

《心理学报》审稿意见与作者回应

题目：汉语复合词语素意义与整词语义整合加工的时间进程

作者：蔡文琦 张向阳 王小娟 杨剑峰

第一轮

审稿人 1 意见：

该研究采用 ERP 技术，考察了汉语双字词的语素意义如何参与并影响复合词的语义整合及其时间进程。采用并列式的复合词，首次在汉语中发现了语义透明度效应。研究有一定的理论意义。

回应：感谢审稿专家的肯定和建议，希望此次修改对文章有很大程度的提升。

意见 1：但该稿件有以下问题，分述如下。摘要。“结果表明在双字词的首字加工中即开始了语素意义与整字意义的整合”。首字的 ERP 结果并不能解释为与整词意义的整合。从实验任务上看，首字呈现了 800 ms，然后空屏 200 ms，然后尾字 800 ms。首字诱发的 ERP 是 -100~700 ms，在这个时间段里，尾字还没有出现，尾字意义尚未通达，更不可能整合双字词的意义。首字的结果，只能说明大脑对复合词和单语素词的首字有不同的反应，被试能够快速地区分复合词与单语素词的首字。并且，由于首字 ERP 的时间内只呈现了一个字，这个效应也可能只是反映了单字的视觉词汇识别过程。幸而在讨论中作者也没有提过首字加工中即开始了语素意义与整字意义的整合。

回应：感谢审稿专家的细致阅读与仔细推敲，这的确是我们初稿写作上的欠缺。此次修改，我们重新对摘要部分进行了整理，对首字加工中出现的语素效应进行了谨慎合理的解释。我们赞同审稿人的意见，首字加工所出现的语素效应反映了单字的视觉识别过程，我们进一步将首字加工出现的语素效应解释为，汉语语素作为独立的表征单元，在早期加工阶段就得到了自动激活，且区别于单语素词中单纯字形加工。我们在讨论部分也指出了这一解释可能存在的不足，首字发现的语素效应还可能是单纯由单字加工差异导致的，需要在对首字更为精确地控制下才能排除这种可能。具体见讨论“4.1 汉语复合词识别加工中语素意义激活过程”小节的第一段文字内容(第 13~14 页)。

意见 2：引言最后一段，“预期将在首字和尾字观测到不同的语素效应和透明度效应”，但是在前文中没有给出可以这样预期的理由。

回应：感谢审稿专家的细致审读。这确实是初稿中存在的问题。此次修改，我们在行文中重新组织了引言最后一段内容，详细论述了本研究要解决问题及实验设计，同时基于复合词的语素加工相关研究结果，提出明确清晰的研究假设。具体见“引言”最后一段(第 4 页)。

意见 3：引言最后一段，“同时记录首字和尾字诱发的 ERP 波幅”，这一描述易造成理解偏差。首字和尾字是依次呈现的，所以并不是“同时”记录首字和尾字的 ERP。

回应：完全赞同审稿专家的意见，也感谢专家对文字表达精确性要求。我们完全接受，并在行文中修改为“分别记录”，详见“引言”最后一段(第 4 页)。

意见 4：被试。样本量没有给出理由。

回应：谢谢审稿专家的宝贵意见。关于样本量的确定，一方面基于 G-Power 软件进行样本量的计算，评估一个中等大小的单因素三水平交互作用的效应量($f = 0.30$, Cohen, 1992), α 设置为 0.05, 统计检验力为 0.95(Faul et al., 2009), 计算得到的样本量为 31 人。另一方面, 结合以往关于复合词加工研究中的样本量(吴建设等, 2020; Wu et al., 2020), 样本量为 25~35 人, 以确保足够的统计效力。本研究实际收集了 34 名被试的数据。对此部分信息已经在正文第 4~5 页进行了详细说明。

参考文献：

- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160.
- Wu, Y., Duan, R., Zhao, S., & Tsang, Y.-K. (2020). Processing ambiguous morphemes in Chinese compound word recognition: Behavioral and ERP evidence. *Neuroscience*, 446, 249–260.
- 吴建设, 常嘉宝, 邱寅晨, Joseph, & Dien. (2020). 汉语复合词视觉识别的时间进程: 基于同形语素的行为与 ERP 证据. *心理学报* 52(2), 113–127.

意见 5：设计和材料。“依照前人文献，透明复合词和不透明复合词的划分采用 7 点评分”，前人文献没给出引用。“如果两个语素意义的评分之和大于 7 则为透明词，”有语病。更重要的是，按这一个定义，如果两个语素中的一个评分很低（例如，小于 2），另一个评分很高（例如，大于 5），那么这个复合词仍然是透明复合词，但我认为这是不合适的。不知道作者的材料中是否有这种情况。作为并列式的透明复合词，更好的定义是否应该为，两个语素的相关程度都大于一个中等以上的值，比如 4。

回应：感谢审稿专家对材料细节的细致审读及仔细推敲。此次修改，我们已添加前人文献，并且进一步明确地阐述了材料筛选的流程。作为并列式透明复合词，我们筛选两个语素的语义透明度值都大于 3.5 的复合词，确定为透明复合词。具体见方法“2.2 设计与材料”部分(第 5 页)。

参考文献：

- Brooks, T. L., & Cid de Garcia, D. (2015). Evidence for morphological composition in compound words using MEG. *Frontiers in Human Neuroscience*, 9, 215.

