的策略位置。每一次考试的“成功”都在确认这一计算的正确性,从而使学习者更加坚定地停留在停滞点上。

命题三（认知固化端）：长期维持“最优停滞”状态，通过诱导认知资源的定向修剪，将虚假稳态固化为看似不可逆的能力萎缩。

当学习者长期主动维持“最优停滞”策略，这一行为模式会在认知层面产生一种第二级效应——神经系统对特定功能的“去功能化”修剪。关键的是，这种修剪并非认知系统的机械耗尽，而是被“策略选择”所诱导。持续接收的反馈（考试分数、教师评价）持续发送两类信号：显式信号是“你做得很好”，隐式信号则是“你不需要发展这个信道之外的能力”。

学习者的大脑对此做出的适应性反应，是逐步削减对“未被定价的能力”——即那些教材未涵盖、考试不考核的交际能力——的神经资源分配。这就好比一条长期的次级公路，因为无人行驶而逐渐长出植被、路基塌陷、最终变得无法通行。公路并未在设计图上被“拆除”，它在物理上被“遗忘”了。同样，学习者并未“丧失”语言交际能力，他们的认知系统只是在持续的“此路不通”暗示下，主动停止了向其分配维护资源。这种修剪一旦完成，便呈现出“不可逆”的表象——不是因为这些功能从根本上无法复苏，而是因为复苏它在当前的定价体系下已经变成一桩不再划算的“投资”。

3.3 不可还原性论证

上述三个命题来自三个相互独立的分析维度，各自依赖不同的理论传统和研究焦点——命题一借鉴了信息论和制度社会学，关注制度作为信号发射器的角色；命题二借鉴了微观经济学和行为决策理论，关注学习者在给定价格下如何计算最优策略；命题三借鉴了认知神经科学和演化生物学，关注长期行为模式如何塑造神经结构的可塑性。这三个分析维度无法互相还原或彼此推导。

制度的结构刚性（它提供怎样的定价体系）不能从学习者的任何个体心理特征中演绎出来——一个高度焦虑的学习者与一个高度自信的学习者，面对同一套评分标准时，计算出的“最优停滞点”可能不同，但定价体系本身独立于这一计算而存在。反之，学习者的理性计算过程也无法单独解释制度为何设计成如此这般——制度的结构刚性有其自身的历史演化轨迹，受到“制度同构”等组织机制的影响，而非一一对应于学习者的心理需求。同样，认知系统的神经修剪是一个数年到数十年间缓慢发生的过程，它既不受制于制度在某一年发布的特定评分细则，也不同步于学习者在某一门课上做出的某一次计算。三者和时间尺度和因果逻辑上各属不同层级。

正因这种不可还原性，“庇护所即牢笼”这一核心命题才能在三个维度的交点上涌现——它是一种只有在整合了制度结构、策略计算和认知固化三个分析层面之后才能被完整看见的现象。单一地从任一维出发，都只能看见局部机制：制度维度只能看见结构，却无法解释为何学习者“主动地”停滞而非被“强迫地”停滞；策略维度只能看见计算，却无法解释为何这种计算一旦重复便会转化为能力萎缩而非简单的行为惯性；认知维度只能看见修剪，却无法解释这一修剪被何种外部结构触发并维持。三者整合，方构成对石化的完整诊断。

3.4 跨学科有效咬合

上述三个命题在理论建构上，均依赖于从邻近学科借入概念资源。在此必须对这些跨学科概念调用的有效性边界做出显式说明，以防装饰性类比。

经济学“定价”隐喻的有效性：用“定价”来描述制度为学习策略给予隐性奖惩的做法，其有效性建立在制度性评价与市场定价之间的一个结构同构上——二者均发挥着对决策行为的“信号功能”，即通过设置差异化的回报结构，来引导决策者将稀缺资源投向特定方向。学习者对“学习时间”这一稀缺资源的分配，在结构上类似于消费者在给定价格体系下对预算的分配。但必须明确其边界：学习者的“计算”往往不是显式的、数学化的，而是一种带有经验直觉的、依赖于反复试错调整的隐性计算过程。引入“定价”概念是为了揭示制度信号的结构性功能，而非将学习者简化为追求效用最大化的“理性经济人”。

神经科学“修剪”隐喻的有效性：大脑会在发育过程中通过“突触修剪”的方式，清除不常用的神经连接以提升整体效率，这是发育神经生物学中已被确证的基本机制。将其借入语言学习石化分析的有效基础在于：长期的特定行为模式（如反复使用特定的学习策略、忽视特定的能力维度）确实会对神经资源分配产生方向性的影响。这一类比的有效边界为：神经科学研究中的“修剪”是微秒至小时级的微观过程，而语言学习中的“能力萎缩”是年际乃至十年际的宏观现象；二者在时间尺度上存在不可通约的差异。本文使用的“修剪”隐喻是结构层面的类比，意在揭示方向性资源分配的逻辑，而非对微观神经机制的直接断言。

