

视听语义一致性影响工作记忆

程嘉 李静静 王爱君

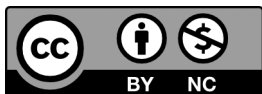
苏州大学心理学系，心理与行为科学研究中心，苏州

摘要 | 视听整合是指视听信息被整合成多感觉事件的过程，能够对工作记忆的编码、保持和提取过程产生影响。随着多感觉研究不断深入，语义一致性被发现能够影响工作记忆，但是其中的认知和神经机制仍不清楚。视听语义一致性是指视觉和听觉两个感官输入之间的内容和意义的一致性，视觉和听觉的信息相互补充和支持，而不是产生冲突或矛盾。近期研究结果表明，语义一致性能够提高人们对视听信息的理解和认知，提升多感觉工作记忆表现。未来的研究需要在认知机制方面进一步探讨视听整合对工作记忆保持过程的影响，进一步探究语义一致性对工作记忆的作用。在神经机制方面，未来的研究需要进一步探讨语义一致中的注意效应对工作记忆提取过程的影响，以此建立更加完善的理论模型。

关键词 | 视听整合；语义一致性；工作记忆

Copyright © 2024 by author (s) and SciScan Publishing Limited

This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/). <https://creativecommons.org/licenses/by-nc/4.0/>



1 引言

人们日常生活中接收到的信息通常来自多个感觉通道。已有证据表明，一些影响视听整合的因素不仅可以提高基本的感知任务，还影响涉及注意、记忆和学习的认知任务^[1, 2]。虽然关于感官之间相互作用的研究在逐渐深入^[3-5]，但是不同感官信息所形成的高阶心理表征影响工作记忆的机制尚不明确。目前研究表明，视听整合可以促进记忆表征的形成^[1, 6]。例如，视听语义一致性效应能够提升工作记忆表现^[7]等，诸多因素被发现能够影响多感觉条件下的工作记忆。

根据以上所述，众多研究已经表明工作记忆可能具有多感觉成分。关于脑成像的研究也揭示了前额叶皮层是视听整合和工作记忆之间重要的神经关联^[8]，并且研究已经发现初级视觉和听觉皮层参与了视听整合与工作记忆^[9]。近年在神经基础研究发现，在语义一致性方面，事件相关电位（event-related potential, ERP）和功能磁共振（functional magnetic resonance imaging, fMRI）的研究发现外侧枕叶皮层（LOC）

基金项目：苏州大学“大学生创新创业训练计划”（202210285015Z）。

通讯作者：王爱君，苏州大学教育学院心理学系副教授，研究方向：认知心理学。

文章引用：程嘉，李静静，王爱君：视听语义一致性影响工作记忆[J]. 中国心理学前沿, 2024, 6(4): 651-663.

<https://doi.org/10.35534/pc.0604074>

在编码阶段对语义一致的视听刺激反应更强烈^[10, 11]。其他研究表明, 涉及左楔前叶和左顶叶下叶的脑区激活可能反映了对语义一致的视听信息的保持^[12]。并且, 谢远俊等人的研究证明了视听语义一致性促进工作记忆提取的神经机制^[13]。

视听语义一致性效应是工作记忆研究中的热点问题, 而多感觉研究方法对于理解工作记忆如何编码、保持和提取信息是必不可少的。近年来, 大量研究开始关注该效应的影响因素和神经机制等问题, 但目前仍缺乏全面且系统地对语义一致性影响工作记忆的认知和神经机制进行综述。本文尝试从工作记忆三大过程的角度, 全面地讨论语义一致性效应的研究现状。首先, 从认知机制的角度阐述语义一致性对工作记忆的影响; 其次, 从多种神经科学技术的角度阐述语义一致性影响工作记忆的神经机制; 最后, 提出未来研究的展望。

2 视听语义一致性效应对工作记忆编码的影响

2.1 视听语义一致性影响工作记忆编码

视听语义一致性是指在编码过程中, 图像和声音从属于同一物体(狗的图片 and 狗的叫声是一致, 狗的图片 and 猫的声音是不一致)。这种语义一致性刺激所构建的多感觉表征可提高视觉识别精度^[14], 并促进后续短时记忆识别^[15]。有关工作记忆的研究发现, 与自然物体相关的语义一致视听编码与单感觉或语义不一致的视听编码相比, 能够加快对工作记忆对象的反应速度^[12, 13]。例如, 布鲁内蒂等人(Brunetti et al., 2017)探讨了3种视听对应条件下(音频/视觉数量、音高/形状和音高/高度)被试的工作记忆表现。结果发现一致的视听刺激条件下的工作记忆与不一致刺激相比表现得更好^[16]。最近, 黑基拉(Heikkilä)等人将口语词和图片配对时, 同样观察到了语义一致性效应对工作记忆编码的促进^[6, 17]。并且, 与语义一致的书面单词配对的声音也能够提升工作记忆识别精度^[18], 而语义不一致刺激在这些研究中均未表现出对工作记忆的促进。根据二元编码理论^[19], 视听刺激以视觉和听觉两种编码进行表征, 然而单感觉刺激只进行一次编码, 因此视听刺激编码的稳定性将优于单感觉刺激^[20], 使得个体能够通过多种编码形式访问有关刺激的更多信息。此外, 语义一致的视听表征可能通过增强个体的认知控制, 促进后续的认知加工^[7, 21]。最近有研究表明 ADHD 患者能够从语义一致的视听刺激中获益更多, 提升辨别任务^[22, 23]以及工作记忆表现^[24]。Cai 等人的研究结果显示, 语义一致的多感觉编码可以改善 ADHD 的认知控制缺陷, 他们可以利用更多的注意资源来建立足够的认知控制, 提升工作记忆表现^[24]。当然, 他们的另一种解释是语义一致编码可能为 ADHD 患者提供最佳的唤醒水平。从上述研究可以看出, 不同类型的视听材料是否能够整合进而改善工作记忆编码是语义一致性研究的重点。值得注意的是, 布鲁内蒂(Brunetti)等人的研究发现不同的视听对应编码条件会对工作记忆产生差异性的影响, 视听数对应条件下的工作记忆存在感觉通道的选择性调制, 而在其他两种情况下则没有^[16]。“数量”这一种类的材料似乎能够对视听整合后的工作记忆产生区别于其他材料的效果。一方面可能是材料性质自身的特殊性, 另一方面可能是视听整合后产生了更加深远的作用进而影响了后续工作记忆的提取。因此, 通过改变视听材料的特性, 能够从不同角度探索视听语义一致性效应如何作用于工作记忆。