意见 6：设计和材料。建议将材料的各种属性和匹配情况用一个表来展示，注明每个属性的平均值及其在条件之间差异的显著性。并且在附件中详细列出材料清单。

回应：感谢审稿专家的建议。此次修改，我们将材料的各种属性以表格的形式展示出来，注明每个属性的平均值以及标准差。条件间匹配的统计结果在正文中已进行了详细报告。考虑到该实验设计和材料同时应用于收集脑功能成像和经颅磁刺激的实验数据，但依据该数据撰写的文章尚未进行投稿，因此，此处我们还不便提供该材料，希望审稿专家体谅。不过该材料会在待投稿文章中提供。

意见 7：实验程序。“当被试按键反应后第二个汉字立即消失，否则将一直呈现 800 ms”。这个描述很容易造成误解，让人理解成：首字和尾字的视觉呈现时间不一样；前者固定呈现 800 ms，而后者在被试反应后立即消失。此描述应修改得更清晰些。

回应：再次感谢审稿专家的细致审读，能够帮助提升文章描述清晰性。我们在行文中已对该描述进行了修改，详见方法“2.3 实验程序”部分(第 6 页)。

意见 8: ERP 数据分析。“波幅超过 $\pm 75\mu V$ 的波移除”应在 ICA 去除伪迹成分之后。因为一些伪迹（如眼电）很容易导致 ERP 波幅超过 $\pm 75\mu V$ ，而眼电是很频繁的，这会导致作者不必要地删除大量有效数据。如果作者确实是这样预处理数据的，我怀疑作者得到的有效试次是严重不足的。作者应报告预处理完之后，每个条件平均还有多少个试次，删除的试次所占比率，以及删除试次比率在条件之间是否有差异。建议作者采用合适的流程重新对 EEG 数据进行预处理。

回应：感谢审稿专家的指正。本实验处理流程参照已发表文献资料(Lo et al., 2019; Maurer et al., 2011; 张瑞等, 2021)。预处理之后，首字、尾字下各条件保留下来的数据平均有效试次均为 98%，已在修改版中进行了详细地报告(第 7 页)。

参考文献：

- Lo, J. C. M., McBride, C., Ho, C. S., & Maurer, U. (2019). Event-related potentials during Chinese single-character and two-character word reading in children. *Brain and Cognition*, 136, 103589.
- Maurer, U., Schulz, E., Brem, S., der Mark, S. van, Bucher, K., Martin, E., & Brandeis, D. (2011). The development of print tuning in children with dyslexia: Evidence from longitudinal ERP data supported by fMRI. *Neuroimage*, 57(3), 714–722.
- 张瑞, 王振华, 王小娟, & 杨剑峰. (2021). 汉字识别中亚词汇语音和语义信息在 N170 上的神经适应. *心理学报*, 53(8), 807–820.

意见 9: ERP 数据分析。“ERP 数据选取刺激呈现点之前 100 ms 到之后 700 ms 的时间窗口”，此描述不清晰。作者分别分析了首字和尾字的 ERP，所以，“刺激呈现点”中的刺激是指首字和尾字吗？并且是首字和尾字分别提取 epoch 吗？

回应：感谢审稿专家的认真审读和指正。“刺激呈现点”中的刺激是指首字和尾字刺激，并且在分析中是首字和尾字分别提取 epoch。我们在行文中已对该描述进行了修改，详见方法“2.5 ERP 数据分析”部分(第 7 页)。

意见 10: ERP 数据分析。对于 ERP 数据中多重比较的校正，Bonferroni 方法过于严格了，可考虑 FDR 校正。

回应：感谢审稿专家的建议。Bonferroni 及 FDR 校正均为多重比较校正的重要方法，并且 Bonferroni 矫正确实比 FDR 校正更为严格。但结合以往研究者对于复合词中语素加工的 ERP 数据多采用 Bonferroni 方法进行多重比较校正，因此，本研究采用 Bonferroni 方法以确保与前人研究的结果和方法具有可对比性。

参考文献：

- Coch, D., Bares, J., & Landers, A. (2012). ERPs and morphological processing: The N400 and semantic composition. *Cognitive, Affective, & Behavioral Neuroscience*, 13(2), 355–370.
- Wu, Y., Tsang, Y.-K., Wong, A. W.-K., & Chen, H.-C. (2017). The processing of homographic morphemes in Chinese: An ERP study. *Language, Cognition and Neuroscience*, 32(1), 102–116.
- 吴建设, 常嘉宝, 邱寅晨, Joseph, & Dien. (2020). 汉语复合词视觉识别的时间进程:基于同形语素的行为与 ERP 证据. *心理学报*(2), 113–127.

意见 11: 结果。应报告行为结果。ERP 结果中，词汇类型的主效应显著之后，多重比较用的 t 检验是对不同半球和分区平均数据做的吗？

回应：感谢审稿专家的建议。行文中我们报告了行为数据的结果(详见第 7 页)，被试在三种词汇类型的平均正确率为 98%，确认他们都认真完成了实验任务。多重比较用的 t 检验是基于不同半球和分区平均数据完成的。

意见 12：结果。图 1，应标出坐标轴的刻度代表什么。应增加不同词汇类型的效应的地形图。

回应：感谢审稿专家的指正。我们已对图 1、2 进行了修改(详见第 8、10 页)。同时添加了不同词汇类型的效应的地形图(详见第 13 页)。

意见 13：讨论。“语素得到自动激活之后，可能在中、后部脑区得到进一步的语义加工”，“后部脑区分布的语素效应可能反映了复合词语义的进一步加工”。这两句推测缺乏根据，作者没有提供文献表明中、后部与语义加工有关。

回应：感谢审稿专家的指正，我们在此的推测确实欠缺充分根据。此次修改，我们在行文中提供了相应的支持文献，来表明中、后部脑区与语素语义加工有关。详见讨论“4.1 汉语复合词识别加工中语素意义激活过程”小节的第二段文字内容(第 14 页)。

参考文献：

- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19(12), 2767–2796.
- Cavalli, E., Colé, P., Badier, J.-M., Zielinski, C., Chanoine, V., & Ziegler, J. C. (2016). Spatiotemporal Dynamics of Morphological Processing in Visual Word Recognition. *Journal of Cognitive Neuroscience*, 28(8), 1228–1242.
- Forgács B, Bohrn I, Baudewig J, Hofmann MJ, Pléh C, Jacobs A. M. (2012). Neural correlates of combinatorial semantic processing of literal and figurative noun noun compound words. *Neuroimage*, 63(3), 1432–1442.
- Wu, Y., Tsang, Y.-K., Wong, A. W.-K., & Chen, H.-C. (2017). The processing of homographic morphemes in Chinese: An ERP study. *Language, Cognition and Neuroscience*, 32(1), 102–116.
- Seghier, M. L. (2013). The angular gyrus: Multiple functions and multiple subdivisions. *Neuroscientist*, 19(1), 43–61.
- Zhao, S., Wu, Y., Tsang, Y.-K., Sui, X., & Zhu, Z. (2021). Morpho-semantic analysis of ambiguous morphemes in Chinese compound word recognition: An fMRI study. *Neuropsychologia*, 157, 107862.