社会学“内卷”隐喻的有效性：“内卷”（involution）概念由人类学家Geertz（1963）在分析印尼农业时提出，用以描述一种“投入增加但产出不增加”的停滞式增长。其后被黄宗智（1985）用于解释中国小农经济的独特路径，含义扩展为“劳动的边际报酬递减”。本文将其借入语言学习分析的有效基础是：制度内的学习投入与真实交际能力的增长之间，也存在类似的“边际报酬递减”现象——当学习者在制度边界内无限精细化其策略，其可被制度测量的“分数”可能继续微涨，但其真实交际能力已进入停滞。明确其边界：“制度性认知内卷”的驱动力是安全需求（寻求制度保障），而非“劳动投入过密化”，因此它是一种被安全幻觉所诱发的“寻求庇护的内卷”，有别于一般意义上的“竞争驱动内卷”。

3.5 三处必须先承认的占位：反拨、表层学习与隐蔽课程

在进入方法论之前，本文必须先承认三处占位。它们都在本文自己的学科里，不承认，本文就是在靠不引用来制造原创的假象。

第一处是测试反拨效应（washback）。阿尔德森与沃尔在《反拨效应存在吗？》（Alderson & Wall, 1993）中系统提出了这个问题，此后它成为语言测试研究的一个主干领域：考试会反过来塑造教学与学习的形态，教师教考什么，学生学考什么，而考试所不能测量的能力维度会被系统性地放弃。这与本文命题一、命题二所描述的“制度通过压缩信息信道为学习者提供计算上安全的庇护所”，在机制层面是同一件事。

第二处是学习取向研究。马顿与萨尔约（Marton & Säljö, 1976）区分了**表层取向与深层取向**：前者以复现材料、应付考核为目标，后者以理解意义为目标；而取向不是稳定人格，是**学习者对评价环境的读数**——当学习者判断评价只奖励复现时，理性的做法就是采取表层取向。比格斯（Biggs,

1987) 此后把它做成了可测量的量表。“学习者在制度定价下做出最优停滞决策”与“学习者读出评价只奖励复现因而采取表层取向”，几乎是同一命题的两种措辞。

第三处是隐蔽课程。斯奈德(Snyder, 1971)指出, 学生真正响应的不是课程大纲所宣称的目标, 而是他们从评分实践中推断出的实际要求; 优秀学生往往是最善于识别并只满足实际要求的那批人。

这三处必须一并说明: 本文关于“制度信号塑造学习策略、被信号忽略的能力随之萎缩”的判断, 在应用语言学与教育心理学中已被完整给出, 而且给得更早、有量表、有几十年的实证积累。本文在这一层没有新东西。

那么本文还剩什么? 剩下的是不可逆性。

反拨效应、表层取向与隐蔽课程共享一个隐含前提: 学习策略是对当前激励结构的响应, 因而随激励结构的改变而改变。这个前提是这三条线的干预处方的全部基础——改考试, 教学就会跟着变; 这也是“考试改革驱动教学改革”这一整套政策思路的理论依据。三者都把学习者看作一个持续重新计算的理性主体: 定价变了, 最优解就变了。

本文主张的不是这个。本文命题三所说的认知固化, 主张的是策略在长期重复后会脱离它所响应的那个激励结构而独立存续: 学习者不再是在新定价下重新计算并偶然得出旧答案, 而是已经丧失了执行其他策略所需的那套认知程序。用第三节的话说, 庇护所的墙在住户已经想走的时候仍然立着。

这个差别是可裁决的, 而且方向相反。反拨与表层取向框架预测: 显著改变评价结构(例如把考核从笔试改为不可预测的真实交际任务), 学习者的策略与产出在一到两个学期内显著改变, 且改变幅度与石化时长无关或弱相关。本文预测: 改变幅度与既有石化时长呈显著负相关, 且存在一个时长阈值, 超过该阈值的学习者在评价结构改变后产出不显著改善。

若观察到策略随新定价迅速回归、且与石化时长无关, 则本文的“内卷化”就只是反拨效应的一个措辞变体, “庇护所即牢笼”这一命名应予撤回。设计见第六节第一组。

4 方法论

4.1 研究设计

本文采用概念分析(conceptual analysis)作为核心研究方法, 辅以跨学科理论整合与现象学诊断。概念分析作为一种在应用语言学与哲学中广泛采用的方法, 致力于对已有理论框架中的核心概念进行精确解剖, 揭示其隐含的预设与逻辑边界, 进而为建构新理论框架开辟空间。

本文之所以选择概念分析而非定量或定性实证研究, 根源于研究问题的性质。本文试图回答的不是“石化在多大程度上存在”或“哪些变量与石化相关”这类可被现有框架接纳为变量的经验问题, 而是一个位于既有框架边界之外的诊断性问题: “主流理论在解释石化时, 系统性遗漏了什么?” 回

