

单通道和多通道下的统计学习跨通道迁移*

唐 溢¹ 赵亚军² 曾清樟³ 张智君³ 吴圣楠¹

(¹重庆行政学院, 重庆 400041) (²西南民族大学教育学与心理学院, 成都 610041)

(³浙江大学心理与行为科学系, 杭州 310022)

摘 要 统计学习是人类从环境中提取规律信息的关键认知能力, 其跨通道迁移的特性与机制仍存理论争议。本研究通过 4 个实验系统考察统计学习在视听通道间的迁移机制。实验 1 采用经典统计学习范式, 验证了个体基于动物图片的视觉统计学习能力; 实验 2 在单一视觉通道学习后, 分别以动物图片和动物声音进行测试, 发现两种条件下统计学习效果无显著差异, 表明存在从视觉到听觉的跨通道迁移; 实验 3 通过在视听双通道同步呈现动物图片与无意义音节刺激流, 发现无论听觉刺激流是否具有统计规律, 视觉向听觉的统计学习迁移效应均稳定存在; 实验 4 进一步在视听双通道同步呈现动物声音与无意义图形刺激流, 结果显示从听觉向视觉的跨通道迁移亦显著。本研究综合表明, 统计学习具备稳定的跨通道双向迁移能力, 支持统计学习具有通道一般性的理论观点, 存在多层平行的统计结构表征系统可能是其认知基础。

关键词 统计学习, 动物图片, 动物声音, 多通道, 跨通道迁移

分类号 B842

1 前言

人类所处的环境中蕴含着丰富的统计规律, 把握客体间的规律性关系是人类认知功能的核心要素之一。统计学习被定义为个体无意识、自动化地提取环境规律的关键能力(Fiser & Aslin, 2001, 2002)。已有研究证实, 统计学习能力独立存在于视觉(Fiser & Aslin, 2005)、听觉(Batterink et al., 2015)和触觉(Conway & Christiansen, 2005)等多个感觉通道, 且各通道的统计学习并非完全独立, 可能存在相互作用。例如, 对典型发展群体与感觉障碍群体的比较研究表明, 多通道信息可通过补偿机制弥补单一通道的统计学习缺陷(徐贵平等, 2020; Kligler & Gabay, 2023)。然而, 对于某一通道中习得的统计规律能否直接在另一通道中表达, 即是否存在跨通道迁移, 目前学界仍存在理论分歧。Frost 等人(2015)从客体特征加工的视角提出抽象规则表征假说, 指出个体在学习客体间规律时可能涉及 4

种特性: 通道特异性、刺激特异性、通道一般性, 以及受通道与刺激双重调节的统计学习。换言之, 统计学习的过程与结果不仅取决于刺激特征, 还受到感觉通道属性与刺激本身特征的双重影响。对于需要依赖高层级信息加工才能提取的刺激间统计关系, 其学习结果的迁移并不依赖于早期感觉区域(低层级)的加工机制, 这表明统计学习在某些条件下可能表现出通道一般性的特点。总体而言, 当前关于“统计学习在何种程度上具有通道特异性或通道一般性”的探讨仍受研究方法的限制。已有关于统计学习迁移的研究主要集中在两个方向: 其一, 关注同一通道内低层级特征之间的迁移(如从形状到颜色, Turk-Browne et al., 2008; Uno et al., 2020), 尚未延伸至跨通道层面; 其二, 关注具有语义信息的客体间迁移(Brady & Oliva, 2008; Emberson & Rubinstein, 2016; Jun & Chong, 2018; Otsuka et al., 2013), 但同样尚未涉及跨通道加工机制。基于上述研究现状, 本研究选用动物图片与动物声音作为实

收稿日期: 2025-03-17

* 国家社会科学基金项目(21XDJ002)。

通信作者: 唐溢, E-mail: tangy436@126.com

验材料,旨在考察统计学习在视听通道间的迁移能力。需要说明的是,尽管两类材料本身具有语义关联,但实验过程中未向被试提供任何语义关联提示,以纯粹探究统计学习在跨通道情境下的发生条件与内在机制。

统计学习是指个体通过追踪刺激流中元素之间的概率关系以提取规律的能力。Saffran 等人(1996)在研究婴儿语言习得机制时首次提出这一概念,发现婴儿能够基于音节进行统计学习。随后的系列研究验证了个体基于不同刺激进行统计学习的能力,例如客体形状与颜色(Turk-Browne et al., 2008)、真实场景(Brady & Oliva, 2008)、真实客体(Emberson & Rubinstein, 2016; Jun & Chong, 2018; Otsuka et al., 2013, 2014)以及名人面孔(唐溢等, 2015)等。经典的统计学习实验包括熟悉阶段与测试阶段。在熟悉阶段,向被试呈现由4个三联体(triplet)构成的刺激流(如ABC、DEF、GHI、JKL),三联体随机呈现,但每个三联体内部的刺激顺序固定不变,被试并不知晓后续将进行测试。在测试阶段,参与者需完成二选一迫选任务,每个试次中呈现两个三联体,一个为熟悉三联体(如ABC),另一个为伪三联体(如AEI),参与者需根据熟悉阶段的体验判断哪一个更熟悉。若被试能显著高于随机水平地区分熟悉三联体与伪三联体,则表明发生了统计学习。

统计学习迁移作为个体将已习得的规律性知识从原始情境扩展至新情境的核心过程,是评估统计学习系统适应性的关键指标。Turk-Browne 和 Scholl (2009)发现,被试在时间序列中习得的顺序关联规律可迁移至空间排列任务,反之亦然。研究者提出,统计学习可能通过调节感觉皮层(如侧枕叶与腹侧枕颞联合皮层)的神经表征结构,增强规律性刺激的知觉突显性,从而实现跨情境、跨任务的泛化性表征(Turk-Browne et al., 2005)。近期研究进一步揭示了统计学习跨特征迁移的现象。例如,Brady 和 Oliva (2008)通过呈现具有类别一致性但视觉特征异质的客体刺激流(如不同桥梁、厨房场景),发现视觉统计学习可在具有相同语义类别、不同视觉特征的客体间保持迁移效应,表明统计结构的表征可能独立于具体视觉特征。值得注意的是,此类迁移存在显著的层级约束:视觉统计学习仅能在具有基本水平(basic-level category)和下位水平(subordinate-level category)概念的不同客体间发生迁移(Emberson & Rubinstein, 2016; Jun & Chong,

2018),而无法在具有相同上位水平(superordinate-level category)概念的不同客体间实现迁移(Otsuka et al., 2014)。

为探究统计学习的跨通道机制,研究者逐步发展出多通道同步呈现范式。早期研究采用视听交替呈现的交叉设计(Conway & Christiansen, 2006),但由于刺激流在时间维度上存在分离,难以揭示真实情境中同步多通道输入的加工特性。石文典等人(2013)采用严格的时间同步呈现技术,以几何图形(视觉通道)与人工语音(听觉通道)为刺激,向被试同步呈现人工语法规则。结果发现,无论视听两通道的语法规则是否一致,均未对单一通道的统计学习效应产生显著影响。基于此,研究者提出统计学习具有通道特异性的观点。随后,以动物图片与听觉语义信息为材料的研究也验证了上述结果(Li et al., 2018),认为不同通道的统计学习是相互独立且平行的机制,不存在跨通道迁移效应。然而,也有研究发现统计学习可能存在通道交互作用。Thiessen (2010)发现多通道学习对单通道学习具有促进效应,且这种效应受个体发展水平(婴儿无促进效应,成人存在促进效应)与刺激源的影响。此外, Mitchel 和 Weiss (2011)通过破坏视听双通道规律的对应性,观察到单通道统计学习效果显著受损,提示多通道加工可能存在规律干扰效应。综上所述,尽管现有研究采用的多通道呈现范式能够有效探究多通道统计学习对单通道学习的潜在影响,但尚难以直接捕捉跨通道迁移的证据。