与此同时, 近年有研究者在多感觉研究中使用了有实际意义的语义刺激, 发现能够对工作记忆产生

与语义一致性效应类似的影响^[25, 26]。例如, 费尔赫斯特 (Fairhurst) 等人利用在展览中的音画导览测试一致性效应是否能够延伸到更复杂的刺激^[25]。结果表明, 现实生活中视听信息的同时呈现可以促进对所见所闻的感觉回忆, 且性别一致的声音和面孔配对能够促进对语言和其他感觉信息的回忆。另外, 视听语义一致性在近年的研究中被发现还能够促进学习记忆的表现^[26, 27]。例如, 杜赛 (Dousay) 比较了图形分别搭配口语单词和印刷单词两者学习效果的好坏, 记忆测试显示利用图形搭配口语的个体得分更高^[27]。由于语义一致视听感觉绑定的主要功能是维护和整合由工作记忆同时存储和处理的感觉信息代码, 还有研究发现这可能在写作中发挥重要作用^[3]。因此, 利用视听整合对工作记忆的促进作用, 能够在设计学习环境和学习应用程序时提供理论指导。正是基于语义一致性对于工作记忆编码的巨大影响, 韦格纳·克莱门斯 (Wegner-Clemens) 等人根据不同的语义特性建立了新的人类语义关系判断数据库^[28]。今后可以依据数据库判断两者是否真正属于语义关系上的一致, 进一步探索视听整合如何影响工作记忆。

值得注意的是, 诸多的研究已经证明目标和刺激之间的不一致性会导致分心并且干扰任务相关表征形成, 进而导致工作记忆表现受损^[7, 29]。然而, 一些研究发现视听一致或不一致条件下的工作记忆表现都超过了单感觉记忆^[30, 31]。例如, 杜阿尔特 (Duarte) 等人发现, 与任务无关的不一致声音刺激可以加快对同时呈现的图片刺激的搜索^[30], 并且 Li 等人也发现声音与视觉目标在语义上不匹配时反应时更快^[32]。重要的是, 在有关空间工作记忆的研究中发现, 无论这些声音在语义上是否与视觉刺激一致, 听有意义的口语或环境声音都能增强对之前看到的视觉物体的记忆^[31]。这些语义不一致刺激对工作记忆的促进效果表明, 无论视听觉刺激是否匹配, 接触有意义的听觉刺激会启动视觉进行更深入或更复杂的语义处理, 超过视听本身的知觉表征^[30]。一方面, 不一致刺激的优势可能来自相对新颖性或感觉系统对不一致时间的相对敏感性^[31]。语义不一致的声音可作为一个关于预测或准备视觉目标的简单提示或警告信号, 不协调的声音可能通过增加个体的警觉性来提高记忆表现^[32]。另一方面, 视听不一致语义刺激编码的增强可能是由违反预期所驱动的^[31]。被试在为识别记忆任务做准备时, 会将更多的注意力分配给视听语义不一致的物体, 而不是与语义一致的物体。重要的是, Yu 等人发现与任务无关的或语义不一致的视觉刺激也为连贯的多感觉表征形成捕获了更多的注意资源^[33]。所以不一致刺激所引发更多注意捕获, 可能会增加将这种语义不一致视觉信息编码为工作记忆的概率^[34]。根据预测编码模型, 由于语义不一致的视听刺激反映了长期系统中对多感觉语义表征的侵犯, 可能会引起原系统的重大更新。这样的更新可能使得个体警觉网络发生改变, 从而使得注意增强或注意模式改变。然而, 这样的促进效果可能存在个体差异^[10]。在现实生活中也可以观察到这样的现象, 当人们沉浸在主要的视觉任务时 (如读书), 可能会认为所有的声音都是一种分心。但在某些情况下, 潜在冲突的声音 (如音乐) 对一些人也可能是有益的。然而, 在不断变化的听觉信息中记住持续呈现的视觉信息, 听觉信息可能确实能够于某些个体上表现出警觉作用, 对工作记忆产生影响。由此看来分心并不总是可以明确预测的, 视听语义不一致刺激不一定就会对个体工作记忆产生干扰。

2.2 视听语义一致性影响工作记忆编码的模型解释

最近, 概念短时记忆 (conceptual short term memory, CSTM) 模型被用来解释视听语义一致性效应。

CSTM 是一种记忆缓冲区, 当一个有意义的刺激被感知, 它会在概念层面上被迅速识别, 并激活长期记忆中相关的信息并与其相互整合, 创造出新的表征。CSTM 中的活动概念之间会形成新的链接, 由此产生的结构可以编码到工作记忆中, 没有被纳入结构的信息很快会被遗忘^[35]。Chen 和 Spence 2010 年提出 CSTM 模型可以利用视听信息进行编码^[36], 这种观点最近得到了大量研究的支持^[6, 17]。黑基拉等人发现当记忆听觉信息时, 一致的视觉刺激通过匹配的信息来改善听觉刺激的编码, 创造了更丰富的多感觉记忆痕迹, 新搭建的稳固链接提升了工作记忆表现^[17]。因此, 语义一致视听刺激中声音或图片提供的连贯信息增强了 CSTM 对视觉表征的编码。

