

《心理学报》审稿意见与作者回应

题目：副中央凹中字 N+2 的预视对汉语阅读眼跳目标选择影响的眼动研究

作者：王永胜 白学军 臧传丽 高晓雷 郭志英 闫国利

第一轮

审稿人 1 意见：

本研究使用边界实验范式操纵字 N+2 的预视可视性，探讨了副中央凹词长信息对眼跳目标选择的影响。我认为本文整体水平较好，研究较有理论价值。但同时也有一些不足之处。

意见 1：首先，希望作者充分考虑到本实验设计上可能存在的局限。本研究中，作者希望考察眼跳目标选择是否受到副中央凹词长影响，此时，应尽力排除句法结构的影响。例如，我发现在例句中：

基础设施水平/低劣/是一些发展中国家的共性。

基础设施水平/低成/了一些发展中国家的共性。

双字词条件下主语、谓语边界在关键词“低劣”之后，而单字词条件下则在关键词中。由于句法结构已被证实可以影响副中央凹信息的提取（例如，Luo, Yan, & Zhou, 2013, JEPLMC），因此属于潜在混淆变量。

回应：首先非常感谢专家的意见，我们也非常的认可专家所提出的问题。Luo et al.(2013)的研究虽然发现语法信息会影响副中央凹信息的提取，但是在他们的研究中存在外显的韵律信息，包含韵律信息的低通过滤过的句子语音信息（实验二），或者标点符号（实验三），而在本研究中并不存在这种外显的韵律信息；其次，在该项研究中，他们在韵律边界前的词上，发现了韵律信息与副中央凹预视的交互作用，而在本研究中，我们分析了边界前词的注视情况，并没有发现词汇类型与副中央凹预视的交互作用 ($F_s < 1, p > 0.05$)，说明本研究中在缺少外显的韵律信息的情况下，副中央凹不同的语法信息加工并没有产生明显的差异，进而说明副中央凹中的不同句法信息在本研究中没有影响副中央凹中的加工。

在我们的实验材料中，为了保证字 N+1 的加工负荷完全相同，从而来探讨字 N+2 预视的作用，我们将双字词条件与两个单字词条件下的首字控制为同一个字。在双字词的词性已定的情况下，那么两个单字词条件下，字 N+2 则可能换成了另一词性的词，例如“安装”（动词）和“安完”（动词+形容词）。从而造成了专家所指出的可能存在的句法混淆变量。而如果将双字词和两个单字词条件都控制为相同的词性，例如“动词”与“动词+动词”，则可能

同样存在专家指出的两种条件下的句法差异，因此这个问题很难避免。在文章中，我们已根据专家的意见，作为本研究的不足，已经进行了说明（P.12）。

意见 2: 另外，字 N+2 的频率并未进行匹配。作者指出，“两个单字词条件下字 N+2 的字频（ $M = 2677.40$ /百万, $SD = 5765.45$ ）显著的高于（ $t = 2.18, p < .05, d = 0.35$ ）双字词条件下字 N+2 的字频（ $M = 1156.20$ /百万, $SD = 2414.15$ ）。我认为，根据本研究的目的是，应操纵的变量仅仅为词长信息；N+2 的字频差异也属于潜在混淆变量。我尚未理解作者为何不进行匹配。

回应: 在 Liu et al.(2014)的研究中，已经发现了副中央凹的加工影响从当前注视词 N 跳出的眼跳长度。在本研究中我们关注的主要问题是，读者只是根据副中央凹中获取信息的多少来选择眼跳目标，还是受到副中央凹信息性质的影响，例如，副中央凹的两个字是否构成词。如果，读者只是根据副中央凹中获取信息的多少决定随后眼跳的长度的话，那么在字 N+1 相同，字 N+2 的加工也不存在加工差异的情况下，将无法检验出差异，因此，我们将两个单字词条件下字 N+2 的频率选择成高于双字词条件下字 N+2 的频率。由于研究的逻辑是用显著的结果说明问题，采用这种操纵，如果从前目标词跳出的眼跳长度在两个单字词条件下字 N+2 的预视大于双字词条件下字 N+2 预视作用，那么我们可以说明副中央凹中信息的多少是影响眼跳目标选择最主要的因素。我们在问题的提出与假设上，对此已经进行了说明。见 P.3 和 P.4。

意见 3: 本文某些地方逻辑不易理解，请作者更详细地阐述。例如，作者指出，“如果读者确实是选择词作为下一次注视的目标，那么我们应该发现双字词条件和两个单字词条件下眼跳长度及随后的注视位置上的差异。”按照我的理解，这段描述应对应读者进入关键区时的眼跳长度和在关键区上的注视点位置。根据表 2，确实存在差异（交互作用显著），但作者为何说“实验的结果并没有发现这种差异”？

如果作者期待*主效应*上的差异，应设计另外一个实验。这是因为，依照目前的实验设计，*只有*在双字关键词、有效预视的这一条件组合下，产生了双字词水平的预视，而其他三种条件都不提供完整的双字词预视。按照 Yan et al. (2010)的假设，这种条件下平均眼跳长度会更长，眼跳目标更远离词首。而此实验结果正发现了这样的交互作用。作者在结果描述时，按照预视效应来解释（“当目标区为双字词词时，预视类型的作用显著.....当目标区为两个单字词时，预视类型的作用不显著”）。虽然此描述正确，但与该文主题（预视词长度是否