意见 14：尾字早期的语素效应和晚期的反转的语素效应，讨论得不够。为什么是这样的模式，为什么与首字的模式不同。

回应：感谢审稿专家的所提出的探索性问题，深入讨论这两个实验发现能够有助于加强本研究的逻辑性。

首先，对于研究结果在尾字早期发现的语素效应，基于检索到相关语素/复合词加工的文献资料，我们认为，这一研究发现可能说明了汉语复合词加工中语素语义得到早期激活的过程；而单语素词仅包含一个语素，无法进行语素分解，当尾字呈现后，结合首字字形进行整词表征加工，所以在尾字加工早期表现出了两类复合词与单语素词的波幅差异。再次，对于研究结果在尾字晚期发现的反转语素效应，我们基于相关的研究证据，同时结合本研究在首字加工及尾字加工早期的研究发现，我们进行了合理推测，该阶段显示出的复合词与单语素词之间的波幅差异，可能不是反映了复合词语素意义的自动激活，而是通达复合词语义之后进一步加工或者语素结构再分析。对于这两部分的详细阐述，详见讨论部分“4.1 汉语复合词识别加工中语素意义激活过程”小节中的第三段内容(第 14 页)以及“4.2 汉语语素意义影响复合词语义整合过程”小节中的第二段内容(第 15 页)。

意见 15a: 讨论。“还有研究者使用认知神经科学技术对比了透明复合词和不透明复合词，但没有发现这两类复合词存在加工时间进程上的差异”。具体是什么认知神经科学技术？

回应: 感谢审稿专家的细致阅读，我们确实在初稿讨论部分没有详细展开介绍这方面文献研究。研究者结合 MEG 技术(Brooks & Cid de Garcia, 2015)、fMRI 技术(Momenian et al., 2021)、ERP 技术(Tsang et al., 2022)，对比了透明复合词和不透明复合词在神经层面上的加工差异，结果均未在神经层面上发现语义透明度效应。此次修改，我们对讨论的该部分内容进行了重新表述和整合。

意见 15b: 这一段是讨论 ERP 表现出了透明度效应，而不是两类词有时间进程上的差异。

回应: 非常赞同审稿专家的建设性意见。在此次修改中，我们在讨论中弱化了两类词有时间进程上的差异，将重点转移在神经生理层面发生的语义透明度效应，深入探讨了语素意义激活如何影响复合词语义整合加工机制。

意见 16: “本研究之所以能够发现复合词加工中的透明度效应，主要的原因是使用并列式复合词，...”这一句说得太肯定。

回应: 完全接受审稿专家的建议，初稿中对本研究结果发现的语义透明度效应结果显著主要归因于使用并列式复合词有些过于肯定，需要更加谨慎。在修改稿中，我们对此进行了修改，更加全面、谨慎地解释了该结果。

意见 17: 整个讨论的思路感觉不够清晰，不过结论的思路是清晰的。建议按照结论的思路重新组织讨论中的论点。

回应: 感谢审稿专家的建设性意见，有助于提升本文讨论部分的逻辑。此次修改，我们重新组织了讨论的论点内容，对其进行详细论述。此次讨论修改，我们主要围绕汉语复合词识别加工中语素意义激活过程及汉语语素意义影响复合词语义整合过程这两部分展开重点论述，详见第 13 至 16 页的讨论部分。

意见 18: 其它小问题：

1. 摘要最后一句，“详细揭示了...时间进程”，不妥；建议改为“揭示了...的精细时间进程”。
2. 摘要。“开始了语素意义与整字意义的整合”， 整字意义-->整词意义
3. 引言“在一项汉语复合词 fMRI 研究中”，有语病。
4. 引言中已提出，拼音文字与汉字在复合词语素整合上有很大的差别。所以该文的标题应指出是汉字中的发现。
5. ERP 数据分析。“采样频率为 1000Hz/导” --> “采样频率为 1000 Hz”。另外，数字和单位之间应该有空格。

回应: 感谢审稿专家的建议。此次修改，我们已经对上述问题进行了修改，也反复通读了全文，还修改了一些类似的小问题。另外，参照审稿意见，文章题目改为“汉语复合词语素意义与整词语义整合加工的时间进程”。

审稿人 2 意见:

该研究通过脑电技术探索语素意义如何参与并影响复合词的语义整合及其时间进程问题，本文行文较为流畅，但还存在一些问题，建议作者仔细修改。

回应: 感谢审稿专家的肯定和建议，希望此次修改对文章有很大程度的提升。

意见 1: 重要概念的定义并不清晰。前言中没有交代“复合词”、“单语素词”、“语素”、“语素效应”以及“语素整合”的定义。这对读者阅读产生困扰，建议作者在适当的位置补充完整。

回应: 感谢审稿专家的宝贵意见。根据您的意见，我们在修改稿中的适当位置增加了“复合词”、“单语素词”等重要概念的描述。具体见修改稿引言的蓝色部分内容。

意见 2: 作者在前言中提到“并列式汉语复合词的特殊性为考察语素意义激活之后的语义整合过程提供了可能”，这句话是怎样理解的？建议作者再阐明一下。

回应: 感谢审稿专家的建议。在此次修改，我们进一步阐明并列式汉语复合词的特殊性为考察语素意义激活之后的语义整合过程。具体解释如下：

在汉语透明复合词中，除了偏正式之外，并列式透明复合词也占据很大比例，这类复合词在语素间的关系结构上与不透明复合词相同，语素间不存在修饰关系。具体来说，并列式透明复合词的首、尾语素具有相同或者相近的语义信息/特征，对词义贡献相同。被试在加工并列式透明复合词时，首、尾语素意义得到激活后，其意义/语义特征信息与整词高度相关，随后语素整合过程所得意义与整词一致，从而快速获取语义。而不透明复合词的语素意义整合所得意义与整词存在冲突，可能干扰或阻碍整词语义的通达。因此，我们可以通过对比并列式透明复合词及不透明复合词来进一步考察语素意义激活之后的语义整合过程。对此问题的具体解释，我们在正文第 3~4 页进行了详细阐述。