此外, 预测编码 (Predictive Coding) 模型最近也被用来解释视听语义一致性效应^[37, 38], 其涵盖了自上而下和自下而上处理。自下而上的处理是直接依赖于外部刺激的处理, 而自上而下的处理是受期望、知识、相关背景等影响的处理^[39]。预测编码框架指出, 高阶脑区为低阶脑区提供自上而下的预测, 并基于处理过的感觉信息不断更新, 这些预测会影响正在进行的感输入的处理。自下而上的感觉处理和自上而下的预测不断地相互协调, 以此解释视听语义一致性效应。高阶脑区接收到一致的信息将激活强大且一致的内部模型, 产生较低的预测误差。如果呈现的信息于感觉通道间不一致, 可能会导致将更大的错误信号传输至感觉皮层, 所以即使在编码表征方面投入了更多的精力也只能形成较弱的表征。此外, 预测和实际感觉输入之间的强烈不匹配将导致内部模型的重大更新, 以此决定整合的优先顺序, 并且处理后的信息会改变预测。因此, 根据预测编码模型的特性, 视听语义一致刺激通过减少表征重组以及更新, 提升了工作记忆编码的效果。

虽然关于工作记忆有不同的理论^[40], 但是现有的工作记忆模型可能无法推广到多感觉环境^[10], 尤其是当概念新颖性与物理熟悉性等发挥作用时^[41]。马突兹等人 (Matusz et al., 2017) 研究结果表明, 尽管刺激与任务无关, 但语义一致性语境带来的益处超过了视觉记忆研究中一些最基本假设 (例如编码-提取一致性假设) 所预测的益处^[42]。因此, 上述的机制与模型的解释通过证明工作记忆可能与一般的多感觉过程相契合, 架起了两者之间的桥梁。对于现有视听整合对工作记忆编码影响的解释还存在激烈的讨论, 学界并没有得出统一的结论, 所以关于视听语义一致性效应的相关框架还需要进一步研究, 以确定工作记忆如何受益于多感觉环境。

3 视听语义一致性效应对工作记忆提取的影响

虽然与任务相关的语义一致视听信息可以提升工作记忆表现, 但与任务无关和不一致的视听信息会干扰其他通道中的任务相关表征^[7, 29]。然而, 这些研究未能得出编码和提取阶段反应时间之间任何显著的相关性, 这表明视听语义一致性效应对工作记忆编码的影响不能解释提取时观察到的所有效应。此外, 巴勒特 (Barutcu) 等人^[1]研究发现, 不一致视听信息的干扰表现在提取阶段而不是编码阶段。这样的结果进一步证明了视听语义一致性对工作记忆编码的影响可能并不能直接推及提取过程。

最近的研究证据已经证明了多感觉表征更加稳定并且更容易提取^[43]。例如, 布鲁内蒂等人 (Brunetti et al., 2017) 的研究结果显示, 当视听语义一致刺激传递大量有关联的信息时, 工作记忆提取阶段是受该一致性影响的特定记忆过程^[16]。有研究表明这样的效应并不是由于编码的促进, 而是由于样本刺激内的完美匹配而产生了某种感知增强, 创建了单个集成内存轨迹^[44]。这样看来, 提取阶段受到的促进

作用可以用冗余效应进行解释：一致的视听语义信息能够提供便利是因为目标提供了信息的跨通道强化。由于语义一致的视听内容在听觉和视觉通道之间是等价的，因此它们可以集成到一个统一的记忆表征中。相反，不一致的视听刺激因为语义内容不一致或相互冲突，可能不会形成新的连贯的表征^[12]。此外，当视听刺激共享相同而非冲突的语义信息时不仅可以促进即时的行为感知表现，还可以加速单感觉工作记忆提取^[12, 38]。例如，于洪涛等人的研究结果表明，只要单感觉成分与多感觉表征相关联，一致的多感觉表征通常可以促进单感觉记忆的提取，这样的结果可能支持工作记忆提取与记忆痕迹再整合的机制密切相关^[38]。与暴露在冗余视听线索下产生的知觉效应类似，语义一致视听信息对工作记忆的促进作用可能取决于两个有效感觉信息源的可用性，而不是对知觉本身更好的回忆。记忆痕迹理论认为，一条单感觉线索可以启动另一条感觉线索的快速再激活，利用这两条线索有利于回忆之前的刺激^[31]。

值得注意的是，最近巴勒特（Barutchu）等人报告当刺激的一致性与手头的任务无关时，一致和不一致的刺激都出现了类似的多感觉增益^[43]。然而，玛丽安等人（Marian et al., 2021）的研究也发现同时呈现的单词和声音都能使物体的记忆更准确，但是只有与物体所处环境相对应的声音能够增强对物体空间位置的提取——尽管声音并没有提供与空间有关的任何线索^[31]。此外，巴勒特等人的研究结果显示，工作记忆提取阶段的听觉噪声和视听不一致性都显著增加了与仅视觉条件相比的错误率^[41]。因此，语义一致性对多感觉加工的影响似乎是任务特异性的，视听刺激的显著性不能解释所有的多感觉效应。视听整合对工作记忆的增益不能推广到所有的记忆任务，语义一致性如何影响工作记忆提取还需要进一步的研究。