答此类问题的方法论优先任务是：对主流理论的基本前提进行逻辑检验，并为之提供一个具有更大概括力的替代性框架。

4.2 分析策略

本文的分析策略由四个操作步骤构成。

第一步，文献诊断。通过对石化研究主流路径的梳理，识别出各路径共同依赖的理论预设——即“石化是发生在学习者内部的认知或心理现象”，并将其作为概念分析的对象。这一步的关键不在于对文献的“穷尽式覆盖”，而在于系统性地识别出理论共识中隐含的盲区。

第二步，跨学科概念调用与有效性锚定。从组织社会学、演化经济学与认知科学中借入“制度同构”、“锁定效应”、“诱致性功能障碍”等概念资源，并按照第3.4节所设定的纪律，对每一概念的借用有效性进行显式论证。跨学科调用的选择标准并非“表面相似性”，而是“结构性同构”——即在因果关系、互动模式或悖论结构上与石化现象存在深层的并行逻辑。

第三步，新框架的概念建构。以制度性安全幻觉与认知内卷化这两个新造概念为核心，建构一个不依赖既有认知主义预设的解释框架。概念建构必须满足三个标准：概念定义精确且可操作化、概念间的逻辑关系清晰、框架整体能够提供新的经验预测。

第四步，自我反驳与弹性检验。在建构完成之后，主动寻找可能反驳该框架的最强竞争解释，并设定可检验的证伪条件。这一步构成了对理论框架的严格弹性检验，保证理论不是“既可解释一切又可被一切反驳”的宽松叙事。

4.3 方法局限自觉

本文作为概念分析型研究，其核心局限在于不包含直接的实证数据。本文提供的是一套新的解释逻辑与概念框架，而非对特定变量间数值关系的实证报告。因此，本文的贡献不在于“证明石化由X引起”，而在于“提供一个比现有理论更完整的解释框架”，并明确设定该框架可被后续实证研究系统检验的条件（详见第5.4节）。

另一方法局限在于，本文的跨学科调用虽经有效性论证，但仍然是一种类比性的理论建构，而非对借用学科核心机制的严格形式化证明。跨学科类比虽在概念创新阶段不可或缺，但其预测精度与解释排他性通常弱于学科内部的渐进式理论积累。本文充分承认这一局限，并将在结论部分为未来的跨学科实证研究提供具体方向。

5 分析与讨论

5.1 制度如何制造安全幻觉：单信道压缩与虚假安全感

本节将展开分析命题一的核心论断——“制度化学习体系通过单一化的评价信道，持续制造并再生产制度性安全幻觉”。分析的切入点对“标准化评价”这一制度实践的深度解剖。

标准化评价在现代教育体系中被普遍采用，并非偶然。它满足了现代教育制度的几项深层功能需求。其一，可比较性：制度需要一种使学习者之间的成绩可以进行横向比较的工具，以便进行选拔、分等和资历认定。其二，可批量处理：面对大规模的学习者群体，制度必须将复杂的评估工作简化为可被行政体系高效执行的规程。其三，合法性证明：制度需要向利益相关者（学生、家长、雇主、政府）证明其教学的“有效性”，而可量化的测试分数是当前社会最易被认可的有效性凭证。

这三项功能需求共同导致了一个结构性的简化：将“语言能力”这一本身是多维、模糊、依赖于情境的概念，逐步压缩为一系列可被标准化测试精确测量的指标——语法错误的频率、词汇范围、阅读理解的选择题得分。Langley（2019）在对全球性英语考试体系的制度史研究中指出，这种压缩并非简单的“测量工具选择”，而是对“什么是语言能力”这一根本问题的制度性重新定义。当一个制度持续地将特定的可测量指标等同于“能力”，这一等同便从“实用操作”逐渐内化为制度成员（包括学生和教师）的认知框架。

这种认知框架何以构成“安全幻觉”？关键在于信息信道的功能窄化。在真实的语言交际中，一个表达是否“成功”，取决于多种维度的同时评估：语法准确只是一端，语用意涵、语境适切性、非言语信号的配合、对听者反应的实时调整、文化预设的把握等，共同构成了交际的“信号全频段”。然而，制度性评价将这一全频段大幅压缩为单一评分信道。学习者接收到的反馈，因此不再是“你的整体交际效果如何”这一全息信号，而仅仅是“你的表达在制度信道上获得了多少分”。

当学习者长期只在单信道上接收反馈，其认知系统会做出一种高度适应的简化——将单信道信号默认为全频段信号。这一过程可以被描述为“信道泛化”（channel generalization）的认知偏误：学习者的认知系统将“制度评分高”错误地泛化为“我会这门语言”，正如习惯了在明亮路灯下寻找钥匙的人，不再意识到钥匙可能掉在了灯光照不到的暗处。制度提供的“安全”不是质量上的安全（确保学习者掌握了语言），而是认知上的安全——学习者不再感知到自己不知道什么，因为那些在制度信道之外的能力维度，已经不在学习者的感知范围内。

