

《心理学报》审稿意见与作者回应

题目：中文词间词和词内词预视加工的差异：词间阴影的作用

作者：关宜韞 宋悉妮 郑玉玮 张颖靓 崔磊

第一轮

审稿人 1 意见：《词边界信息对中文词间和词内预视加工的影响》一文研究视角较为新颖，对文献把握较为充分，在数据分析和理论解释方面较为翔实。但是，该文对词内和词间预视效应的研究未能完全排除其他因素的干扰；同时，在论据论述以及问题提出方面还存在不足；另外，该文数据分析的理据不充分。建议退稿。

意见 1：该文行文不通顺句子较多。如：在Yang, Wang, Xu和Rayer (2009)的研究在凝视时间上发现了68ms的预视效应。Vasilev和Angele(2017)通过元分析中也再次证实中文的预视效应在凝视时间上比拼音文字的预视效应大约大了10ms。如：这两类复合词都属于一个语言单元，从而排除了词语的词素属性对预视加工的影响，结果发现无空格复合词的预视效应显著地大于有空格复合词的预视效应。从而验证了Hyönä等的假设，说明空格导致读者的注视点距离预视词更远，从而导致词间词的预视效应小于词内词的预视效应。一句话的意义还没有完结，使用句号；而且，不断使用同一个连接词“从而”。再如：本研究采用了120对目标词，词内词，即双字词的尾字是研究的目标字，而对于词间词，双字词的首字是研究的目标字，控制词内词的尾字与词间词的首字相同。阅读中几重意思重叠，读者需要详细的分析，才能厘清楚逻辑关系。“母语为汉语，无阅读障碍实验结束后，被试可以获得一份小礼品作为实验报酬。”“不再”应该是“不在”。

回应：谢谢审稿人的细心阅读和指正，原文确实存在句子表述方面待完善的细节问题，除了审稿人提出的这一系列表述和细节问题，我们另已对全文内容进行了多次挑剔性的阅读和内容校正，并用红色显示，敬请审稿人审阅。

“Vasilev 和 Angele(2017)通过元分析中也再次证实中文的预视效应在凝视时间上比拼音文字的预视效应大约大了 10ms。”修改为“Vasilev 和 Angele(2017)通过元分析也再次证实中文的预视效应在凝视时间上比拼音文字的预视效应大约大了 10ms。”

“结果发现无空格复合词的预视效应显著地大于有空格复合词的预视效应。从而验证了Hyönä等的假设，”已经将句号改为逗号，并且将“从而”删除，具体修改如下“结果发现无空格复合词的预视效应显著地大于有空格复合词的预视效应，验证了Hyönä等(2004)的假设，”

“本研究采用了 120 对目标词，词内词，即双字词的尾字是研究的目标字，而对于词间词，双字词的首字是研究的目标字，控制词内词的尾字与词间词的首字相同。”修改为“本研究采用了 120 对词间词和词内词，如前所述，词间词是指注视词和预视词属于不同的词；而词内词是指注视词和预视词属于同一个复合词，所以本研究控制每对词间词（“纸盒”）的首字和词内词（“稿纸”）的尾字相同，即本研究的预视目标字。”

“母语为汉语，无阅读障碍实验结束后，被试可以获得一份小礼品作为实验报酬。修改为“母语为汉语，无阅读障碍。实验结束后，被试可以获得一份小礼品作为实验报酬。”“不再”应该是“不在”。抱歉，这两个词语在全文均没有找到。

意见 2：文中第五段引入“P-O-F”效应，介绍单纯词存在此类效应，此类效应与预视效应同一？同时，复合词和短语未能出现此类效应，那么是否表明复合词没有出现预视效应。作者没有

交待。另外，作者指出，CUI等认为单纯词的首字对尾字预期性高，那么，作者所使用的材料（复合词）出现预示效应，是否和CUI等研究结果不同。而且，作者是否有效评估本文所使用材料（依据作者提供的案例，应属偏正结构），首字对尾字是否存在不同的预测，或者说偏正结构的材料，首字和尾字加工的时间是否相同。这是否是造成研究结果的一个重要干扰因素。作者在讨论中指出，ZANG等的研究表明读者会依据首字的位置频率来调节预视效应的大小，是否是本研究应该控制首字、尾字位置概率的佐证。另外，CUI等的研究发现45%的短语被认定为复合词，作者指出本研究要对词边界进行严格清晰的界定，这一界定是指？作者仅仅通过隐含边界区分了词内词和词间词，并未对复合词的属性进行严格的界定？

回应：首先，“P-O-F”效应与“预视效应”不同。很多实验研究均证实，在阅读过程中，读者除了对当前词（焦点信息）进行注视加工外，还会对右侧的词（预视信息）进行一定的预视加工，预视加工获得的信息越多，阅读效率就会越高，我们称这种预视加工可以获知信息的现象为“预视效应”（Rayner, 1998）。P-O-F效应是指预视加工获得的信息对当前注视加工产生的即时影响。两种效应反应的加工水平和加工阶段均不同，但是两者都属于预视加工的范畴。因此，Cui等(2013)的研究中发现,复合词和短语没有出现P-O-F效应，但是存在预视效应，恰与本研究的结果一致，两个研究结果可以相互印证。

其次,关于审稿人提出实验材料中的首字和尾字的加工时间是否相同。本研究发现词内词(双字词的尾字)比词间词(双字词的首字)加工时间长,但是本研究主要探讨的词间词和词内词的预视效应的差异,即词间不一致预视——一致预视 vs 词内不一致预视——一致预视,因此词间一致预视和词内一致预视加工时间的差并不会影响我们的实验结果。至于审稿人提出的偏正结构的材料,首字和尾字加工的时间是否相同。在本研究中,虽然材料举例中的目标词是偏正结构,但是并不是所有目标词均是偏正结构,还有其他结构的词语,如并列结构。但是我们觉得审稿人提出的这个问题本身确实非常重要,通过查阅相关文献,我们发现对语法结构的加工特点和加工时程进行研究的文章很少,请参阅如下文献。

Cui, L., Cong, F., Wang, J., Zhang, W., Zheng, Y., & Hyönä J. (2018). Effects of grammatical structure of compound words on word recognition in Chinese. *Frontiers in Psychology-Language Sciences*.9:258.

