

《心理学报》审稿意见与作者回应

题目：中文词类信息在副中央凹中的加工

作者：鹿子佳; 符颖; 张慢慢; 臧传丽; 白学军

第一轮

审稿人 1 意见：

意见 1：论文自检报告“1.创新性贡献”中，“这为将来补充和完善眼动控制模型相关假设”一句很空洞，不清楚哪个模型的假设有此需要。

回应：非常感谢外审专家的宝贵意见。当前阅读中眼动控制的两大模型(序列注意转移模型和平行分级加工模型)都是基于拼音文字的特点而建立的，并且它们对于高水平信息(语义、句法等)的加工方式和进程仍在完善阶段，并不能很好地解释副中央凹高水平词汇信息的加工。因此，作者认为模型或许可以从灵活性和跨语言的普适性的角度进行完善和发展。作者对此部分进行了重新组织，修改后的内容如下：

“（1）本研究首次利用边界范式探讨中文词类信息在副中央凹的加工情况，此前仅在拼音文字中进行过此项研究，并不清楚中文读者能否在副中央凹提取到词类信息。

（2）这不仅对于中文读者副中央凹加工深度给出了实验证据，而且对基于拼音文字建立的两类眼动控制模型(序列注意转移模型和平行分级加工模型)的灵活性和跨语言的普适性提出了挑战。

（3）作者认为可以从字形与词类关联程度的角度解释不同语言间，词类在副中央凹加工的不同结果，以及建议未来模型的发展可以考虑从此角度。”

意见 2：论文自检报告“5.计划与实际的样本量”中，“这个计算方向反了。问的不是你用 120 人的检验力。问你根据效应量等计算的样本量。计划是多少，最后是多少？您用 120 人一定是足够的。但是有一点浪费吧！所以，被试不足是问题，浪费也是问题。”

回应：非常感谢外审专家的宝贵意见。这里没有表达清楚，给专家阅读造成不便，我们表示歉意。根据效应量($d = 0.43$)计算的样本量是 120 人，实际收到的有效数据也是 120 人。另外，有关被试人数过多的问题，首先是因为实验材料数量不够多(共 45 句实验句，单因素三水平的实验设计下，每种条件 15 句)；其次是在目标词上的跳读率在 50% 以上，也就是注视时间的数据只有不到 50%，但目前眼动的 power 计算时并没有考虑进跳读的情况。因此，本实验的人数才会多了一些。已对表述进行修改。

意见 3：摘要中，“加工词类信息，就一定在眼动指标上有差异？如果加工了，对理解没有影响，不可以吗？请作者思考！”

回应：非常感谢外审专家的宝贵意见。读者在注视中央凹词汇的时候可以从副中央凹(通常所说的余光)获得部分信息，但是受到视敏度等的影响，加工程度不如中央凹单词。我们通过

边界范式操纵副中央凹词的词类信息，如果读者能在副中央凹加工到词类信息，那么会反应在某些眼动指标上(早期或后期，注视时间或注视概率)；如果不能产生任何影响，这说明在副中央凹短暂的加工时间中，词类信息并没有起到作用，它可能发生于加工的后期阶段。我们在讨论4.1中补充了词类信息加工进程的内容，来自中央凹加工和ERPs的证据发现，中文词类加工的时间进程不早于语义。

意见 4：摘要中，“词类信息不影响的结果就支持了这个假设吗？”

回应：非常感谢外审专家的宝贵意见。这里没有表达清楚，抱歉给专家阅读造成不便。这里的“假设”指的是模型中关于高水平信息加工的观点，是指模型的对阅读行为的假设。已删除了此句。

意见 5：实验方法“2.1 被试”中，性别差异是否会产生额外的影响？

回应：非常感谢外审专家的宝贵意见。并未有研究发现，性别差异会显著影响中文读者阅读时的注视行为，以往的研究也有男女生数量不平衡的情况。例如，臧传丽等人(2013)考察副中央凹的信息量对朗读和默读的调节作用时，45名被试中男生有4人；张慢慢等人(2020)探讨快读与慢速读者的中央凹加工对副中央凹预视的影响时，快速与慢速组的被试均为28人，其中男生分别为8人和6人。

参考文献：

臧传丽，张慢慢，岳音其，白学军，闫国利. (2013). 副中央凹信息量对中文朗读和默读的调节作用. *心理与行为研究*, 11(4), 444–450.