意见 3a: 实验材料部分，建议作者增加适当的参考文献支持透明词与不透明词的筛选。另外，词频是根据哪个数据库？三种类型词汇的字频是否也进行了相应匹配？

回应: 感谢审稿专家的细致审读。我们已经增加相关文献来支持透明复合词与不透明复合词的筛选(Brooks & Cid de Garcia, 2015)。本研究双字词频来自《现代汉语通用词表》(2003)。考虑到复合词、单语素词的首、尾字频很难进行匹配，同时参考相关材料匹配情况(Brooks & Cid de Garcia, 2015; Coch et al., 2012; Lee et al., 2021)，我们并未对三种类型词汇的字频进行匹配。我们在讨论部分也指出了首字发现的语素效应还可能是单纯由单字加工差异导致的，需要在对首字更为精确地控制下才能排除这种可能，比如字频。详见讨论“4.1 汉语复合词识别加工中语素意义激活过程”小节的第一段文字内容(第 13~14 页)。

意见 3b: 再者，作者是否考虑了词汇的歧义性，比如“风流”，它有两种含义，歧义性可能会影响大脑的神经反应。

回应: 非常感谢审稿专家提供的建设性意见，给我们后续研究带来很大启发性。尽管本研究结果支持了复合词加工中同时存在着语素分解再整合、以及直接整词通达两条途径，但对于这两条途径的机制及神经基础还需要进一步深入研究。歧义性作为重要词汇语义变量，能够调节词汇语义通达过程，得到了越来越多研究证据支持。未来的研究我们考虑将词汇歧义性纳入研究变量，对复合词加工过程背后的神经机制进行探究。详见讨论“4.4 局限与展望”小节的文字内容(第 16 页)。

意见 4: 实验材料和设计与本研究的研究问题是怎样对应的？本研究的目的是探讨语素语义何时被激活、以及如何影响整词语义整合的时间进程问题。实验材料是三种不同类型的词汇，那么，如何通过比较复合词与单语素词的语素效应，以及透明词和不透明词之间的差异（语义透明度效应）解决本研究的问题？建议作者在前言中重新梳理思路。

回应: 感谢审稿专家的仔细推敲。本研究结合 ERP 技术，分别记录首字和尾字依次呈现时诱发的 ERP 波幅，深入揭示出汉语复合词识别中语素意义何时被激活以及如何影响复合词语义整合的时间进程。在首字加工中，通过具体对比两类复合词与单语素词加工，能够揭示

出汉语语素激活语素意义激活过程；在尾字加工中，对比两类复合词与单语素词的加工，能够揭示出汉语复合词的加工路径。更重要的是，通过对比透明复合词、不透明复合词在尾字加工上的差异，能够揭示出在汉语复合词识别中，语义透明度如何调节语素意义与整词语义整合过程。

此次修改，我们重新对引言内容进行了调整，尤其是最后两段内容，从而本研究的研究问题与所采用实验材料与设计更好地对应。详见正文第 3~4 页。

意见 5：与上个问题类似，本文对于结果的预期似乎比较模糊。作者只表述了“预期实验结果能够同时得到语素效应和透明度效应”、“预期将在首字和尾字观测到不同的语素效应和透明度效应”，什么样的结果能够揭示语素义的激活影响整词语义整合的作用机制？建议作者再明确一下预期。

回应：感谢审稿专家的细致审读。这确实是初稿中存在的问题。此次修改，我们在行文中重新组织了引言最后一段内容，详细论述了本研究要解决问题及实验设计，同时基于复合词的语素加工相关研究结果，提出明确清晰的研究假设。详见“引言”最后一段内容(第 4 页)。

意见 6：数据分析部分，作者采用数据驱动的方式确定时间窗，是否有相应文献支持？另外，所有实验刺激呈现两次，那么第一次呈现是否会对第二次呈现产生影响？

回应：采用数据驱动的方式确定时间窗在之前的研究文献中广泛采用，此次修改稿中我们添加了相应的文献支持。本研究每个条件下 30 个刺激材料，考虑脑电分析叠加试次少，很难得出稳定、有意义的结果，我们将实验刺激呈现了两次，同时保证了相同刺激的出现间隔至少一个 run，尽可能控制刺激重复性在本实验中的干扰作用。

参考文献：

Bemis, D. K., & Pykkänen, L. (2011). Simple composition: A magnetoencephalography investigation into the comprehension of minimal linguistic phrases. *Journal of Neuroscience*, 31(8), 2801–2814.

Kim, S., & Pykkänen, L. (2019). Composition of event concepts: Evidence for distinct roles for the left and right anterior temporal lobes. *Brain and Language*, 188, 18–27.

张瑞, 王振华, 王小娟, & 杨剑峰. (2021). 汉字识别中亚词汇语音和语义信息在 N170 上的神经适应. *心理学报*, 53(8), 807–820.

意见 7：结果部分建议精简，比如“分区和半球有交互作用”的结果并非本研究的主要结果，建议精简描述后续的简单效应分析结果。

回应：感谢审稿专家的意见，在初稿中分区和半球有交互作用”的结果阐述确实有点拖沓，不够简洁。此次修改我们精简描述了分区×半球交互作用显著下的简单效应分析后续统计分析结果以及主效应显著下的事后检验分析结果，具体数据结果信息以附表的形式报告。在正文中对附表信息进行了介绍。

意见 8：讨论部分，本文讨论分为两部分，第一部分标题“双字词语素整合加工的时间进程”与第二部分标题“复合词语素意义整合的时间进程”的含义似乎是一样的。另外，建议增加有关该研究的创新性贡献和理论贡献。建议作者重新梳理讨论的逻辑。

回应：非常感谢审稿专家提出的关于讨论部分的建议，能够帮助提升文章的研究意义。在此次修改中，我们重新梳理了讨论的逻辑内容，并对讨论这两部分标题内容做出了调整。本研究对比了并列式透明复合词、不透明复合词的神经加工过程，结果发现了语义透明度调节了