影响眼跳)不贴近。建议作者在描述实验结果时更好地回应全文主旨,同时在讨论时更贴切实验设计。

类似地,作者指出“双字词条件下的注视位置上应该与两个单字词条件存在显著的差异。但是研究结果并没有发现这种差异,这说明副中央凹中是否完成词切分可能并不影响随后眼跳目标的选择。”这一描述并不正确,因为如上文所述,在双字关键词、掩蔽预视条件组合下,被试不能在副中央凹加工词长信息;因此作者不应期待实验的主效应。

回应:非常感谢专家指出我们之前逻辑上的一些不足之处。正如专家所指出的,如果读者选择词为眼跳目标,我们之前确实希望获得目标区词汇类型的主效应,忽略了在假字预视条件下,改变了副中央凹中的词长信息,这会明显的削弱词汇类型的主效应。但是我们在比较目标预视下双字词与两个单字词下的眼跳长度,仍然没有发现显著地差异。也可以说明我们的观点,在汉语阅读中读者并不是选择词作为眼跳目标的。

在上一个问题的回答中,已经指出本研究的主要目标是来检验副中央凹中信息的性质是否会影响随后的眼跳目标,因此我们分别比较了双字词条件下两种预视类型的差异和两个单字词条件下两种预视类型的差异。并在随后的讨论中,根据专家的建议进行了讨论(P.11)。

意见 4: 建议引文应注意全面性。根据国际惯例,应引用已知的最先提出某一问题的研究。例如,作者指出,“如果破坏了词间空格的信息,那么对读者的注视时间和注视位置都会产生显著的影响(Perea, & Acha, 2009)”。就我所知,破坏空格对眼跳目标选择的问题在 90 年代已有研究。

回应: 已根据专家的建议,引用了之前的相关研究。修改如下:

“如果破坏了词间空格的信息,那么对读者的注视时间和注视位置都会产生显著的影响(Morris, Rayner, & Pollatsek, 1990; Rayner, Fischer, & Pollatsek, 1998)。”

意见 5: 作者指出“由于这个过程发生在注意被抑制的眼跳过程中,读者一般不会意识到这种呈现的变化(白学军等,2011)”。请作者同时引用 saccadic suppression 这一概念的原创研究。

回应: 已根据专家的建议,引用了 Rayner (1975) 的研究。

意见 6: 请作者在这一陈述中引用知觉广度的研究:“这与汉语阅读中眼跳的长度与右侧知觉广度的大小相近的事实相契合。”知觉广度的研究与 N+2 的预视效应高度相关:研究发现,中文读者在阅读中可以得到第 N+3 (Inhoff & Liu, 1998, JEPHPP) 甚至 N+4 (Yan, Zhou, Shu,

& Kliegl, 2015, JEPLMC) 个汉字的信息。

回应：已根据专家的建议，引用了之前的相关研究。修改如下：

“在汉语知觉广度的研究中，例如，Inhoff 和 Liu(1998)的研究发现汉语阅读中右侧知觉广度的大小为 3 个字，而向右侧眼跳长度为 2 个字到 2 个半字，可见汉语阅读中右侧的知觉广度与向右的眼跳长度比较接近，间接说明了 Wei 等人（2013）理论的可能性（闫国利，巫金根，臧传丽，白学军，2010; Inhoff, liu, 1998）。”

意见 7：请作者在描述“词 N+2 的预视受副中央凹词 N+1 加工负荷的调节，当词 N+1 为高频词时，读者可以从词 N+2 上获取预视效应”这一结果时，同时考虑引用 Yang, Rayner, Li, and Wang (2012, Reading & Writing)，本文是对 Yang et al., (2009, JEPHPP)的延续，这两篇论文的研究结果与 Yan et al. 2010 高度一致。

回应：已根据专家的建议，引用了之前的相关研究。修改如下：

“已有研究结果发现，词 N+2 的预视受副中央凹词 N+1 加工负荷的调节，当词 N+1 为高频词时，读者可以从词 N+2 上获取预视效应（Yan, Kliegl, Shu, Pan, & Zhou, 2010; Yang, Wang, , Xu, & Rayner, 2009; Yang, Rayner, Li, & Wang , 2012）”

意见 8：作者在自检报告中指出，“国内外的研究当中尚没有操作副中央凹当中的词汇类型以及词 N+2 的预视，来探讨汉语阅读中的眼跳选择策略的研究”。就我所知，Luo, Yan, & Zhou (2013, JEPLMC)一文报告了预视对眼跳落点的影响。请查阅并考虑引用该文。

回应：已根据专家建议，在自检报告中进行了修改，并且已将该文作为了本研究的参考文献。

审稿人 2 意见：

总体而言，这是一个不错的实验，结果也有点意思，修改好之后值得发表。但是本文中对前人的研究结果，及关于中文阅读过程中眼跳目标选择的最新理论的论述存在一系列的错误或者理解有问题：

意见 1：对阐述 Yan 等人（2010）的理论，及他们认为可以支持他们的证据理解不到位。Yan 等人认为首次落点位置可以根据在这个单词上的注视点数量细分下去。在只有一个注视点的情况，这个分布的中心（最高点）在词中心；而有多个注视点的情况下，这个分布的中心（最高点）在词的前部。他们据此提出，中文的眼跳目标选择具有旁中央凹词的切分，成功则跳向中心，不成功则跳向前部。这一理论只有它直觉上的可能性，但是缺乏确定无疑的证据。

