

## 《心理学报》审稿意见与作者回应

题目：联结再认中双语者第二语言记忆优势效应

作者：刘贵雄，贾永萍，王余娟，买合甫来提·坎吉，郭春彦

---

### 第一轮

#### 审稿人 1 意见：

本研究借助行为和 ERP 技术探讨联结再认中维、汉双语者第二语言(L2)记忆优势效应。总的结果表明联结再认中双语者存在 L2 记忆优势效应，同时整合编码能够同时促进双语者 L2 和 L1 基于熟悉性的联结再认。本研究具有一定的理论和实践价值，但仍存在以下问题：

**意见 1：**实验假设部分，“在高整合条件下，由于个体在加工项目间关系时对认知资源的需求降低，受到其自身的激活水平影响较小；相反，L2 中来自无关情景节点的干扰小。”这句话表述不完整，表达的意义不明确，请补充表述。

**回应：**感谢审稿专家的意见。

我们对实验假设的表述进行了修改（见 P<sub>3</sub>），修改部分在正文中已用绿色标记。修改如下：在高整合条件下，两个项目被加工成一个项目或一个整体概念，个体在加工项目间关系时对认知资源的需求降低；而由于 L2 词汇的弱联结，与其相关的情景节点个数少，L2 词汇受到来自无关情景节点的干扰小。

**意见 2：**实验材料部分请用表格呈现出评定材料信息的具体的平均值和标准差。此外，应补充对新材料的具体介绍。

**回应：**感谢审稿专家的意见。

（1）我们把实验材料的评定结果在正文中用表格呈现（见 P<sub>3</sub>），修改部分已用绿色标记。

修改如下：L1( $t(25) = 93.260, p < 0.001$ )和 L2( $t(25) = 89.640, p < 0.001$ )材料中复合词和无义词在整合程度上均存在显著差异。在词频上，L1( $t(25) = 1.120, p = 0.278$ )和 L2( $t(25) = 0.370, p = 0.606$ )；在熟悉度上 L1( $t(25) = 1.290, p = 0.208$ )和 L2( $t(25) = 0.640, p = 0.412$ )均无显著差异。

（2）对新材料的具体介绍，我们做了修改：每组材料包含 12 对复合词、12 对无义词。

其中 8 对复合词和 8 对无关词用于学习, 剩余 8 对(4 对复合词和 4 对无关词)作为测试“新”的材料。修改部分(见 P<sub>3</sub>)已用绿色标记。

**意见 3:** 实验程序部分, “学习阶段: 8 对复合词、8 对无关词均以随机方式呈现在屏幕中央……”, 实验材料部分介绍每组材料有 12 对, 这里为什么变成了 8 对? 此外, 实验材料部分介绍一共有 13 组材料, 那么正式实验中每个被试都要完成这些材料吗, 请补充说明。另外, 学习阶段和测试阶段中间有没有间隔时间, 请报告清楚。

**回应:** 感谢审稿专家的意见。

(1) 每组材料包含 12 对复合词、12 对无关词, 复合词和无关词各 8 对用于学习。其中复合词和无关词各 4 对用于测试“旧”, 另 4 对复合词和 4 对无关词经过重新组合用于测试“重组”。例如, 学习“劳动-模范”和“日期-人民”经过重组后形成“劳动-人民”(复合词)和“日期-模范”(无关词), 原词的位置不变。未学习过的 4 对复合词和 4 对无关词作为测试“新”的材料。修改部分(见 P<sub>3</sub>)已用绿色标记。

(2) L1 和 L2 实验任务各包含 13 个 block, 实验为: 语言 2 (L1、L2) × 关系 2 (复合词, 无关词) × 反应 2 (旧, 重组) 三因素被试内实验设计。正式实验中每个被试完成所有实验任务, L1 和 L2 实验任务通过被试间平衡。完成全部实验任务约 2 小时 40 分钟。