本研究共包含4个实验,结合跨通道迁移范式与多通道统计学习范式,系统考察真实客体场景下统计学习的跨通道迁移效应。

实验1基于经典视觉统计学习范式(Otsuka et al., 2013),以动物图片构建视觉刺激流,旨在验证基于动物图片的视觉统计学习能力。

实验2引入跨通道迁移范式,熟悉阶段保持视觉单通道呈现,测试阶段分别采用动物图片和动物声音进行测试,通过比较视觉-视觉与视觉-听觉条件的统计学习效果,探查视觉统计学习的跨通道迁移效应。

实验3进一步采用多通道学习范式,旨在分离通道特异性学习与跨通道迁移效应。在熟悉阶段,同步呈现视觉(动物图片)与听觉(无意义音节)刺激流,并设置对比条件:实验3a中听觉刺激流无统计规律,实验3b中听觉刺激流包含与视觉刺激同步的统计规律。测试阶段采用动物图片、动物声音及

无意义音节(仅实验 3b)进行检测。该设计与实验 2 形成递进,实验 2 仅考察视觉单通道的统计学习,而实验 3 在视觉刺激基础上增设听觉刺激流,主要有两个目的:第一,检验不同感知通道的统计学习是否具有通道独立性。基于 Saffran 等人(1997)关于统计学习不受另一通道无关任务干扰的发现,实验 3a 在呈现视觉刺激流的同时加入随机听觉刺激流,一方面考察多通道同步呈现条件下无关听觉刺激是否干扰视觉统计学习,另一方面作为基线条件,与实验 3b(呈现具有统计规律的听觉刺激流)进行对比,以检验视听双通道统计学习是否存在相互影响(Li et al., 2018; Thiessen, 2010)。第二,考察在多通道统计学习条件下,从视觉向听觉的统计学习跨通道迁移与听觉通道自身统计学习之间的关系。实验 2 已发现统计学习可从视觉向听觉迁移,实验 3 进一步探讨在发生跨通道迁移的同时,若听觉通道也进行统计学习,是否会影响迁移效果,从而揭示单通道统计学习与跨通道迁移的发生机制。具体推论如下:(1)若实验 3b 中基于动物声音的统计学习结果(作为迁移指标)相较于实验 3a 出现显著变化,表明跨通道迁移过程受到听觉通道统计学习的调节,两者可能为交互过程;(2)若实验 3b 与实验 3a 中动物声音的统计学习结果无显著差异,则说明单通道统计学习与跨通道迁移可能为两个平行的过程。

实验 4 旨在检验多通道呈现条件下,统计学习从听觉向视觉跨通道迁移的能力,并与实验 3 共同验证迁移的双向性。实验采用多通道同步呈现范式,与实验 3b 的关键区别在于:熟悉阶段的听觉刺激为动物声音,视觉刺激为无意义图形。测试阶段,若动物图片的统计学习结果(作为从听觉向视觉迁移的指标)显著,则表明存在从听觉向视觉的跨通道迁移;若不显著,则说明该方向的迁移能力有限。此外,通过对无意义图形的测试,可进一步考察在跨通道迁移过程中,若视觉通道同时进行其他统计学习(如对无意义图形的规律提取),是否会对迁移效果产生影响。该设计与实验 3b 共同为揭示单通道统计学习与跨通道迁移之间的机制关系提供对比证据。

总体而言,本实验设计参考了 Seitz 等人(2007)提出的多通道模型,通过系统操纵统计规律输入的通道组合方式,拟从以下三个方面推动统计学习迁移研究的发展:第一,将统计学习迁移研究从单一通道扩展至视觉-听觉跨通道范畴,为理解统计学习的迁移机制提供新的实证依据;第二,采用真实

客体的动物图片与动物声音作为实验材料,相较于传统人工符号更贴近真实环境中的多通道信息交互,有助于揭示统计学习在自然认知过程中的适应机制;第三,通过多通道统计学习范式,实现了对单通道学习与跨通道迁移效应的有效分离,为探索人类处理多通道复杂规律信息的认知过程提供了新的方法路径。

2 实验 1: 基于动物图片的视觉统计学习

2.1 方法

2.1.1 被试

采用 G*Power 3.1 软件进行样本量估计(Faul et al., 2007)。设定 $\alpha = 0.05$, 统计检验力 $(1 - \beta) = 0.8$, 以大效应量估计($d_z = 0.8$)为参考,计算得出所需最低样本量为 14 人。最终从浙江大学招募在校大学生 20 名(女性 10 名,男性 10 名),年龄 18~23 岁($M = 20.45$ 岁, $SD = 1.69$ 岁)。所有被试均为右利手,视力或矫正视力正常(自述)。所有被试未参加过该研究的其他实验。实验在浙江大学心理与行为科学系伦理委员会的指导下进行。实验前均签署知情同意书,实验结束后被试获得相应报酬或课程学分。

2.1.2 刺激与设备

本研究采用 Hocking 等人(2013)建立的标准化动物声音材料库,共包含 110 种生态效度较高的动物声音样本。

听觉刺激材料:招募 41 名健康大学生(筛选标准与正式实验组相同)对全部动物声音样本进行规范化的命名任务及自信心评估。使用 E-Prime 3.0 编写评价程序,要求被试依次听取随机呈现的动物声音,并通过按键进行实时命名反应,随后在 5 点 Likert 量表上对判断信心进行评分(1 = 完全猜测,5 = 确定无疑)。基于命名准确率($M = 95.24\%$, $SD = 8.37\%$)与自信心得分($M = 4.58$, $SD = 0.31$)的双重验证,最终选取 12 种真实动物声音(如青蛙鸣叫、马嘶等)作为正式听觉刺激。所有声音文件均通过 Adobe Audition CS6 进行统一标准化处理,保存为 WAV 格式(采样率 44.1 kHz,时长 1000 ± 5 ms)。

视觉刺激材料:选取 12 张动物图片,与上述 12 种动物声音一一对应。其中 11 张图片选自张清芳与杨玉芳(2003)修订的中文版 Snodgrass-Vanderwart 线条图库;因图库中缺少乌鸦图片,特邀两名美术专业研究生参照 Snodgrass 画法补绘完成。所有图片经 Photoshop CC 统一处理为黑色线

条图(分辨率 300 dpi), 并标准化为 $5.2^\circ \times 5.2^\circ$ 视角(对应 470×470 像素), 详见图 1。为控制被试对材料的熟悉度, 在招募阶段要求潜在被试完成一项网络问卷匹配任务: 先播放一个动物声音, 从 12 张图片中选择现实中与之匹配的动物图片, 能完全正确匹配所有动物声音与图片(正确率 100%)者, 方可参与正式实验。

使用 Matlab R2020a 编写刺激呈现程序, 并调用 Psychophysics Toolbox 3.0.18 进行同步控制。视觉刺激通过 ViewSonic G76f+ CRT 显示器呈现(分辨率 1024×768 , 刷新率 85 Hz), 使用 ColorCal II 进行 Gamma 校正(亮度 50 cd/m^2 , 灰度值 128)。听觉刺激通过耳机输出, 构建双耳声场(频率响应范围 18~28000 Hz)。被试眼睛与屏幕距离固定为 57cm。

2.1.3 实验设计与程序

本研究采用 Otsuka 等人(2013)提出的经典视觉统计学习范式, 包括熟悉阶段与测试阶段, 实验流

程见图 2。

熟悉阶段: 以动物图片作为刺激材料, 每张图片呈现在屏幕中央。刺激持续时间为 300 ms, 刺激间隔(inter-stimulus interval, ISI)为 700 ms, 刺激起始异步(stimulus onset asynchrony, SOA)为 1000 ms。12 张图片构成 4 个固定顺序的三联体(如 ABC), 各三联体在保持内部顺序不变的条件下, 经伪随机排列组合形成连续刺激流。刺激流遵循以下控制条件: (a)同一三联体不连续重复(如 ABCABC), 且任意两个三联体也不连续重复(如 ABCGHIABCGHI); (b)三联体首尾位置的图片共重复 96 次(如 ABCCGHI 或 ABCGGHI), 重复次数在各三联体间平衡。具体流程见图 2A。为引导被试注意刺激流并避免其主动探求规律, 主试仅告知被试执行重复图片监测任务(Otsuka et al., 2013): 若当前图片与前一图片完全相同, 需尽快按空格键反应。本阶段持续约 20 分钟, 所有被试均未获知后续测试相关信息。