Jia, X., Wang, S., Zhang, B., & Zhang, J.X. (2013). Electrophysiological evidence for relation information activation in Chinese compound word comprehension. *Neuropsychologia*. 51(7), 1296–1301.

Liu, P., and McBride-Chang, C. (2010). Morphological processing of Chinese compounds from a grammatical view. *Appl. Psycholinguist*. 31, 605-617.

再次，审稿人关于位置频率的意见非常正确，诚然，一个字很可能就携带着其作为一个单字词或者双字词的信息，因而导致加工不同，这已被Zang等人的研究证实，当目标字作为双字词首字的概率比目标字作为单字词的更大时，会产生更大的预视效应。但是与Zang等人的研究不同，本研究操控的变量是，该字作为双字词的首字和尾字，如若想控制掉位置频率的影响，则最好的控制是该字在首字和尾字出现的概率均为50%，虽然难度较大，但是我们尽力控制。在本研究中，目标字作为双字词首字的概率的平均值是39%；目标字作为双字词尾字出现的概率为42%，两者差异不显著， $t(119) = .84, p > .05$ 。这说明目标字作为双字词的首字和尾字的位置频率不会对研究结果产生差异性的影响。该内容已加入到研究方法的实验设计与材料部分并用红色标注。

最后，多谢审稿人的认真审阅，关于句子词切分评定相关过程，已经加在实验设计与材料部分。24名不参加正式实验的某大学学生参与了评定。评定时发给每位学生一份卷子，里面包含了120个实验句子。要求被试仔细阅读句子，然后用竖线对句子中的词语进行切分(Yan

et al., 2010)。结果发现，被试对词间词和词内词的边界进行切分的一致性分别为85.9%和87.2%，符合本研究的实验要求，该内容已加入到句子词切分评定部分并用红色标注。

意见3：在文献梳理方面，作者大量介绍印欧语言和中文在预视效应方面的差异，认为中文的预视效应更大，在凝视时间上有68ms的预视效果。但是，作者在前文中（第三段）介绍，Hyönä等发现，读者能够获得80ms的预视效应？那么，Hyönä等的研究是一个特例？而作者要证实的假设均基于Hyönä等的研究。

回应：正如审稿人所说，Hyönä等发现，读者能够获得80ms的预视效应，这并不是一个特例，在后续研究中已被广泛证实，究其原因就是Hyönä等将边界放在了词内，但是究竟是因为空格因素造成的还是词语是否属于同一语义单元造成的尚不清楚的，这也是本研究的研究目的。但是为了避免审稿人提出的该理解上可能存在的混淆，我们将该句子修改为“研究者发现，英文词间词的预视效应大约是30~50ms(Rayner, 2009)”。

意见4：作者主要想证实词内和词间预视效应的不同，并且在文献中已经指出Rayner等的研究证实大写字母吸引了被试的注意，增加了预视效应。但是，作者依旧不在设计中避免这一干扰因素，而使用黄色阴影来增强预视效应？“在本研究中我们采用黄色阴影来标记边界信息，以突出边界信息，从而最大化地增加边界信息对预视加工可能产生的影响。”那么，实际的情况是否是黄色阴影促进了预视效应的产生？依据文章的结果，黄色阴影未能增强预视效果，是否说明这一假设错误？而且，作者指出“同时该方法不会对不同类型词语的词边界掩盖产生不同的影响”，即阴影切分不会对边界掩盖产生不同的影响？既然Cui等的研究已经证实这一加工特征，那么，作者操作这一因素的意义何在？

回应：是的，如审稿人所说，我们采用黄色阴影来标记边界信息是为了最大化地增加边界信息对预视加工可能产生的影响，可是，即便如此，本研究仍发现，词间阴影并未对预视加工产生影响，由此得出研究结论，词间阴影产生的词边界信息并不能促进预视加工。在刚刚发表的一篇文章上，Zhou等(2018)采用分辨率最高的红色和绿色作为词边界信息交替呈现，以增加词边界信息的区分度，结果同样发现词边界信息不会影响注视时间。与本研究结果一致。但是，Cui等人(2014)发现用空格来标记词边界可以促进单字词的预视加工，但是对双字词的预视加工无影响，所以研究者推测，这一结果可能是由于空格给单字词和双字词带来了不同的词边界掩盖效应导致的。对于单字词，加入空格边界信息后，单字词的左右两边均无词边界掩盖；但是对于双字词，即便加入空格，双字词的两个字仍是临近出现的，对于目标字（即双字词的首字），左边是空格，无边界掩盖，可是右边仍是一个字，仍然存在边界掩盖。所以本研究证实了Cui等人(2014)的研究假设，在本研究中采用阴影作为词边界信息，在提供边界信息的同时，不会改变目标词的边界掩盖，结果发现，词边界信息不会影响目标词的预视加工。综合分析本研究 and Cui等人(2014)的研究结果，我们认为，词边界的标记方式不同会导致对预视加工产生不同的影响。

Zhou, W., Wang, A., Shu, H., Kliegl, R., & Yan, M. (2018). Word segmentation by alternating colors facilitates eye guidance in chinese reading. *Memory & Cognition*, 6, 1-12.

意见5：作者在材料部分指出，控制词内词和词间词的词频。词频的来源是什么？同时，作者在第二部分数据分析时对整词进行分析，那么词内词和词间词的笔画数是否应该匹配。

回应：谢谢审稿人的仔细审阅。本研究中的词频采用的是人民日报版，并且控制词间和词内的词频差异不显著。本研究也匹配了词间和词内的笔画数，词内词的平均笔画数为16.25；词间词的平均笔画数为16.68，两者的差异并不显著。相关内容已加入到研究方法的实验设计与材料部分，并且用红色标注（P8表1）。

意见6: 作者在实验材料的评定部分指出,除了实验句子外,作者还加入了36个语义不合理的句子。这些不合理的句子如何构建?同时,在实验程序部分,作者指出,每名被试阅读156个句子,120个实验句子,30个填充句子,6个练习句子,练习句子的来源?依据作者的介绍推测,6个句子均为不合理句子?