汉语复合词语素意义与整词语义整合的神经加工过程，这是本研究最大的研究亮点。此外我们还增加了理论贡献，详见第 13 至 16 的讨论部分。

意见 9: 作者似乎遗漏掉两篇与本研究内容相关的文献，“词汇识别中歧义语素语义加工: ERP 研究. 心理学报”以及“Morpho-semantic analysis of ambiguous morphemes in Chinese compound word recognition: An fMRI study.”*Neuropsychologia*,”建议作者阅读。

回应: 感谢审稿专家的建议。我们已认真阅读这两篇非常有价值的文献，研究者分别利用了 ERP、fMRI 技术，并结合掩蔽启动范式，探究了汉语歧义语素的神经加工过程。我们将重要研究发现作为有力的支持证据添加在文中相应位置。具体表述如下：

首先，歧义语素的 ERP 研究结果得出，在词汇识别中歧义词素语义可以激活，并对目标词汇加工的早期加工阶段发挥作用。该结果支持了语素在早期加工阶段就得到分解激活的观点，在文中引言部分第四段位置进行了具体引用，详见第 2 页。再次，歧义语素的 fMRI 研究结果发现了同形异义条件在分布式脑网络中具有显著的神经启动效应，同形同义条件在双侧颞中回、左侧额上回均有显著的启动效应。该研究结果说明了语素语义加工并非单一脑区负责，而是由众多脑区共同负责加工。本研究的研究结果发现也支持了这一观点，因此我们在文中讨论部分“4.1 汉语复合词识别加工中语素意义激活过程”小节中进行了具体引用，详见第 14 页。

意见 10: 自检报告中的第一个问题，对本研究的创新性贡献，似乎理由不充足。

回应: 感谢审稿专家的意见。本次修改对自检报告中的问题重新进行了论述，突出了本研究的特色和亮点。具体表述如下：

本研究采用 ERP 技术，构建了透明复合词、不透明复合词及单语素词三类双字词材料，探讨了汉语复合词语素意义与整词语义整合加工的时间进程。结果揭示了在复合词识别中，语素作为独立表征单元，在早期加工阶段就得到了自动激活；语义透明度调节了复合词语素意义整合的早期加工过程。总的来说，本研究揭示了双字词语素意义整合加工的精细时间进程。

意见 11: 细节问题：

1. 一些句子的表达不够通顺，建议作者仔细核对。如，引言的第二段“尽管上述理论模型存在一定的分歧，但是都肯定了语素在复合词加工中的具有区别于整词的表征并能够影响整词的语义通达加工”。如，引言的第五段“尽管有研究表明语素意义在复合词加工会得到自动激活，但是语素意义激活后的整合加工过程还缺乏深入地研究”。

2. 建议摘要部分再重新整理一下。

回应: 谢谢审稿专家的意见。此次修改，我们已经对文章语句通顺性的问题进行了修改，也反复通读了全文，修改了一些类似的小问题。对于文章摘要部分我们也做了进一步整理。

第二轮

审稿人 1 意见:

作者在修改稿中很好地解决了我提出的问题。只有几个小建议：

意见 1: 正文中没有明确表述“两个语素的语义透明度值都大于 3.5 的复合词，确定为透明复合词”。

回应：感谢审稿专家的细致审读。此次修改，我们已在正文中添加了这部分内容：“本实验选取首、尾词素的透明度评分均 ≥ 3.5 的词汇确定为并列式透明复合词，并且首、尾词素的透明度评分不存在显著差异($t(58) = -0.443, p = 0.660$)。”见正文第 5 页“实验材料”部分。