Li, Liu 和 Rayner (2011) 主要是在“攻击”Yan 等人 (2010) 得到的数据模式并不能用来证明他们提出的理论。他们用了两个方法：1) 提供另一种更简单的解释，正如你论文中引用的那样“他们认为并不是由于在副中央凹中词切分情况导致不同的注视位置结果，而是由于偶然的因素注视点落在了词的开始或者中心位置，导致了不同的注视情况。如果读者的注视点落在了词的中心位置，由于注视词中心时对词汇的加工效率最高，那么可能只需要一次注视则能完成对该词的加工，因此只对该词产生了一次注视；如果读者的注视点落在词开始的位置，那么一次注视不能完成对该词的加工”。2) 用一种简单的“恒定长度”的模拟也能产生 Yan 等人观察到的数据模式。因此，Li 等人认为 Yan 等人提供的证据不可靠。但是值得说明的是，Li 等人提出的两种方法只是用来说明 Yan 等人提供的证据来自于数据分析的方法（详见 Li 论文）。只要用这种数据分类的方法，就会观察到 Yan 等人得到的结果（另见 Ma, Li & Pollatsek, in press），不管中文读者用任何眼跳策略。但是 Li 等人并不认为他们提出的策略是对的。他们在文章中有清楚的写明，所以他们肯定不会同意你所说的他们认为中文眼跳目标用了上述两种策略。另外，虽然 Yan 等人的理论正确的可能性基本上没有，但是客观来讲，目前也没有确定无疑的证据直接 falsify 他们的理论。正如你们没有发现词的类型（双字词和两个单字词）对单次注视的落点有影响一样。暗示 Yan 等人的理论可能不对。但是科学的游戏是用 positive 的结果证否，null results 并不是很强的证据。多种原因可能导致没有差异。

Wei 等人(2013)提出了基于加工的眼跳策略（认为在汉语阅读中，读者会估计一次注视时所能加工的字数，而下一跳的眼跳目标为能够获得新信息的位置）已经被 Liu, Reichle & Li (2014)排除，应该清楚的写明。Liu 等人(2014)认为是旁中央凹加工的难度影响了 Outgoing 眼跳长度，中央凹加工的难度只是调节了旁中央凹的加工。也就是中文阅读过程中眼跳长度的控制并不是根据中央凹难度的估计，旁中央难度直接决定了眼跳长度。

回应：首先非常感谢专家所指出的对之前一些文献理解的不足之处，如此详细的阐述之前研究之间的关系。根据专家的详细的讲解，我们对本文的文献综述部分进行了修改，见 P.2-P.3。

意见 2：该论文显然是对 Liu 等人研究的 follow up。所以在问题提出这一部分要对 Liu 等人的理论做更详细的介绍。这样才能准确地提出你们的研究问题。比如，你们说的“虽然 Liu 等人(2014)的研究表明了副中央凹中的加工对汉语阅读中眼跳的影响，但是仍没有说明汉语读者在随后的眼跳中是选择词还是选择字作为眼跳目标”显然是没有正确理解 Liu 等人对中

文眼跳目标选择这一重要理论问题的思路。中文的眼跳目标选择显然既不是基于词，也不是基于字的。这一“抄袭”英文的概念对理解中文的眼跳目标选择没有任何好处。基于中文的特点，旁中央的字和词加工都会影响眼跳长度，但是现有证据没有表明眼睛跳向任何特定的位置。也就是说旁中央字和词的加工难度影响了整个旁中央加工的激活程度，进而影响了眼跳长度。而不应该用基于字或基于词这一简单粗暴的分类方法来归类。这是因为，如果是基于词的，那就要回答跳向词的什么位置（前部，中心或者别的地方），基于字的就要回答字的什么特征决定了中文阅读过程中眼睛跳向它们。显然这是 *misleading* 的。

回应：根据专家的建议，我们已经对 Liu 等人(2014)的研究做了详细地介绍，并且在该研究的基础上指出本研究的主要问题是，获取的副中央凹预视信息的性质是否会影响随后的眼跳目标的选择。修改如下：

“研究表明，当前注视词的加工调节副中央凹中的加工(Henderson & Ferreira, 1990)，那么 Wei 等人(2013)的研究中所发现的当前注视词的特征影响随后的眼跳长度，则可能是由于中央凹注视词的特征影响了副中央凹信息加工，进而影响了随后的眼跳长度。Liu, Reichle 和 Li (2014)对这一问题进行了研究。他们操纵了中央凹注视词的词频以及副中央凹信息的可见性。他们预测：如果中央凹注视词与副中央凹同时影响接下来的眼跳长度，那么中央凹注视词频与副中央凹信息的可见性的交互作用显著，即使在副中央凹信息不可见的条件下，词频的效应也应该显著；如果，中央凹注视词的特征对接下来眼跳长度的影响是通过对副中央凹信息的调节，进而影响接下来眼跳长度的话，那么当副中央凹信息不可见时，词频的效应应该不显著。而研究结果发现，只有当副中央凹信息可见时，词频效应显著，而当副中央凹信息不可见时，词频效应不显著。表明中央凹注视词的特征对接下来眼跳长度的影响，是通过对副中央凹信息的调节，进而影响随后的眼跳长度的。”