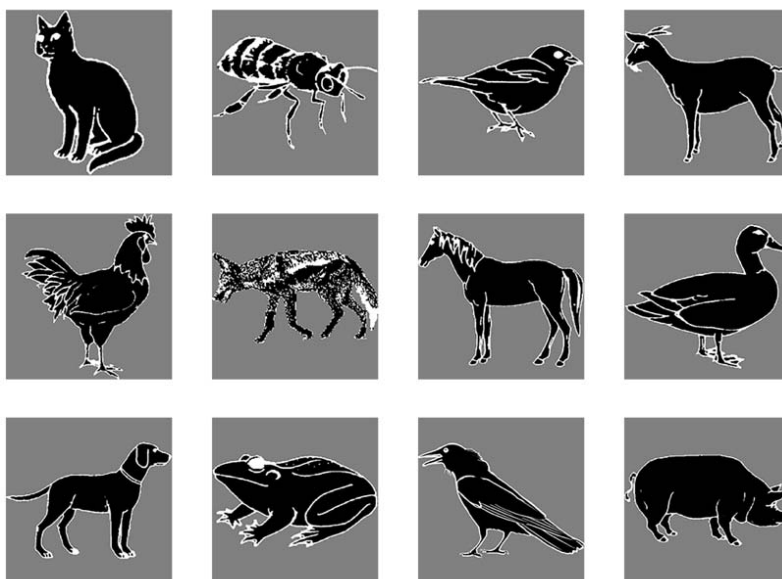


图 1 实验 1 的动物图片材料

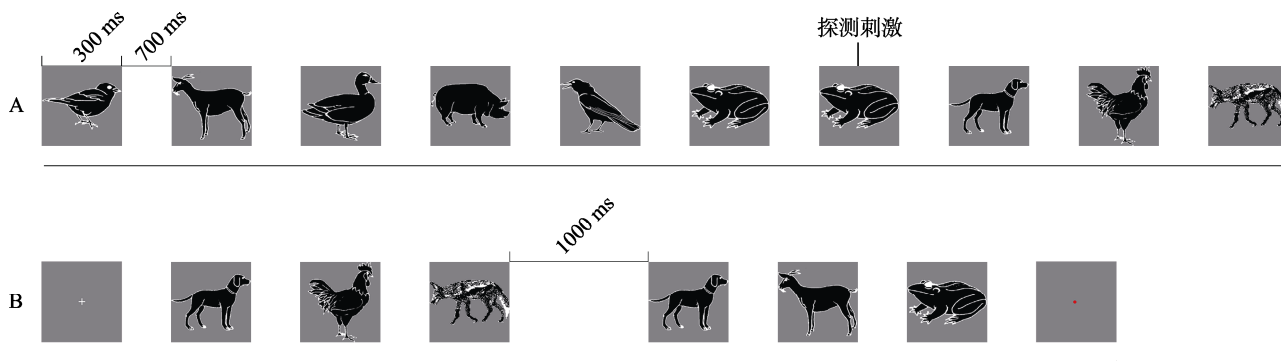


图 2 实验 1 熟悉阶段(A)和测试阶段(B)的刺激示例

测试阶段：熟悉阶段结束后立即进行二项迫选任务(two-alternative forced-choice, 2AFC)。每个试次开始前呈现 1000 ms 的注视点“+”，随后依次呈现两个三联体：一为熟悉三联体(即熟悉阶段中出现的统计结构，如 ABC)，另一为伪三联体(由熟悉阶段中的刺激构成但未曾出现，如 AEI)。两个三联体呈现完毕后，屏幕中央出现红色提示点，要求被试使用方向键判断哪一个三联体更具熟悉感，“←”为先呈现者，“→”为后呈现者，按键后提示点消失。共 32 个试次，按完全随机顺序呈现，每个熟悉三联体分别与 4 个伪三联体各配对两次。本阶段持续约 5 分钟。

实验控制：为排除外显策略的影响，所有被试在实验结束后接受开放式访谈，报告其对刺激流中统计规律的觉察程度及可能采用的认知策略(Brady & Oliva, 2008)。数据收集采用双盲设计，主试与被试均不了解熟悉三联体与伪三联体的具体组合方式。

2.2 结果与讨论

在熟悉阶段，所有被试顺利完成重复探测任务，平均正确率为 92.40% ($SD = 6.76\%$)，对非重复刺激的虚报率仅为 0.20%。该结果表明，被试在实验过程中有效维持了对主要任务的注意力，未将认知资源主动用于探测潜在的统计规律。事后主观报告显示，所有被试均未能外显报告刺激流中隐藏的统计规律，也未主动采用任何外显记忆策略。

测试阶段结果显示(图 4)，被试对熟悉三联体与伪三联体的正确辨别率平均值为 74.22% ($SE = 4.00\%$)。单样本 t 检验表明，该成绩显著高于随机

猜测水平(50%)， $t(19) = 6.02$ ， $p < 0.001$ ，Cohen's $d = 1.35$ 。这一发现成功复现了前人关于真实客体统计学习的研究结果(Otsuka et al., 2013)，进一步验证了个体能够基于动物图片进行视觉统计学习。

3 实验 2：视觉统计学习的跨通道迁移

3.1 方法

3.1.1 被试

采用 G*Power 3.1 软件进行样本量估计。参数设定与实验 1 一致($\alpha = 0.05$, $1 - \beta = 0.8$, $d_z = 0.8$)，计算得出最低样本量为 22 人。最终从浙江大学招募在校大学生 23 名(女性 14 名，男性 9 名)，年龄范围 18~23 岁($M = 21.57$ 岁， $SD = 1.47$ 岁)。均为右利手，视力或矫正视力正常(自述)。所有被试未参加过该研究的其他实验。实验的知情同意及报酬方式同实验 1。

3.1.2 刺激与设备

动物图片、动物声音与实验 1 相同。

实验设备与实验 1 相同。

3.1.3 设计与程序

熟悉阶段的流程与实验 1 完全相同(图 3A)。

测试阶段分别采用动物图片(图 3B)和动物声音(图 3C)进行。动物图片测试的设置与实验 1 一致。在动物声音测试中，告知被试所呈现的声音与第一阶段出现的动物图片具有一一对应关系，并要求他们依据自身直觉，参照第一阶段所体验的图片顺序，对声音刺激的顺序熟悉性作出判断。按键反应方式与实验 1 相同。两种测试材料的呈现顺序在被试间进行平衡。

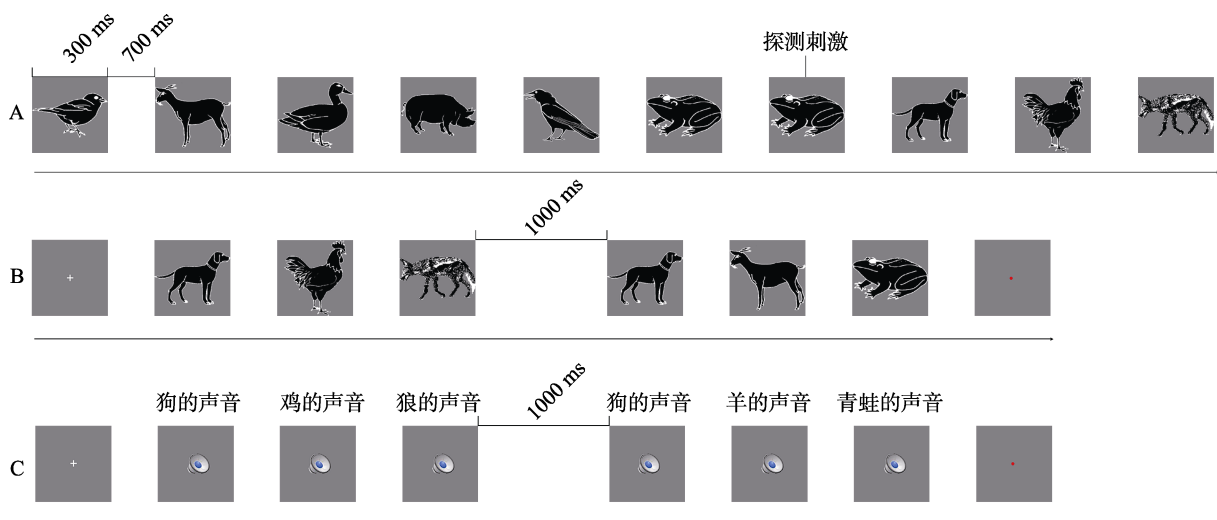


图 3 实验 2 熟悉阶段(A)和测试阶段动物图片(B)与动物声音(C)测试的刺激示例

3.2 结果与讨论

被试在熟悉阶段的重复探测任务中正确率为 92.24% ($SD = 5.38\%$), 对非重复刺激的误报率极低 (0.26%), 表明其注意力集中于当前任务, 未主动探测潜在规律。事后主观报告显示, 所有被试均未能外显报告刺激流中隐藏的统计规律, 也未主动采用任何外显记忆策略。