张慢慢，臧传丽，徐宇峰，白学军，闫国利. (2020). 快速与慢速读者的中央凹加工对副中央凹预视的影响. *心理学报*, 52(8), 933–945.

意见 6：实验方法“2.3 实验材料”中，目标词和预视词的描述较为混淆。

回应：非常感谢外审专家的宝贵意见。已经对其进行了修改，修改后内容如下，文中用红色字体标注。另外，对目标词和预视词的语音和语义相关性进行了评定和补充说明，同见2.3。

“例子中，条件1为一致预视条件，目标词和预视词都是“虾”；条件2为词类不违背预视，预视词“狼”与前半句(今天我准备炸些……)无法构成语义合理的连贯句，但它的句法类型是合理的(根据前半句推测即将出现的词为名词，而“狼”是名词)；条件3为词类违背预视，预视词“割”与前半句无法构成语义合理的连贯句，且句法类型也不合理(根据前半句推测即将出现的词为名词，而“割”是动词)。”

意见 7：实验结果“3.1”之前最后一段中，分析数据时，如果最大模型不能拟合，随机斜率的取舍顺序和标准是怎样的？

回应：非常感谢外审专家的宝贵意见。已经对其进行了补充，修改后内容如下，文中用红色字体标注。

“首先移除项目的相关，然后移除项目的斜率，如果还不能拟合成功，再按顺序移除被试的相关和斜率，直至拟合成功。”

意见 8：讨论部分“4.1 中文词类信息的加工”中，“中文曾有研究支持这一结论”表述有误。

回应：非常感谢外审专家的宝贵意见。已删除此不合适的表述，重新组织了语言，文中用红色字体标注。

意见 9：结论中，“读者无法从副中央凹提取到句法类型信息(词类信息)”本研究证明了读者并不能在副中央凹提取单字名词的词类信息，但对于单字动词呢？这一结论能否直接推论到全部词类信息还有待考虑。另外，“为各眼动控制模型日后的修订和扩展提供了正字法会促进关联密切的高水平词汇信息加工的新视角”这也仅是作者的推断，实验并未操作正字法信息，因此这一结论需谨慎对待。

回应：非常感谢外审专家的宝贵意见。对于第一个问题，作者也认为将结论归纳为“读者无法从副中央凹提取到句法类型信息(词类信息)”过于笼统和草率，修改为“读者无法从副中央凹提取到单字名词和动词的词类信息”。本研究采用的目标词均为名词，句法违背条件下的预视词均为动词，准确的说是无法意识到本该出现一个名词的位置上出现了一个动词的冲突。预测这种冲突在动、名词中会大于其余词类不一致的情况，例如形容词-名词的冲突，这值得进一步研究。对于第二个问题，作者非常认可专家的建议，作者的推测放在结论中并不合适，所以将其删除。

.....

审稿人 2 意见：

意见 1：该论文采用边界范式探讨中文读者是否能够在注视点右边的副中央凹加工词类信息。实验使用单字名词作为目标词，匹配了三种预视条件：一致预视、词类不违背预视、词类违背预视。结果发现词类不违背和违背两种条件下没有显著差异，因此，作者认为中文读者并不能在副中央凹上获取词类信息，该结果支持了 E-Z 读者模型。文章最大的问题是，数据并不如作者所说的那样，词类不违背和词类违背两种条件下没有显著差异。实际上，这两个条件在首次注视时间和回视路径时间上有显著差异，虽然凝视时间上的差异不显著，但是两条件之间的差异有 16 毫秒，且差异相检验 t 值为 1.65，趋向边缘显著。如果作者用贝叶斯分析来说明这两个指标上的显著差异实质上是不显著的，那应该提供文中所有显著差异的贝叶斯分析。而实际上，贝叶斯因子的分析要大于 3 才有意义（非作者说的大于 1）。