(3) 正式实验包含: 学习-干扰-测试阶段三阶段, 三阶段之间没有间隔时间。学习结束后, 被试进行持续 20s 的三位数倒减 3 运算, 然后进入测验阶段。修改部分(见 P<sub>4</sub>)已用绿色标记。

**意见 4:** 结果报告部分, 行为数据未对被试的总体正确率进行说明, 同时反应时分析时是否删除了极端值? 标准是什么? 请补充说明。

**回应:** 感谢审稿专家的意见。

(1) L1 联结再认中, 总体正确率为 0.858, 标准差为 0.138; L2 联结再认中, 总体正确率为 0.803, 标准差为 0.151。考虑到数据分析中主要分析两种语言在“旧”和“重组”的正确率上的差别。因此行为数据中分别报告了双语者在 L1 和 L2 联结再认中“旧”和“重组”的正确率。

(2) 被试反应时介于 920.54~2276.42ms 之间, 平均反应时为 1525.03ms, 标准差为 324.21ms, 被试的反应时均在  $\pm 2.5$  个标准差以内, 符合被试反应特点, 因而在反应时分析时未删除极端值。

意见 5: 表 1 中 FA 重组是什么指标? 正文中都未提到?

回应: 感谢审稿专家的意见。

每个 block 中学习 8 对复合词和 8 对无关词, 其中 4 对复合词和 4 对无关词经过重组后形成 4 对复合词和 4 对无关词作为“重组”项目。例如, 学习“劳动-模范”和“日期-人民”经过重组后形成“劳动-人民”(复合词)和“日期-模范”(无关词), 原词的位置不变。FA<sub>重组</sub>为“重组”项目虚报为“旧”的概率, 用来计算 Pr 值, Pr<sub>重组</sub>等于“旧”集中率-FA<sub>重组</sub>。结合专家意见, 我们在表格下进行了标注说明(见 P<sub>3</sub>)。修改部分已用绿色标记。

意见 6: 在报告结果部分 (1) 3.1.1 正确率部分: “进一步分析结果显示, 复合词中“旧”和“重组”存在显著差异”建议表述为“进一步简单效应分析结果显示, 复合词中反应为‘旧’的正确率显著大于反应为‘重组’的正确率……” (2) 3.1.2 反应时部分: “进一步检验结果显示, 在复合词  $F(1, 17) = 7.660, p < 0.05, \eta_p^2 = 0.311$  和无关词  $F(1, 17) = 8.920, p < 0.01, \eta_p^2 = 0.344$ , L2 反应时均快于 L1。”建议表述为“进一步简单效应分析结果显示, 在复合词  $F(1, 17) = 7.660, p < 0.05, \eta_p^2 = 0.311$  和无关词  $F(1, 17) = 8.920, p < 0.01, \eta_p^2 = 0.344$ , L2 反应时均快于 L1, 但是在无关词上两种语言的差异更大。” (3) 3.1.3 鉴别指标部分: 未对语言的主效应结果进行报告, 请把不显著的结果也要报告出来。ERP 结果报告也要遵循这个原则。 (4) 请报告出所有 p 值的具体大小, 而不是用诸如  $p > 0.05$  或  $p < 0.05$  表示。

回应: 感谢审稿专家的意见。

(1) 3.1.1 正确率部分(见 P<sub>5</sub>)已修改为: 进一步简单效应分析结果显示, 复合词中反应为“旧”的正确率显著大于反应为“重组”的正确率,  $F(1, 17) = 17.552, p = 0.001, \eta_p^2 = 0.508$ 。无关词中“旧”和“重组”差异不显著,  $F(1, 17) = 2.013, p = 0.314$ 。

(2) 3.1.2 反应时部分(见 P<sub>5</sub>)已修改为: 进一步简单效应分析结果显示, 在复合词  $F(1, 17) = 7.660, p = 0.013, \eta_p^2 = 0.311$  和无关词  $F(1, 17) = 8.920, p = 0.011, \eta_p^2 = 0.344$ , L2