测试阶段结果如图 4 所示。被试对动物图片三联体与伪三联体的正确辨别率 ($M = 74.87\%$, $SE = 4.10\%$) 显著高于随机水平 (50%), $t(22) = 6.09$, $p < 0.001$, $Cohen's d = 1.27$; 对动物声音三联体与伪三联体的正确辨别率 ($M = 71.33\%$, $SE = 4.60\%$) 也显著高于随机水平, $t(22) = 4.64$, $p < 0.001$, $Cohen's d = 0.97$ 。配对样本 t 检验表明, 视觉与听觉测试的统计学习效应量无显著差异, $t(21) = 1.16$, $p = 0.260$, $Cohen's d = 0.24$ 。为进一步评估零假设的支持强度, 使用 JASP 进行贝叶斯因子分析, 进一步支持动物图片与动物声音条件间无显著差异 ($BF_{01} = 2.53$), 微弱强度的零假设支持证据。

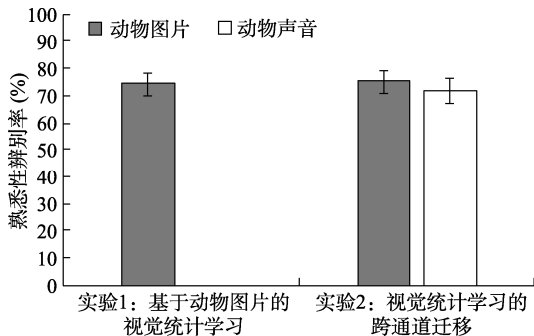


图 4 实验 1 和实验 2 测试阶段三联体熟悉性判断情况 (误差线: 为标准误差值)

上述结果表明, 被试在单一视觉通道习得的统计规律, 能够迁移至听觉通道, 并表现为跨通道间无显著差异的学习效应。这一结果为统计学习的通道间可迁移性提供了行为证据。

4 实验 3: 多通道条件下统计学习从视觉到听觉的跨通道迁移

4.1 方法

4.1.1 被试

采用 G*Power 3.1 软件进行样本量估计。参数设定与实验 1 一致 ($\alpha = 0.05$, $1 - \beta = 0.8$, $d_z = 0.8$), 得出实验 3a 和 3b 各组至少需 22 名被试。最终从

浙江大学招募 18~27 岁在校大学生, 实验 3a 共 23 名 (女性 12 名, 男性 11 名) ($M = 22.39$ 岁, $SD = 2.56$ 岁), 实验 3b 共 23 名 (女性 16 名, 男性 7 名) ($M = 22.13$ 岁, $SD = 2.61$ 岁)。均为右利手, 视力或矫正视力正常 (自述)。所有被试未参加过该研究的其他实验。实验的知情同意及报酬方式同实验 1。

4.1.2 实验材料与设备

视觉刺激材料沿用实验 2 的动物图片。

听觉刺激分为两类: (1) 动物声音材料: 沿用实验 2 的动物声音。(2) 无意义音节材料: 参照 Saffran 等人 (1996) 的实验, 选取 12 个无意义音节 (/pa/, /bu/, /bi/, /da/, /do/, /go/, /ku/, /la/, /pi/, /ro/, /ti/, /tu/), 通过 Neospeech_Liang TTS 语音合成软件 (v3.4.1) 生成。所有声音材料由专业男声播音员标准化录制, 经 Adobe Audition CS6 统一处理为 WAV 格式文件 (采样率 44.1 kHz, 时长 1000 ± 5 ms)。

刺激呈现时程与实验 2 保持一致 ($SOA = 1000$ ms)。实验设备与实验 2 相同。

4.1.3 设计与程序

实验采用多通道统计学习范式。

熟悉阶段: 在视听双通道同步方式呈现刺激流。视觉刺激为动物图片 (呈现 300 ms, $ISI = 700$ ms), 听觉刺激为无意义音节, 其呈现时间与视觉刺激完全匹配。本阶段的关键操作为统计规律的通道配置: 在实验 3a 中, 仅视觉刺激流包含统计规律, 听觉刺激流为随机呈现; 在实验 3b 中, 视觉与听觉刺激流包含相同的统计规律, 如图 5A 所示。

测试阶段: 实验 3a 使用动物图片 (图 5B) 和动物声音 (图 5C) 作为测试材料; 实验 3b 在此基础上增加无意义音节 (图 5D) 作为第三种测试材料。测试程序与实验 2 完全一致。

4.2 结果与讨论

在熟悉阶段, 被试顺利完成重复探测任务, 实验 3a 的正确率为 92.57% ($SD = 6.27\%$), 误报率为 0.23%; 实验 3b 的正确率为 94.20% ($SD = 6.92\%$), 误报率为 0.10%。结果表明, 两组被试均能有效维持任务注意力。事后主观报告显示, 所有被试均未能外显报告刺激流中隐藏的统计规律, 也未主动采用任何外显记忆策略。

测试阶段结果如图 6 所示。在实验 3a (多通道非同步统计规律) 中, 被试对动物图片三联体的正确辨别率 69.43% ($SE = 3.90\%$), 显著高于随机水平,