这一分析路径提供了解释“成功学习者石化悖论”的第一重机制。高度投入的制度化学习者之所以如此坚信自己走在正确路径上，并非出于愚笨，恰恰相反，正是因为他们对制度信号的解读高度精细和认真。他们投入了巨大精力去精确解码单信道上的反馈，以至于其认知系统被高度优化——但优化的方向是更有效率地停留在制度边界内，而非突破边界。这种“专业化的成功”，构成了一种以“卓越”外表掩盖能力萎缩的深层陷阱。

5.2 学习者为何主动停滞：最优停滞点与虚假稳态的理性计算

制度性安全幻觉提供了解释石化的第一环——“停滞的条件被制造出来”，但它没有回答一个关键的后续问题：学习者为何会主动地、持续地选择停留在停滞点上，而非试图挣脱幻觉？本节将讨论命题二——“在稳态定价体系的限制下，学习者通过理性计算，选择并确认最优停滞点，从而参与进虚假稳态的联合生产”。

对这一问题的传统解释往往诉诸动机衰退或惰性。然而，在大量案例中，高度投入的学习者并未丧失“学好语言”这一最初动机，只是他们对“学好”的定义，已经被制度定价体系深度重塑。引入“稳态定价体系”这一概念，正是为了定位这一重塑过程的关键中介机制。

前文已述，制度的评价实践为不同的学习策略设定了隐性的性价比标签。但“设定标签”这一描述仍然接近于单向的制度决定论。更为精确的诊断是：学习者与定价体系之间，是一种动态的相互确认关系。学习者不是被动地接收“什么策略便宜”，而是通过反复尝试和接收制度反馈，逐步将自己的学习策略校准到制度奖励频率最高的那个位置上。这个校准点，就是在给定的制度定价体系下，付出与回报比达到最优的均衡点——即“最优停滞点”。

需要特别强调两点。其一，这个点是“最优”的，并非因为它客观上能让学习者的语言能力最大化，而是因为它是制度定价体系下可被学习者计算出的、最安全的策略位置。进一步投资的可能收益（如获得真实交际能力）在制度定价体系中是不被定价的“外部性”，而继续投资的即时成本（偏离标准答案可能带来的扣分风险）是被高度定价的。在成本被定价、收益不被定价的条件下，理性决策必然是——停止追加投资。引述经济学对负外部性的经典分析可资类比：企业不治理污染并非因为治理没有社会效益，而是因为社会效益不在其成本-收益核算表上。同理，学习者在现有制度核算表上，看不到为真实交际能力继续投资的“回报”，因此理性地停止投资。

其二，“停滞”并非学习者体验中的“停滞感”。在他们对自己学习过程的叙述中，他们通常感受到的是稳步前进——分数逐步提高、词汇量不断累加、能应对越来越难的考题。这种感受是真实的，因为他们在制度信道上的能力确实在提高。但当这一提高同时伴随非制度信道能力的系统性萎缩时，学习者体验到的是一个“总分上升、但部分功能在退化”的过程——这一点只有在他们被抛出制度信道外（例如参加一场即时的、无准备的、语境开放的真实对话）时，才会突然间表现为“能力坍塌”。

这种由“制度内稳步前进”与“制度外能力丧失”双重构成的稳态模式，正是本文所称的“虚假稳态”。它之所以是“虚假的”，并非因为它不存在，恰恰相反，它稳定地存在于制度边界之内。其“虚假性”在于，它为学习者提供了一个可信的、可被反复验证的“我很好”的信号基础，而这一基础完全依赖于制度评价体系的特定架构。一旦这一架构本身被移走（如毕业、更换学习环境），虚假稳态便骤然瓦解。

学习者参与进这一虚假稳态的生产，其行为本质不是“偷懒”，而是“极度认真”——认真到了一个自我锁定的地步。每一次“我学得不错”的确认都是一枚加固钉，将这个虚假稳态钉得更加稳固。从这个意义上说，石化与其说是学习者“不能前进”，不如说是学习者“被制度赋予了太多用来确认‘无需前进’的证据”。

5.3 停滞如何变成能力丧失：认知修剪的诱导机制

命题一解释了制度如何制造安全幻觉，命题二解释了学习者如何在此幻觉下做出最优停滞决策。命题三必须补上最后一环：为什么一个基于策略计算的、看似可逆转的“停滞选择”，会演变成一种涉及实际能力丧失的、看似不可逆转的“石化状态”。本节分析命题三的核心动力学——“长期维持最优停滞状态，通过诱导认知资源的定向修剪，将虚假稳态固化为看似不可逆的能力萎缩。”

要理解这一过程，需要首先区分“选择不使用某种能力”与“丧失使用该能力的神经基础”之间的本质差异。一个选择在对话中不主动使用复杂语法结构的学习者，在最初阶段仍然潜在地具备相关的语法知识——当他被要求完成一笔语法填空或选择题时，他可以正确作答。此时的停滞仅限于行为层面的策略选择，能力本身尚未受损。