4 视听语义一致性效应影响工作记忆的神经机制

4.1 视听语义一致性影响工作记忆编码的神经机制

通过运用 ERP 和 fMRI 方法发现，相对于单独呈现的视觉图像，双侧枕叶皮层（LOC）对最初伴随着语义一致声音的自然视觉物体反应更加强烈（Matusz et al., 2017）^[10]。更具体地说，在 ERP 研究中，该研究发现语义一致的视听刺激使得可区分的大脑网络在第一个 60–135ms 时间段内就已经变得活跃。近年，经马突兹等人（Matusz et al., 2020）进一步研究发现，LOC 对先前语义一致条件下的多感觉环境中所呈现的视觉刺激反应明显更强烈^[11]。LOC 这种在早期较短时间内能够对语义一致刺激做出强烈反应，表明语义一致性刺激可能通过增强 LOC 的反应强度增强工作记忆编码。因此，LOC 对语义一致视听刺激的强烈反应可能是促进工作记忆编码的原因之一。此外，在神经振荡领域， θ 振荡在工作记忆不同类型信息的编码过程中已经被证明十分重要^[45, 46]。来自多感觉工作记忆研究的证据表明， θ 振荡使来自不同感官输入的情景记忆得到了更好的整合^[47]。最近，一项研究考察了被试在工作记忆任务中编码复杂视听物体时头皮电极神经信号的振荡特征，并绘制了脑功能连接图。该实验结果表明， θ 波振荡（4–8 Hz）在多个皮层区域显著分布，包括前额叶（如额上回）、顶叶（如楔前叶）、颞叶（如颞下回）和枕叶（如楔前叶）皮层^[13]。此外，进一步的研究发现，与单通道条件下物体的编码相比，这些皮层区域在 θ 振荡频率区域的跨皮层神经连接呈显著状态^[48]。这样的研究表明，通过 θ 波活动所产生的局部振荡和区域间连接在视听信息编码过程中起着重要作用，并且有助于将同时呈现的听觉和视觉信息

整合为一个完整的感受表征以加强视听工作记忆的编码,促进多感觉工作记忆痕迹的形成。因此,工作记忆中的视听信息编码过程是一个涉及多个脑区的复杂认知过程,并通过区域间 θ 连通性反映出来(前额叶皮层与包括颞、顶叶和枕叶区域在内的后皮层的连接增加),并且局部 θ 振荡及其长程神经连接是工作记忆中视听信息编码的潜在机制。

此外,一种观点认为相比于语义不一致的刺激,语义一致的视听刺激更能够增强注意捕获^[12, 13],进而促进工作记忆编码。陆兆华等人最近还发现左颞区与空间和语义信息的整合相关,其激活会影响与注意相关的大脑区域的激活^[49]。研究表明,早期视听整合在刺激呈现后40–160 ms内增强了注意相关区域,与刺激的空间和语义信息相关的视听整合在刺激呈现后200–280 ms内增强了目标刺激注意相关区域,晚期视听整合在刺激呈现后320–520 ms增强了决策相关区域的激活^[49]。神经水平的相关研究表明,与不一致的视听刺激相比,语义一致的视听整合唤起的颞上沟、颞中回和后颞区激活更强^[50]。由此看来,视听刺激之间的语义一致性加工涉及更高水平的认知加工,多感觉语义一致性被认为是决定注意对工作记忆影响程度的一个重要因素。另一种观点认为不一致的视听刺激更容易被注意所捕获^[14, 51],但是并不能促进工作记忆的编码和提取^[32]。最近,有研究在右侧颞顶叶交界处(TPJ)和控制点(顶点)上采用抑制性重复经颅磁刺激(rTMS)对语义不一致性的影响进行研究^[52, 53]。结果表明,rTMS在右TPJ上有效地减少了不一致语义对目标检测的影响,但是语义不一致性刺激很难被识别。此外,右侧TPJ上的rTMS也会影响物体的识别,所以需要重新评估右侧TPJ在加工物体信息时发挥的作用^[52]。然而,其他证据表明,与图式不一致的信息有时会被记住得更好,这可能是因为它的新颖性^[31]。然而,这种不一致的促进作用似乎会在较长的记忆延迟中消失^[54]。

综上所述,以上研究在更深层次的领域拓展了多感觉影响工作记忆的神经机制,深入探讨了在工作记忆任务中编码复杂视听物体时的神经振荡,初步揭示了神经振荡的来源及其神经连通性。然而,如今对于究竟是语义一致还是语义不一致的刺激更容易引起注意捕获依旧存在争论。此外,即使一些研究发现语义不一致刺激更容易捕获注意,但是仍产生更差的工作记忆效果的内在神经机制也没有确切的说法^[14, 51]。因此,还需要大量的实验证明视听语义不一致刺激对工作记忆产生影响的内在机制。

4.2 视听语义一致性促进工作记忆保持的神经机制

在保持语义一致的复杂视听信息过程中,常见的脑区激活主要位于左顶叶皮质,包括左楔前叶和左顶叶下叶^[12]。这些区域不仅负责物体的语义处理^[55],而且长期以来一直被视为多感觉皮层汇聚区^[56]。因此,涉及左楔前叶和左顶叶下叶的脑区激活可能反映了对语义一致的视听信息保持的促进作用。考恩(Cowan)等人在fMRI实验中将视觉刺激、听觉刺激或视听刺激作为工作记忆的负荷,观察到在工作记忆保持阶段左侧顶内沟的一个区域对视听双通道刺激反应增强^[57]。在随后进一步的研究中,研究证据表明工作记忆保持过程中,存在包括左侧顶叶内沟在内的一般神经网络,该神经网络对视听刺激的反应增强^[58]。最近,谢远俊等人采用fMRI和连接分析探究了工作记忆中语义一致与不一致的视听刺激保持^[12]。他们发现在维持语义一致的视听信息时,左顶叶皮层有共同的脑区激活(如左侧角回、边缘上回和楔前叶)。此外,有证据表明左顶上叶包含视觉神经元,并被不同类型的视觉刺激激活^[59]。人类左顶叶上叶的病变损害视觉空间表征^[60]。同时,左颞叶中叶皮层与听觉语义记忆有关^[61],并在复杂的