意见 2：地形图应尽量与其对应的 ERP 波形图放在同一张图中。

回应：感谢审稿专家的认真指正。我们已将地形图、ERP 波形图做了调整，并放在了同一张图内，见正文第 8 页(图 1)、第 10 页(图 2)。具体调整如下：

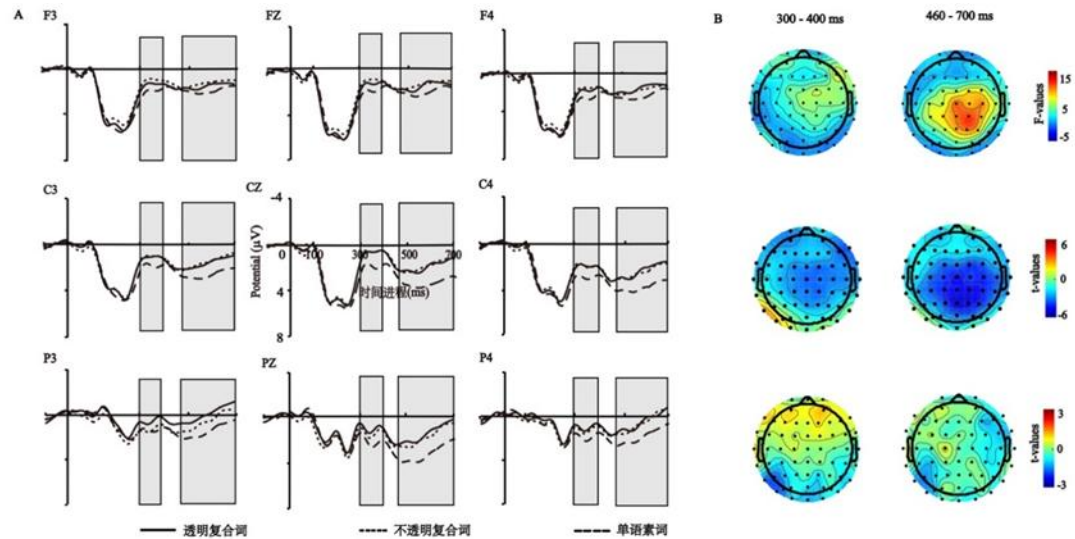


图 1 A. 首字加工中不同词汇条件诱发的平均波幅变化；B. 词汇类型主效应(上)、语素效应(中)、语义透明度效应(下)的地形分布图。

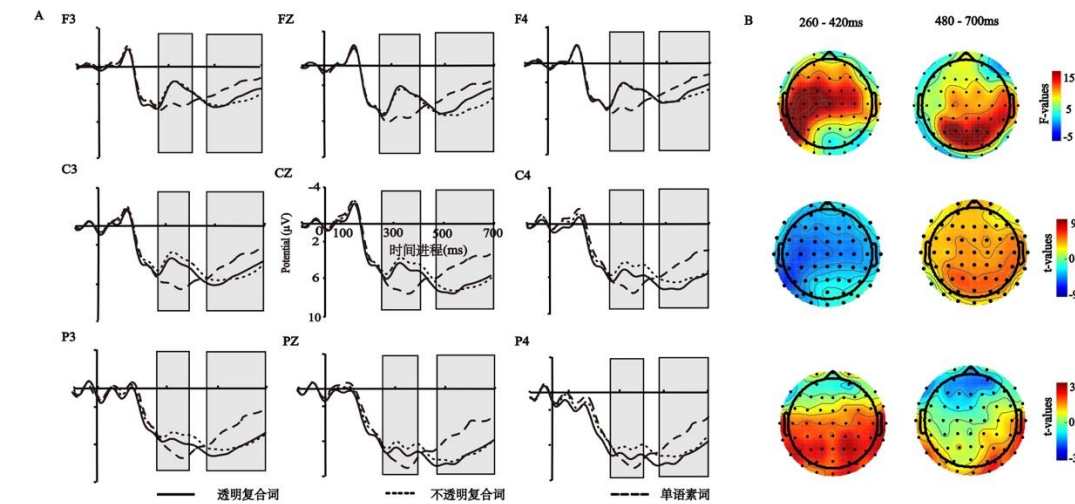


图 2 A. 尾字加工中不同词汇条件诱发的平均波幅变化；B. 词汇类型主效应(上)、语素效应(中)、语义透明度效应(下)的地形分布图。

意见 3：图 4 的第一行地形图的 colorbar 取值范围可以考虑加大一些，以显示出不同区域的变化。

回应：谢谢审稿专家的宝贵意见。本次修改将地形图重新做了调整，将原先语素效应的地形图取值范围[-3, 3]进行了适当加大，使其更加清晰地显示出不同区域的变化。在首字加工中，语素效应的地形图取值范围具体调整为：[-6, 6]；在尾字加工中，语素效应的地形图取值范围具体调整为：[-9, 9]。详见正文第 8 页(图 1)、第 10 页(图 2)。

意见 4：引言“而在汉语透明复合词中，除了偏正式之外，并列式透明复合词也占据很大比例”，这句话可能存在语病。

回应：再次感谢审稿专家的细致审读，能够帮助提升文章表述的流畅性。此次修改已经对该部分文字表述进行了修改，见“引言”部分(第 3 页倒数第二段)。具体表述为：“英语复合词大都是偏正式结构，而汉语复合词则包含了并列式、偏正式等多种构词结构(Kuo et al., 2006; Liu et al., 2010)。”

参考文献：

Kuo, L., & Anderson, R. C. (2006). Morphological awareness and learning to read: A cross-language perspective. *Educational Psychologist, 41*, 161–180.

Liu, P. D., & McBride-Chang, C. (2010). Morphological processing of Chinese compounds from a grammatical view. *Applied Psycholinguistics, 31*(4), 605–617.

.....
审稿人 2 意见：

本文距离上次修改有了很大进步，但还存在一些表述或者说逻辑上不通顺的地方，希望作者仔细推敲。

回应：感谢审稿人的肯定，希望此次修改对文章有更大程度的提升。

意见 1：作者在结论中写到“首次在神经生理上发现了汉语复合词加工中的语义透明度效应。”本文应该不是首篇文章在神经生理上发现汉语复合词加工中语义透明度效应。首先，2008 年的一篇文章“An ERP study of the semantic transparency and semantic combinability effects in reading Chinese”采用 ERP 技术发现了语义透明度效应。其次，2021 年 Lee 等人的“The neural basis of compound word processing revealed by varying semantic transparency and morphemic neighborhood size”也发现了相比于汉语透明复合词，汉语不透明复合词在左侧额中回、双侧额下回、前运动辅助区的激活程度更大，即存在汉语复合词的语义透明度效应。另外，2022 年发表在 *Psychophysiology* 杂志的“An ERP megastudy of Chinese word recognition”也表明语义透明度影响词汇通达。所以，本文并不是首篇在神经生理上发现汉语复合词加工中语义透明度效应。总的来说，创新性有些不足，有关双字词语素意义整合加工的精细时间进程的研究在上述列举的两篇文章中都有体现，那么本文的创新点不是很突出。

回应：感谢审稿专家的细致阅读和仔细推敲。

我们承认，原版稿件的确存在这方面的不足。首先，我们原稿件的“首次”表述不够严谨，尤其是遗漏了对 2022 年最新文献的分析。此次修改中，不仅去掉不正当的表述，还对相关文献进行了仔细分析，更加清晰的理清了本研究的创新和贡献。其次，由于原稿件遗漏了 Tsang 等人(2022)的最新工作错误地以为是“首次”探讨，在考虑到 Tsang 等人(2022)的研究时就显得创新性不足。

本次修改，我们对 Tsang 等人(2022)的研究进行了深入分析(详见第 4 页第 2 段)，明确指出其方法局限没能探讨语素与整词语义的整合加工机制：一方面，使用混合效应模型对多变量进行标注，考察首、尾语素的语义透明度效应发生的时间窗口，透明度效应反映的是与

所有词汇平均水平相比的高低差异，并不是相对与不透明语素的比较；另一方面，没有同时考察两个语素的透明度。从而本研究明确提出针对上述不足，结合汉语复合词的特点，能够对汉语复合词的整合加工过程进行深入细致地考察。

参考文献：

Tsang, Y., & Zou, Y. (2022). An ERP megastudy of Chinese word recognition. *Psychophysiology*, 59(11), e14111.