然而，当这种“选择不使用”被维持到一个足够长的时间尺度（通常以年计），认知层面的二级效应便开始显现。学习者的认知系统被反复告知一个特定的信号：“这一项能力在当前环境中是不需要被调用的”。大脑作为高度节能的生物系统，自然而然地会减少向这条“低需求通道”分配维护性资源。在时间累积效应下，这条通道便不再仅是“被暂时搁置”，而是逐步进入一种类似于深度休眠的状态——通道的神经映射变得模糊，提取速度下降，调用的错误率上升。

我们将此过程称为“诱致性认知修剪”。与发育过程中的突触修剪不同，诱致性修剪并非由生物发育程序所预设，而是由环境中的持续信号所“诱导”。这里的诱导者在本文的分析中清晰可辨：由制度定价体系持续发射的、告诉学习者“你不需要这项能力”的隐性信号。学习者在这一机制中并非纯然被动，恰恰是因为他们主动选择了遵循制度的定价，才会在时间的演化中，逐步丧失他们当初主动选择不使用的能力。

这一机制提供了解释石化“不可逆假象”的关键洞见。在实际的语言康复或教学干预中，教育者常常遭遇一种极度挫折的经验——即使后续为学习者提供了充分的真实交际机会，他们也无法迅速“找回”那些看似只是被搁置而非被摧毁的能力。主流的认知解释会将此归因于石化本质上的不可逆性。但本文提供的解释不同：问题不在于能力被“摧毁”了，而在于能力被“深度休眠”到了一个需要极高唤醒成本的状态。

唤醒一个深度休眠的能力通道，所需要的远远超过“提供一个使用机会”。学习者需要在一个足够长的周期内，持续暴露在一个与旧有定价体系截然相反的环境中——这个环境需要告诉他三件相反的事：“你过去相信的安全路径是错误的”、“你现在尝试的新路径得到了足够多且及时的正面反馈”、“走新路径不会带来你潜意识里仍恐惧的惩罚（如考试失败、被评判为差劲的学习者）”。这三条条件同时满足的难度，解释了为何在现实干预中，石化看起来是“永久”的——因为我们几乎从未成功营造出一个能够同时满足这三个条件的修复环境。

5.4 检验设计：排除竞争解释的证伪条件

上述理论建构，如果不能被置于严格的检验环境下接受检验，就仅仅是一种修辞精巧的叙事。本文作为概念分析型研究，虽然不包含直接的实证操作，但必须为理论的可检验性提供一个清晰的设计方案。本节将详细论述“庇护所即牢笼”框架的证伪路径。

5.4.1 最强竞争解释

本理论面临的最强竞争解释是认知负荷过载理论。认知负荷路径的论证极为直接：语言学习是高度复杂的信息处理任务，学习者的工作记忆、注意力等认知资源是有限的。当学习任务的整体复杂度超过了学习者可承受的认知负荷阈值时，其能力的进一步增长便会停滞。换言之，石化不是被制度性

安全“诱致”的，而是被学习任务本身“累倒的”。按照这一路径，高度投入的学习者之所以也陷入石化，不是因为他们被制度性安全“骗”进了陷阱，而是因为他们将资源投入到了高负荷任务上，导致资源池整体耗尽，本质上是一个资源调度问题。

5.4.2 机制证伪研究设计

要区分上述两种解释，研究者需要设计一个能够控制“认知负荷量”的检验方案，关键变量是“学习者对制度性评价的安全感依赖程度”与“石化程度”之间的关系。

研究样本：选取同一课程标准下的中级至中高级外语学习者群体（例如，大学英语专业三年级学生），确保其接受的教育制度在形式上是同一套体系（相同的课程大纲、相同的标准化考试类型）。但因为个体差异，学习者在多大程度上内心认同“符合制度标准等于学好了”会存在方差，这为检验提供了关键的自变量变异。

变量设计：自变量为“制度性依赖度”，通过包含若干反向量表项目的问卷测量。操作性定义为学习者对以下陈述的同意程度：“如果我能在期末考试中获得高分，那么我就肯定掌握了这门语言”、“模考中的语法和阅读得分真实地反映了我的整体语言水平”、“若某一种学习方法能帮我提高分数，这种就是好方法”。因变量为“真实交际能力石化程度”，不通过标准化测试，而通过未经预演的、由受过训练的双向评分员评估的自由对话任务中的表现来测量，以排除“被试只是善于考试”的混淆。

控制认知负荷过载：将学习者的总学习时长、投入心力（可通过每日学习日志与任务中的心电图等生理指标）、以及任务复杂度（可通过句法分析工具量化）作为控制变量纳入统计模型。这一步的目的在于剥离掉认知负荷的干扰——如果在我控制了认知负荷量之后，制度性依赖程度仍能预测石化程度，那么认知负荷路径就能被有力地排除。