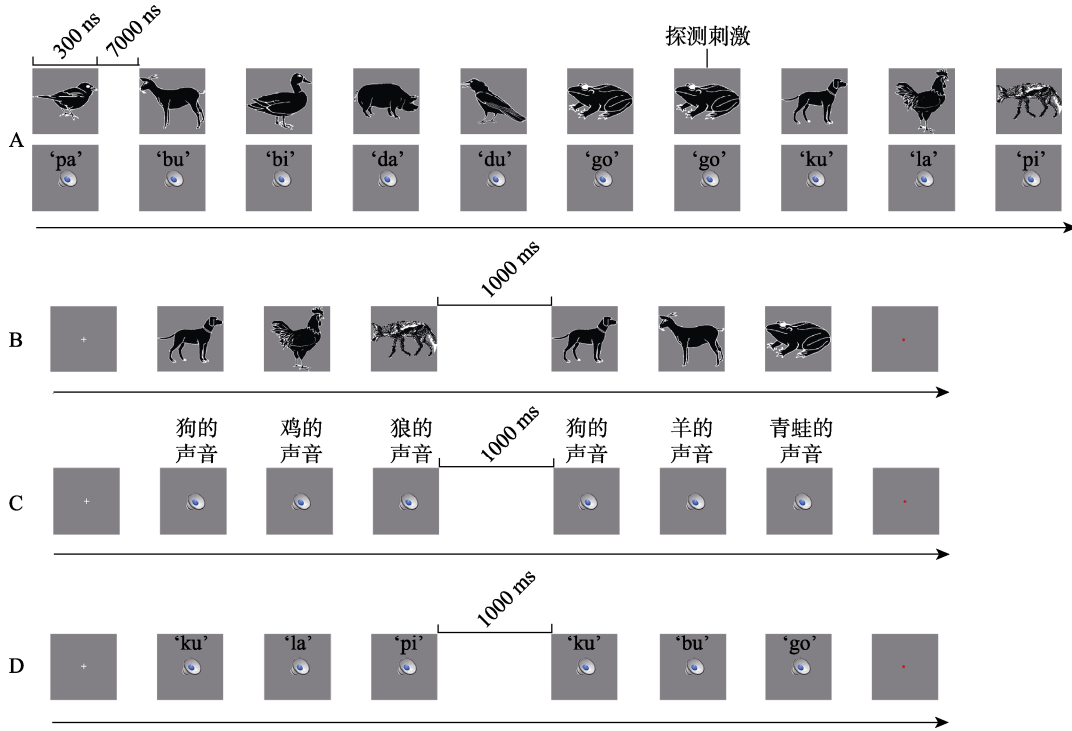


图 5 实验 3b 熟悉阶段(A)和测试阶段动物图片(B)、动物声音(C)、无意义音节(D)测试的刺激示例

$t(22) = 4.95, p < 0.001, \text{Cohen}' d = 1.03$; 对动物声音三联体的正确辨别率 65.49% ($SE = 4.70\%$), 亦显著优于随机水平, $t(22) = 3.32, p = 0.003, \text{Cohen}' d = 0.69$ 。配对样本 t 检验表明, 动物图片与动物声音条件间差异不显著, $t(22) = 1.57, p = 0.131, \text{Cohen}' d = 0.33$ 。使用 JASP 进行贝叶斯因子分析, 进一步支持动物图片与动物声音条件间无显著差异($\text{BF}_{01} = 1.57$), 微弱强度的零假设支持证据。以上结果表明, 在多通道非同步输入条件下, 视觉统计学习能够稳定地向听觉通道迁移。

正确辨别率分别为 69.29% ($SE = 4.30\%$)、66.85% ($SE = 4.20\%$)和 62.50% ($SE = 3.50\%$), 均显著高于随机水平(动物图片: $t(22) = 4.48, p < 0.001, \text{Cohen}' d = 0.93$; 动物声音: $t(22) = 3.97, p < 0.001, \text{Cohen}' d = 0.83$; 无意义音节: $t(22) = 3.54, p = 0.002, \text{Cohen}' d = 0.74$)。配对样本 t 检验显示, 三类材料间的差异均不显著(动物图片-动物声音: $t(22) = 0.92, p = 0.370, \text{Cohen}' d = 0.19$; 动物声音-无意义音节, $t(22) = 1.19, p = 0.249, \text{Cohen}' d = 0.25$; 动物图片-无意义音节, $t(22) = 1.59, p = 0.127, \text{Cohen}' d = 0.33$)。使用 JASP 进行贝叶斯因子分析, 进一步支持动物图片与动物声音条件间无显著差异($\text{BF}_{01} = 3.14$, 中等强度), 动物声音与无意义音节条件间亦无显著差异($\text{BF}_{01} = 2.45$, 微弱强度), 为零假设提供了支持证据。这些结果表明, 在多通道同步输入条件下, 个体不仅保留了各通道的统计学习能力, 也维持了从视觉向听觉的跨通道迁移。

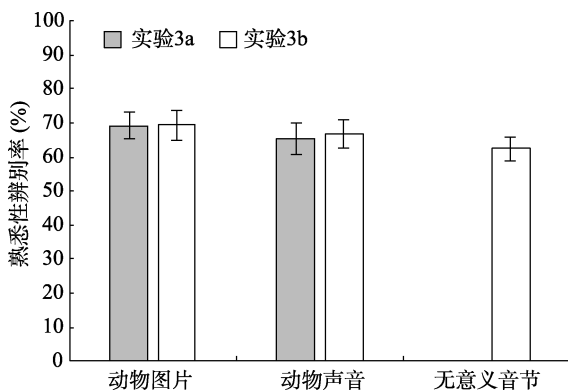


图 6 实验 3 测试阶段三联体熟悉性判断情况(误差线: 为标准误差值)

在实验 3b (多通道同步统计规律)中, 被试对动物图片、动物声音及无意义音节三联体熟悉性的

对实验 3 数据进行 2 (刺激特征: 动物图片/动物声音) \times 2 (统计规律呈现模式: 多通道同步/多通道非同步)双因素方差分析。结果显示, 交互效应不显著, $F(1, 44) = 0.17, p = 0.685$; 刺激特征主效应未达显著水平, $F(1, 44) = 3.04, p = 0.088$; 统计规律呈现模式的主效应亦不显著, $F(1, 44) = 0.01, p = 0.916$ 。这些结果表明, 多通道统计规律的同步性对跨通道迁移效应无显著调节作用。

进一步整合实验2和实验3数据,进行2(刺激特征:动物图片/动物声音) \times 3(统计规律呈现模式:单通道/多通道同步/多通道非同步)的双因素方差分析。结果显示,交互效应不显著, $F(2, 66) = 0.08$, $p = 0.924$;刺激特征主效应显著, $F(1, 66) = 4.32$, $p = 0.042$, $\eta_p^2 = 0.007$;而统计规律呈现模式的主效应未达显著水平, $F(2, 66) = 0.57$, $p = 0.567$ 。表明多通道刺激的呈现方式对跨通道迁移效果无显著影响。

综合以上结果可知,在多通道统计学习条件下,学习者不仅能够同时保持对动物图片视觉刺激流和无意义音节听觉刺激流的统计学习能力,还成功实现了从视觉到听觉的跨通道迁移。这一发现印证了统计学习跨通道迁移效应在复杂多通道情境中具有稳定性,为统计学习的通道一般性特征提供了更贴近自然认知场景的实证证据。

5 实验4:多通道条件下统计学习从听觉到视觉的跨通道迁移

5.1 方法

5.1.1 被试

采用G*Power 3.1软件进行样本量估计。参数设定与实验1一致($\alpha = 0.05$, $1 - \beta = 0.8$, $d_z = 0.8$),计算得出最低样本量为22人。最终从浙江大学招募在校大学生25名(女性14名,男性11名),年龄范围18~28岁($M = 22.88$ 岁, $SD = 2.48$ 岁)。均为右利手,视力或矫正视力正常(自述)。所有被试未参加过该研究的其他实验。实验的知情同意及报酬方式同实验1。

5.1.2 实验材料与设备

听觉刺激材料:沿用实验3的动物声音。

视觉刺激材料:(1)动物图片材料:沿用实验3的动物图片。(2)无意义图形材料(图7):12个无

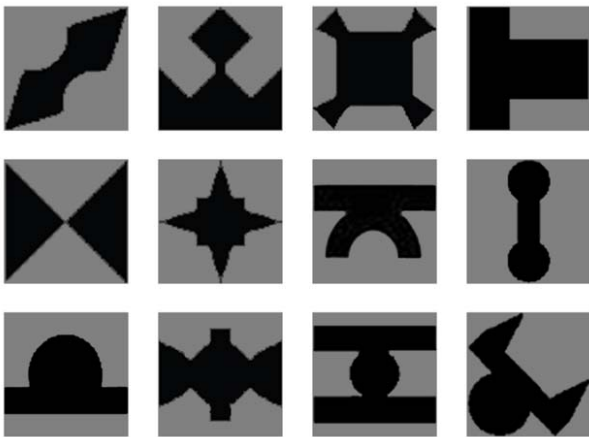


图7 实验4的无意义图形材料

意义图形选自Zhao等人(2013)、Turk-Browne等人(2005)考察视觉统计学习的实验材料,颜色为黑色,大小与动物图片相同。

刺激呈现时程与实验3b保持一致。实验设备与实验3相同。

5.1.3 设计与程序

熟悉阶段采用多通道统计学习范式,在视听双通道同步呈现刺激流。视觉刺激为无意义图形,呈现时间为300ms,刺激间隔为700ms;听觉刺激为动物声音,呈现时间为1000ms。两个通道的刺激流包含同步的统计规律,其生成方式与实验3b相同。

测试阶段分别使用动物声音、动物图片和无意义图形作为测试材料,测试程序与实验3b保持一致。

5.2 结果与讨论

在熟悉阶段,被试顺利完成重复探测任务,正确率为98.17% ($SD = 2.27\%$),误报率仅为0.11%。表明被试的注意力能够有效集中于任务要求。事后主观报告显示,所有被试均未能外显报告刺激流中隐藏的统计规律,也未主动采用任何外显记忆策略。