回应: 在实验材料的合理性评定中,不合理的句子构建是通过用不符合句子语义的一个真字替换了合理句子中的一个字。在实验中,练习句子的结构和字数设置与目标句子相同,均为实验者编制,被试无法区分实验句和填充句。6个练习句子与填充句相同,为合理句子,用于在实验开始之前让被试熟悉实验程序。

意见7: 作者在结果与分析部分介绍,剔除4名看到边界变化在25%以上的被试,作者没有介绍如何确定这些被试看到变化的比例。

回应: 根据Eyelink眼动仪的Dateviewer数据分析软件,从里面可查到边界变化的时间,以及在那时读者的注视位置,如边界在读者注视目标词之后发生,即被试看到了边界变化,则该项目数据需要删除。看到边界变化的句子数量除以总的句子数量,如果超过25%,则该被试的所有数据删除。句子修改为:“剔除4名看到边界变化超过总体句子数量25%以上的被试”并用红色标注。

意见8: 作者指出“实验操作的隐含边界位于词内词的两词素之间,而位于词间词的两词素之前,因此我们将词内词的尾字和词间词的整词作为预视加工的兴趣区进行比较”,如此比较的意义何在?前、后两类兴趣区不同数据的比较结果作者没有介绍。增加兴趣区的范围是否还有可比性?

回应: 根据前人研究,对于词内词,因为边界位于两词素之间,因此我们将尾字定义为预视词进行数据分析。根据Yang等人(2009)的研究发现,预视可以对词N+2进行加工。在本研究中,词间词的边界位于该双字词之前,并且n+1和n+2属于同一个语义单元,因此,读者更有可能对这两个词作为一个整词进行加工。基于此,对于词间词,我们采用Cui等人(2014)的数据分析方法,以双字词的首字和双字词整词分别作为兴趣区进行分析,以对结果有更清晰的展现。两种分析方法的结果和结论一致,更说明了研究结果的稳定性。

意见9: 作者在整篇文章都在探讨预视效应,而且作者在表格中也呈现了这一结果,那么作者为何没有就不同条件下的预视效应直接进行差异检验。

回应: 预视效应的计算方法是不同预视条件和一致预视条件之间的差异。如果只呈现预视效应,读者就不能看到正常阅读条件下(一致预视条件)不同句子类型和不同词语类型之间的阅读差异;也不能看到预视效应的差异是来自于一致预视条件,还是不一致预视条件,从而不能真正反映阅读过程的特点和差异。该方法也是现在阅读认知加工领域普遍采用的方法,详见Rayner等(2014)、Yan等(2009)。

意见10: 在13页结果汇报中,作者指出“词语类型和预视类型的交互作用显著, $F_1(1,94) = 11.63, p < .01, \eta^2_p = .11; F_2(1,113) = 8.23, p < .01, \eta^2_p = .07$,词内词的预视效应(42ms)显著大于词间词(24ms)。”“词语类型和预视类型的交互作用显著, $F_1(1,94) = 15.06, p < .001, \eta^2_p = .14; F_2(1,113) = 11.79, p < .01, \eta^2_p = .09$,词内词的预视效应(60ms)显著大于词间词(36ms)。”交互作用是指?

回应: 词语类型包含词间词和词内词两个水平;预视类型包含一致预视和不一致预视两个水平。两者交互作用显著,说明不同词语类型条件下,不同预视类型的差异趋势是不同的,而

不同预视类型之间的差异即为预视效应。即文中表述为“词内词的预视效应(42ms)显著大于词间词(24ms)”、“词内词的预视效应(60ms)显著大于词间词(36ms)”。

.....

审稿人 2 意见：《词边界信息对中文词间和词内预视加工的影响》研究过程规范，有一定的创新性，但存在一些问题需要修改或说明。

意见1：摘要的条理不清晰，请按投稿要求写。

回应：根据心理学报的要求，摘要的内容包括研究目的、方法、结果、结论，尽可能不用“本研究”、“我们”等第一人称，限 200 字以内。谢谢审稿人的建议，我们对文章的摘要进行了重新的组织和描述。本研究主要探讨两个问题，一是语言单元对预视加工的影响，二是词间阴影和非词阴影对词间和词内的影响，摘要按照两个问题分别写出目的、方法、结果、结论。由于第一个问题介绍背景太多，造成了条理不清晰，所以将背景信息删除；由于第二个问题的目的和方法的顺序写错了，可能造成了摘要的条理不清晰。将目的和方法的顺序调换，另外将本研究中的第一人称用“实验”进行了替换。并且用蓝色显示。

意见2：问题的提出部分太长，读者不容易把握其中的逻辑关系，最好分小标题来写；问题和问题提出混在一起叙述，读起来有重复之感；在问题提出部分的最后，最好有一段话，明确要研究的问题是什么，为什么要研究。

回应：谢谢审稿人的意见，确实如此，我们已经依据审稿人的建议，在前言描述的每个部分增加了小标题，并且在每个问题提出之后都将问题和研究目的进行总结以更清晰。增加之后用蓝色显示。

意见 3：对重要观点或关键词缺少必要解释，如什么是“边界掩盖”？为什么“在词之间增加边界信息会减少词之间的边界掩盖”？为什么“在词之间增加边界信息会减少词之间的边界掩盖，这将导致更快的词语识别，从而增强预视加工”？什么是“边界掩盖的干扰作用”？

回应：非常感谢审稿人的意见，原文确实存在句子表述方面待完善的细节问题。

第一，边界掩盖已增加相应的解释：词与词之间无明确的边界信息，则称之为词与词之间边界的掩盖（Drieghe et al., 2005, 2017）。

第二，“增加边界信息会减少词之间的边界掩盖”表述不准确，已经修改为“增加空格会减少词之间的边界掩盖”。

第三，“在词之间增加边界信息会减少词之间的边界掩盖，这将导致更快的词语识别，从而增强预视加工”原文表述不准确，因此将原文改为“在词之间增加边界信息会减少词之间的边界掩盖，这将导致更快的词语识别（Drieghe et al., 2005）”。

第四，“词边界掩盖的干扰作用”表达不够准确，已经修改为“边界信息的掩盖”。

意见 4：预视效应与预视时间是什么关系？

回应：研究发现预视效应随着预视时间的增长而增大,这是因为预视时间越长，可以对预视词进行更多的加工，从而获得更多的预视信息(Li et al., 2017; Yan, Risse, et al., 2012; Yan, 2015)。例如，Yan 等（2012）的研究用预视词前的单一注视时间作为预视词的预视时间，结果发现，随着预视时间的增长，预视效应增大。

Yan, M. (2015). Visually complex foveal words increase the amount of parafoveal information acquired. *Vision research* 111, 91-96.