声音处理任务中被激活^[62]。值得注意的是,谢远俊等人研究发现在保持不一致的复杂视听信息时,左顶叶上叶和左颞中叶被共同激活^[12]。并且在保持语义一致和语义不一致的复杂视听对象时,观察到由双侧角回和左额中回组成的共同联合脑区被激活。因此,这些大脑区域可能在保持感觉信息时发挥超越感觉储存本身抑或是主动注意的作用。基于上述发现,语义不一致视听刺激的表征并没有形成统一的表征。然而,有研究已经发现,通过 θ 波活动产生的局部振荡和区域间连接可能有助于将同时呈现的视听信息统一为一个感觉呈现^[48]。因此,这种统一的表征可能是导致多感觉工作记忆益处的一个关键因素,而语义不一致刺激通过阻碍统一表征的形成,以损害记忆编码而无法在保持过程中得到良好的存储。

4.3 视听语义一致性影响工作记忆提取的神经机制

诸多研究发现,当需要记忆的视觉刺激最初伴随着语义一致的听觉刺激时,后续的感官记忆识别总体上更好^[10, 16]。虽然语义一致性对于随后的视觉工作记忆和听觉工作记忆都有促进作用^[11],但行为研究已经证明,语义一致的视听刺激呈现方式呈现为听觉记忆带来的多感觉益处比视觉记忆更强^[64]。与此相一致的是,不同的大脑区域网络和机制似乎协调了听觉和视觉记忆的多感觉益处^[65]。例如,Myers等人(2017)在刺激呈现后35-85 ms发现了与先前多感觉语义一致语境下听觉工作记忆相关的ERP差异^[65]。尽管他们的研究表明右侧颞上皮层(rSTC)、右侧顶叶内皮层(rIPC)、右侧枕下皮层和左侧额叶皮层支持了听觉记忆中的多感觉诱导益处,但是重要的一点是,rIPC和rSTC这两个区域的活动调节方式与观察到的行为受益模式一致。也就是说,与单独呈现的声音相比,它们对先前语义一致的声音表现出了抑制反应。所以根据以上的实验结果我们可以发现,语义一致的听觉工作记忆表现可能会因为某种抑制机制而表现出更差的记忆效果,这与在视听语义一致条件下听觉工作记忆获益更多的观点相矛盾。然而,目前大部分神经机制研究中所观察到的多感觉对工作记忆所产生的益处都与视觉记忆有关。因此,对于多感觉条件下的视觉和听觉工作记忆的是否受益程度存在差异还需要大量的实验去探索。另外,这些实验结果已经表明语义一致性对于视觉和听觉信息的记忆改善效果是有差异的^[7, 64],那么这种由于提取通道所导致的差异是否会在不同类型的个体呈现不一样的表现同样值得进一步研究。

此外,谢远俊等人2017年的研究发现,语义一致性的视听工作记忆编码条件下的视觉工作记忆提取更快^[13]。进一步的标准化低分辨率脑电磁断层扫描(sLORETA)结果显示,后顶叶皮层(PPC)可以发挥中枢执行作用,将视觉空间画板和语音回路中最初加工的感觉信息整合为统一的视听觉表征,从而导致更快的视觉工作记忆的提取。在随后的进一步研究中,谢远俊等人2019年的研究结果表明在刺激呈现后的236~530 ms内,顶枕电极和前额电极(从392 ms开始)出现视听信息的超附加效应,视听刺激诱发的ERP大于单感觉刺激诱发ERP的总和^[12]。此外,局部一致性(regional homogeneity, ReHo)是一种基于静息态功能磁共振成像的分析方法,能够反映全脑体素在局部区域脑功能活动状态的同步性,已有研究表明ReHo与工作记忆表现相关^[66, 67]。最近的一项研究探讨了视听工作记忆及其与静息态功能性磁共振成像(fMRI)局部一致性(ReHo)的关系^[48]。结果表明在语义一致视听刺激条件下,高行为表现(反应时快)组执行控制网络(ECN)的ReHo值增加,默认网络(DMN)和生理盐水网络(SN)的ReHo值降低。此外,ECN节点(如额叶中回和额叶上回)的静息状态ReHo值与记忆表现相关。以上结果表明视听语义一致视听条件下的工作记忆提取优于单感觉条件下的工作记忆提

取,并且可能与ECN等不同功能网络有关。如上文所述,前额叶皮层的损伤与工作记忆的缺失有关^[68],且前额叶皮层被认为是多感觉表征的组织单位^[66]。谢远俊等人的研究中的高行为表现组的ECN节点的ReHo值相较于低行为表现组增强^[48]。并且,这些区域的ReHo值越高,视听工作记忆任务的反应时越快。前额叶皮层中额中回和额上回的ReHo值与视听工作记忆的反应时之间的相关性可能表明,前额叶皮层与处理工作记忆中复杂的视听信息有关。因此,前额叶皮层可能在视听工作记忆中发挥重要作用。