回应：非常感谢外审专家的宝贵意见。

首先，对于目标词上 FFD 差异显著和 GD 边缘显著的结果，作者进行了重新分析。在此次分析中，我们删除了从目标词上发生回视的试次。这是因为考虑到，如果读者眼跳到目标

词后立即产生了回视,那么上一次注视时对副中央凹单词的加工可能不充分,不能很好地反映对目标词的预视加工。根据 E-Z 读者模型,当眼跳进入到不可变阶段,即使对当前加工的单词出现了困难,也无法取消这次眼跳,只能按计划完成这次眼跳再进行回视进行重新加工。而对中央凹存在加工困难的情况下,注意资源就会较晚、较少地转移到副中央凹单词上,这就导致读者较少甚至没有对副中央凹单词进行正常的加工,因此应排除这种情况。另外,除了在文中报告的指标之外,作者也分析了 SFD、TFD,结果条件 23 间都为无显著差异(SFD: $b = -0.03$, $SE = 0.02$, $t = -1.45$, $p = 0.147$, 95% CI = $[-0.07, 0.01]$; TFD: $b = -0.03$, $SE = 0.02$, $t = -1.32$, $p = 0.187$, 95% CI = $[-0.07, 0.01]$)。本实验被试量为 120 人,统计检验力高于 0.8,因此作者认为条件 23 间差异不显著的结果较为可靠。重新分析发现,相比于词类不违背预视(条件 2),词类违背预视(条件 3)时,回视入概率在数值上更低(边缘显著: $b = -0.28$, $SE = 0.16$, $z = -1.77$, $p = 0.077$, 95% CI = $[-0.59, 0.03]$),在其余眼动指标上差异均不显著($|t/s| < 0.90$, $ps > 0.372$; $|z/s| < 0.85$, $ps > 0.396$),贝叶斯因子在所有指标上均大于 10,有较强证据支持了虚无假设。

其次,作者补充了贝叶斯因子结果的解释:“贝叶斯因子 1~3 有较弱的证据支持虚无假设,大于 3 有中等强度证据支持虚无假设,大于 10 有较强证据支持虚无假设;0.33~1 有较弱的证据支持备择假设,小于 0.33 有中等强度证据支持备择假设,小于 0.1 有较强证据支持备择假设。”(胡传鹏 等,2018)

最后,作者补充了对三个兴趣区(目标前字、目标词、目标后字)的所有指标的贝叶斯因子结果。详见结果部分 3.2~3.6,修改的内容用红色字体标出。

参考文献:

胡传鹏,孔祥祯, Wagenmakers, E. -J., Ly, A., 彭凯平. (2018). 贝叶斯因子及其在 JASP 中的实现. *心理科学进展*, 26(6), 951-965.

意见 2: 预视控制带来的效应还溢出到目标后区上,后区一致和词类不违背条件上的差异超过 130 毫秒,且词类不违背和词类违背条件之间也有显著差异, $t=2.11$ 。这跟先前关于预视加工的研究结果也不一样,一般情况下,预视控制带来的影响不会延续到后区,即使有效应也不会如此强。但是作者并没有解释这个效应。实际上,这些数据对验证本文的数据是否可靠有重要参考意义。

回应: 非常感谢外审专家的宝贵意见。

首先,再次感谢专家对讨论部分提出的宝贵意见,引发了作者更深入的思考,根据专家的意见,作者对讨论中结果解释的部分进行了整体的修改,在讨论 4.1 中以红色字体标出。具体解释如下:

“一致预视效应溢出到目标后字上且效应较强的原因可能有三:第一,溢出效应的观点(Kliegl et al., 2006; Pollatsek et al., 2008),即是由对目标词的不完全加工引起,导致在目标后词上继续之前未完成的加工过程,体现了对目标词后期的整合困难;第二,本研究采用的目标词为单字词,且词频较高,三种条件