意见 3：摘要第二行。“随后的眼跳”后加“长度”。

回应：已经根据专家的建议，进行了修改、添加。

意见 4：摘要第三和四行。“当副中央凹中为双字词时，字 N+2 为目标预视时随后的眼跳长度显著地长于假字预视条件，随后的注视位置更靠近词的中心位置”。这一表述有错误。眼跳长度更长，并不一定表明随后的注视位置更靠近词的中心位置。因为，你也可以说它靠近词的后部，句子的末尾，下一个词，*anyway*，后面的任何一个位置。这并不是无理取闹。因为你这样写很容易误导出一个结论，即有目标预视时，眼睛刻意跳向“词中心”。这显然

是不妥的。这也是这个行业常犯的错误。Yan 等人就用这个来说事。

回应：已根据专家的建议进行了修改，修改为：“随后的注视位置距离词首更远”

意见 5: P.3. 关于边界范式会不会影响眼动行为，请引用更权威的论文（如：Slattery, Angele, & Rayner, 2011 或者 Rayner 更早期的研究）。在删除 later-display 的数据时也参照 Slattery, Angele, & Rayner, 2011 的建议。即跨过边界之后，落在 target 上，10ms 之后还没变过来的都要删了。另外，由 fixation（微眼跳）引起 display change 的也要删了。

回应：已根据专家的建议引用了 Rayner（1975）的研究。

在数据筛选方面，在边界变化之前有微眼跳引起的边界变化的数据我们已经删除；对于注视点已经越过边界 10ms 之后的数据，也已经删除。数据筛选中“边界变化提前或延迟”已经包括了这两部分。

意见 6: P.3. “安装/安完”这一例子不好。“安完”作为两个单字词有点勉强。对照你后面提供的材料，大部分都比这个好。请修改，以免误导读者。

回应：根据专家建议，我们将例子换为“古董店被歹徒抢砸/抢劫后重新安装了防盗装置”。

意见 7: P.5. 整个实验回答问题的比例太少，有可能不能有效的检测被试是否认真阅读或者理解了。一般的比例是三分之一或二分之一。你们至少要报告，在各种条件下，正确率没差异。而不是一个笼统的百分数。

回应：根据专家建议，我们对被试回答问题的正确率进行了检验，没有发现各个实验条件下存在显著地差异，并在文中进行了说明。专家指出的回答问题的比例较少的问题，可能确实存在。在我们实验之前，我们详细的给被试进行了说明，需要他们在阅读时认真的阅读并理解句子。此外，参加我们实验的被试，可以获得期末考试中平时成绩一定的加分，所加的分数=10×回答问题的正确率，因此我们认为被试在我们的实验中认真阅读了句子。修改如下：

“被试回答问题的准确率为93%，各种条件下，被试回答问题的正确率不存在显著的差异， $F_s < 1, p_s > 0.05$ ，表明被试认真阅读了实验句子。”

意见 8: 请参照英文论文，详细定义报告的因变量。你不能假设所有读者都知道，比如凝视时间。

回应：根据专家的建议，我们已经给出了因变量的定义，修改如下：

“在目标区的分析上,我们首先分析了注视点从前目标区跳到目标区的眼跳长度以及注视点的起跳位置。注视位置的分析指标有平均首次注视位置(在目标区上的第一次注视但不管在该目标区上总共有多少次注视),单次注视中首次注视位置(目标区只有一次注视时,注视点的注视位置),多次注视中首次注视位置(目标区有两次以上的注视时,首次注视的注视位置)。在注视位置的分析中,以字为单位进行,即注视点落在目标区上字的位置。

同时我们也对目标区的注视时间指标进行了分析,分析的指标主要有:首次注视时间(first fixation duration),是指在首次通过阅读中某兴趣区内的首个注视点的注视时间,而与兴趣区内注视点的个数无关;凝视时间(gaze duration),是指从首次注视点开始到注视点首次离开当前兴趣区之间的持续时间;单次注视时间(single fixation duration),是指在第一遍阅读中,目标区内只有一次注视时的注视时间;跳读率(skip),是在在第一遍阅读中,目标区被跳读的比率(闫国利,熊建萍,臧传丽,余莉莉,崔磊,白学军,2013)。除了对两个字的标区进行了分析之外,我们还分别分析字 N+1 和字 N+2 的注视时间的指标。”

意见 9: P.6. 注视位置不能最后报告像素值。你可以按照 Yan 的方法算,但最好再除以每个字的像素值。因为心理学不是物理,大家最后关心的是相对于字的相对值。而不是让读者自己去算。表 2 都要换算成以字为单位。

回应: 根据专家建议,我们已经将像素值转化为字单位,并且以字为单位进行了数据的分析。

意见 10: 如果会做,最好都用 linear mixed model 来分析所有的数据,避免项目和被试分析不一致的情况和避免 anti-conservative。LMMs 模型里面的 random 效应设定参看(Barr, Levy, Scheepers, & Tily, 2013)。分类数据(如 skipping 比率)必须要用 GLMM 分析(参看 Jaeger, 2008)。

回应: 已根据专家建议对数据重新进行了分析,使用 R 语言环境下的 LMM 分析了数据。由于分析方法的不同,使用 LMM 对数据进行的分析,与之前使用重复测量方差分析时指标的平均数与标准差在数值上存在一些差异,但基本的数据模式是相似的。