Yan, M., Risse, S., Zhou, X., & Kliegl, R. (2012). Preview fixation duration modulates identical

and semantic preview benefit in Chinese reading. *Reading and Writing*, 25, 1093 – 1111.

意见 5: “在英文的研究中，虽然可以通过在词间词和词内词都存在空格的情况下，来考察两类词预视加工之间的差异，以控制空格对词间词和词内词的预视加工的影响相同，但是由于其语言特性的限制，不能完全排除空格对词间词和词内词预视加工的影响”，其中“不能完全排除空格对词间词和词内词预视加工的影响”指的是什么？

回应: 因为在英文中不存在没有空格的词间词，所以研究者不能直接比较没有空格的词间词和词内词预视加工之间的差异，而只好退而求其次选择有空格的词间词和词内词。虽然词间词和词内词之间均有空格，但是空格的存在也是一种干扰因素，所以我们说“不能完全排除空格对词间词和词内词预视加工的影响”。如果审稿人认为这样的表述太过复杂，我们可以改为“不能排除空格对词间词和词内词预视加工的影响”。

意见 6: “阴影所提供的词边界信息”，阴影是在目标词下面还是在非目标词下面？如果在非目标词下面，目标词的边界怎样体现？非目标词下面的阴影设定的规则是什么？

回应: 阴影是标记在非目标词下面。如论文中的图 1 所示，通过将不同的词用阴影间隔标出。为了避免阴影对目标词的加工产生额外的影响，因此本研究均是在前目标词和后目标词上标记阴影以产生词边界信息，而没有标记在目标词上。

意见 7: 目标词加粗，本身就有边界信息，在一个句子中同时以加粗和阴影表示边界信息，如果边界信息起作用，怎样分离是加粗带来的还是阴影带来的？

回应: 感谢审稿人的意见。目标词在正式实验中并没有加粗。实验举例只是为了突出我们的目标词而加粗了。已经在“实验设计与材料”中增加相关的说明并用蓝色标注。

意见 8: 词间词和词内词没有评定，不能保证词间词就是词间词，词内词就是词内词，也就是说，两类词划分或选定的标准没有说明。

回应: 首先，如前所述，词间词是指注视词和预视词属于不同的词；而词内词是指注视词和预视词属于同一个复合词，所以本研究控制每对词间词（“纸盒”）的首字和词内词（“稿纸”）的尾字相同，即本研究的预视目标字。”

其次，我们控制目标字和相邻的词均不能组合成词，避免对阅读过程中的词切分产生混淆。比如，“看到稿纸已经”，“稿”不能与之前词组合成词；“纸”也不能与之后词组合成词。

再次，我们对词切分进行了评定，关于句子词切分评定相关过程，已经加在实验设计与材料部分。24 名不参加正式实验的某大学学生参与了评定。评定时发给每位学生一份卷子，里面包含了 120 个实验句子。要求被试仔细阅读句子，然后用竖线对句子中的词语进行切分 (Yan et al., 2010)。结果发现，被试对词间词和词内词的边界进行切分的一致性分别为 85.9% 和 87.2%，符合本研究的实验要求。

意见 9: 各眼动指标与预视效应的关系是什么

回应: 因为越过边界后，预视词将变为目标词，因此我们采用反应早期阅读加工的眼动指标 (Rayner, 1998)；另外，也可参考研究预视效应的文献 (Yang, 2013; Rayner & Schotter, 2014; Schotter et al., 2015)。

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422.

Rayner, K., & Schotter, E. R. (2014). Semantic preview benefit in reading English: The effect of initial letter capitalization. *Journal of Experimental Psychology: Human Perception and*

Performance, 40, 1617–1628.

Schotter, E. R., Lee, M., Reiderman, M., & Rayner, K. (2015). The effect of contextual constraint on parafoveal processing in reading. *Journal of Memory & Language*, 83, 118-139.

Yang, J. (2013). Preview effects of plausibility and character order in reading Chinese transposed words: evidence from eye movements. *Journal of Research in Reading*, 36, 18-S34.

.....

审稿人3意见:文章借助中文视觉特点,检验了预试位置和词边界显示对于词汇加工的影响,发现预试边界位于词内位置时较词间位置时预试效应更大,但词边界提示并无促进作用。这一结果为汉语阅读的眼动模式提供了新证据。论文逻辑清晰,方法可靠,表述较为清楚。但也有几个小问题还需要进一步考虑。

意见1:不同的词(between word词间词):实际上这是指边界位置,建议改名称,如词内边界和词间边界条件。

回应:感谢审稿人的意见,诚如审稿人所言,词间词是指边界位于词与词之间,即词间边界条件;而词内词是指边界位于复合词的两个词素之间,即词内边界条件。但是在本文中用阴影标记词边界信息,该“边界”一词与上述的“边界”文字相同,但是内涵不同,所以容易混淆。另外,为了保持与国际上的表述相一致,请参看 Juhasz, Pollatsek, Hyönä, Drieghe and Rayner (2009)。故此,再三斟酌之后,仍采用了词间词和词内词的表述。

意见2:边界位置和空格之间存在较强的联系。目前的论文对于两者的联系有待加强。

回应:的确如此,目前考察词边界信息对阅读影响的研究中,大多采用词间空格作为边界,如Drieghe等(2005)、Sheridan等(2016)。正如Cui等人(2014)的研究中,也是采用空格考查了词边界信息对预视加工的影响。但是如作者所言,用空格对预视加工进行考察,词间空格在提供词边界信息的同时,也减少了词边界掩盖。所以这种方法不能直接说明空格条件下产生的更大的预视效应是由于增加的词边界信息导致的,还是由于减少的词边界掩盖所导致的。所以,在本研究中,为了排除在增加词边界信息的同时,对不同词的预视加工产生不同边界掩盖作用,故而采用了词间阴影的方法(Bai et al., 2008)。本研究得到了与词间空格方法不同的结果,因此,我们认为,不同词边界的标记方式不同会导致对预视加工产生不同的影响。也由此提醒研究者,在考察词边界信息的作用时,要分清到底是增加的词边界所导致的,还是采用的词边界标记方式所导致的,这分属于不同的认知加工过程。

意见3:图1实验条件名称应改成中文;图2、图3“词内词”的内用了镜像字。表1和表2的标题相同,应调整以便区分。

回应:这些确实是作者的疏忽,文中已修改,谢谢审阅人指正!