5 未来研究展望

视听整合对工作记忆的影响反映了相应的多感觉因素对编码、保持和提取过程的影响。本文对工作记忆产生影响的多感觉语义一致性研究进行了系统性梳理,通过本文的综述,不但能够系统地了解视听语义一致性影响工作记忆的认知神经机制,同时也能对语义一致性相关研究提供借鉴。目前有关工作记忆的多感觉研究还存在一些局限和不足。比如目前语义一致性影响工作记忆的行为研究仅限于编码和提取阶段,对于保持阶段的研究较为缺乏。亟待解决的问题是将语义一致性的研究延伸至保持过程,探究保持阶段的视听整合效应相对于其他加工过程的共性以及差异,为语义一致性甚至多感觉整合的相关研究提供更为全面的实证支持。

虽然诸多研究都证明了视听语义一致性效应有助于工作记忆表现的提升^[6, 12, 69],但视听语义一致性对工作记忆的影响仍不明确。目前有关视听语义一致性效应的研究都无法排除干扰问题,即语义一致性效应可能是由于不一致和非语义(噪声)刺激的干扰引起的。并且有研究表明,不一致刺激的干扰表现在工作记忆的提取阶段而不是编码阶段^[11]。未来的研究需要关注实验中语义一致、不一致以及非语义配对的试次比例,考察不一致和非语义刺激对于语义一致性效应的影响,进而能够深入研究干扰效应作用于工作记忆的具体阶段。

此外,目前虽然有研究表明了语义不一致刺激会导致工作记忆表现受损^[7, 29],但是同样有研究发现语义不一致刺激可能凭借新颖性^[31]捕获注意^[33],从而促进工作记忆表现。然而当概念新颖性发挥作用时^[41],现有的工作记忆模型可能无法推广到多感觉情境^[10]。因此,未来的研究需要在排除刺激新颖性影响的前提下,借助预测编码模型进一步考察视听语义不一致刺激在工作记忆过程中所引起的注意效应等,以此进一步完善工作记忆的理论模型。

参考文献

- [1] Barutcu A, Sahu A, Humphreys G W, et al. Multisensory processing in event-based prospective memory [J]. *Acta Psychologica*, 2019 (192): 23-30.
- [2] Pahor A, Collins C, Smith-Peirce R N, et al. Multisensory Facilitation of Working Memory Training [J]. *Journal of Cognitive Enhancement*, 2021, 5 (3): 386-395.
- [3] Davies S J, Bourke L, Harrison N. Does audio-visual binding as an integrative function of working memory influence the early stages of learning to write? [J]. *Reading and Writing*, 2020, 33 (4): 835-857.
- [4] Li Q, Wu Q, Yu Y, et al. Semantic congruency modulates the effect of attentional load on the audiovisual integration of animate images and sounds [J]. *i-Perception*, 2020, 11 (6): 2041669520981096.
- [5] Lunn J, Sjoblom A, Ward J, et al. Multisensory enhancement of attention depends on whether you are already

- paying attention [J]. *Cognition*, 2019 (187) : 38–49.
- [6] Heikkilä J. Benefits of audiovisual memory encoding across the life span [J]. Department of Psychology and Logopedics, Faculty of Medicine University of Helsinki, Finland, 2018.
- [7] Yu H, Wang A, Li Q, et al. Semantically congruent bimodal presentation with divided-modality attention accelerates unisensory working memory retrieval [J]. *Perception*, 2021, 50 (11) : 917–932.
- [8] D’Esposito M, Postle B R. The cognitive neuroscience of working memory [J]. *Annual Review of Psychology*, 2015, 66 (1) : 115–142.
- [9] Henschke J U, Oelschlegel A M, Angenstein F, et al. Early sensory experience influences the development of multisensory thalamocortical and intracortical connections of primary sensory cortices [J]. *Brain Structure and Function*, 2018, 223 (3) : 1165–1190.
- [10] Matusz P J, Wallace M T, Murray M M. A multisensory perspective on object memory [J]. *Neuropsychologia*, 2017 (105) : 243–252.
- [11] Matusz P J, Wallace M T, Murray M M. Multisensory contributions to object recognition and memory across the life span [J]. *Multisensory Perception*, 2020: 135–154.
- [12] Xie Y J, Li Y Y, Xie B, et al. The neural basis of complex audiovisual objects maintenances in working memory [J]. *Neuropsychologia*, 2019 (133) : 107189.
- [13] Xie Y, Xu Y, Bian C, et al. Semantic congruent audiovisual integration during the encoding stage of working memory: An ERP and sLORETA study [J]. *Scientific Reports*, 2017, 7 (1) : 5112.
- [14] Almadori E, Mastroberardino S, Botta F, et al. Crossmodal semantic congruence interacts with object contextual consistency in complex visual scenes to enhance short-term memory performance [J]. *Brain Sciences*, 2021, 11 (9) : 1206.
- [15] Aizenman A M, Gold J M, Sekuler R. Multisensory integration in short-term memory: Musicians do rock [J]. *Neuroscience*, 2018 (389) : 141–151.
- [16] Brunetti R, Indraccolo A, Mastroberardino S, et al. The impact of cross-modal correspondences on working memory performance [J]. *Journal of Experimental Psychology: Human Perception and Performance*, 2017, 43 (4) : 819.
- [17] Heikkilä J, Alho K, Tiippana K. Semantic congruency improves recognition memory performance for both audiovisual and visual stimuli [J]. *Multisensory research*, 2017, 30 (7/8) : 763–781.
- [18] Heikkilä J, Alho K, Tiippana K. Semantically congruent visual stimuli can improve auditory memory [J]. *Multisensory Research*, 2017, 30 (7/8) : 639–651.
- [19] Paivio A. *Mental representations: A dual coding approach* [M]. Oxford university press, 1990.
- [20] Kassim A A, Rehman R, Price J M. Effects of modality and repetition in a continuous recognition memory task: Repetition has no effect on auditory recognition memory [J]. *Acta psychologica*, 2018 (185) : 72–80.
- [21] Fitzhugh M C, Whitehead P S, Johnson L, et al. A functional MRI investigation of crossmodal interference in an audiovisual Stroop task [J]. *PloS one*, 2019, 14 (1) : e0210736.
- [22] McCracken H S, Murphy B A, Burkitt J J, et al. Audiovisual Multisensory Processing in Young Adults with Attention-Deficit/Hyperactivity Disorder [J]. *Multisensory Research*, 2020, 33 (6) : 599–623.
- [23] Mc Cracken H S, Murphy B A, Glazebrook C M, et al. Audiovisual multisensory integration and evoked potentials in young adults with and without attention-deficit/hyperactivity disorder [J]. *Frontiers in Human*