意见 11: P. 10. “如果读者确实是选择词作为下一次注视的目标,那么我们应该发现双字词条件和两个单字词条件下眼跳长度及随后的注视位置上的差异。但是实验的结果并没有发现这种差异,说明读者不是选择词作为注视目标的。” 你是指没有词的类型的主效应? 这个再正常不过了,在计算主效应的时候,加入了假字预览,会降低主效应的 power。你们发现的

交互作用，即双字词条件下，假字预视和正常预视的差异清楚表明了词的作用。

回应：我们已经将此处删除，并且根据专家的建议对我们的数据结果重新进行了讨论。

意见 12: P.11. 引入 Li, Rayner 和 Cave (2009) 模型的意图不清楚。该模型最大的特点就是词切分和识别是同一过程。也就是说词切分有没有值得商榷。词识别的心理机制不是计算机处理。即使有个所谓的切分过程，也没有证据或者理由表明中文的眼跳是基于此的。这个要不删掉，要不就详细阐述清楚。

回应：我们之前所想表达的是在副中央凹的加工中被试可以获取字 N+2 的预视信息，如果被试同时加工了副中央凹中字 N+1 和字 N+2，那么被试则可能在副中央凹中完成了词切分。此处，我们已经删除。

意见 13: P. 12. “我们认为，对于汉语阅读中眼跳目标选择策略的研究中，应该同时考虑当前注视词和副中央凹中的加工特征，以及二者可能存在的交互作用。汉语读者可能同时运用当前注视词的加工信息和副中央凹中的加工信息来选择随后的眼跳目标，两种信息对于汉语读者的眼跳目标选择来说都产生重要的影响。”这是 Liu, Reichle, & Li (2014) 的观点。你们至少应该加上引用。你们的实验提供了新的数据进一步支持了这一观点。这也是你们研究的意义。

回应：根据专家建议，我们对此处进行了修改，并说明了我们的研究是对 Liu 等人 (2014) 研究的进一步的深入。修改如下：

“Liu 等人 (2014) 认为对于汉语阅读中眼跳目标选择策略的研究中，应该同时考虑当前注视词和副中央凹中的加工特征，以及二者可能存在的交互作用。汉语读者可能同时运用当前注视词的加工信息和副中央凹中的加工信息来选择随后的眼跳目标，两种信息对于汉语读者的眼跳目标选择来说都产生重要的影响。本研究的结果为这一观点提供的新的实验数据的支持，并且说明了不仅副中央凹中信息的多少会影响随后的眼跳目标的选择，副中央凹中获取的预视信息的特征也会影响随后的眼跳目标的选择。”

第二轮

审稿人 1 意见：

我认为作者在修改稿中整体上比较好地回答了我提出的问题。有个别之处我觉得或有不

妥，请作者斟酌。

意见 1: 关于问题 1 (关于句法结构的影响):

首先，文献方面，作者认为由于 Luo et al.(2013) “存在外显的韵律信息”，“而在本研究中并不存在这种外显的韵律信息”，因此其结果与本研究性质不同，本研究“在缺少外显的韵律信息的情况下，副中央凹不同的语法信息加工并没有产生明显的差异”。但是事实上，使用普通的阅读理解任务，已经有实验证实位于副中央凹视野的语法信息可以得到加工 (Kim, Radach, Vorstius, 2012)。

其次，分析方面，作者“分析了边界前词的注视情”，希望说明副中央凹中的不同句法信息不存在影响。但事实上，副中央凹信息影响可以存在于不同区段。作者仅仅分析边界前词，通过不存在 POE 效应来做推论不存在影响，似乎不够全面。

最后，统计方面，作者认为“并没有发现.....交互作用”的结果说明句法信息“没有影响副中央凹中的加工”，这一论述在统计学上不够严谨。

基于上述原因，我认为作者对该问题的思考有待进一步加强。

回应: 首先非常感谢专家所提出的宝贵意见，在本研究中确实可能存在句法结构变量的混淆。专家指出在普通阅读理解任务中读者可以在副中央凹视区对句法信息进行加工 (Kim, Radach, Vorstius, 2012)。在 Kim 等人 (2012) 的研究中虽然发现在副中央凹视区可以对句法信息进行加工，但这可能是由于韩语的语言特点所造成的，韩语与汉语不同，典型的句法顺序为 subject-object-verb，正如 Kim 等人 (2012) 所指出的，这种句法顺序可能造成句子的主语与宾语的混淆，因此在韩语中存在 case markers 来标记出主语和宾语等，例如，“O1”和“가”为主语的 case markers，“을”和“를”为宾语的 case markers 等。因此读者可以使用这些 case markers 快速地识别单词在句子中的成分，而且这种识别在句子的理解中具有十分重要的作用。可能正是由于存在 case markers，才使得韩语读者可以在副中央凹视区中获得句法方面的信息。而在汉语中，不存在这种标记词在句子中句法作用的较为外显的信息。虽然在 Luo, Yan 和 Zhou (2013) 发现汉语读者可以获取副中央凹中的句法信息，但是正如我们所指出的，在该研究中存在外显的韵律信息。而在本研究中不存在较为外显的句法信息，因此本研究中句法结构变量的混淆可能是不存在的，或者作用是较小的。

