

《心理学报》审稿意见与作者回应

题目：中文读者词汇视觉信息获取速度的发展——来自消失文本的证据

作者：闫国利，刘妮娜，梁菲菲，刘志方，白学军

第一轮

审稿人 1 意见：

意见 1：作者在提出问题时说：“然而在信息加工的时间方面，即提取或编码视觉信息速度上是否存在差异还有待进一步验证。”，和前面提到的已知：“随着读者生理年龄或阅读技能的不断提高，阅读所需要的总阅读时间和平均注视时间减少”似乎存在矛盾。阅读时间和平均注视时间减少，似乎一定程度上就相当于提取或编码视觉信息速度的提高。建议对问题的描述更加精细化，使得已知和问题的关系更加清晰。

回应：按照您的建议，为了使得已知内容和研究问题之间的关系更加清晰，我们在前言介绍中，重新组织和增加了相关信息。具体见前言 3~5 段落的蓝色标注内容。

意见 2：有些表述不够稳妥。例如，根据总阅读时间的统计结果，作者得出结论说“消失文本操作对读者总阅读时间的影响只表现在 20ms 条件上，对大于 20ms 的条件没有影响”；“也就是说 40ms 及其以上的消失文本条件并不干扰读者的总阅读时间”。该结论不是很妥当。应该是“在大于 20ms 的实验条件中，消失文本操作的影响没有达到显著水平”，“干扰效应不显著”。类似的表述不一一列举。

回应：非常感谢您提出该类问题，我们按照您的建议，在原文中找到了类似不稳妥表达的地方，并逐一做了修改。例如：

- 1) 摘要中：“结果发现：文本呈现时间大于 40ms 时，消失文本操作对儿童和成人的阅读理解率和总阅读时间的影响均没有达到显著水平。”
- 2) 讨论第一段中：“然而在进一步的事后检验中发现：导致总阅读时间的增加主要来自于 20ms 消失文本条件，也就是说在大于 20ms 的实验条件中，消失文本操作的影响没有达到显著水平。”
- 3) 结论中：“当文本呈现时间大于 40ms 时，儿童和成人的阅读理解率和总阅读时间均没有受到显著的干扰…”

意见 3：鉴于本研究的焦点是儿童和成人的对比，建议研究的关键假设集中在儿童和成人的关键差异上，目前聚焦程度有些不够。

回应：参考您提出的建议，我们对假设部分进行了修改，具体见原文前言倒数两段部分的蓝色标记内容。

审稿人 2 意见：《中文读者词汇视觉信息提取速度的发展——来自消失文本的证据》一文，研究方法恰当，得出了有价值的结果，但存在一些问题：

意见 1：实验材料的评定：对句子的难度和通顺性评定，结果分别报告了小学生和大学生的平均数以及两者的比较，没有对得分低的句子进行修改，直接作为实验材料，不管对实验结果有无影响，这样做都缺乏严谨性。

回应：感谢审稿专家的意见。我们有一个问题没有表述清楚：我们最终所使用的材料是反复评定和修改 3 遍后的结果，每一次都会对得分低于 3 分的句子进行修改。因此，研究所使用

的句子在每个被试评定上都是大于 3。除此，之所以选择 3（而不是更高分）作为判断标准，是因为本实验材料的特殊性：双字词句子。由于部分几个句子缺少“的”，“了”（但不影响句子意思表达）的方式可能让句子稍显不自然。例如：“农民普遍反映今年庄稼长势良好。”，读者可能更习惯：“今年的庄稼”。从而影响了整体评定的结果。

意见 2: 被试人数：5 组实验材料，每名被试只阅读其中一组，25 名大学生分配到 5 组，每组平均只有 5 个人，虽然是被试内设计，人数还是偏少，至少要与小学生人数相当；30 名小学生可以接受，因为到小学做实验的难度很大。