意见4:眼动实验结果较多,建议整理成表格方便阅读。

回应:感谢审稿人的建议,结果已经用表格表示。

意见5:讨论第一段简写,前言简写,知识介绍性内容略写。

回应:感谢审稿人的意见,已经将讨论第一段不重要的信息删除。

意见6:空格作为一个自然的书写属性可能与底色对词的标记加工存在不同的机制,在加入底色的情况下可能带来信息凸显度的差异,引发额外策略。

回应：的确如审稿人所述，凸显的信息会对阅读产生不同的影响。Rayner和Schotter (2014)的研究发现当预视词是大写字母时的预视效应显著地大于预视词是小写字母时的预视效应。在刚刚发表的一篇文章上，Zhou等(2018)采用分辨率最高的红色和绿色作为词边界信息交替呈现，以增加词边界信息的区分度，结果同样发现词边界信息不会影响注视时间。另外，Cui等(2014)的研究中也发现空格对双字词的预视加工无影响。上述研究的结果与本研究的结论一致。

如前回答审稿人一的问题时所述，我们采用黄色阴影的底色信息来标记边界信息是为了最大化地增加边界信息对预视加工可能产生的影响，可是，即便如此，本研究仍发现，词间阴影并未对预视加工产生影响，由此得出研究结论，词间阴影产生的词边界信息并不能促进预视加工。

Zhou, W., Wang, A., Shu, H., Kliegl, R., & Yan, M. (2018). Word segmentation by alternating colors facilitates eye guidance in Chinese reading. *Memory & Cognition*, 6, 1-12.

意见 7：前言中对于中英文的预视做了很多对比，但是讨论中对此问题回应较少，应考虑调整。

回应：感谢审稿人的意见，本研究的目的是想为中文和印欧语系预视效应的差异性提供可能性的解释，结果发现边界信息无影响，这说明两种语系预视加工的差异有更深层次的原因，有待于继续挖掘。已将这一结果的讨论增加在第 19 页，并用绿色字体标注。

第二轮

审稿人2意见：《词边界信息对中文词间和词内预视加工的影响》的作者已经根据审稿意见进行了认真修改，同意发表，但还有几个问题需要斟酌。

意见1：实验是用眼动仪做的，在摘要的方法中应有所体现。

回应：谢谢审稿人的细心阅读和指正，原文摘要确实没有提到眼动仪，已经将“利用眼动轨迹记录法并结合边界范式”增加到原摘要中。

意见2：采用词间阴影做为词边界信息，研究题目中为何不直接使用“词间阴影”一词。空格也是一种“词边界信息”。

回应：谢谢审稿人的意见，确实题目用“词间阴影”更加合适，已经修改。

意见3：研究题目是《词边界信息对中文词间和词内预视加工的影响》，作者预期词边界信息对词间的预视效应的影响可能要大于对词内的预视效应的影响，结论是词间阴影（即词边界信息）对预视加工并没有产生促进作用，根据全文的描述，可以理解为词间阴影（即词边界信息）对预视加工并没有产生影响，是一个阴性结论；而且，研究结论是词间阴影对预视加工并没有产生影响，那么对词间词和词内词的区分有无必要。因此，需要考虑研究题目是否合适。

回应：本研究将目标词分为词间词和词内词的主要目的是探讨词语的属性特征对预视效应的影响，即注视字与预视字属于或不属于一个语言单元是否会影响预视加工。由于英文语言特性的限制，研究者只能通过词间词和词内词都存在空格的情况下，来考察两类词语预视加工之间的差异，以控制注视词和预视词的距离相同，但是不能在排除空格影响的条件下，对两类词语的预视加工进行考察。而中文文本可以排除空格的影响对两类词的预视加工进行探

讨。这也是本研究的主要目的和创新点。研究结果发现词内词的预视效应显著地大于词间词，验证了 Hyönä 等(2004)的假设，说明词语的语言属性会影响读者的预视加工。但是如审稿人所言，词间阴影对预视加工并没有产生影响，是一个阴性结论，题目确实不是特别恰当，考量后，我们将题目修改为“中文词间词和词内词预视加工的差异：词间阴影的作用”。

意见4：正文末尾的第二个结论“词间阴影对预视加工并没有产生影响（概括之后）”和第三个结论“词间阴影对词间词和词内词的预视加工没有差异（原句）”需要斟酌。既然词间阴影对预视加工不产生影响，那么何谈词间阴影对词间词和词内词预视加工有没有差异。

回应：谢谢审稿人提出这个建议。确实，从数据结果来看，结论二、三分属于不同结果表达出的不同结论；但是从整体表达的语言逻辑来看，确实应该修正，所以我们将这两个结论合二为一，以更清晰更全面地说明研究结果。修改如下：“词间阴影对词间词和词内词的预视加工均没有显著影响，说明中文读者的词切分是智能化的过程，读者并不需要借助词间阴影所标示的词边界信息确定词的边界，因而对读者的预视加工并没有产生促进作用。该结果支持 Li 等(2009)提出的词切分和词汇识别的模型，目标字不论位于首字或者尾字，都是同时被加工的，读者在加工前目标词时已经提前完成了目标词的词切分。”具体见正文第21-22页的橙色部分。

意见5：“每个汉字所占用的空间大小相同”的说法欠准确，例如“繁”和“一”能不能说所占的空间大小相同。

回应：谢谢审稿人的指正，为了使表达更加准确，在正文第2页将原句子修改如下“每个汉字所占用的宽度相同。”并用橙色表示。

Liversedge, S. P., Zang, C., Zhang, M., Bai, X., Yan, G., & Drieghe, D. (2014). The effect of visual complexity and word frequency on eye movements during chinese reading. *Visual Cognition*, 22(3-4), 441-457.