- Neuroscience, 2019 (13) : 95.
- [24] Cai B, Cai S, He H, et al. Multisensory Enhancement of Cognitive Control over Working Memory Capture of Attention in Children with ADHD [J] . Brain Sciences, 2023, 13 (1) : 66.
- [25] Fairhurst M T, Scott M, Deroy O. Voice over: Audio-visual congruency and content recall in the gallery setting [J] . PloS one, 2017, 12 (6) : e0177622.
- [26] Pearson H C, Wilbiks J M. Effects of Audiovisual Memory Cues on Working Memory Recall [J] . Vision, 2021, 5 (1) : 14.
- [27] Dousay T A. Effects of redundancy and modality on the situational interest of adult learners in multimedia learning [J] . Educational Technology Research and Development, 2016 (64) : 1251-1271.
- [28] Wegner-Clemens K, Malcolm G L, Shomstein S. How much is a cow like a meow? A novel database of human judgements of audiovisual semantic relatedness [J] . Attention, Perception, & Psychophysics, 2022, 84 (4) : 1317-1327.
- [29] Lorenc E S, Mallett R, Lewis-Peacock J A. Distraction in Visual Working Memory: Resistance is Not Futile [J] . Trends in Cognitive Sciences, 2021, 25 (3) : 228-239.
- [30] Duarte S E, Ghetti S, Geng J J. Object memory is multisensory: Task-irrelevant sounds improve recollection [J] . Psychonomic Bulletin & Review, 2022: 1-14.
- [31] Marian V, Hayakawa S, Schroeder S R. Cross-Modal Interaction Between Auditory and Visual Input Impacts Memory Retrieval [J] . Frontiers in neuroscience, 2021: 887.
- [32] Li Q, Yu Y, Liu Y, et al. Whether attentional loads influence audiovisual integration depends on semantic associations [J] . Attention, Perception, & Psychophysics, 2022, 84 (7) : 2205-2218.
- [33] Yu H. Study on Benefits of Semantically Audiovisual Interaction on Unisensory Working Memory [J] . 2022.
- [34] Santangelo V. Forced to remember: When memory is biased by salient information [J] . Behavioural Brain Research, 2015 (283) : 1-10.
- [35] Potter M C. Conceptual short term memory in perception and thought [J] . Frontiers in Psychology, 2012 (3) : 113.
- [36] Chen T, Michels L, Supekar K, et al. Role of the anterior insular cortex in integrative causal signaling during multisensory auditory-visual attention [J] . European Journal of Neuroscience, 2015, 41 (2) : 264-274.
- [37] Quak M, London R E, Talsma D. A multisensory perspective of working memory [J] . Frontiers in human neuroscience, 2015 (9) : 197.
- [38] Yu H, Wang A, Zhang M, et al. Semantically congruent audiovisual integration with modal-based attention accelerates auditory short-term memory retrieval [J] . Attention, Perception, & Psychophysics, 2022: 1-10.
- [39] Eysenck M W. Psychology: an integrated approach [M] . Harlow: Addison Wesley Longman, 1998.
- [40] Adams E J, Nguyen A T, Cowan N. Theories of working memory: Differences in definition, degree of modularity, role of attention, and purpose [J] . Language, speech, and hearing services in schools, 2018, 49 (3) : 340-355.
- [41] Reggev N, Bein O, Maril A. Distinct neural suppression and encoding effects for conceptual novelty and familiarity [J] . Journal of Cognitive Neuroscience Journal of Cognitive Neuroscience, 2016 (28) : 1455-1470.
- [42] Baddeley A, Eysenck M W, Anderson M C. Memory [M] . Psychology Press. 2009.