回应: 同意审稿专家的看法，在被试量上我们可能没有做到足够大。之所以成人选择 25 人，是因为参考了同类研究中被试的数量。Blythe et al., 2011；2009 研究中，儿童和成人分别是 16 名，其中被试内文本呈现条件包含 4 个水平；Joseph et al., 2009 同样的发展研究，成人和儿童被试各 20 名，其中被试内变量也是 4 个水平结合。本研究的被试内变量是 5 个水平，每个被试阅读 75 句正式实验句，即每个条件下 15 句。同时研究结果中差异显著性较高，结果较为稳定。其次，成人和儿童数量不完全相同，因为我们先做的成人实验，按计划成功收集了 25 名被试的数据。在接下来儿童实验中，考虑到儿童无法按要求完成实验的可能性要比成人高，所以在小学预约的被试比预设的多。但实际中，儿童被试出乎意料都完成得很不错（正确率较高，追踪丢失的数据很少，校准也都达到了 0.2 的较高标准）。因此，最后我们将成功收集的 30 名儿童数据均纳入了分析。

虽然我们选择 25 人参考了相关国外的研究，我们觉得审稿专家提出的建议非常好，在今后的研究中，我们一定会注意这个问题。

意见 3: 因变量的确定：对整句和每个词进行分析，而没有确定和分析关键词，这与这类研究的作法有区别，不知原因是什么。

回应: 我们出于以下原因而没有设定关键词：首先，同类研究所设置的目标词，主要是在消失文本条件下，通过词频效应验证眼动的认知控制理论（例如：所注视文本的视觉信息在非常短暂时间后消失，不能再提供信息。但读者在该区域的注视时间能达到与正常注视接近（200ms 左右），特别是注视时间同样受到词汇属性（词频）的影响。），然而这一观点已经得到大量研究的支持。因此，本研究没有进行考察。

其次，Blythe et al., (2009) 的研究中考察词频效应还有另一个目的是：词频效应反映词汇识别的难易程度，因此在消失文本条件下，是否存在词频效应反映读者能否产生正常的词汇通达。然而，这项研究发现，词频效应显著存在，但与文本呈现时间没有交互作用，即词频效应并不受消失文本操作的调节。另一项拼音文字研究（Rayner et al., 2011）也得出类似的结果。在中文中，我们有另一篇考察词频效应的研究（正在进行中）中，读者的阅读效率明显受到极短消失文本操作干扰的条件下，也出现了与正常文本阅读同样大小的词频效应量。鉴于此，我们认为，采用词频效应并不能有效说明在文本短暂呈现时间内，读者能够获得到足够的文本信息。此外这样的操作还会使实验设计和统计处理变得复杂。综上三方面的原因，我们没有设置词频变量。

此外，为回答本研究的问题，我们采用阅读效率的重要指标：总阅读时间、阅读理解率。以及回视次数，这样的基于整体分析的重要指标来反映所研究的问题。同时，为考察消失文本操作是否对读者早期加工产生影响，我们也采用 First fixation duration, Gaze duration 等指标进行了基于局部的分析。参考有关研究，在数据处理上没有分析句子的首词和尾词，对于余下的双字词，每个双字词为一个目标进行计算，最后求所有目标区在各项指标上的平均值。

意见 4: 数据结果与分析：选择呈现时间为 20ms/40ms/60m/80ms 的依据是什么？“导致总

阅读时间的增加主要来自于 20ms 消失文本条件，也就是说 40ms 及其以上的消失文本条件并不干扰读者的总阅读时间”，30ms 的条件会怎样？“考虑到实验仪器本身 10~15ms 的延迟，文本真正呈现给读者的时间应该是操作时间与仪器延迟时间之和。因此，本研究中读者在阅读短暂呈现 50~55ms 以上的消失文本时总阅读时间几乎不受影响。”仪器为什么会有 10~15ms 的延迟，以英文为材料的研究中，是否也考虑了仪器延迟的时间？如果仪器没有延迟，就不能得出“读者在非常短暂的文本呈现时间内能够快速提取到足够的中文视觉信息，这表现出一定的跨语言一致性”的结论。而且，若考虑 10~15ms 的延迟，文中其余所有涉及时间的结论都要照改，不知作者是否考虑过这个问题。

回应：非常感谢您提到这个问题，这也是我们一直思考，并努力在解决的问题。

任何呈现随眼动变化技术（gaze-contingent change paradigm）都会存在这样一个延迟。如果需要计算出精确时间，就需要考虑显示器的刷新率。消失文本操作中，当眼睛注视一个词开始计时，在预设时间后消失，呈现新屏（空白）。如果预设时间结束时，显示器刷新扫描刚好已经通过了当前文本行，那么这时将产生最长的延迟，但如果扫描刚好在当前文本的上一行，那么便能最快的呈现新屏。由此可见，延迟时间在每一次屏幕变化时刻都可能不同，即时刻处在一个变化范围中。如何算出每个变化的延迟时间，将其精确算入我们的数据中，可能今后还需要继续去解决。也因此，本研究并没有给出确切的时间。