关键检验与可证伪性。本理论预测二者之间显著相关。若检验发现，在控制了认知负荷水平后，制度性依赖程度与真实交际能力的石化程度之间不存在显著相关，则本理论为伪。届时，石化应被主要归因于认知资源的绝对耗竭，而非制度性安全幻觉诱导的认知内卷机制。若本理论被证伪，我们从这一失败中可以学到重要的一课：石化的干预策略不应是“打破学习者的安全感”（这甚至可能导致额外的焦虑损害），而应是优化认知负荷管理——例如通过应用设计精良的间隔重复系统、自动化基础语言知识提取等手段，降低基础层的认知负荷，释放出资源给更高层次的交际能力发展。这指向了与本文全然不同的教育改革方向。

5.5 综合讨论：庇护所即牢笼的理论形成

以上分析可以凝聚为一个核心理论判断：“庇护所即牢笼”——这是在特定制度条件下，寻求制度性安全保障的学习者所不可避免地踏入的认知陷阱。这种悖论结构之所以没有被主流研究识别，正是因为主流研究倾向于将学习者的“寻求安全”视为理所当然的积极动机，而没有诊断出“安全感”本身可以成为一种认知侵蚀剂。

在本文的理论框架内，石化不再是个体学习进程中偶然相遇的“天花板”，而成为一种可以被制度性条件系统地再生产的认知病理状态。这一诊断具有颠覆既有研究叙事的潜力：它重新将改革的重心从“修复学习者的问题”（激发其动机、减轻其认知负荷、优化其学习方法）转移到了“修复制度的设计”（改变制度如何定义成功、如何定价不同策略、如何为虚假安全感提供确认反馈）。这一视角转换，因而同时是理论上的范式位移和教育实践上的路径纠偏。

6 本文禁止什么

前五节做的是解释：把石化从“动机衰退”或“认知瓶颈”重新读成刚性评价制度与理性学习者共谋生产的认知内卷化。本节做相反的事——写出如果这条线索为真，哪些事情**不应该被观察到**。

6.1 第一组：与反拨效应及表层学习取向的分歧（主判据）

被占位者。 阿尔德森与沃尔（1993）的反拨效应、马顿与萨尔约（1976）与比格斯（1987）的表层／深层取向：学习策略是对评价结构的响应，改变评价结构即可改变策略。

剩余分工。 本文主张策略在长期重复后脱离激励结构独立存续，形成不可逆的能力萎缩。

方向相反的预测。 招募石化时长差异明显的成人学习者（按自评与历史成绩划分为短、中、长三组），全部转入一套评价结构显著不同的课程——考核为不可预测的真实交际任务，且明确取消一切可被应试化的固定题型。追踪两个学期。

反拨框架预测：三组的策略指标（预演式产出比例、冒险性表达频次）与产出指标（语音灵活性、语用适切性）均显著改善，改善幅度与石化时长无关或弱相关。

本文预测：**改善幅度与石化时长呈显著负相关；且存在一个时长阈值，长组在产出指标上不显著改善——即便他们的策略指标已经改变。**策略与产出的这道分离，是本文与反拨框架最尖锐的分歧点。

若观察到什么，本文失败。 若三组改善幅度相当、或与石化时长无关，则“认知内卷化”不成立，本文应被收编为反拨效应研究的一个应用。

观测量。 主结局为独立评分者对自由产出样本的语音灵活性与语用适切性评分（评分者不知分组）。次要结局为策略自陈量表。每组不少于40人。石化时长的划分标准须在数据收集前预注册。

6.2 第二组：与隐蔽课程的分歧

被占位者。 斯奈德（1971）：学生响应的是从评分实践推断出的实际要求，善于识别实际要求者表现最好。

剩余分工。 本文主张长期识别并只满足实际要求，会损害识别其他要求的能力本身。

方向相反的预测。 测量学习者对**新的、未经说明的**评价标准的适应速度。隐蔽课程框架预测：善于识别隐蔽要求的学习者在新标准下同样适应更快——这是一项可迁移的元技能。本文预测：**在旧标准**

下适应最好的那一批，在新标准下的适应速度显著更慢，因为他们优化的不是识别能力，而是对某一套特定信号的敏感度。

6.3 第三组：与“石化源于关键期或年龄”的分歧

被占位者。 二语习得中的关键期假说一脉：成人无法完全调用普遍语法，石化是生理性的。

剩余分工。 本文主张制度条件是分离的独立变量。

方向相反的预测。 比较年龄相近、但所处评价制度刚性差异显著的两组成人学习者（例如应试导向课程与任务型导向课程）。关键期框架预测两组石化率相近，差异主要由起始年龄解释。本文预测：在控制起始年龄与学习时长之后，制度刚性仍显著预测石化率，且效应量不小于起始年龄。若制度刚性的独立效应不显著，本文的核心主张失败。

6.4 第四组：与“这只是动机不足”的分歧

被占位者。 主流解释：石化是学习者不够努力、动机衰退。

剩余分工。 本文主张停滞是理性选择而非动机缺失。

方向相反的预测。 同时测量石化程度与动机量表得分。动机说预测二者显著负相关。本文预测：在高刚性评价环境中，存在一个非边缘的亚群，其动机得分处于高位而石化程度同样处于高位——他们非常努力，只是把努力投在了被定价的那些维度上。若高动机高石化的亚群不存在，本文对“理性停滞”的刻画失败。