下跳读率均超过了 50%，向目标词的眼跳有可能错误落到了目标后字上，溢出的一致预视效应由目标词的加工差异引起；第三，可能是由预视代价(preview cost)导致的，有研究发现预视效应包括预视效益和预视代价(Vasilev et al., 2020)，当边界变化，与目标词有某些相关属性的预视词会促进对目标词的加工，但是不合理的预视词也会阻碍加工，且越不合理的预视造成的加工代价越大，本实验中条件 2 下的预视词与目标词存在语义上的不合理，这会带来加工代价，可能导致注视时间的增长和溢出效应的变大。

另外，词类违背和词类不违背条件在目标后字上晚期指标上表现出一定的差异，但回视路径时间和回视出比率的贝叶斯因子分别为 3.52、1.76，至少较弱证据支持了虚无假设。在前人对于副中央凹的研究中较少有人关注晚期指标，作者分析这一指标的目的在于初步考察对副中央凹词汇的操纵是否会对后期整合产生影响，未来还需要更多的证据进一步来探讨，研究结果仍以目标词上的前期指标为主。”

参考文献:

- Angele, B., Slattery, T. J., & Rayner, K. (2016). Two stages of parafoveal processing during reading: Evidence from a display change detection task. *Psychonomic Bulletin & Review*, 23(4), 1241–1249.
- Kliegl, R., Hohenstein, S., Yan, M., & McDonald, S. A. (2013). How preview space/time translates into preview cost/benefit for fixation durations during reading. *Quarterly Journal of Experimental Psychology*, 66(3), 581–600.
- Vasilev, M. R., Yates, M., Prueitt, E., & Slattery, T. J. (2020). Parafoveal degradation during reading reduces preview costs only when it is not perceptually distinct. *The Quarterly Journal of Experimental Psychology*, 74(2), 254–276.
- White, S. J., Rayner, K., & Liversedge, S. P. (2005). Eye movements and the modulation of parafoveal processing by foveal processing difficulty: A reexamination. *Psychonomic Bulletin & Review*, 12(5), 891–896.

意见 3：目标词上一致预视和其他两种不同预视的条件差异太大，如比较一致和词类不违背两种条件下的差异，首次注视时间上为 62 毫秒，凝视时间上超过 80 毫秒，回视路径时间上的差异超 120 毫秒。这是否是真的效应，还是因为边界变化带来的差异，作者需要检查删除边界变化发生在注视点中的试次。

回应：非常感谢外审专家的宝贵意见。我们重新检查了数据的剔除流程，确保在剔除数据的过程中没有出现错误。我们删除了由于系统误差(例如电脑显示器没有及时刷新)导致边界提前变化(眼睛还没越过边界时预视词就变为了目标词)和边界变化延迟(眼睛已经越过了边界但是图片没有及时变化)的试次，共剔除大约 13.3% 的数据。

其次，本实验中一致预视效应是比较一致预视条件和语义不合理预视词条件，在操纵了预视词和目标词语音、语义都无关之后，后者可看做是一个无关预视。有关一致预视和无关预视的预视效应的大小，在许多研究中数值都比较大。例如，Yan 等人(2009)在首次证明中文语义副中央凹预视效应的研究中，一致预视和无关预视的差异大小在 FFD 上约为 40ms，GD 上约为 80ms；Yan 等人(2012)在探讨中文词汇和亚词汇副中央凹语义预视效应的研究中，一致预视和无关预视的差异大小在 FFD 上为 41ms，GD 上为 77ms。这可能也是由预视代价导

致的。前面回答意见 2 的时候有提到，本实验中条件 2 下的预视词与目标词存在语义上的不合理，这会带来加工代价，导致注视时间的增长。对于晚期指标 GO-PAST，作者分析这一指标的目的是想看，对副中央凹词汇的操纵是否会对后期整合产生影响。结果发现了 130ms 左右的一致预视差异，这说明预视代价同样影响后期整合阶段，越不合理的预视词可能会导致更多的回视。当前副中央凹预视的研究较少关注这一指标，未来需要进一步的研究。