其次，这样的延迟在拼音文字研究中也不例外，由于所使用显示器刷新率不同，延迟时间变化范围也都不尽相同。拼音文字中最早确定时间的研究显示，操作 50ms 时不影响阅读效率（Rayner et al.,1981）；Blythe et al.,（2009）研究中发现操作 40ms 条件也不影响阅读效率（这与早期研究并不矛盾）。粗略加上延迟时间，拼音文字中读者所需要文本呈现时间不超过 60ms；在本研究中，同样发现 40ms 开始不影响读者阅读效率，加上延迟时间，大概在 50~55ms。基于这样的对比，研究者认为，两种书写语言下，读者都能够在早期注视非常短的时间内内容获取到足够的视觉信息，这表现出跨语言的一致性。但两种书写语言系统下，读者在具体时间需要上可能还无法直接说完全一致。因此本研究所做一些结论可能还不够严谨，我们在语言的表达进行了如下的修改，将具体时间作出的“这与拼音文字的研究结果一致”的观点进行了删除。保留了，“读者在非常短暂的文本呈现时间内能够快速获取到足够的中文视觉信息，这表现出一定的跨语言一致性”的观点。”

最后需要回答的是，之所以选择 20ms/40ms/60m/80ms 来操作，主要是基于以往中文中采用消失文本范式研究以及结合拼音文字研究的结果。例如：闫国利等人（2007）（消失文本条件下认知控制的眼动研究）发现，80ms 不影响正常阅读，且出现词频效应；而刘志方等人（2011）.（汉语阅读中眼跳目标选择单元以及词汇加工方式：来自消失文本的实验证据），的研究中，消失文本条件选择的是 40ms，发现对成人读者总阅读时间没有产生显著影响；Yan, G. L. et al(2010)(Chinese readers' eye movements when reading disappearing text.)同样发现，成人读者 40ms 呈现时间同样对总阅读时间没有产生明显干扰。当前研究发展的的问题，我们设定 80ms 为最大值，20ms 为最小值。在保证时间间隔相同的情况下，减少变量水平太多带来统计上的复杂，选择了每间隔 20ms 为一个水平。其次，考虑到延迟本身会带来大约 10ms 的误差，所以，目前没有对更精确的时间进行考察。

意见 5：需要细化的问题：

（1）作者在前言中写道，“拼音文字的研究发现，当被注视词呈现至少 60ms 时，读者能够正常阅读和理解文本。需要说明的是，60ms 不影响读者的正常阅读，并不代表读者在 60ms 内能完成对所注视词的全部加工，而只是在短暂的呈现时间内能够获取到词汇识别所需的足够视觉信息。既然在 60ms 内只能获取词汇识别的视觉信息，就不能说此时读者能够正常阅读（reading）和理解（comprehension）文本，很可能读者只是能够跟读（reading after

words)。

(2)前言中多次提到“提取或编码视觉信息”，该研究探讨的是提取还是编码视觉信息？编码和提取是信息加工的不同加工阶段。

(3)让参加实验的大学生和小学生进行句子通顺性和难度评定，而不是另外的大学生和小学生，与常规做法不同，是基于什么样的考虑？

(4)在考虑“词长”影响读者获取视觉信息的速度这个问题时，是不是也应该同时考虑词频？在英语中，长词的词频一般低于短词，在汉语中，使用双字词控制了词长，并不能保证词频相近。

(5)英语的词长和汉语的词长是否是同一个概念？英语词由字母组成，字母本身没有意义，汉语词由汉字组成，汉字是有意义的；英语字母和单个汉字是否可以对应。

(6)考虑到读者层次，前言中，对再注视和回视可以简要说明，以便较快理解“中国读者的回视比率比美国读者高，而美国儿童比中国儿童的再注视比例高，但是两国成人读者之间没有差异。”实验部分，因变量指标的界定，如平均注视时间、跳读率、回视时间、单次注视的登陆位置、多次注视的首次登陆位置等。