Zang, C., Liversedge, S. P., Bai, X., & Yan, G. (2011). Eye movements during Chinese reading. In S. P. Liversedge, I. Gilchrist, & S. Everling. (Eds.), *The Oxford handbook of eye movements* (pp. 961 - 978). Oxford: Oxford University Press.

意见6：在讨论部分：“另一方面，正如 Bai 等人所言，读者每天阅读的文本都是非边界文本，词间阴影的句子是读者所不熟悉的，但是读者并没有出现阅读的障碍，反而表现出跟熟悉的正常文本相同的阅读时间，这表明了词边界信息的促进作用。”读者阅读不熟悉的词间阴影的句子跟熟悉的正常句子的时间相同，是表明词边界信息对阅读有促进作用还是没有抑制作用？

回应：谢谢审稿人的意见，这一表述确实比较微妙，虽差之分毫，但可能谬之千里。我们引用 Bai 等人(2008)对该结果的解释来说明，加入词间阴影的句子是读者所不熟悉的，所以对阅读应该会产生一个抑制作用，但是研究结果表明，词间阴影的句子跟熟悉的正常文本阅读时间相同，所以据此推断，词边界信息是存在促进作用的。但是为了表达的精确性，已将句子改为“这表明词边界信息可能存在促进作用。”详见正文第18页红色部分。谢谢审稿人的批评指正。

审稿人3意见：作者已经很好地回应了我之前所提的问题。建议发表。

回应：感谢审稿专家对本文的肯定。

第三轮

审稿人4意见：《词边界信息对中文词间和词内预视加工的影响》一文从中文词的特点出发，利用眼动技术尝试验证几种不同假设的正确性，为中文和印欧语系预视效应的差异性提供了可能性的解释。实验设计基本合理，数据分析详实，有较好的理论和应用价值，但有以下几个问题需要进一步澄清和说明：

意见1：在第二自然段，文章写到“串行和并行眼动控制模型对阅读过程中的预视加工均进行了详细的描述和预测(Engbert et al., 2005; Reichle et al., 2009)”。 “串行”和“并行”分别指的什么？

回应：串行模型认为注意是按序列分布的，即注意一次只能聚焦在一个词上，只有在当前注视词加工完成之后，注意才能转移到下一个词的加工，并开始对下一个词进行加工(Rayner, White, Kambe, Miller, & Liversedge, 2003; Reichle, Pollatsek, & Rayner, 2006)；而并行模型认为注意在空间上是平行分布的，可以对视觉区域内的所有词进行加工，因此视觉范围内所有词语的属性特征均会对视觉加工产生即时的影响(Engbert, Nuthmann, Richter, & Kliegl, 2005; Richter, Engbert, & Kliegl, 2006)。详见P1第二段紫色部分。

意见2：“比较了中文阅读中单纯词、复合词、短语的预视加工”。虽然作者在之后对其做出了解释性说明，但考虑到本刊的作者主要来自心理学领域，请分别给出上述三种词类的一个具体示例，以便读者理解。

回应：谢谢审稿人的细心阅读和宝贵建议，增加示例确实会有助于理解，现已在原文中增加具体示例：单纯词“玫瑰”、复合词“灯塔”、短语“斜塔”。详见P3第11行到13行紫色部分。

意见3：作者在对审稿人一一的回复意见中指出“在本研究中，虽然材料举例中的目标词是偏正结构，但是并不是所有目标词均是偏正结构，还有其他结构的词语，如并列结构”。除偏正结构和并列结构外，实验材料是否还包含其他类型的结构？作者需要思考的是，目标词结构的差异性是否会导致不同的预视效应？语言结构类型是否是实验的一个混淆变量？例如，某些结构是否有可能促使被试产生更大的预期，从而导致更大的预视效应。

回应：谢谢审稿人的宝贵意见和建议。诚如审稿人所言，中文复合词按照语法结构的不同可以分为五种类型：偏正、并列、动宾、动补、主谓结构，不同语法结构复合词的预视加工特点很可能是不同的；正因为如此，所以在此研究中，目标词语法结构的选取我们采用了随机化的处理，目标词除了偏正结构和并列结构，也有其他结构类型的词语，比如“酒酿”、“女儿”、“专业”等。但是，如前所述（在第一次修改意见中对第一位审稿专家的第二个问题的回复），通过查阅文献，我们发现对语法结构的加工特点和加工时程进行研究的文章很少，目标词语法结构的差异性是否会导致不同的预视效应这一问题本身确实非常重要，值得研究者进一步探讨和验证（请参阅如下文献）。

Cui, L., Cong, F., Wang, J., Zhang, W., Zheng, Y., & Hyönä J. (2018). Effects of grammatical structure of compound words on word recognition in Chinese. *Frontiers in Psychology-Language Sciences*.9:258.

Jia, X., Wang, S., Zhang, B., & Zhang, J.X. (2013). Electrophysiological evidence for relation information activation in Chinese compound word comprehension. *Neuropsychologia*. 51(7), 1296 - 1301.

Liu, P., and McBride-Chang, C. (2010). Morphological processing of Chinese compounds from a

grammatical view. *Appl. Psycholinguist.* 31, 605-617.

意见4: “被试每读完三个句子后, 根据所读句子的内容对一个问题进行判断”, 请给出一个具体的例子;

回应: 在读者阅读完三个句子之后内, 会出现一个要求被试判断“是”或者“否”的阅读判断题。例如实验句子:

“老板惊慌地瞥见稿纸已经不在桌上了。”

“评委们称赞字画精巧与细致。”

“解放军利用台风的掩护顺利挺进敌区。”

之后出现问题句子: “解放军撤离敌区。”

然后要求被试做“是”或者“否”的判断, 目的是确保被试认真地阅读句子。

意见5: “被试对词间词和词内词的边界进行切分的一致性分别为85.9%和87.2%”, 差异不显著, 符合本研究的实验要求”, 请补充一致性检验的Kappa系数。此外, “本研究的实验要求”具体指什么?