意见 2：前言中提到的多层激活模型和彭聃龄的 IIC 模型与讨论中提到的 Taft 的 lemma 模型以及混合表征模型并没有对应上。建议前后照应。

回应：感谢审稿专家的细致阅读，完全接受审稿专家的建议。

此次修改，我们在引言部分增加了对于上述认知理论模型的系统介绍，并提出结合认知神经科学技术手段，深入揭示模型内在加工机制的必要性。通过对该部分内容的修改，不仅做到了前后照应，也阐明了从事 ERP 研究的必要性。见“引言”部分(第 2 页)。

参考文献：

Taft, M. (2003). Morphological representation as a correlation between form and meaning. In E. Assink & D. Sandra (Eds.), *Reading complex words* (pp. 113–137). Amsterdam: Kluwer.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology Section A*, 57(4), 745–765.

意见 3：关于透明词和不透明词的筛选标准。“本实验采用的是并列式复合词，尽可能平衡了首、尾语素的透明度评分，在透明复合词中，首、尾语素的透明度评分不存在显著差异($t(58) = -0.443, p = 0.660$)”，意思是并列式复合词中每个语素的透明度评分均大于 3.5 么？

回应：感谢审稿专家对材料细节的仔细推敲。

本研究将两个词素的语义透明度值都大于 3.5 的复合词，确定为并列式透明复合词。此次修改，我们已在正文中添加了这部分内容，见正文第 5 页“实验材料”部分。具体表述如下：

本实验选取首、尾词素的透明度评分均 ≥ 3.5 的词汇确定为并列式透明复合词，并且首、尾词素的透明度评分不存在显著差异($t(58) = -0.443, p = 0.660$)。

意见 4：前言中的一段表述“最新的一项汉语复合词研究中，研究者结合 ERP 技术，使用被动 oddball 范式，也没有发现被试在加工透明复合词与不透明复合词的差异(Tsang et al., 2022)。而在一项 fMRI 研究中，结果却发现了汉语不透明复合词比汉语透明复合词更多激活了左侧额中回、双侧额下回、前运动辅助区(Lee et al., 2021)。以上结果不一致可能是由不同文字体系的语素特点及复合词构词结构引起的。”这两篇文章均为汉语复合词却得到不一样的结论，复合词构词结构是怎样引发结果的不一致？

回应：感谢审稿专家的仔细推敲，能够帮助我们重新思考并整理文章表述内容，进一步提升文章论述的逻辑性。

我们承认，原稿件在上述文献综述上存在逻辑问题。首先，Lee 等人(2021)的研究是 fMRI 实验，而 Tsang 等人(2022)的研究是 ERP 实验。其次，这里 Tsang 等人(2022)的研究是听觉词汇加工，Lee 等人(2021)的研究是视觉词汇，不同模态下词汇识别过程存在差异也很正常。我们原稿件确实存在简单罗列文献，没有细致分析问题。

此次修改，我们对语义透明度效应的认知神经研究进行了细致分析(详见引言第 4 页第 1~2 段)。首先，从 fMRI 的大脑功能定位的角度介绍了透明度效应涉及的脑区功能激活，反

映了复合词语素整合涉及了语素分解、语素义表征通达、以及语素义整合等复杂的加工过程；然后，结合目前 ERP 研究中听觉与视觉复合词加工的不一致结果，进而指明 Tsang 和 Zou(2022)研究还有待深入揭示的问题。

通过此次修改，不仅使文献结果更加条件清晰，也让本研究的创新和贡献更加明确。

参考文献：

- Boylan, C., Trueswell, J. C., & Thompson-Schill, S. L. (2017). Relational vs. attributive interpretation of nominal compounds differentially engages angular gyrus and anterior temporal lobe. *Brain and Language*, 169, 8–21.
- Gao, F., Wang, R., Armada-da-Silva, P., Wang, M., Lu, H., Leong, C., & Yuan, Z. (2022). How the brain encodes morphological constraints during Chinese word reading: An EEG-fNIRS study. *Cortex*, 154, 184–196.
- Tsang, Y., & Zou, Y. (2022). An ERP megastudy of Chinese word recognition. *Psychophysiology*, 59(11), e1411.
- Zou, L., Packard, J. L., Xia, Z., Liu, Y., & Shu, H. (2016). Neural correlates of morphological processing: Evidence from Chinese. *Frontiers in Human Neuroscience*, 9, 714.

第三轮 编委复审

审稿人 2 意见：同意发表。

编委复审意见：

作者根据审稿专家的问题做了认真修改，审稿专家认可作者的修改，修改稿的质量有了很大提升，达到了《心理学报》的要求，建议接受发表。我阅读后发现文章中存在一些问题，建议作者进行修改。

意见 1：实验程序（2.3）部分出现了中英文混合表达的情况，将其中的英文修改为中文表达。

回应：感谢编委专家的宝贵意见。此次修改，我们已将“2.3 实验程序”部分中的“trial”统一修改为“试次”，“run”统一修改为“部分”。见正文第 6 页“2.3 实验程序”的蓝色字体部分。

意见 2：图 1A 中 Fz 点的平均波形图的-100 ms 至 0 ms 之间的波形看起来有问题，放大看都处于 x 轴以下，检查一下是否在做图的过程中波形发生了偏移。

回应：感谢编委专家的细致地审读。经过核对，我们在使用 Illustrator 软件作图过程中产生了波形偏移，这是我们的疏忽。此次修改中，我们纠正了上述错误，将 ERP 波形图做了调整，见正文第 8 页(图 1)。

意见 3：“实验共分为 6 个 run，每个 run 包括 40 个 trials，其中每个条件 10 个 trials”，这里的每个条件包括了假词以及真词的条件，需要明确说明。

回应：感谢编委专家的建设性意见。此次修改已经对该部分文字进行了调整，力求文字表达的精确性。具体表述为：“实验共分为 6 个部分，每个部分包括 40 个试次，其中透明复合词、不透明复合词、单语素词以及假词条件各 10 个试次”。(见正文第 6 页“2.3 实验程序”)