回应：(1)该观点引用于 Rayner et al.(2011)的研究。本研究作者的理解是：虽然 60ms 只能获取到词汇识别的视觉信息，但读者在该词上的注视时间与正常文本接近(200ms~350ms)，所以在文本消失之后剩下注视时间里能够完成对所注视词的全部的词汇加工。从而在总阅读时间和理解率指标上反映出正常的加工和理解。

(2)非常感谢您提出这个问题，我们在用词上存在疏忽。首先，这里的“提取”可能容易被理解为“retrieval”，而实际想表达的意思是“获取”(capture or extract)。所以本研究探讨的问题是：为保障正常词汇通达，读者在注视的早期阶段获取足够视觉信息的速度。据此，对文章中所有涉及该概念的地方均统一修改为“获取”。

(3)这里是我们信息呈现不完整的问题。事实上，评定材料和参加眼动实验的是不同的被试，与常规做法一样。鉴于信息呈现的疏漏，我们在文章中做了补充说明（见正文实验材料部分蓝色内容）。

(4)的确存在您提到的这个问题，我们也进行过考虑。但参考相关拼音文字的研究，他们并没有发现：被注视词的词频会调节读者获取文本信息的速度。同时，我们在中文的考察中（撰写中），也没有发现词频与文本呈现时间之间的交互作用。关于这个问题，在我们后续的研究中也还在进行验证。

(5)英语的词长对眼动行为有一定影响。而汉语则没有相应的词长方面的相关研究。我们认为，无论是汉语词长还是英语词长，它们对眼动行为都存在一定的影响。比如，词长，其占的物理空间就大，词长短，其占物理空间就小。词的长度会影响到眼跳计划，影响眼跳的落点位置和眼跳距离的大小。因此控制词长可以最大限度的减少由词长引起的无关影响（例如：登陆位置，眼跳距离，注视时间）。

但是，中文词长与英语的词长也存在着重要不同，英语词长度不同，所包含的字母数不同，而中文的词长不同，除了包含词中所含的汉字数不同外，每个汉字的笔画数、汉字的字频也是不同的，就此而言，汉字的词长应该涉及到比英语单词更多的因素。因此，将英语词长与汉字词长简单地进行类推，的确有不妥之处。

非常同意您的观点。就此我们在文中对相关描述也进行了修改（见前言第7段的蓝色内容）。

(6)参考您的建议，我们在文中增加了所有指标的说明。

前言中：“但是在回视比率（指阅读中从后面区域回视到当前兴趣区的比率）以及再注视比率（首次阅读中兴趣区被多次注视的概率）上，两种语言儿童之间出现发展性的差异。”

结果部分的指标介绍：“整体分析，以整句为兴趣区对阅读过程中所有注视和眼跳行为

进行分析。具体包括如下指标：总阅读时间（兴趣区所有注视点的时间总和）、总注视次数（兴趣区所有注视点的次数的总和）、平均注视时间（兴趣区所有注视点的持续时间的平均值）、回视次数（从当前注视词向之前某个词语进行眼跳的次数的总和），再注视比率（首次阅读中，兴趣区被多次注视的频率与该兴趣区被单一注视和多次注视的频率之和的比值）以及跳读率（首次阅读中，兴趣区被跳读的频率与该兴趣区被跳读和被注视的频率之和的比值）；(2)局部分析，以词为单位进行分析(排除词首和词尾)。具体包括如下指标：首次注视时间（首次通过阅读中某兴趣区内的首个注视点的注视时间）、凝视时间（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）、回视时间（所有回视到当前兴趣区的注视时间之和）以及登陆位置（指注视点所处的位置）。”

意见 6：信息不完整的地方：

- (1) 对“语言熟悉假说”没有必要的介绍或出处。
- (2) 没有明确写出研究目的，可能给读者带来不必要的理解困难，而且也不符合写作规范。
- (3) 注视时间小于 80ms 或者大于 1200ms，需标注参考文献。
- (4) 实验程序不够清晰，如被试怎样开始实验、是否自定步调等都没有交代。
- (5) 分析的是哪只眼睛的数据？