专家认为“副中央凹信息影响可以存在于不同区段。作者仅仅分析边界前词，通过不存在 POE 效应来做推论不存在影响”不够全面。从我们的数据结果可以看出，不仅在边界前词上没有发现目标区词汇类型与预视的交互作用，在目标区以及字 N+1 与字 N+2 的注视时

间指标的分析上也没有发现两因素的交互作用,只是在目标区注视位置的相关指标上发现了两因素的交互作用。但是也存在一种可能,句法结构信息不对注视时间指标产生影响只对注视位置指标产生影响。但是,我们认为这种可能性极小。首先,在 Luo 等人(2013)的研究中,不仅在注视位置上发现了句法信息(韵律信息)的作用,在注视时间的指标上也发现了句法信息(韵律信息)的作用;其次,在拼音文字的眼动控制中“where”主要受低水平信息的影响,而不受语义信息的影响(Rayner, 2009),因此在汉语阅读中句法结构信息影响注视位置的相关指标而不影响注视时间相关指标的可能性应该极小。

虽然,在本研究句法结构作为混淆变量的可能性较小,但是我们作为研究的不足,在讨论中进行了说明。如下

“在本研究中目标区双字词条件下和两个单字词条件下,目标区在整个句子中的句法成分可能不同。而已有研究发现,句法结构可以影响副中央凹信息的提取(Kim, Radach, Vorstius, 2012; Luo, Yan, & Zhou, 2013),这可能造成了句法结构成为本研究的一个混淆变量。在本研究中如果句法结构影响了随后的副中央凹视区加工,那么在随后的注视时间各指标上应该发现目标区词汇类型与预视类型的交互作用,但是在目标区及字 N+1 与字 N+2 的分析中,没有发现两因素的交互作用,而只是在目标区注视位置的相关指标上发现了两因素的交互作用。在拼音文字的眼动控制中“where”主要受低水平信息的影响,而不受语义信息的影响(Rayner, 2009),因此在汉语阅读中句法结构信息影响注视位置而不影响注视时间的可能性较小。虽然在本研究中没有发现目标区词汇类型与预视的交互作用,但是并不能完全排除句法结构信息对副中央凹加工产生了影响,并对随后的注视位置选择产生了影响。”

意见 2: 关于问题 3:

我应先说明一下,我对眼跳目标选择的问题并无任何偏见。只要实验设计合理,解释得当的结果,我都非常认可。所以我对问题 3 从统计适用性以及结论合理性两方面谈一下我的看法。

由于表 2 中提到的眼跳长度和首次注视位置的分析皆存在显著交互作用,结合表 1 的平均数来看,在双字词有效目标预视条件下的眼跳更远,注视点更靠近词中心。统计上,这两个显著的交互作用已经说明,对眼跳目标选择存在差异。我认为更符合理论主题的检验为:双字词有效目标预视 vs (单字词有效目标预视、单字词无效目标预视、双字词无效目标预视)。这样的检验才真正回答了词长(双字词 vs 单字词)是否影响眼跳目标选择的理论问题。

作者提到“在比较目标预视下双字词与两个单字词下的眼跳长度,仍然没有发现显著地

差异。”我认为这样的统计和结论存在两方面问题：首先，其数据量较少，统计力有限。其次，作者对非显著结果的解释不够严谨，不显著的检验结果不能支持或反对虚无假说。具体来说，如果不显著，不能作为支持或反对眼跳目标理论。当然，最后关于模型的使用，以及结果的解释，我只是提出建议，请作者自行定夺。

回应：根据专家的建议，为了检验副中央凹视区中的字能否构成词是否影响随后的眼跳目标选择，我们进行了副中央凹双字词预视与两个单字词（包括双字词假字预视、两个单字词目标预视、两个单字词假字预视）预视的比较，在眼跳长度及平均首次注视位置、单次注视位置上发现了显著地的差异，双字词条件下的眼跳长度更长，注视位置更远离词首靠近词的中心（见 p.8-9）。

同时也对我们的讨论部分进行了相应的修改（p.11），如下：

“首先，副中央凹中词长信息对随后眼跳目标选择产生显著地影响。在对副中央凹双字词预视与两个单字词（包括双字词假字预视、两个单字词目标预视、两个单字词假字预视）预视进行比较时，发现副中央凹中的两个字能构成词时，随后的眼跳长度更长，平均首次注视位置以及单次注视位置更远离词首而靠近词中心的位置。这说明读者在副中央凹的加工中可能已经对副中央凹的字能否构成词进行了一定的识别，进而影响了随后的眼跳目标选择。”

我同时拜读了审稿专家 2 的建议，由于我自己对某些问题也有粗浅的认识，所以希望以我个人有限的知识，进行一些补充说明。

意见 3：首先，关于 Yan et al. (2010)和 Li et al. (2011)之争。对于专家 2 指出的 Li et al. (2011)

“用一种简单的“恒定长度”的模拟也能产生 Yan 等人观察到的数据模式”的这一论述，我觉得尚有待商榷。Yan et al. (2010)中的恒长眼跳模拟发现，随机模型确实可以*部分地*生成其观察到的数据模式，但*并非全部*；因此，基于这样的结果，Yan et al. 指出，中文的眼跳目标选择不能由随机模型来解释。请审稿专家 2 和作者参见该文于 717-719 页之描述。同时，请作者考虑修正文章中对相关问题的描述（如第二页）。