反应时均快于 L1，但是在无关词上两种语言的差异更大。

(3) 3.1.3 鉴别指标部分补充了语言的主效应结果（见 P<sub>5</sub>），已修改为：语言主效应不显著， $F(1, 17)=0.212$ ， $p=0.651$ ， $\eta_p^2=0.012$ 。

(4) 文中除了 P 值为 0 的修改为： $p<0.001$ ，其余的都报告了  $p$  值的具体大小。

以上四个部分的修改在正文中已用绿色标记。

**意见 7：**讨论部分很有必要针对项目记忆和联结记忆的区别对本研究结果进行讨论，突显本研究的研究价值。

**回应：**感谢审稿专家的意见。

结合专家的建议，我们对讨论部分从项目记忆和联结记忆的区别的视角进行了修改。修改参见正文（见 P<sub>9</sub>），修改部分在正文中已用绿色标记。

项目再认中 L2 记忆优势可归因于 L2 词汇-语义的弱联结。由于弱联结，来自无关情景节点的干扰较小，击中率高；同样由于弱联结，L2 概念节点的激活水平较低，虚报率低(Francis & Strobach, 2013)。根据 SAC 模型，信息提取成功与否取决于情景节点的激活水平。与项目再认相比，联结再认中，情景节点的激活除受概念节点的激活水平、与其竞争的情景节点数目影响之外，还取决于项目间关系的激活水平，而后者又受到原概念节点激活水平的影响。原概念节点激活水平越高，其项目间关系的激活也越强。由于弱联结，L2 概念节点的激活水平低于 L1，L2 两个项目之间关系的激活水平也相应地低于 L1，存在编码上的劣势(Kuperman & Van Dyke, 2013; Cohn, Emrich, & Moscovitch, 2008)。因此，提高 L2 联结再认的关键在于削弱原概念节点激活水平对项目间关系编码的影响。整合编码在一定程度上弥补了这种编码劣势。通过整合编码，两个或多个项目被编码成一个项目，并非记忆原单个概念节点，原概念节点激活水平对项目间关系编码的影响减弱(Graf & Schacter, 1989)，这与单个项目的提取类似，同样由于 L2 词汇-语义的弱联结，使得整合后的联结再认击中率高而虚报率低，表现出 L2 记忆优势。

在低整合条件下，提取项目间关系的线索为单个项目信息，项目间关系编码受到原概念节点激活的影响较大。概念节点的激活水平越低，项目间关系的激活水平也越低。其联结再认成绩也相应地越低。例如低频词的联结再认成绩低于高频词（Clark, 1992）。

参考文献: Cohn, M., Emrich, S. M., & Moscovitch, M. (2008). Age-related deficits in associative memory: The influence of impaired strategic retrieval. *Psychology and Aging*, 23, 93-103.

意见 8: 结论部分, 第 3 条不属于结论。

回应: 感谢审稿专家的意见。

第 3 条是研究的意义, 不属于结论。结合专家意见, 在正文中我们考虑予以删除 (见 P<sub>10</sub>)。

.....

审稿人 2 意见:

本文研究立意较为创新, 以少数民族学生习得国家通用语言的规律这一热点问题为内容, 通过使用 ERPs 技术, 探究联结再认中维?汉双语者第二语言(L2)记忆优势效应。研究设计较为合理, 结果较为可信, 从实际应用的角度, 本研究结果对第二语言词汇教学方法具有一定的启发意义。但还有几个细节需予以解答或说明。

意见 1: 实验材料中的复合词语义结构都是偏正式复合词吗? 如果是, 结果是否可以推至其他类型的复合词结构 (如动宾式、补充式或并列式)?