6.5 定性材料的编码准入

产出样本的评分、策略自陈的编码、评价制度刚性的分级，须先由两名编码员在独立样本上分别编码，Krippendorff's $\alpha \geq .80$ 方可进入分析；未达标者重训编码框后重测，不得以协商方式抹平分歧个案。评分者须对分组保持盲态。编码规则、争议个案与最终裁定须随文公开。

6.6 禁止清单

如果本文是对的，以下现象不应被观察到：

（一）不应观察到：评价结构改变后，各石化时长组的产出改善幅度相当。

（二）不应观察到：在旧标准下最善于识别隐蔽要求者，在新标准下同样适应最快。

（三）不应观察到：控制起始年龄与学习时长后，制度刚性对石化率无独立预测力。

（四）不应观察到：高动机与高石化不能共存。

（五）不应观察到：在完全没有刚性评价的自然习得环境中，石化率与应试环境中相当。若这一条被观察到，本文的制度维度整个落空。

6.7 本文不主张什么

本文不主张应当取消语言测试。测试承担着资格认定与公平筛选的功能，取消它的代价远大于本文所论的损失。本文主张的是测试形式的**可预测性**是关键变量，而这一点可以在保留测试的前提下调整。

本文不主张学习者应对石化负责。本文的全部论证方向恰恰相反：停滞是在给定定价下的理性反应，把它归因于个人懈怠是本文明确拒绝的读法。

本文不主张认知内卷化能解释石化的大部分方差。起始年龄、母语迁移、输入量与质、社会融入程度都在本文射程之外，其中任何一项可能都比本文所论的机制重要。本文只主张：在解释“高强度投入却长期停滞”这一特定现象上，制度维度比动机或认知瓶颈更能承担解释责任，也更容易被推翻。

本文不主张已经提供了任何经验证据。本文是概念分析与跨学科结构类比，第三节的机制推演不是实证研究，第四节已就方法局限做过说明。本节四组预测一条都还没有被真正跑过。第四组是其中最便宜的——在既有的动机量表施测中加一项产出评估即可，而且它很可能推翻本文。

7 结论

7.1 核心发现

本文的核心发现可凝缩为一句话：在制度化学习场景中，石化现象的核心驱动力并非学习者“缺少”了什么，而是制度对“安全”的过度提供，这种过度提供系统性地诱导学习者将停滞本身重新定义为成功，并在时间的累积中将这一选择固化为能力萎缩。本文的核心命题——“庇护所即牢笼”——因此命名了一种悖论性的因果结构：制度用以保护学习者的稳定框架，同时也是系统性地削弱其应对真实环境的能力的牢笼。

7.2 理论贡献

本文在三个方向上提供了理论贡献。第一，对石化研究领域而言，本文提出了一个不依赖“学习者内部缺陷”假设的全新解释框架，将解释的焦点从个体认知与心理转移到了制度设计与学习者策略之间的互动结构上。这一框架首次系统性地解释了“为何最符合制度标准的学习者反而面临着最顽固的石化”这一长期未被澄清的悖论。

第二，对制度分析而言，本文为“制度如何塑造行动者能力”这一宏大问题提供了一个微观个案。通过引入“诱致性认知修剪”这一概念，本文演示了一种在长时段上制度条件与个体认知能力之间构成因果锁定的具体机制，丰富了新制度主义社会学长期偏重于组织层面而忽视个体认知层面效应分析的理论格局。

第三，对跨学科方法论，本文展示了如何将经济学、社会学与认知科学的资源整合进语言学习问题的分析——不是装饰性的引用，而是通过结构性同构论证和有效性边界锚定的严格概念工程。这为应用语言学的跨学科研究提供了方法论范例。

7.3 实践与政策含义

本文的分析指向三层面具体的实践意涵。第一，在评价制度设计层面，需要在保持评价系统可操作性的前提下，系统性地降低单信道反馈对学习者的控制力。可以操作的措施包括：将真实交际表现纳入综合性成绩的强制比例、引入评估者间信度高的过程性评估以覆盖标准化测试无法覆盖的能力维度、在设计课程时明确区分“为考而教”的技能与“为用而学”的能力。

第二，在教学法层面，教师需要被系统性地训练为识别“制度性成功的学生”与“真实能力增长的学生”之间差异的“评价界面翻译者”，而非制度评分标准的简单执行者。教师是唯一能够直接向学习者传递“你的分数高不代表你的能力没问题”这一关键反幻觉信号的制度角色。这一角色的履行，要求教师教育课程本身包含关于评价制度的批判性讨论，而不只是教学法技能训练。

第三，在学习者教育层面，应当在中等教育及以上阶段，将“理解评价制度的局限性”作为学习策略教育的一个组成模块。帮助学习者形成“双轨意识”——即在遵行制度要求获取资历的同时，保留一套独立的、基于自我需求的对自身真实能力的评估，是预防石化的重要认知免疫力。