回应: 为了确保本研究中, 词边界条件下关于词边界的切分符合读者的认知, 所以进行了句子词切分的一致性评定。我们要求两组各12名被试分别对词间词和词内词的句子进行词边界的切分, 结果显示, 词间词和词内词进行词边界切分的一致性分别为85.9%和87.2%, 一致性水平较高 (Yan et al., 2010), 并且词切分的一致性水平在词间词和词内词之间没有显著差异 ($t = 1.42, p > .1$)。此外, 我们对词切分的一致性系数进行了统计, 分别计算了词间和词内的克隆巴赫 α 系数, 结果显示词间的一致性系数 $\alpha = .60$, 词内的一致性系数 $\alpha = .64$, 符合在心理测量学中对材料的要求(戴海琦, 2015), 详见P10第二段。

戴海琦, 心理测量学(2015), 高等教育出版社, 39-40.

意见6: 在创新性方面, Zhou等(2018)采用分辨率最高的红色和绿色作为词边界信息交替呈现, 以增加词边界信息的区分度, 结果同样发现词边界信息不会影响注视时间。本研究与其不同之处是什么? 建议将这篇新文献放在综述部分进行评述。

回应: Zhou等(2018)采用正常阅读的实验范式考察词边界信息是否会影响到目标词的识别和加工; 而本研究的重要目的是探讨词边界信息对预视加工的影响, 并且该影响在词间词和词内词之间是否有差异。因此, 本研究的部分结果可以与Zhou等(2018)的结果相互印证, 但是本研究探讨了更多关于预视加工的问题, 这是在Zhou等(2018)的研究中未曾涉及的。

意见7: 文章需要说明采取黄色而非其他颜色(或方式)标记词边界的理据性。比如Zhou等(2018)的理由在于红色和绿的分辨率高。作者在回复审稿人二的意见“在加入底色的情况下可能带来信息凸显度的差异, 引发额外策略”时说, “词间阴影并未对预视加工产生影响, 由此得出研究结论, 词间阴影产生的词边界信息并不能促进预视加工”。然而, 加入底色并不一定是促进效应, 也有可能是抑制作用, 比如某些颜色可能激发消极的情绪效价, 从而对实验结果造成影响, 因此需要作者对这一操作方式的合理性做出进一步的说明。

回应: 本研究采用黄色作为底色, 是因为在Bai等(2008)的研究中, 灰色的词间阴影并没有显著减少读者对目标词的阅读时间; 为了最大化地增加边界信息对预视加工的影响, 因此, 本研究中采用更有可能引发读者注意力的黄色阴影(高分辨度、高亮度)作为词边界信息。可是, 即便如此, 本研究仍发现, 词间阴影并未对预视加工产生显著的促进影响。此外, 审稿人指出某些颜色对情绪效价的影响问题, 我们查阅后发现, 关于情绪对普通文本阅读影响的研究非常少, 因此我们没找到可借鉴的依据, 有待于以后相关研究的深入探讨。

意见8: 据《心理学报》要求, 英文摘要的字数应为400-500字, 目前篇幅略短, 需要进一步扩充, 尤其需要简要回顾已有研究及其研究空缺, 而不是直接论述自己的实验

回应: 感谢审稿人的提醒, 已经在英文摘要部分增加了先前的相关研究, 并且字数已经达到要求, 再次感谢审稿人的提醒。

意见9: 英文摘要表达方面不够地道, 需要邀请英语本族语者修改和加工, 例如“we adopted a word highlighting sentence and nonword highlighting sentence to compare with normal sentence to test whether”一句话当中出现了两个to, 不符合英语表达习惯。又如, Most of all, word highlighting can provide the word boundary information meanwhile it can manipulate the length of the sentence一句实际上可以使用not only...but also进行连接

回应: 已经修改相关表述, 并且请英语较好的把关, 谢谢审稿人的建议。

意见10: 文章的可读性仍待提高, 需要进一步润色和斟酌, 力求文从字顺。例如“词边界信息是否可能会对预视加工产生影响仍在争论中, 而词边界信息对词间词和词内词的预视加工是否会产生不同的影响尚无研究可解答, 这是本研究最重要的研究目的”等表达略显啰嗦重复。

回应: 感谢审稿人的认真审阅和宝贵建议, 我们已经对相关的句子进行修改, 并且对全文的文字表述进行了精心的检查和修正, 力求更加精炼和流畅, 敬请审稿人审阅指正。感谢!

第四轮

审稿人4意见: 《词边界信息对中文词间和词内预视加工的影响》一文的作者已经根据审稿意见进行了认真修改, 质量较前一稿有了很大提升, 研究的理据性也有所增强, 但还有几个问题需要进一步思考和斟酌(按文章顺序给出):

意见1: 中文摘要当中尽量不要出现文内引用如Li等人(2009), 可直接说“支持的是某理论或模型”;

回应: 感谢审稿人的意见, 已经在摘要部分进行了修改。

意见2: 引言部分不够简洁, 大幅度超过《心理学报》对该部分篇幅3500字的规定。参考文献也略多(54条), 超过50条的规定;

回应: 感谢审稿人的意见, 引言部分由于先前研究与本研究的联系较多, 考虑到读者不同的背景知识, 因此尽可能详细的介绍与本研究有直接关系的研究, 另外我们对引言的逻辑进行了多次修改, 目的是为了在字数比较多的情况下读者能够迅速而准确的理解实验逻辑。在保证读者能够准确理解本研究的情况下, 根据字数要求尽力对本研究不重要的引言以及参考文献部分进行了删减。

意见3: 文章有不少“我们”等主观性较强的第一人称叙述, 请改正;

回应: 已经对全文进行了检查并依次修改, 感谢审稿人的提醒。

意见4: “短语也是由两个字组成的, 两个字也均独立成词, 但是两个字无法组成一个词(例

如“斜塔”)”。请作者思考所举实例是否恰当,一方面,“斜”是否可以单独成词;另一方面,“斜塔”是否为非词;