回应: 感谢审稿专家的意见。

本文实验材料中复合词是由两个双字名词构成, 比如“希腊字母”、“新闻记者”等。在语义结构上是偏正式或并列式。考虑到动宾式复合词的语义结构中包含动词和名词, 可能存在 SPT 效应, 所以实验材料中未涉及动宾式。谢谢专家给予我们的启发, 未来研究可以进一步探索动宾式复合词的联结再认记忆。

意见 2: 在低整合条件下 L2 和 L1 记忆效果差异不显著, 这一结果与低频词的联结再认成绩低于高频词 (Clark, 1992)的观点不一致的原因除了实验设计上的差别之外, 还可能由于不能将 L2 词汇等同于双语者 L1 中的低频词。

回应：感谢审稿专家的意见。

在低整合条件下 L2 和 L1 记忆效果差异不显著，这一结果与低频词的联结再认成绩低于高频词（Clark, 1992）的观点不一致。一方面缘于，在本研究中，整合编码是一个重要变量。Clark(1992)未涉及整合编码，因而低频词的联结再认成绩低于高频词。以往研究结果也表明，在整合编码条件下，熟悉性支持联结再认。另一方面，专家的意见给了我们很大启发。事实上，少数民族学生学习汉语和外国人学习中文的确不同。对于外国人学习中文，L2 词汇可能类似于双语者 L1 中的低频词。但是，对于少数民族学生学习汉语，尤其是熟练双语者，L2 词汇未必等同于双语者 L1 中的低频词。

结合专家的意见对讨论进行了修改（见 P<sub>9</sub>），在正文中已用绿色标记。

修改如下：二是 L2 词汇可能未必等同于双语者 L1 中的低频词。虽然有研究者认为，L2 词汇类似于双语者 L1 中的低频词(Gollan, Montaya, & Sera et al., 2008)，但对于中国少数民族学生，汉语作为第二语言也是其国家通用语言，L2 的接触程度可能并不等同于低频词。

意见 3：结合本研究的结果，建议可以对 SAC 模型进行修订。

回应：感谢审稿专家的意见。

结合专家建议，我们在讨论部分增加了“整合编码策略可能会影响情景节点激活水平”的观点（见 P<sub>9</sub>）。正文修改部分已用绿色标记。

在 SAC 模型中，情景节点的激活受到概念节点的激活水平以及与之相关情景节点数目两方面影响。本实验中，高整合条件下联结再认表现出 L2 记忆优势效应，而在低整合条件下 L2 和 L1 记忆效果差异不显著。这一结果表明，在再认记忆中，情景节点的激活水平可能还受到整合编码策略的影响。

---

## 第二轮

审稿人 1 意见：

作者对文章的修改使文章得到很大的提升，但是仍然存在以下问题：

意见 1：前言部分最后一段关于实验假设的表述还是不够详细和具体，建议在该处明确被试要完成实验任务是什么，交代清楚“复合词”和“无关词”与“高整合条件”与“低整合条件”的关

系，不然直接对高低整合条件下的结果进行推理假设显得唐突，显得思路很不清楚，让人理解起来困难。

**回应：**感谢审稿专家的意见。

我们对实验假设的表述进行了修改（见 P<sub>3</sub>），第一，补充了实验任务。第二，补充了复合词、无义词与高低语义整合之间的关系。第三，对实验假设提出的逻辑进行了梳理和完善。修改部分在正文中已用黄色标记。

修改如下：实验中，双语者分别完成 L1 和 L2“学习-测试”任务。在 L1 任务中，学习和测验阶段均为维吾尔语词语；在 L2 任务中，学习和测验阶段均为汉语词语。学习阶段要求被试学习混合呈现的复合词和无义词，测验阶段要求被试区分“旧”、“重组”或“新”。在复合词中，两个词语自身存在语义关联，其语义整合程度比无义词强，两个词语能够较容易地被加工成一个整体概念，个体在加工项目间关系时对认知资源的需求降低。并且由于 L2 词汇的弱联结，其受到来自无关情景节点的干扰也较小。据此我们假设，在复合词中熟悉性支持联结再认，并表现出联结再认的 L2 记忆优势效应。在无义词中，两个词语被加工成一个整体概念较困难，加工项目间关系时对认知资源的需求也较高，而且由于 L2 词汇的弱联结，其概念节点自身的激活水平较低，因此在无义词中熟悉性可能不支持联结再认。