意见 4：行为结果三种条件下的平均正确率要分条件进行报告，三种条件之间是否存在显著

差异？

回应：这是我们的疏忽。因为实验中的按键反应是为了确保被试能够保持注意并认真完成任务，只需要被试对 25 % 的试次(假词条件)作出反应，三种真词条件下，被试不需要作任何反应。实验结果也发现三种真词条件下(透明复合词、不透明复合词和单语素词)，被试几乎都能够正确地做出“否”反应的判断，三个条件平均仅有 1.4 试次(2.34 %)被试错误地作出了按键反应。所以我们就在正文中仅报告了被试行为数据的总体结果，未对真词的行为数据进行分条件详细统计分析。

此次修改，我们接受了专家的建议，分条件报告了被试的错误反应并对条件间差异进行了统计分析，见正文第 8 页“3.1 行为结果”的蓝色字体部分。具体表述如下：

“所有 32 名被试都能成功地完成真假词判断任务，被试在三种真字条件下错误地作出按键反应的平均试次仅为 1.41 次(2.34 %)。其中，在透明复合词条件下的平均错误试次为 1.53 次(2.55 %)，在不透明复合词条件下的平均错误试次相对最高为 2.31 次(3.85 %)，而在单语素词条件下的平均错误试次仅为 0.38 次(0.63 %)。对三种条件的错误率进行单因素重复测量的方差分析，结果发现词汇类型的主效应显著($F(2, 62) = 12.978, p < 0.001, \eta^2 = 0.295$)，被试在单语素词条件下的错误率要显著低于透明复合词($t(31) = -3.966, p = 0.001, \text{cohen's } d = -0.826$)和不透明复合词($t(31) = -3.956, p = 0.001, \text{cohen's } d = -0.919$)，在透明复合词条件下的错误反应也要略低于不透明复合词条件($t(31) = -2.310, p = 0.083, \text{cohen's } d = -0.500$)。”

意见 5：ERP 结果的报告部分需要报告每种条件下有效试次的平均值。

回应：完全接受编委专家的建议。此次修改，我们已将“预处理完成后，首字、尾字下各条件平均有效试次比例为 98%。”替换为：“在首字下，透明复合词、不透明复合词和单语素词条件保留的有效试次数相同，都保留了 97.91 % 的有效试次。”以及“在尾字下，三种实验条件保留的有效试次也完全相同，都保留了 98.43 % 的有效试次。”见正文第 8 页、第 10 页的“3.2 ERP 结果”。

意见 6：本研究中构成词的两个语素是分开呈现的，与两个语素同时呈现的条件相比，其时间进程可能不同，尤其是对尾字加工的影响可能不同。另外，每个语素呈现时间是 800 ms，一般地，人们通达一个汉字意义的时间在 400 ms 左右，因此这样的呈现时间与同时呈现两个语素的加工可能产生显著不同。建议作者讨论研究中所采用的呈现方法对结果可能产生的影响。

回应：感谢编委专家的建议。此次修改，我们详细讨论了研究中所采用的呈现方法对于结果可能产生的影响(详见第 16 页的讨论部分)，指出连续呈现刺激方式在探究汉语复合词内部语素意义与整词语义整合加工中的优势，同时说明了连续呈现刺激方式相比于同时呈现方式，可能会加大了语素分解再整合加工与直接整词加工这两条加工通路之间的竞争，最后我们也承认，实验设置 800 ms 的刺激呈现时间使得被试对首字和尾字都进行了较为充分的加工，在 800 ms 之内是否有更复杂的加工过程，以及对于这两种通路加工机制及其神经基础还需更多研究来进行细致和深入地探究。具体表述如下：

“本研究在连续呈现刺激方式下深入探讨了汉语复合词内部语素语义与整词语义整合加工。连续呈现刺激方式在短语语义整合加工(Bemis & Pykkänen, 2011; Zhang & Pykkänen, 2015; Ziegler & Pykkänen, 2016)研究中得到了应用，能够很好探究语义整合加工的大脑机制。以往研究者通过对比启动词与目标词之间共享语素条件与无关条件产生的语素启动效应，考察了汉语语素在复合词加工中的激活变化(Wu et al., 2020; 赵思敏 等, 2017)。尽管结果发现在不同时间窗内语素影响了整词语义通达，但是对于语素分解及整合过程并未得到细致地分离，尤其是语素如何参与并影响复合词语义整合过程。本研究将连续呈现刺激方式与

ERP 技术相结合, 结果分别在首、尾字的加工过程中发现语素效应和语义透明度效应, 说明了汉语复合词的语素整合存在着语素分解、语素义通达以及语素义整合等复杂的加工过程。值得注意的是, 相比于同时呈现, 连续呈现方式可能加大了语素表征与整词表征加工之间的竞争。同时, 需要承认一点, 本研究的刺激呈现时间为 800 ms, 首字和尾字都得到了充分加工, 这不仅使得语素意义表征得到了充分激活, 同时可能扩散激活更多的语义信息, 进而影响到整合加工。未来的研究可以通过对比不同的刺激呈现方式或者结合词汇歧义性等语言学属性来进一步展开研究, 对复合词加工过程背后的神经机制进行更加细致和深入地探讨。”

参考文献:

- Zhang, L., & Pykkänen, L. (2015). The interplay of composition and concept specificity in the left anterior temporal lobe: An MEG study. *Neuroimage*, 111, 228–240.
- Ziegler, J., & Pykkänen, L. (2016). Scalar adjectives and the temporal unfolding of semantic composition: An MEG investigation. *Neuropsychologia*, 89, 161–171.
-

第四轮 主编终审

主编终审意见: 该文经专家多轮审理, 作者对专家提出的意见进行了详尽的回复, 基本上回答了审稿专家和编委的问题。应该说, 我仍然觉得这个实验设计生态效度偏低, 在加工第一个字时, 读者可能是当单字词来处理的, 因此理论上第一个字的控制应该更为严格, 这个设计才能成立。同时, 非常慢速的呈现和特殊的任务也可能是解释当前结果的重要原因。由于整个实验缺乏生态效度, 所获得的结果是否可以说明真实的多字词加工问题值得商榷。但鉴于作者已经在讨论部分对此不足进行了讨论, 同时编委和审稿专家也同意发表, 因此我也同意作出同意发表的结论。