回应:在本研究中,关于短语的举例是Cui et al.(2013)一文中的界定方法和材料:在人民日报的词频表中,可以查到“斜”、“塔”,但是查不到“斜塔”,因此“斜”和“塔”可以被称为词,而“斜塔”不是词。根据短语的定义:由语法上能够搭配的词组合起来的没有句调的语言单位,又叫词组,所以在他们的研究中,将“斜塔”定义为短语。

Cui, L., Drieghe, D., Yan, G., Bai, X., Chi, H. & Liversedge, S. P. (2013). Parafoveal processing across different lexical constituents in Chinese reading. *Quarterly Journal of Experimental Psychology*, 66, 403-416

意见5:表1中,各项数字分别代表什么指标(如频次还是频率等),需明确给出。

回应:感谢审稿人的提醒,已经将对应的单位添加在表格中。

意见6:1中“词间词和词内词的词汇属性统计”中“句子合理性判定”与“句子词切分评定”并不属于词汇属性,请作者注意词汇属性和句法属性的区分。另外,此处已在表格中给出句法合理性等相应指标,但在之后的2.4节又出现了“实验材料的判定”,出现了“结果”在“方法”之前的顺序颠倒,请考虑将这两部分合二为一,保证文章的连贯性;

回应:如审稿人所言,合理性判定与词切分评定确实不属于词汇属性,而且合理性的结果确是在材料评定之前,为保证文章的连贯性,因此将合理性和词切分评定结果从词汇属性的表格中去除。

意见7:文章“讨论”部分某些段落过长,影响了行文的流畅性。如“再次”和“最后”部分,建议按照逻辑层次,进一步划分小自然段;

回应:感谢审稿人的建议,已经根据逻辑层次将该部分进一步划分成小自然段。

意见8:文章仍有部分表达不够文从字顺。例如,“Drieghe等认为,词与词之间无明确的边界信息,则称之为词与词之间存在边界的掩盖。在词之间增加边界信息会减少词之间的边界掩盖,这将导致更快的词语识别(Drieghe et al., 2005,2017)”。该句不够流畅,应该修改为,“Drieghe等人(2005,2017)认为,在词之间增加边界信息会减少词之间的边界掩盖(即词与词之间无明确的边界信息),这将导致更快的词语识别”。

回应:感谢审稿人的修改,已经按照审稿人的意见进行了修改,并且对全文相似问题进行了检查并修改。

意见9:英文摘要仍需邀请合格的英语母语者进一步校对和润色,里面仍有不少语法错误,如“there was no differences”以及不地道的英语表达,例如,Some researchers used compound words and phrases to further explore this question, and they found that there was no differences in preview effects between within words and between words(Juhasz et al., 2009)。英语中习惯使用短句表达,后半句话可以另起一句,改为:The results showed that there were no differences…。又如,“and researchers thought”表达过于主观,可以改成this is possibly because...(参考文献),

回应:感谢审稿人的建议,已经按照审稿人的意见进行了改正,并且找英语专业的同行对英文摘要进行了修改。

第五轮

审稿人4意见：经过几轮大的修改，作者已经较好地回应了审稿人的问题，建议发表。

但仍需注意的是，目前的中文摘要写作不够规范，前后两个“结果表明”给人感觉是做了两个毫不相关的实验。请作者参考以往学报上的文章进行调整，先介绍研究问题，然后阐述研究范式，最后分点列举研究结果。

回应：感谢审稿人的宝贵意见和肯定。针对审稿人提出的摘要问题，我们已经参阅审稿人的建议和学报相关文章进行了认真的修改，详见P1摘要部分。

第六轮

编委复审：文章经过几轮修改。但却存在较多问题：

意见1：作者修改后应该将修改后的文本呈现出来，而不是保留修改的痕迹，删去的应该接受删去，加入的也要接受加入。现在的文本根本没法看，不知道作者是留下还是修改。作者这样做显得态度上不够认真。

回应：首先，我们要说抱歉。因为误解了杂志的要求，以为每一次的修改痕迹都要保留，因此从第一稿每一个变化都保持着修订格式，由此给主编和审稿人的阅读带来麻烦，深表歉意。本次修订，为了更加全面和清晰，我们保留了文章的两个版本格式。在前面一个版本中，我们只保留了这次修改的痕迹，其他几次的修改已经改为接受模式；在后面一个版本中，我们将之前根据审稿人的重要建议做出的修改，分别用不同颜色进行了区分标注，并在之前的审稿意见中指出了其具体的位置，以方便主编和审稿人查阅。最后，代表本研究的所有作者再次向主编和审稿人表达诚挚的歉意，我们保证会认真认真更认真。

意见2：行文上问题很多。基本上不符合汉语的标点符号规定，该用句号的用逗号，该有逗号的没有。病句太多，如“实验结果发现”，实验者发现才对。但是后面应该有逗号。建议作者找中文水平比较好的帮助修改。不然，即使研究有价值，也不宜在学报上发表。修后再审。

回应：确实如此，我们诚恳接受批评。我们现已根据审稿人的意见通篇进行了检查和修改，将所有的“实验结果发现”、“结果发现”改为了“实验结果显示”、“实验结果表明”、“统计结果表明”，将全文的标点符号进行了检查并修改。关于行文的流畅性等方面，作为重点着力进行了检查、修订和完善，敬请查阅和指正。真诚感谢主编和审稿人给予的机会和帮助！

第七轮

编委复审：论文经过几轮修改，水平有了很大提高，建议刊用。

回应：感谢编委专家对本文的肯定。

主编终审：《词边界信息对中文词间和词内预视加工的影响》一文借助中文视觉特点，检验了预试位置和词边界显示对于词汇加工的影响，研究视角较为新颖。本研究总体设计合理，数据处理科学规范，结果可信。最终的研究结果说明，预试边界位于词内位置时较词间位置时预试效应更大，但词边界提示并无促进作用，这一结果为汉语阅读的眼动模式提供了新证据，有较好的理论和应用价值。研究者根据几位外审专家的意见进行了认真修改，修改后的文章条理清晰，在结构和表述上更加完善和严谨。

回应：感谢主编对本文的肯定。