**意见 2：**行为数据报告部分，在正确率和反应时两个因变量上请补充报告所有的两因素交互作用结果以及三因素交互作用结果，不显著也应该报告。在反应时结果部分，简单效应分析结果是否存在错误？从具体数值来看，复合条件下 L2（1276+1655=2931）与 L1（1352+1572=2924）应该是不显著；无义词条件下 L2（1495+1863=3358）应该是慢于 L1（1589+1583=3172）。而文中报告确不是这样，请作者核实。在鉴别指标结果部分，不要只是简单的报告主效应显著不显著，而要表述出谁大谁小。另外，建议在行为数据分析部分补充“Pr 联结值越大鉴别能力越强”。

**回应：**感谢审稿专家的意见。

由于作者的疏忽，行为数据中的反应时确实存在错误，作者在呈现反应时数据时误把汉语中无义词中旧和重组的反应时作为维吾尔语中复合词数据。我们对表格中行为数据进行了修正（见 P<sub>5</sub>），第一，重新处理了实验中的行为数据，并进行逐项数据比对，结果发现，数据分析结果不存在问题，但误把汉语中无义词中旧和重组的反应时作为维吾尔语中复合词数

据。第二，报告了行为数据分析所有的数据分析结果（见 P<sub>5-6</sub>）。修改部分在正文中已用黄色标记。

**意见 3：**文中讨论和结论中的“整合编码能促进熟悉性在维吾尔语联结再认中发挥作用，熟悉性支持联结再认具有跨语言的普遍性”以及“整合编码能够同时促进双语者 L2 和 L1 基于熟悉性的联结再认”，句中“熟悉性”如何体现？研究中具体哪一结果体现了“熟悉性”？实验假设部分以及整个实验设计部分都未对“熟悉性”进行任何的表述，请对这一问题进行补充说明。

**回应：**感谢审稿专家的意见。

我们对上述问题的表述进行了修改（见 P<sub>10</sub>），第一，补充了“熟悉性”和 ERPs 成分之间的关联。第二，补充研究结果如何体现了整合编码能促进熟悉性支持 L1 和 L2 联结再认。第三，进一步补充了语义整合效应对双语者 L2 和 L1 的联结再认的影响程度不同。熟悉性能够独立支持 L2 联结再认；而 L1 联结再认需要凭借熟悉性和回想共同起作用。修改部分在正文中已用黄色标记。

修改如下：实验结果从双语的角度证明了语义整合能够促进基于熟悉性的联结再认。如前言所提及，刺激后 300~500ms，额区旧/重组效应(或 FN400 效应)反映基于熟悉性的提取；刺激后 500~800ms，顶区旧/重组效应(或 LPC 效应)与回想相关 (Zheng, Xiao, & Broster et al., 2015a; Kamp, Bader, & Mecklinger, 2016)。在复合词中，双语者在 L2 和 L1 中都诱发 FN400 效应，表明熟悉性支持两种语言联结再认。在无关词中，两种语言都只诱发了 LPC 效应，表明无关词中联结再认主要凭借回想，熟悉性不支持联结再认。复合词中两个词语之间语义高度关联，其语义整合程度比无关词强。复合词中存在 FN400 效应，而无关词中仅存在 LPC 效应表明语义整合促进了双语者两种语言基于熟悉性的联结再认。维吾尔语和汉语文字是透明性差异比较大的两种文字。维吾尔语是一种透明文字，在维吾尔语中一个字素和音素之间是规则的一一对应关系。汉语是一种不透明文字，在汉语中，通常一个音节可以代表多个汉字的发音 (Liu & Cao, 2016)。在高整合(复合词)条件下，两种透明性差异比较大的文字在 250~450ms 时间窗口都诱发了 FN400 效应，此结果表明熟悉性支持联结再认记忆具有跨语言的普遍性。