测试阶段结果显示,被试对动物声音、动物图片及无意义图形三联体熟悉性的正确辨别率分别为68.25% ($SE = 3.27\%$)、65.75% ($SE = 4.31\%$)和68.13% ($SE = 4.55\%$),均显著高于随机水平(动物图片: $t(24) = 5.57$, $p < 0.001$, $Cohen' d = 1.11$;动物声音: $t(24) = 3.65$, $p = 0.001$, $Cohen' d = 0.73$;无意义音节: $t(24) = 3.98$, $p < 0.001$, $Cohen' d = 0.79$)。

配对样本 t 检验表明,三类测试材料间的差异均不显著(动物声音-动物图片: $t(24) = 0.77$, $p = 0.445$, $Cohen' d = 0.15$;动物声音-无意义音节: $t(24) = 0.03$, $p = 0.974$, $Cohen' d = 0.01$;动物图片-无意义音节: $t(24) = -0.69$, $p = 0.492$, $Cohen' d = -0.14$)。使用JASP进行贝叶斯因子分析,进一步支持动物声音与动物图片($BF_{01} = 3.61$)、动物声音与无意义音节($BF_{01} = 4.74$)条件间无显著差异,均为中等强度的零假设支持证据。上述结果表明,在多通道同步输入条件下,个体不仅保持了各通道的统计学习能力,也实现了从听觉向视觉的跨通道迁移。

6 讨论

本研究通过4个实验,系统考察了统计学习跨通道迁移的特点及其具体机制。实验1采用经典统计学习范式,验证了个体基于动物图片的视觉统计学习能力,成功复现了已有研究的结果(例如

Otsuka et al., 2013)。

实验 2 采用视觉学习—听觉测试的跨通道迁移范式,发现在仅通过视觉通道学习统计规律后,该规律不仅能在原通道中保持,也能迁移至听觉通道,且两个通道间的学习效果无显著差异。这一发现突破了以往研究多局限于单一通道内迁移的框架(如 Emberson & Rubinstein, 2016; Jun & Chong, 2018),在行为层面证实了统计学习具备跨通道迁移的可能性。

实验 3 引入多通道同步学习范式,系统操纵视听通道中的统计规律配置,实验 3a 中仅视觉刺激流具有统计规律,听觉刺激流为随机序列;实验 3b 中视听刺激流则包含同步的统计规律。结果发现:第一,与单通道学习条件相比,多通道输入并未对视觉或听觉统计学习产生显著干扰,该结果与 Li 等人(2018)及 Mitchel 和 Weiss (2011)的发现一致,支持各感觉通道具备相对独立的统计学习能力;第二,无论听觉刺激流是否包含统计规律,视觉向听觉的跨通道迁移效应均稳定存在,表明跨通道迁移过程可能与单通道统计学习平行进行,互不依赖。这一发现拓展了 Otsuka 等人(2013)关于视觉统计学习中存在图片与语义信息平行加工的观点,提示人脑可能存在专门负责跨通道统计信息整合的平行加工机制,为构建多层级的统计学习认知模型提供了新的实验依据(Frost et al., 2015)。

实验 4 进一步考察从听觉向视觉的跨通道迁移能力。结果发现,在听觉通道习得的统计规律可成功迁移至视觉通道。结合实验 3 的结果,本研究共同构成了统计学习具备通道双向迁移能力的完整证据链,表明跨通道迁移并非特定通道的优势效应,而可能是一种更具普遍性的认知机制。尤其重要的是,实验 4 中视觉通道对无意义图形的统计学习与实验 3b 中听觉通道对无意义音节的学习结果模式高度一致,进一步支持了“平行加工”推论,即人脑可能同时存在两套相对独立的统计学习系统,一套处理通道内统计结构,另一套负责跨通道规则的抽象与迁移。

本研究结果支持统计学习具有通道独立性的观点。尽管有研究表明多通道学习可能因个体发展水平或刺激特性而产生通道促进效应(Thiessen, 2010),但本研究的实验 3 发现,在多通道输入条件下,无论听觉刺激是否具有统计规律,视觉统计学习效应均保持稳定;实验 3b 中双通道统计学习的效果与实验 2 的单通道学习亦无显著差异,提示各

通道的统计学习机制具备功能独立性。这一发现与采用人工语法范式(石文典 等, 2013; Li et al., 2018)及统计学习范式(Seitz et al., 2007)的研究结论相吻合。值得注意的是, Otsuka 等人(2013)在研究视觉统计学习向语义信息迁移时发现,在被试间设计中,基于图片和基于语义的统计学习效应量存在数值差异,但在控制被试间变异后,该差异不再显著。本研究实验 2 与实验 3 均采用被试内设计,结果与 Otsuka 等人(2013)的实验 3 高度一致,进一步支持统计学习可能具备通道或领域一般性。未来研究可引入倒序三联体测试等范式(Otsuka et al., 2013),深入探讨跨通道迁移过程中不同通道间统计结构表征的具体异同。

本研究系统揭示了统计学习在视听通道间具备双向跨通道迁移能力,且该效应在多通道条件下依然稳定。当前,跨通道迁移研究在不同认知领域中仍存争议:在物体识别、类别学习及时间知觉等领域,迁移是否发生尚不一致(Casale et al., 2012; Von Der Emde & De Perera, 2020; Xiong et al., 2022),部分研究甚至报告难以发生迁移(Edmunds et al., 2020; Lengyel & Fiser, 2019; Sun et al., 2023)。统计学习领域亦存在类似理论分歧:一方面有研究强调通道或刺激特异性对迁移的制约(Han & Reber, 2022; Otsuka et al., 2014; Turk-Browne et al., 2008),另一方面也有研究观察到跨刺激的迁移能力(Brady & Oliva, 2008; Kirkham et al., 2002)。针对上述争议,理论解释上主要存在两种竞争性观点:其一主张抽象的元认知机制(amodal mechanism),认为大脑存在超通道(supramodal)或元通道(metamodal)系统(Pascual-Leone & Hamilton, 2001; Ricciardi & Pietrini, 2011),能通过任务导向的功能重组实现跨通道表征,人工语法学习研究为该观点提供了支持(Shanks et al., 1997; Tunney & Altmann, 2001);其二则倡导多层平行加工机制,认为感知加工与统计学习属于序列过程(唐溢 等, 2015; Siegelman & Frost, 2015),其迁移能力受神经层级调节,低层级感觉皮层处理特异性信息,高层级联合皮层支持跨通道泛化(Angelaki et al., 2009; Eberhardt et al., 2014; Pavlovskaya & Hochstein, 2011; Zaltz et al., 2020)。本研究结果更倾向于支持多层平行加工机制。首先,实验 3b 和实验 4 中,被试在无意义音节与无意义图形上均表现出显著的统计学习,且该学习并未干扰跨通道迁移过程,说明多种统计学习机制可并行运作。其次,本研究采用的视听材料具有

本体一致性,可能同时诱发低层级特征加工与高层级语义加工,从而既形成通道特异性表征,也构建支持跨通道迁移的高层级抽象表征。神经影像学研究为此提供了佐证,发现当学习任务需整合多维信息时,内侧前额叶—前侧海马网络被显著激活,负责构建动态统计表征以支持复杂学习的泛化与迁移(Liu et al., 2023)。

值得深入探讨的是,被试在本研究中究竟依赖何种高层级特征实现跨通道迁移?候选机制主要包括语义信息与客体本体表征。唐溢等人(2015)通过操纵面孔图片的呈现时长与方向,发现当语义信息加工受阻时,统计学习迁移效应消失,提示语义信息在迁移中具有关键作用。后续研究进一步表明,视觉统计学习可在共享基本水平或下位水平概念的客体间迁移,但在仅共享上位水平概念的客体间则无法迁移(Emberson & Rubinstein, 2016; Jun & Chong, 2018; Otsuka et al., 2014),再次凸显语义信息的重要性。然而,也有研究采用无意义图形提出,客体本体可能是迁移的另一驱动因素。Turk-Browne 等人(2008)发现,视觉统计学习更倾向于以客体整体为加工单元,而非孤立特征,表明基于客体本体的统计结构表征同样具备迁移潜力。本研究由于实验材料同时包含语义信息与客体本体特征,尚无法明确区分二者在跨通道迁移中的相对贡献。未来研究可通过选择性抑制语义加工或操纵客体完形加工,进一步澄清跨通道迁移所依赖的高层级特征机制。