再次，我们在实验后要求被试完成一份问卷，询问的问题包括：被试是否意识到了有字的变化；这样的情况所占的百分比及大约的数量；尝试回忆并写下变化的字。由于被试会反问实验句的总数，所以我们以被试意识到字的变化百分比为准。将被试估计的边界变化的百分比作为协变量在线性混合模型中进行分析，结果发现，被试估计边界变化的百分比平均为 20% (range = 0 ~ 80%, $SD = 17.5\%$)，但与一致预视效应没有交互作用(Mirault et al., 2020)。因此，我们在主要分析时没有加入这一变量。

参考文献:

- Mirault, J., Yeaton, J., F Broqua, Dufau, S., & Grainger, J. (2020). Parafoveal-on-foveal repetition effects in sentence reading: a co-registered eye-tracking and electroencephalogram study. *Psychophysiology*. <https://doi.org/10.1111/psyp.13553>
- Yan, M., Richter, E.M., Shu, H., & Kliegl, R. (2009). Chinese readers extract semantic information from parafoveal words during reading. *Psychonomic Bulletin & Review*, 16(3), 561–566.
- Yan, M., Zhou, W., Shu, H., & Kliegl, R. (2012). Lexical and sub-lexical semantic preview benefits in Chinese reading. *Journal of Experimental Psychology Learning Memory & Cognition*, 38(4), 1069–1075.

意见 4: 正如作者前言所引用，中英文都有不少研究表明读者能够从副中央凹获取语义信息，甚至文本的合理性信息，这该如何解释，读者不能在副中央凹上分辨词类信息？中文的语义整合和词类加工是什么关系？

回应: 非常感谢外审专家的宝贵意见。根据您给出的建议，作者在讨论 4.1 中分别对“为什么中文读者无法在副中央凹提取到词类信息，而英文读者可以；为什么英文读者无法在副中央凹提取到语义信息，而中文读者可以”和“中文的语义整合词类加工的关系”两个问题进行了详细的解释。在 4.1 中用红色字体标出。

简而言之，中文字形与词类联系不强，因此中文读者无法从副中央凹获得词类信息，而英文词形与词类联系较强，因此英文读者可以从副中央凹获得词类信息；两类语言系统字形与语义的联系与词性恰好相反，因此效应也相反。另外，中文语义和句法加工的关系，有来自自动和 ERP 的证据证明，对中文来说，不论是加工时间上还是影响力上，词类信息的加工并不优先于语义加工。

意见 5: 在写作上，前言部分条理较为清楚，但是讨论上，难以把该结果和支持 E-Z 眼动模型联系起来，作者是支持中文读者能在副中央凹上获取高水平信息，还是不支持，这里就论

述得不清楚。

回应：非常感谢外审专家的宝贵意见。根据您的意见，作者在讨论 4.2 中对本研究结果及其对 E-Z 读者模型的支持进行了更详细的解释。对于中文高水平信息在副中央凹的加工情况，并不能一概而论，语义可以被加工，但是词类信息不能。这也对眼动控制模型的灵活性和跨语言的普适性上提出了挑战，也在 4.2 中进行了解释。修改内容在文中用红色字体标出。

意见 6：作者引用 E-Z 眼动模型也能解释副中央凹上的高水平信息加工，因此该结果并不能对模型进行验证。

回应：非常感谢外审专家的宝贵意见。这里没有表达清楚，给专家阅读造成不便，我们表示歉意。作者在此处引用 E-Z 读者模型的“强迫注视”的观点，目的是为了说明，“强迫注视”的观点并不能作为 E-Z 读者模型兼容高水平信息在副中央凹被加工的合理解释，仍然倾向支持否定的观点。前言中提到：“Schotter 等人(2014)利用计算机模拟 E-Z 读者模型阅读文本时发现，仅有 8%的试次中将副中央凹单词加工到了 L_2 阶段，说明高水平信息在副中央凹提取到的情况非常少。”作者把这句话从前言改到了讨论中，使逻辑更加流畅、表达更加简明。我们对讨论中模型的解释和启示进行了重写，详见 4.2 中红色标注。

.....