回应：非常感谢您具体指出了这些信息不完整的地方。参照您的建议，我们在文中逐一增加了相应的信息（正文相关的蓝色标注内容）。

- (1) 对“语言熟悉假说”没有必要的介绍或出处。（见：前言第 3 段开始部分）
- (2) 没有明确写出研究目的，可能给读者带来不必要的理解困难，而且也不符合写作规范。（见：前言倒数第 2 段开始部分）
- (3) 注视时间小于 80ms 或者大于 1200ms，需标注参考文献。（见：3 结果的第 2 段）
- (4) 实验程序不够清晰，如被试怎样开始实验、是否自定步调等都没有交代。（见：2.5 程序部分）
- (5) 分析的是哪只眼睛的数据？（见：2.5 程度的（3））

意见 7：需要重新考虑的问题：

- (1) 既然明确写出了研究假设，就需要表述清楚以下两个问题：

第一，如果出现其一结果，说明什么问题；如果出现其二结果，说明什么问题；该研究预期出现哪种结果。

第二，实验操作了 5 种呈现时间，根据该研究设想，成人和儿童在哪些时间上会表现出差异。

(2) 在讨论部分，“本研究认为再注视在阅读过程中不仅限于单一作用，而上文提到的三种假设：提取视觉信息的作用、避免新信息进入而干扰当前词汇加工的作用和仅仅是一种阅读习惯，都同时存在。”从上文的分析中，似乎得不出这样的结论。

(3) 对被试内因素呈现时间的 5 种条件下各因变量指标的两两比较采用了 t 检验，是参考了 Blythe et al. (2011) 的做法，在这种情况下，t 检验是不合适的。Spss 软件里有多重比较，通常对应的是 F 检验。

回应：(1) 参考您提出的建议，我们对假设部分进行了修改，具体见原文前言倒数两段部分的蓝色标记内容。

(2) 恩，同意您提出的这个问题。在当前研究中，我们确实无法对每种猜测做具体的验证。为此，本研究尝试结合数据结果，分别对每种假设做出有限的支持。首先，当消失文本时间不能满足读者获取到足够视觉信息时（20ms,40ms），再注视概率降低，读者的回视

次数显著增加，这支持再注视可能与获取文本视觉信息有关；其次，当消失文本操作时间大于读者所需要的时间时（60ms 和 80ms），再注视概率依然相比正常文本显著降低，但此时回视并没有增加，这可能说明，再注视不仅是为了再次获取视觉信息，而也有可能是读者在阅读发展中保留下的一种习惯，这与 Hazel 等人(2011)的研究结果一致。最后，研究中所有消失文本条件下读者再注视概率降低的同时，平均注视时间都显著增加（尤其是反映早期加工的首次注视时间），这符合文中提到的第三种猜测——避免新信息进入的干扰。

然而我们也觉得，这种选择性的解释并不严谨。因此，为揭示再注视在阅读中的作用，可能需要接下来进一步专门的研究去做具体的实验设计，从而提供直接的证据。鉴于文章中对上述想法的阐述不够清晰，我们也做了修改，见讨论 4.1 中第 3 段的蓝色标注内容。

(3) 参考您给出的建议，我们重新使用 spss-repeated measures-contrast 里面两两多重比较的 simple 和 repeated 两种方法分别进行了各消失水平与参考水平的比较，以及相邻近水平的连续比较。并在文章中进行了相应修改。（数据结果在此不逐一呈现，具体见原文结果部分的蓝色标记内容。）

最后，非常感谢两位审稿专家给我们提出的宝贵意见，我们针对您的意见和建议进行了较为细致的回复。如果专家有进一步的建议，敬请赐教，我们一定认真修改。

第二轮

审稿人 1 意见： 修改稿较好回应了我的问题，建议接受。

审稿人 2 意见

意见 1： 作者已经回答了大部分问题，还需要补充统计检验效力和效果大小值。

回应： 按照您的建议，已经在文章中“3 结果”部分补充了统计检验力和效果大小值（见蓝色标注内容）。同时参考学报以及其它一些文章，通常只呈现效果量大小，因此有个小疑问想征求您的建议，即本文是否也可以只呈现两者中的一个。

意见 2： 另外，关于“但参考相关拼音文字的研究，他们并没有发现：被注视词的词频会调节读者获取文本信息的速度”这个回答，我不能同意，可以参考 Rayner & Duffy(1986)、Staub, White, Drieghe, & Rayner (2010)、Liu & Perfetti (2003) 等等的研究。