总体而言,本系列实验在理论建构方面取得三方面突破:第一,将统计学习迁移的研究范畴从单通道内拓展至跨通道层面;第二,通过多通道同步输入范式,为各通道具备独立统计学习系统的观点(Seitz et al., 2007)提供了实证支持;第三,为统计学习具备通道一般性提供了证据,提示可能存在于一个与通道内学习平行的、支持跨通道迁移的抽象规则表征机制。

本研究仍存在一定局限。首先,实验材料主要集中于动物客体,虽有效验证了跨通道迁移的基本假设,但结论能否推广至其他类别客体(如工具、人造物等)尚待进一步检验。未来研究可拓展材料范围,以系统考察统计学习跨通道迁移的普适性。其次,本研究虽表明存在高层级统计结构表征,但尚未能明确其具体形式。后续可结合神经影像技术与计算建模方法,深入分离低层级特征与高层级抽象特征在迁移中的贡献机制。

7 结论

(1)个体对真实动物客体的统计学习表现出跨通道双向迁移的特性;(2)视觉与听觉通道的统计学习具有相对独立性;(3)单通道统计学习与跨通道迁移可独立地平行且同时发生,支持存在多层次统计结构表征体系的观点。

参 考 文 献

- Angelaki, D. E., Gu, Y., & DeAngelis, G. C. (2009). Multisensory integration: Psychophysics, neurophysiology, and computation. *Current Opinion in Neurobiology*, 19(4), 452–458. <https://doi.org/10.1016/j.conb.2009.06.008>
- Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language*, 83, 62–78. <https://doi.org/10.1016/j.jml.2015.04.004>
- Brady, T. F., & Oliva, A. (2008). Statistical learning using real-world scenes: Extracting categorical regularities without conscious intent. *Psychological Science*, 19(7), 678–685. <https://doi.org/10.1111/j.1467-9280.2008.02142.x>
- Casale, M. B., Roeder, J. L., & Ashby, F. G. (2012). Analogical transfer in perceptual categorization. *Memory & Cognition*, 40(3), 434–449. <https://doi.org/10.3758/s13421-011-0154-4>
- Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 31(1), 24–39. <https://doi.org/10.1037/0278-7393.31.1.24>
- Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychological Science*, 17(10), 905–912. <https://doi.org/10.1111/j.1467-9280.2006.01801.x>
- Eberhardt, S. P., Auer, E. T., & Bernstein, L. E. (2014). Multisensory training can promote or impede visual perceptual learning of speech stimuli: Visual-tactile vs. visual-auditory training. *Frontiers in Human Neuroscience*, 8, 829. <https://doi.org/10.3389/fnhum.2014.00829>
- Edmunds, C. E. R., Inkster, A. B., Jones, P. M., Milton, F., & Wills, A. J. (2020). Absence of cross-modality analogical transfer in perceptual categorization. *Open Journal of Experimental Psychology and Neuroscience*, 1, 3–13. <https://doi.org/10.46221/ojepn.2020.8639>
- Emberson, L. L., & Rubinstein, D. Y. (2016). Statistical learning is constrained to less abstract patterns in complex sensory input (but not the least). *Cognition*, 153, 63–78. <https://doi.org/10.1016/j.cognition.2016.04.010>
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/bf03193146>
- Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, 12(6), 499–504. <https://doi.org/10.1111/1467-9280.00392>
- Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 458–467. <https://doi.org/10.1037/0278-7393.28.3.458>

- Fiser, J., & Aslin, R. N. (2005). Encoding multielement scenes: Statistical learning of visual feature hierarchies. *Journal of Experimental Psychology: General*, 134(4), 521–537. <https://doi.org/10.1037/0096-3445.134.4.521>
- Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, 19(3), 117–125. <https://doi.org/10.1016/j.tics.2014.12.010>
- Han, Y. C., & Reber, P. J. (2022). Implicit sequence learning using auditory cues leads to modality-specific representations. *Psychonomic Bulletin & Review*, 29(2), 541–551. <https://doi.org/10.3758/s13423-021-02001-z>
- Hocking, J., Dzafic, I., Kazovsky, M., & Copland, D. A. (2013). Nessti: Norms for environmental sound stimuli. *PloS One*, 8(9), e73382–e73382. <https://doi.org/10.1371/journal.pone.0073382>
- Jun, J., & Chong, S. C. (2018). Visual statistical learning at basic and subordinate category levels in real-world images. *Attention Perception & Psychophysics*, 80(8), 1946–1961. <https://doi.org/10.3758/s13414-018-1566-z>
- Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, 83(2), B35–B42. [https://doi.org/10.1016/s0010-0277\(02\)00004-5](https://doi.org/10.1016/s0010-0277(02)00004-5)
- Kligler, N., & Gabay, Y. (2023). A cross-modal investigation of statistical learning in developmental dyslexia. *Scientific Studies of Reading*, 27(4), 334–354. <https://doi.org/10.1080/10888438.2023.2166413>
- Lengyel, G., & Fiser, J. (2019). The relationship between initial threshold, learning, and generalization in perceptual learning. *Journal of Vision*, 19(4), 28. <https://doi.org/10.1167/19.4.28>
- Li, X. J., Zhao, X. D., Shi, W. D., Lu, Y., & Conway, C. M. (2018). Lack of cross-modal effects in dual-modality implicit statistical learning. *Frontiers in Psychology*, 9, 146. <https://doi.org/10.3389/fpsyg.2018.00146>
- Liu, Z. Y., Liao, S. Y., & Seger, C. A. (2023). Rule and exemplar-based transfer in category learning. *Journal of Cognitive Neuroscience*, 35(4), 628–644. https://doi.org/10.1162/jocn_a_01963
- Mitchel, A. D., & Weiss, D. J. (2011). Learning across senses: Cross-modal effects in multisensory statistical learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(5), 1081–1091. <https://doi.org/10.1037/a0023700>
- Otsuka, S., Nishiyama, M., & Kawaguchi, J. (2014). Constraint on the semantic flexibility in visual statistical learning. *Visual Cognition*, 22(7), 865–880. <https://doi.org/10.1080/13506285.2014.923548>
- Otsuka, S., Nishiyama, M., Nakahara, F., & Kawaguchi, J. (2013). Visual statistical learning based on the perceptual and semantic information of objects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(1), 196–207. <https://doi.org/10.1037/a0028645>
- Pascual-Leone, A., & Hamilton, R. (2001). Chapter 27 the metamodal organization of the brain. In *Progress in brain research* (Vol. 134, pp. 427–445). Elsevier. [https://doi.org/10.1016/S0079-6123\(01\)34028-1](https://doi.org/10.1016/S0079-6123(01)34028-1)
- Pavlovskaya, M., & Hochstein, S. (2011). Perceptual learning transfer between hemispheres and tasks for easy and hard feature search conditions. *Journal of Vision*, 11(1), 8. <https://doi.org/10.1167/11.1.8>
- Ricciardi, E., & Pietrini, P. (2011). New light from the dark: What blindness can teach us about brain function. *Current Opinion in Neurology*, 24(4), 357–363. <https://doi.org/10.1097/WCO.0b013e328348bdfb>
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926–1928. <https://doi.org/10.1126/science.274.5294.1926>
- Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, 8(2), 101–105. <https://doi.org/10.1111/j.1467-9280.1997.tb00690.x>
- Seitz, A. R., Kim, R., van Wassenhove, V., & Shams, L. (2007). Simultaneous and independent acquisition of multisensory and unisensory associations. *Perception*, 36(10), 1445–1453. <https://doi.org/10.1068/p5843>
- Shanks, D. R., Johnstone, T., & Staggs, L. (1997). Abstraction processes in artificial grammar learning. *Quarterly Journal of Experimental Psychology Section A*, 50(1), 216–252. <https://doi.org/10.1080/713755680>
- Shi, W. D., Li, X. J., Wang, W., & Yan, W. H. (2013). Comparison of implicit learning effect between multisensory and unisensory. *Acta Psychologica Sinica*, 45(12), 1313–1323. <https://doi.org/10.3724/SP.J.1041.2013.01313>
- [石文典, 李秀君, 王维, 严文华. (2013). 多通道与单通道的内隐学习效应比较. *心理学报*, 45(12), 1313–1323.]
- Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, 81, 105–120. <https://doi.org/10.1016/j.jml.2015.02.001>
- Sun, X. W., Yao, L. S., Fu, Q. F., & Fu, X. L. (2023). Multisensory transfer effects in implicit and explicit category learning. *Psychological Research-Psychologische Forschung*, 87(5), 1353–1369. <https://doi.org/10.1007/s00426-022-01754-z>
- Tang, Y., Zhang, Z. J., Zeng, M. M., Huang, K., Liu, W., & Zhao, Y. J. (2015). Visual statistical learning based on the visual feature and semantic information of famous faces. *Acta Psychologica Sinica*, 47(7), 837–850. <https://doi.org/10.3724/SP.J.1041.2015.00837>
- [唐溢, 张智君, 曾攻媚, 黄可, 刘炜, 赵亚军. (2015). 基于名人面孔视觉特征和语义信息的视觉统计学习. *心理学报*, 47(7), 837–850.]
- Thiessen, E. D. (2010). Effects of visual information on adults' and infants' auditory statistical learning. *Cognitive Science*, 34(6), 1093–1106. <https://doi.org/10.1111/j.1551-6709.2010.01118.x>
- Tunney, R. J., & Altmann, G. T. M. (2001). Two modes of transfer in artificial grammar learning. *Journal of Experimental Psychology: Learning Memory and Cognition*, 27(3), 614–639. <https://doi.org/10.1037//0278-7393.27.3.614>
- Turk-Browne, N. B., Isola, P. J., Scholl, B. J., & Treat, T. A. (2008). Multidimensional visual statistical learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(2), 399–407. <https://doi.org/10.1037/0278-7393.34.2.399>
- Turk-Browne, N. B., Junge, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, 134(4), 552–564. <https://doi.org/10.1037/0096-3445.134.4.552>
- Turk-Browne, N. B., & Scholl, B. J. (2009). Flexible visual statistical learning: transfer across space and time. *Journal of Experimental Psychology: Human Perception and Performance*, 35(1), 195–202. <https://doi.org/10.1037/0096-1523.35.1.195>
- Uno, K., Asano, M., Kadowaki, H., & Yokosawa, K. (2020). Grapheme-color associations can transfer to novel