此外，在复合词中，L2 仅诱发了 FN400 效应；L1 同时诱发了 FN400 效应和 LPC 效应。结果表明，语义整合对双语者 L2 和 L1 中联结再认的影响程度不同。熟悉性能够独立支持



L2 联结再认；而 L1 联结再认需要熟悉性和回想共同起作用。

---

### 第三轮

编委复审意见：

论文经过几轮修改有了较大进步，但还存在以下问题：

**意见 1：** 文章不要用第一人称写法，尽量不要用“我们”之类的表述，可以用“笔者”。

**回应：** 感谢审稿专家的意见。

我们对上述问题的表述进行了修改（见 P<sub>3</sub>）。同时，我们对全文进行了检查，均未出现类似情况。

**意见 2：** 文章的行文需要提高，流畅性与可读性较差，标点符号运用存在问题，如“认为”、“表明”之类的表达后面应该有逗号，摘要中有一些句子不完整，不该有逗号的反而有了。建议找文字水平高者阅读一下，修改文章。

**回应：** 感谢审稿专家的意见。

本文在行文过程中的确存在不少的问题。参考专家的意见，我们组织了人员对上述问题进行了讨论和修改。第一，对语言表达的流畅性，句和句之间、段落与段落之间的连贯性进行了检查和修改。第二，对全文的标点符号进行了检查和修改。

**意见 3：** 除 P 值外，数字保留两位小数。

**回应：** 感谢审稿专家的意见。

文中除 P 值保留三位小数之外，统计量及效应量都保留两位小数。

**意见 4：** 研究的理论意义、实践价值说的不清楚，理论贡献是什么，研究结果对双语学习的具体启示是什么，看了不知所云。

**回应：** 感谢审稿专家的意见。

我们对上述问题的表述进行了修改（见 P<sub>12</sub>）。修改部分在正文中已用黄色标记。

修改如下：长期以来，人们认为双语者的 L2 关系加工存在编码不充分的缺陷(Buchler & Reder, 2007)。本研究却发现，整合编码能够弥补双语者这类加工缺陷，促进 L2 的联结再认，表现出 L2 记忆优势。这一结果为我国少数民族学生有效地习得国家通用语言提供了重要的认知神经科学依据。常用的整合编码策略，如交互表象(Diana, Yonelinas, & Ranganath, 2008)、知觉特征(Yonelinas, Kroll, & Dobbins, 1999)以及语义关联(Ahmad & Hockley, 2014)等，都可以用于帮助少数民族学生学习和掌握国家通用语言。

**意见 5：**文章的排版也不规范。正文一般用 5 号宋体，行距 1.5。

**回应：**感谢审稿专家的意见。

按照学报的要求，正文用 5 号宋体，1.5 倍行距。

**意见 6：**文章的结论应该进一步修改，突出文章的主要发现。建议修改后再审。

**回应：**感谢审稿专家的意见。

我们对文章结论的表述进行了修改（见 P<sub>13</sub>）。第一，突出了，在语义整合的条件下，双语者在联结再认中存在第二语言记忆优势效应。第二，突出了，整合编码不仅促进了双语者第一语言中基于熟悉性的联结再认，而且引发了第二语言联结再认中的熟悉性加工。修改部分在正文中已用黄色标记。

修改如下：

(1) 在语义整合的条件下，双语者在联结再认中存在第二语言记忆优势效应，即在联结再认中，双语者仅凭借熟悉性就能成功提取联结关系，而且提取速度快于第一语言。

(2) 整合编码不仅促进了双语者第一语言中基于熟悉性的联结再认，而且引发了第二语言联结再认中的熟悉性加工。结果表明，整合编码促进熟悉性支持联结再认的效应不仅存在于单一语言中，而且在双语中也存在。

---

#### 第四轮

**主编终审意见：**

经过几轮的修改，几位审稿专家都认为文章质量提升很多，达到发表水准。我个人看完文章后，同意专家意见。建议录用发表。