- [43] Barutchu A, Spence C, Humphreys G W. Multisensory enhancement elicited by unconscious visual stimuli [J] . *Experimental brain research*, 2018, 236 (2) : 409–417.
- [44] Del Gatto C, Brunetti R, Delogu F. Cross-modal and intra-modal binding between identity and location in spatial working memory: The identity of objects does not help recalling their locations [J] . *Memory*, 2016, 24 (5) : 603–615.
- [45] Esmaeili V, Diamond M E. Neuronal correlates of tactile working memory in prefrontal and vibrissal somatosensory cortex [J] . *Cell reports*, 2019, 27 (11) : 3167–3181.
- [46] Proskovec A L, Heinrichs-Graham E, Wilson T W. Load modulates the alpha and beta oscillatory dynamics serving verbal working memory [J] . *NeuroImage*, 2019 (184) : 256–265.
- [47] Clouter A, Shapiro K L, Hanslmayr S. Theta phase synchronization is the glue that binds human associative memory [J] . *Current Biology*, 2017, 27 (20) : 3143–3148.
- [48] Xie Y, Li Y, Duan H, et al. Theta oscillations and source connectivity during complex audiovisual object encoding in working memory [J] . *Frontiers in human neuroscience*, 2021 (15) : 614950.
- [49] Lu Z, Li Q, Gao N, et al. Time-varying networks of ERPs in P300-speller paradigms based on spatially and semantically congruent audiovisual bimodality [J] . *Journal of Neural Engineering*, 2020, 17 (4) : 046015.
- [50] Beauchamp M S, Lee K E, Argall B D, et al. Integration of auditory and visual information about objects in superior temporal sulcus [J] . *Neuron*, 2004, 41 (5) : 809–823.
- [51] Mack A, Clarke J, Erol M, et al. Scene incongruity and attention [J] . *Consciousness and cognition*, 2017 (48) : 87–103.
- [52] Ortiz-Tudela J, Martín-Arévalo E A B, Lupiáñez J. Semantic incongruity attracts attention at a pre-conscious level: Evidence from a TMS study [J] . *Cortex*, 2018 (102) : 96–106.
- [53] Ortiz-Tudela J, Milliken B, Jiménez L, et al. Attentional influences on memory formation: A tale of a not-so-simple story [J] . *Memory & Cognition*, 2018 (46) : 544–557.
- [54] Furtak M, Doradzińska Ł, Ptashynska A, et al. Automatic attention capture by threatening, but not by semantically incongruent natural scene images [J] . *Cerebral Cortex*, 2020, 30 (7) : 4158–4168.
- [55] Cornelissen K, Laine M, Renvall K, et al. Learning new names for new objects: Cortical effects as measured by magnetoencephalography [J] . *Brain and language*, 2004, 89 (3) : 617–622.
- [56] Kitada R, Sasaki A T, Okamoto Y, et al. Role of the precuneus in the detection of incongruity between tactile and visual texture information: a functional MRI study [J] . *Neuropsychologia*, 2014 (64) : 252–262.
- [57] Cowan N, Li D, Moffitt A, et al. A neural region of abstract working memory [J] . *Journal of Cognitive Neuroscience*, 2011, 23 (10) : 2852–2863.
- [58] Li D, Christ S E, Cowan N. Domain-general and domain-specific functional networks in working memory [J] . *Neuroimage*, 2014 (102) : 646–656.
- [59] Galletti C, Fattori P, Kutz D F, et al. Arm movement-related neurons in the visual area V6A of the macaque superior parietal lobule [J] . *European Journal of Neuroscience*, 1997, 9 (2) : 410–413.
- [60] Eckert M A, Hu D, Eliez S, et al. Evidence for superior parietal impairment in Williams syndrome [J] . *Neurology*, 2005, 64 (1) : 152–153.
- [61] Kraut M A, Pitcock J A, Calhoun V, et al. Neuroanatomic organization of sound memory in humans [J] .

- Journal of cognitive neuroscience, 2006, 18 (11) : 1877–1888.
- [62] Ashtari M, Lencz T, Zuffante P, et al. Left middle temporal gyrus activation during a phonemic discrimination task [J] . Neuroreport, 2004, 15 (3) : 389–393.
- [63] Brunetti R, Indraccolo A, Del Gatto C, et al. Are crossmodal correspondences relative or absolute? Sequential effects on speeded classification [J] . Attention, Perception, & Psychophysics, 2018 (80) : 527–534.
- [64] Thelen A, Talsma D, Murray M M. Single-trial multisensory memories affect later auditory and visual object discrimination [J] . Cognition, 2015 (138) : 148–160.
- [65] Myers N E, Stokes M G, Nobre A C. Prioritizing information during working memory: beyond sustained internal attention [J] . Trends in cognitive sciences, 2017, 21 (6) : 449–461.
- [66] Chen X, Zhang Q, Wang J, et al. Cognitive and neuroimaging changes in healthy immigrants upon relocation to a high altitude: a panel study [J] . Human Brain Mapping, 2017, 38 (8) : 3865–3877.
- [67] Cheng H, Sun G, Li M, et al. Neuron loss and dysfunctionality in hippocampus explain aircraft noise induced working memory impairment: a resting-state fMRI study on military pilots [J] . BioScience Trends, 2019, 13 (5) : 430–440.
- [68] Barbey A K, Koenigs M, Grafman J. Dorsolateral prefrontal contributions to human working memory [J] . Cortex, 2013, 49 (5) : 1195–1205.
- [69] Wang A, Qian Q, Zhao C, et al. Modal-based attention modulates attentional blink [J] . Attention, Perception, & Psychophysics, 2022: 1–11.

Audio-visual Semantic Coherence Affects Working Memory

Cheng Jia Li Jingjing Wang Aijun

*Department of Psychology, Research Center for Psychological and Behavioral Sciences, Soochow University,
Suzhou*

Abstract: Audiovisual integration refers to the process in which visual and auditory information is integrated into a single multisensory event, which can affect the encoding, maintenance, retrieval process of working memory. Along with the continuously go deep into of the multisensory research, multiple factors have been found to exert influences on working memory, but the cognitive and neural mechanisms of the effect of audiovisual integration on working memory remain unclear. Audiovisual semantic congruency refers to the congruency of content and meaning between the visual and auditory sensory inputs. Visual and auditory information complement and support each other, rather than conflict or contradict each other. Recent studies show that semantic congruency can improve multisensory working memory performance. Future studies needs to further explore the impact of audiovisual integration on working memory maintenance in cognitive mechanism, and further explore the effect of semantic congruency on multisensory working memory by comprehensively considering the properties of materials and attention effects, so as to extend researches to people of different ages. In terms of neural mechanism, future studies should enrich researches on the impact of attention on the working memory retrieval process, and explore the role of different brain regions on multisensory working memory, so as to establish a more complete theoretical model.

Key words: Audiovisual integration; Semantic congruency; Working memory