7.4 研究局限

本研究作为概念分析性研究，其主要局限在于未能对“制度性依赖”与“石化程度”之间的因果关系提供经验数据的直接支持。虽然我们在第5.4节构建了完整的证伪检验设计，但这一设计本身尚未被执行为一项实证研究。此外，本文虽然涉及“诱致性认知修剪”的论断，但其心理机制和神经基础的详细论证仍停留在隐喻和理论推演的层次，缺乏直接的神经影像学或纵向追踪证据。最后，本文主要聚焦于高度制度化的二语/外语学习场景，其对非正式学习环境中石化现象的解释力有待进一步论证。

7.5 未来研究方向

基于本文的理论框架，至少可生长出五条相互关联的研究纲领。

第一，证伪执行。执行本文5.4节设计的实证检验方案，在控制认知负荷变量的前提下，系统检验制度性依赖与真实交际能力石化程度的剂量-反应关系。这是所有后续研究的基础先决任务。

第二，教育体制比较研究。选取在“评分的多维性”上存在显著差异的两个或多个可比较的教育体制，通过对学习者的长期追踪，检验“制度评价维度越单一，石化越易发生且越顽固”这一衍生假设。

第三，修复干预研究。设计并执行旨在“逆向解除诱致性修剪”的干预方案。例如，对已产生石化的学习者给予密集的、免于评分的、高度正反馈的自由交际训练。关键因变量应是“石化的可逆程度”。这一研究纲领可以测试本文“石化是可逆的深度休眠而非永久损伤”的核心预期。

第四，无意识防御测量。如果学习者的认知系统确实将制度性安全内化，那么当被置于可能挑战其制度性成功状态的情境（如在没有准备的情况下接受真实交际能力评估）时，他们应表现出更强的防卫性反应。可以设计实验，调用类似社会心理学中使用的内隐联想测验、生理指标（皮肤电传导、心率变异）等工具，测量制度性高依赖者在面对“你可能没学好”这类反馈时的无意识防御水平。

第五，微观生成机制研究。进行小样本的、高密度的纵向追踪研究（例如对一个班级进行每两月一次的持续三年以上的真实交际能力记录），用以捕捉学习者从“策略性停滞”向“结构性删除”过渡的关键时期。这使得研究能够描述“修剪”发生的时间窗口和临界点，这将是检验诱致性修剪假说的最严格方式。

参考文献

- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115–129.
- Arthur, W. B. (1989). Competing technologies, increasing returns, and lock-in by historical events. *The Economic Journal*, 99(394), 116–131.
- Biggs, J. B. (1987). *Student Approaches to Learning and Studying*. Australian Council for Educational Research.
- DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147–160.
- Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and motivation in second-language learning*. Newbury House.
- Geertz, C. (1963). *Agricultural involution: The process of ecological change in Indonesia*. University of California Press.
- Han, Z. (2004). Fossilization in adult second language acquisition. *Multilingual Matters*.
- Han, Z., & Odlin, T. (Eds.). (2006). *Studies of fossilization in second language acquisition*. Multilingual Matters.
- Higgs, T. V., & Clifford, R. (1982). The push toward communication. In T. V. Higgs (Ed.), *Curriculum, competence, and the foreign language teacher* (pp. 57–79). National Textbook Company.
- Krippendorff, K. (2004). *Content Analysis: An Introduction to Its Methodology* (2nd ed.). Sage.
- Langley, P. (2019). The institutional making of proficiency scales: A global-historical perspective. *Language Testing*, 36(4), 505–527.
- Lardiere, D. (2007). *Ultimate attainment in second language acquisition: A case study*. Lawrence Erlbaum Associates.
- Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I—Outcome and process. *British Journal of Educational Psychology*, 46(1), 4–11.
- Pierson, P. (2000). Increasing returns, path dependence, and the study of politics. *American Political Science Review*, 94(2), 251–267.

- Powell, W. W., & DiMaggio, P. J. (Eds.). (1991). *The new institutionalism in organizational analysis*. University of Chicago Press.
- Schumann, J. H. (1978). The acculturation model for second-language acquisition. In R. C. Gingras (Ed.), *Second-language acquisition and foreign language teaching* (pp. 27–50). Center for Applied Linguistics.
- Schwartz, B. (2016). *Why we work*. Simon & Schuster.
- Schwartz, B., & Sharpe, K. (2010). *Practical wisdom: The right way to do the right thing*. Riverhead Books.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, 10(3), 209–231.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
- Snyder, B. R. (1971). *The Hidden Curriculum*. Alfred A. Knopf.
- White, L. (2003). *Second language acquisition and Universal Grammar*. Cambridge University Press.
- 束定芳 (2004). 《外语教学改革: 问题与对策》. 上海外语教育出版社.
- 黄宗智 (1985). 《华北的小农经济与社会变迁》. 中华书局.