回应：已根据专家建议，进行了修改，修改如下：

“但是研究者并不认为汉语读者是采用恒定步幅的策略来选择眼跳目标的（Li et al., 2011），同时虽然这种眼跳策略可以对注视位置分布进行较好的模拟，但是并不能解释全部的数据结果，例如跳读率随预测性的增加而提高，也不能很好的解释词长、词频与注视可能性之间的关系（Yan, et al., 2010）”

我认为 Yan et al. (2010)和 Li et al. (2011)之争，是对同样结果的不同解释，都有其合理一面。正如前文提及的，我个人对中文眼跳目标选择的问题并无任何理论偏见，同时也尊重专家 2 的观点；但，正如专家 2 所指出的那样，目前 Li 组尚没有直接证据来推翻 Yan 等人的观点，在这种情况下，正如专家 2 本人指出的“科学的游戏是用 positive 的结果证否”，专家 2 认为“Yan 等人的理论正确的可能性基本上没有”的证据为何？类似地，专家 2 指出“中文的眼跳目标选择显然既不是基于词，也不是基于字的”这一论述中用了“显然”一词，似乎在暗示已经盖棺定论，其证据又为何？同时，基于字、词或随机的眼跳模型，这几种说法在学术上都有前人提及，我尚未理解专家 2 为何简单将这些分类斥为“粗暴”。

关于结果的描述，我个人认为“随后的注视位置更靠近词的中心位置”之类的描述并无不妥：这仅是对观测到现象的描述，对于同样的现象完全可以有不同的理论解释。因此，专家 2 认为“这一表述有错误”的论述似有不妥：表述本身确实无误（参照表 2）。专家 2 指出的“Yan 等人就用这个来说事”从科学角度似乎有欠严谨。在 Yan et al. (2010)一文中，所列出的实验结果远多于注视点位置分布。我不要求作者采用某种特定方式表述实验结果，我认为更重要的是推理严谨。

意见 4: 作者的观点“对于汉语阅读中眼跳目标选择策略的研究中，应该同时考虑当前注视词和副中央凹中的加工特征，以及二者可能存在的交互作用。汉语读者可能同时运用当前注视词的加工信息和副中央凹中的加工信息来选择随后的眼跳目标，两种信息对于汉语读者的眼跳目标选择来说都产生重要的影响”是否应归属于 Liu et al. (2015)，我个人认为有待商榷。如作者所述，“当前注视词的加工调节副中央凹中的加工”这一结论在 Liu 文之前就有中、英文研究涉及。如，眼跳距离、注视点位置受到加工负荷的影响(Inhoff, 1989; Hirotni, Frazier, & Rayner, 2006)，也有中文的研究得到类似的结论（如 Luo et al., 2013）。可能 Liu 文确实有提及该观点，但我想指出，从广义角度来看，该观点或许存在更久远的历史。

回应: 已根据专家建议，进行了修改，修改如下：

“在拼音文字的研究中发现，不仅副中央凹的加工影响眼跳目标的选择，中央凹注视词的加工也会对眼跳目标的选择产生影响(Hirotni, Frazier, & Rayner, 2006; Inhoff, 1989; White, & Liversedge, 2006)，类似地在汉语阅读的研究中同样发现了副中央凹与中央凹的加工都会影响眼跳目标的选择（Li et al., 2011; Luo et al., 2013; Wei et al., 2013; Liu et al., 2014）。因此在之后的研究中应该同时考虑中央凹和副中央凹中的加工特征，以及二者可能存在的交互作

用。汉语读者可能同时运用当前注视词的加工信息和副中央凹中的加工信息来选择随后的眼跳目标。本研究的结果为这一观点提供了新的实验数据支持，并且说明了不仅副中央凹中信息的多少会影响随后的眼跳目标的选择，副中央凹中获取的预视信息的特征也会影响随后的眼跳目标的选择。”

由于我对专家 2 的审稿意见阐述了若干条粗浅的个人见解，为避免误会以及更好地学术交流，我对审稿意见进行签名：

（根据规定，编辑部删除了审稿人的签名，请理解！）

审稿人 2 意见：

该文经作者认真修改后，质量已大幅提高。为提高质量，提供如下参考：

意见 1: p.6. 在进行 LMM 分析时，仅“指定被试和项目作为交叉随机效应”是不够的。而应该采用最大随机效应设置，即考虑到固定效应的 slope 在被试和项目上也有变异（参看，Barr, Levy, Scheepers, & Tily, 2013）。仅指定被试和项目作为交叉随机效应的最简单模型不能很好避免 anti-conservative，往往结果可靠性还不如常规的 F1 和 F2。不过，考虑到该文报告的显著性都较大，采用最大化随机效应可能也不会改变该文的主要结果。所以不强求。尤其是考虑到国内 LMM 和 R 使用都还处于初级阶段，要求太高，不利于推广。

回应: 本研究中所使用的数据分析方法与 Zang 等人（2015, just-accepted, QJEP）的研究中的数据分析方法相同，在分析中使用“full”随机效应模型，在随机效应的截距与效率的计算中包括了被试与项目的效应以及两者的交互作用。