回应： 非常感谢您对本研究提出诸多全面的，建设性的问题和建议，我们从中受益匪浅。

在认真阅读了您推荐的几篇文献之后，我们觉得之前的回答存在不妥之处。首先，对于“但参考相关拼音文字的研究，他们并没有发现：被注视词的词频会调节读者获取文本信息的速度”的概括表述不准确。已有研究只是发现，在不同的消失文本条件下都存在显著的词频效应，只是不同消失文本条件下词频效应量没有差异，即文本呈现条件与词频之间不存在显著交互作用（Blythe et al., 2009；本文作者在另一项实验在中文阅读中也发现了同样的结果）。此外，关于为什么在不同消失文本条件下词频效应量没有差异，我们觉得这也是一个值得深入探讨的问题。

最后对您的工作再次表示诚挚感谢！

第三轮

审稿人 1 意见： 有以下建议：

意见 1： 实验设计方面，作者谈到希望在中文中可以找到只使用双字词组成的通顺句子，这

种使文本中所有词都具有相同长度的做法在任何拼音文字的研究中都很难实现,因此他们采用中文词长的这种独特性在中文阅读中继续探究儿童和成人获取视觉信息速度的问题。但事实上这个切入点有些问题:英文词长之所以是一个重要因素,主要是因为词与词之间有空格,词的物理形状特征非常显著。而在中文中,很多研究已经表明,字是重要的知觉基本单位,因此,双字词从句子中切分出来是否能在那么短暂的呈现时间内完成都是问题;而且,全部由双字词组成的句子并不常见,这种设计首先可能带来额外的策略影响,利用这种过于规则的材料所获得的描述性结果价值有限。其次,相比于西文,由于缺乏词界,笔划、部件、字频、词频可能是影响早期加工更重要的因素,而作者完全不考虑这个因素应该说是一个缺陷,作者至少也应该报告这些基本的数据的情况。事实上,由于太多变量混杂其中,整个研究偏于描述性的分析,在结果的理论解释方面较为牵强。作为一项描述性研究,作者应该考虑使用高阶的统计方法,将混杂在其中各种变量对眼动发展模式的某一效应进行分析,结果才会更有价值。

回应: 非常感谢审稿专家提出的这些重要问题,对本研究具有很好的启示。

第一,关于使用双字词句子的问题,我们有如下几点原因。

(1) 双字词组成的句子不常见,这个问题的确存在。我们在编制材料的时候也考虑到了这个问题,因此尽量使造出来的句子看起来常见并且读起来通顺。通过最后的5点通顺性评定(儿童为 $M=3.56$ ($SD=0.26$), 大学生为 $M=3.64$ ($SD=0.19$)), 发现结果在可接受范围内。

(2) 选择双字词句子进行研究,是为了尽可能减少由词长带来的各种混淆效应以及词切分不一致的干扰。具体来说,正常语料句子中包含单字词,双子,三字,甚至是四字词,研究发现不同长度的词会影响读者的登陆位置,眼跳幅度,注视时间等,因此这些变量也都可能影响读者对文本呈现时间的需要。采用词长固定的句子进行研究,能够最大限度的减少词长不同而引起无关变量产生的干扰。其次,词切分的很多研究发现,不同的中文读者对词的边界的划分存在不一致性 (Hoosain, 1992; Tsai, McConkie, & Zheng, 1998; Miller et al., 2007; Liu, Li, Lin & Li, 2013)。如,很多人把“美丽的”划分为一个词,而另一些人把它划分为“美丽”和“的”两个词。Peng 和 Chen (2004)在研究中发现了一个有趣的现象:大多数的中文读者倾向于将单音节词和双音节词连在一起,形成一个词,并将此现象命名为“单词扩展假设”。Liu 等人 (2013) 采用词切分任务更为直接地考察了中文读者的词切分情况。结果发现,大多数情况下,被试并不认为助词,数词,量词等是单一的词单元,他们倾向于将功能词和实词联系在一起,形成单一的词单元。而助词,数词和量词等功能词大多为单字词。鉴于中文阅读中词单元的不确定性,研究采用规则的双字词句子,按照双字词单元进行消失操作,能在一定程度上减少词切分差异带来的影响。

(3) 采用相同词长的句子进行阅读加工研究的做法,在以往中文阅读的同类研究中也存在,例如:

刘志方, 张智君, 赵亚军. (2011). 汉语阅读中眼跳目标选择单元以及词汇加工方式: 来自消失文本的实验证据. *心理学报*, 43, 1-11;

张智君, 刘志方, 赵亚军, 季靖.(2012). 汉语阅读过程中词切分的位置: 一项基于眼动随动显示技术的研究.*心理学报*, 44, 1, 51-62;

刘志方, 闫国利, 张智君, 潘运, 杨桂芳(2013). 中文阅读中的预视效应与词切分. *心理学报*, 45, 6, 641~625;

(4) 您提到“字是重要的知觉基本单位,因此,双字词从句子中切分出来是否能在那么短暂的呈现时间内完成都是问题”。根据新近的中文词切分的观点 (Li, Rayner, & Cave, 2009), 词汇切分和词汇识别是一个统一的过程。即一个词被切分出来的同时也就被识别了。本研究从句子理解率和阅读效率的结果来看,即使最短的消失文本操作对读者句子的理解率和阅读效率的影响也非常的小,这表明读者的正常阅读加工并没有受到影响。进而说明双字

词的切分也可能没有受到影响。

然而，使用规则的双字词句子，很可能存在您提到的额外策略的变化，我们也将进一步考察相同词长句子阅读可能会与正常语料的异同，从而考察和弥补本研究可能会存在的策略变化。此外，在当前研究中我们会比较谨慎地下结论。

第二，关于字、词熟悉的问题，您提到“相比于西文，由于缺乏词界，笔划、部件、字频、词频可能是影响早期加工更重要的因素，……”。这些额外变量对实验结果的确有可能造成影响，这也是我们之后做进一步研究的方向。在这个实验中，我们做的是初步的研究，因此在一个实验中，难对实验句中的每个字词都控制笔画/频率，字频和词频，但是本研究的分析全部都是基于整句，即使是局部分析，也是基于整句中所有词（除首尾词）的均值。同时我们控制了整句的难度，保证了儿童和成人所阅读的句子在难度上不存在差异。这些方法在一定程度上控制了上述变量带来的影响。对于这些字词属性变量是否影响读者对视觉信息的需要，以及如何影响，这些都需要在今后研究中逐一进行深入的考察。

最后，按照您的建议，我们在当前文章中增加了所有双字词的词频和笔画数的基本情况信息：“双字词的平均词频为：255.30 次/百万，(0.03~11080.02)，平均笔画数为 15.21(4~27)。”

意见 2：本研究的主要变量是词汇呈现时间，分别为呈现 20ms,40ms,60ms,80ms 后消失。然而作者对变量的操纵非常不准确，因为实验中显示器存在 10~15ms 的延迟。这个延迟时间是比较长的。作者在程序设置时，或者应该将延迟时间考虑在设计的呈现时间内，或者选用高刷新率的显示器。目前的这种作法不准确，作者应该至少将系统的延迟时间考虑进去。

回应：本实验采用的显示器 ViewerSonic P225f 属于 CRT 显示器，其响应时间非常短(大概 2ms)。屏幕刷新产生的延迟时间在 0~10ms。综合响应时间和刷新延迟时间，平均延迟时间为 6ms (最小为 2ms，最大为 12ms) (重新请计算机专业人士给出的估计值，之前文章中的 10~15ms 已经被修改)。然而，因为延迟时间并不恒定，始终处于动态变化之中(2ms~12ms)，我们难以确定每一次屏幕转换的准确延迟时间，因此并没有将其加入设定的具体时间中。为了便于读者的理解，我们将在文章中通过脚注的方式对延迟的问题进行说明。具体见文章“2.2 实验设计”中的脚注内容。

此外，鉴于延迟时间属于随机变量系统存在于各个条件中，本研究的主要目的是比较儿童和成人对文本呈现时间需要的差异，因此文章中使用的条件对研究的结果和结论不产生实质影响。但是，如何计算每个变化的延迟时间，将其精确纳入我们的数据中，是今后还需要继续认真解决的一个问题。

意见 3：除消失文本范式外，消失掩蔽范式也是此类问题研究中常用的方法，并且该方法可以更有效避免视觉后像的影响，作者似应在文中介绍一些消失掩蔽范式研究的结论，比如如果用消失掩蔽范式，当文本呈现时间大于 40ms 时，消失文本操作对儿童和成人阅读理解率和总阅读时间的影响如何，并侧重说明本研究为何选择消失文本而不是消失掩蔽范式。