审稿人 3 意见：

本研究采用眼动技术考察了副中央凹加工的词类效应。研究问题较深入，研究方法适当。然而由于目前论文存在以下问题，建议作者修改。

意见 1：研究的理论意义不清楚。尽管作者引用了副中央凹加工的两个注意模型，但是作者研究的词类加工与两个模型的关系并没有表述清楚。在对两个模型的介绍中，也完全没有介绍关于词类加工的任何假设。

回应：非常感谢外审专家的宝贵意见。根据专家的意见，作者在前言和讨论 4.2 中，分别对模型的部分进行了修改，增加了模型对词类加工的相关假设。

简而言之，词类信息是句法信息的其中一种。E-Z 读者模型认为句法信息在后期整合阶段(I)进行，而在副中央凹加工的短暂时间内，时间不允许加工到句法的水平，因此 E-Z 读者模型不支持高水平信息在副中央凹中的加工。而 SWIFT 模型认为，高水平信息的加工不受时间的限制，只与激活程度有关。如果当中央凹单词加工难度较低，副中央凹获得的注意资源就越多，此时对其加工的程度也越深，可以加工到句法的程度。因此，SWIFT 模型支持高水平信息在副中央凹中得以加工和整合的观点。

具体内容请见前言和讨论 4.2 中模型介绍的部分，在文中以红色字体标出。

意见 2：关于词汇的控制还需要再补充一些信息。比如，作者质疑以往类似研究对词汇的控制，那么作者应该强调本研究在词汇控制上的改进。从作者提供的例子来看，词汇在声调上

并没有很好的控制。

回应：非常感谢外审专家的宝贵意见。当前拼音文字中有研究者进行过词类在副中央凹提取的研究，但还没有在中文中探讨过这一问题。我们的研究作为初步探索，匹配了目标词和预视词的词频、复杂性等已被证明会对注视时间产生较大影响的变量。当前对于副中央凹的研究有关于语音的研究，但还没有深入到探讨声调在副中央凹的变化能否被读者意识到，这是未来的研究方向。根据专家给出的建议，我们额外评定了目标词和预视词词对的语音和语义相关性，在无法消除它们差异的情况下，确保条件 23 中的预视词与目标词在语音和语义上完全显著不相关，保证这两个变量造成的影响大致相当。补充内容在 2.3 实验材料的第一段中，以红色字体标出。

意见 3：在对实验材料的评定上也需要进行补充。由于研究考察的是读者对词类违背的即时反应，那么对词汇进行整句的通顺性评定显然是不适合的。作者需要评定词汇与其之前句子片段的通顺性，比如“今天我准备炸些虾”。

回应：非常感谢外审专家的宝贵意见。根据专家给出的建议，我们评定了目标词(或预视词)与之前句子片段的通顺性，对三种条件下的句子片段都进行了评定。补充内容在 2.3 实验材料的最后一段中，以红色字体标出。

意见 4：作者只对目标词上的词类效应使用贝叶斯分析方法检测，但是没有对后目标词上的词类效应再使用贝叶斯分析。由于目标词和后目标词上使用了不同的分析方法，作者对其产生的结果进行比较和推论都是不适当的。

回应：非常感谢外审专家的宝贵意见。作者补充了对三个兴趣区(目标前字、目标词、目标后字)的所有指标的贝叶斯因子结果及对其的解释。详见结果部分 3.2~3.6，修改的内容用红色字体标出。另外，在讨论 4.1 中对结果进行了进一步的解释。

第二轮

审稿人 3 意见：

作者对评审意见做了比较合适的回答和修改，满意。

编委意见：

这篇稿件经过作者的修改，在理论创新、实验设计、结果呈现等方面质量均有所提升，基本解决了评审专家提出的问题，达到学报发表的水平。所以，建议接受发表。

主编意见：同意责编的判断，建议发表。