意见 2: p.9-10.统计结果 3.2 部分（即字 N+1 与字 N+2 注视结果分析）可读性不高，意图也不明显。唯一可能相关的只有讨论部分一句话，“字 N+2 的预视会对接下来字 N+1 的加工产生显著的影响”。作者应该在在报告这部分结果之前清楚表明它的目的是什么。能说明什么问题，哪怕就是为了跟前人对比之类的。

回应: 根据专家的建议进行了相应的修改，修改如下（p.9, p.10）

“已有研究发现操纵词 N+2 的加工会影响词 N+1 的加工（Kliegl, Risse, & Laubrock, 2007; Radach, Inhoff, Glover, & Vorstius, 2013; Yang et al., 2009; 白学军, 王永胜, 郭志英, 高晓雷, 闫国利, 2015），因此本研究中分析了字 N+1 的注视情况，以探讨副中央凹的字能否构成词是否影响这种延迟的副中央凹-中央凹效应。”

“为了检验能否在副中央凹中对字 N+2 进行预视加工，分析了字 N+2 的注视情况。”

意见 3: p. 8. “表明目标词类型对从前目标区到目标区的眼跳长度没有产生显著性地影响”最好删除。因为词的类型和预视条件的交互作用已经表明词的类型可以影响眼跳长度。只是没有主效应而已。没有主效应其实很容易理解: (1) 假字预视条件下不可能有词的类型效应; (2) 目标预视条件下没有词的效应是 power 不够而已 ($t=1.65$, 已经非常接近 $p=0.05$ 的 t 值), 主要原因是第二个字的频率在两个单字词条件下高于一个双字词的。可能会拉平词的类型效应。虽然有没有词的类型效应并不能改变这篇论文的贡献, 还是希望作者可以理解到位。

回应: 根据专家建议, 已将此处删除。并且根据专家建议我们对副中央凹双字词预视与两个单字词 (包括双字词假字预视、两个单字词目标预视、两个单字词假字预视) 预视的比较, 在眼跳长度及平均首次注视位置、单次注视位置上发现了显著地的差异, 双字词条件下的眼跳长度更长, 注视位置更远离词首靠近词的中心 (见 p.8-9)。

意见 4: p.9-10. 表 5 和 7 标题中, “格” 改成 “各”

回应: 已经对此处错误进行了修改。

意见 5: 删除参考没用的参考文献, 如 “Li, Rayner, & Cave, 2009”。请仔细检查参考文献部分。

回应: 根据专家建议对参考文献进行检查, 添加或删除了相应的参考文献。

第三轮

审稿人 1 意见:

经过第二次修改, 我认为稿件在文献综述的完整性、数据分析的合理性、结果解释的逻辑性等方便都有明显改进。同时, 作者很好地回答了我提的审稿问题。

本次修改中, 我对作者的论述有一些小建议。

意见 1: 关于句法信息加工的问题。我同意作者提出的观点, 即, Kim et al. (2012) 和 Luo et al. (2013) 的实验中都使用了较为外显的句法线索 (case markers 或韵律结构), 这些外显的信息可能更有利于读者利用句法信息。但, 我们并不知道对句法信息的利用是否*仅仅*依靠外显线索; 因此, 如果认为在没有外显句法信息情况下不进行句法加工, 这样的论断似乎在逻辑

辑推论上有所欠缺。因此我建议作者对类似“本研究中句法结构变量的混淆可能是不存在的，或者作用是较小的”的论断稍做修改。我相信，如果以更开放的形式来讨论句法结构可能带来的影响，将更有利于后续实验研究的开展。

回应：非常感谢专家指出我们在逻辑上的欠缺。根据专家的建议，我们对之前的论断进行了修改，指出句法信息是否对本研究的结果产生影响并不明确，并且在缺少外显的句法线索时读者是否能够从副中央凹中获取句法相关的信息并对随后的眼跳目标的选择产生影响，需要今后进一步的研究。

意见 2：本文作者指出“在拼音文字的眼动控制中 *where* 主要受低水平信息的影响，而不受语义信息的影响 (Rayner, 2009)，因此在汉语阅读中句法结构信息影响注视位置而不影响注视时间的可能性较小”。对这一论断，我认为可能存在文献上和逻辑上的一些偏差。首先，目前关于眼跳目标选择这一研究主题，研究者确实更强调低水平信息的影响；但事实上，存在研究证实了高水平信息（如频率、语素结构等）也可以参与眼跳目标选择（例如，可以参见 Yan et al., 2014, *Cognition* 一文的文献综述部分，以及该文的实验结果）。因此，建议作者在引述 Rayner (2009)这一文献时，同时考虑来自另一方面的证据。其次，众所周知，中、英文存在诸多差异，与本文主题有关的方面涉及到眼跳目标选择、副中央凹信息提取等等，因此以英文阅读的结论来推论汉语加工，在逻辑上是否靠得住？

回应：根据专家的建议，我们引用了关于高水平信息对眼跳目标选择的影响 (Yan, Zhou, Shu, Yusupu, Miao, Krügel, & Kliegl, 2014)；并且引用了之前汉语研究注视位置的相关研究 (Ma, & Li, 2015; Shu et al., 2011; Zang, Liang, Bai, Yan, & Liversedge, 2002; 臧传丽, 孟红霞, 白学军, 闫国利, 2013)，确实也发现注视位置主要受低水平信息的影响。我们增加了相关的文献，以增加我们逻辑推理上的可靠性。