- graphemes when synesthetic colors function as grapheme "discriminating markers". *Psychonomic Bulletin & Review*, 27(4), 700–706. <https://doi.org/10.3758/s13423-020-01732-9>
- von der Emde, G., & de Perera, T. B. (2020). Cross-modal sensory transfer: Bumble bees do it Stored sensory input permits two sensory channels to exchange and compare information. *Science*, 367(6480), 850–851. <https://doi.org/10.1126/science.aba8519>
- Xiong, Y. Z., Guan, S. C., & Yu, C. (2022). A supramodal and conceptual representation of subsecond time revealed with perceptual learning of temporal interval discrimination. *Scientific Reports*, 12(1), 10668. <https://doi.org/10.1038/s41598-022-14698-6>
- Xu, G. P., Fan, R. L., & Jin, H. (2020). The cognitive and neural mechanisms of statistical learning and its relationship with language. *Advances in Psychological Science*, 28(9), 1525–1538. <https://doi.org/10.3724/SP.J.1042.2020.01525> [徐贵平, 范若琳, 金花. (2020). 统计学习的认知神经机制及其与语言的关系. *心理科学进展*, 28(9), 1525–1538.]
- Zaltz, Y., Kishon-Rabin, L., Karni, A., & Ari-Even Roth, D. (2020). Practice makes transfer imperfect: Evidence from auditory learning. *Ear and Hearing*, 41(6), 1470–1482. <https://doi.org/10.1097/aud.0000000000000860>
- Zhang, Q. F., & Yang, Y. F. (2003). The determiners of picture-naming latency. *Acta Psychologica Sinica*, 35(4), 447–454. <https://journal.psych.ac.cn/xlxb/CN/> [张清芳, 杨玉芳. (2003). 影响图画命名时间的因素. *心理学报*, 35(4), 447–454.]
- Zhao, J., Al-Aidroos, N., & Turk-Browne, N. B. (2013). Attention is spontaneously biased toward regularities. *Psychological Science*, 24(5), 667–677. <https://doi.org/10.1177/0956797612460407>

Cross-modal transfer of statistical learning under unimodal and multimodal learning conditions

TANG Yi¹, ZHAO Yajun², ZENG Qingzhang³, ZHANG Zhijun³, WU Shengnan¹

¹ Chongqing Academy of Governance, Chongqing 400041, China)

² College of Sociology and Psychology, Southwest University for Nationalities, Chengdu 610041, China)

³ Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou 310022, China)

Abstract

Statistical learning (SL), defined as an unconscious and automatic ability to extract regularities from the environment, has been shown to operate across multiple sensory modalities, including vision, audition, and touch. Although SL exhibits a certain degree of modality independence, these processes are not entirely isolated and may interact. From the perspective of object-feature processing, Frost et al. (2015) proposed the abstract rule representation hypothesis, suggesting that individuals may rely on four types of characteristics when learning inter-object regularities: modality specificity, stimulus specificity, modality generality, and SL jointly modulated by both modality and stimulus. However, theoretical disagreement remains regarding whether statistical regularities learned in one modality can be directly expressed in another—namely, whether cross-modal transfer occurs. Existing research on SL transfer has mainly focused on two areas: (a) transfer between low-level features within the same modality (e.g., from shape to color), which has not been extended to the cross-modal level, and (b) transfer between objects with semantic information, which has also not addressed cross-modal processing mechanisms. Against this backdrop, the present study used animal pictures and animal sounds as materials to examine the cross-modal transfer of SL between the visual and auditory modalities.

This study included four experiments that integrated the cross-modal transfer and multimodal SL paradigms to investigate cross-modal transfer of SL in realistic object contexts systematically. Experiment 1 constructed a visual stimulus stream using animal pictures to verify visual SL. Experiment 2 employed a cross-modal transfer paradigm in which participants were visually familiarized only with animal pictures and then tested with either animal pictures or animal sounds. By comparing performance between visual–visual and visual–auditory conditions, the experiment evaluated whether visual SL transfers across different modalities. Experiment 3 used a multimodal learning approach to separate modality-specific learning from cross-modal transfer. It aimed to (a) examine whether SL in the visual modality is independent of that in the auditory modality, and (b) investigate the relationship between visual-to-auditory transfer and auditory SL under multimodal learning conditions. Experiment 4 assessed the transfer of SL from audition to vision and, together with Experiment 3, examined the bidirectionality of cross-modal transfer.

Results showed that Experiment 1 successfully validated visual SL with animal pictures, confirming previous findings (e.g., Otsuka et al., 2013). In Experiment 2, learned statistical regularities through visual unimodal exposure persisted within the visual modality and also transferred to the auditory modality, indicating comparable learning effects across both senses. Experiment 3 revealed that multimodal input did not significantly interfere with unimodal visual or auditory SL, aligning with studies by Li et al. (2018) and Mitchel and Weiss (2011), and supporting the idea that SL operates relatively independently across sensory modalities. Furthermore, regardless of whether the auditory stream contained statistical regularities, the visual-to-auditory transfer effect remained robust, suggesting that cross-modal transfer can occur alongside unimodal SL. Experiment 4 confirmed that statistical regularities learned through audition could transfer to vision. Together, these experiments offer converging evidence for bidirectional cross-modal transfer of SL, indicating that it is not modality-specific but instead reflects a general cognitive mechanism.

In summary, the study presents three main conclusions: (1) SL of real animal objects shows bidirectional cross-modal transfer; (2) SL in visual and auditory modalities is fairly independent; and (3) unimodal SL and cross-modal transfer can occur independently in parallel and simultaneously, supporting the idea of a multilevel statistical representation system.

Keywords statistical learning, animal picture, animal sound, multimodal, cross-modal transfer