回应：在同类问题的研究中，研究者们一开始使用的是消失掩蔽范式(我们通常也成为移动掩蔽范式)(Ishida & Ikeda, 1989; Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981; Slowiaczek & Rayner, 1987)，发现当注视词呈现 50~60ms 后被掩蔽，读者能够正常阅读和理解文本。后来采用消失文本范式的研究发现，被注视词呈现 60ms 后消失，阅读加工同样不受影响。(Blythe, Häikiö et al., 2011; Blythe, Liversedge et al., 2009; Liversedge et al., 2004; Rayner, Liversedge, & White, 2006; Rayner, Liversedge, White, & Vergilino-Perez, 2003;)。此后，研究者们开始采用消失文本范式进行同类问题的研究。但是有研究者认为，读者在阅读消失文本时有可能使用被消失词所形成的视觉后像，而掩蔽实验中，掩蔽能够防止视觉后像

的产生。为此，Liversedge et al. (2004) 比较消失和掩蔽两种范式下的阅读，发现非常相似的实验结果。这表明采用消失文本范式所得到的效应并不是由于加工视觉后像产生。因此消失文本范式和消失掩蔽范式都能够用于同类问题的研究。

然而，如果使用消失掩蔽范式会存在一个选择掩蔽材料的问题。因为不同掩蔽材料对阅读会产生不同的干扰效应。如何寻找到即能够起到掩蔽作用，又对阅读的干扰较小的掩蔽材料是较为重要的问题。在拼音文字的研究，掩蔽材料的使用较为统一。多数研究者均使用“X”来掩蔽拼音文字中的字母。这种掩蔽材料的选择主要是针对拼音文字的特点而定的，“X”符号简洁而有不具有任何意义，对阅读干扰较少。在中文的研究中，掩蔽材料的选择也需要考虑到中文自身的一些特点。但是目前就掩蔽材料的使用还没有一个统一的标准。鉴于此，本研究为避免掩蔽材料选择的问题，也采用消失文本范式进行当前问题的研究。

非常感谢专家给出的建议，我们在文章前言中对相应内容进行了补充，详见“前言”第2段蓝色标注内容。

意见 4: 作者在数据分析部分，重在比较成人和儿童在不同呈现时间之间的变化（比如在分析平均注视时间和回视次数时），而同一呈现时间下成人和儿童之间是否存在差异，也是非常有价值的，需要作者增加这方面的数据分析结果。此外，建议作者增加报告不同呈现方式下的正确率，并做统计检验，确保被试在不同呈现方式下对实验句子的理解程度没有显著差异。

回应: 您的建议非常好。我们在文章中增加了相关数据和分析结果。

首先对阅读理解率作了如下修改：“55 名被试在阅读理解题目上的总平均正确率为 91.5%，具体见表 1。其中，儿童和成人之间差异不显著 $F(1,53)=0.02, p>0.05, \eta^2=0.113$ ；文本呈现条件之间的差异不显著： $F(4,212)=1.15, p>0.05, \eta^2=0.021$ ；二者交互作用亦不显著： $F(4,212)=1.04, p>0.05, \eta^2=0.019$ 。表明儿童和成人都能很好的理解文本，且文本呈现方式不影响读者的阅读理解率。”

表 1 儿童和成人在五种文本呈现方式下问题理解正确率的平均数和标准差 (%)

	控制	80ms	60ms	40ms	20ms
儿童	88.1 (14.3)	88.3 (14.0)	95.6 (9.1)	91.6 (14.8)	93.3 (12.7)
成人	92.1 (14.1)	92.6 (11.8)	92.7 (16.7)	87.7 (17.2)	93.1 (15.6)

其次，对四个交互作用显著的指标：平均注视时间，再注视概率，回视次数和首次注视时间，增加了相同文本呈现条件上儿童和成人的差异比较。所得到的眼动行为差异与已发现的儿童和成人眼动行为差异的研究结果较为一致，例如，在各个文本呈现条件上，儿童的平均注视时间更长，回视更多等。由于这不是本研究重点要探讨的问题，我们通过脚注对其分析结果进行呈现。详见“3.1 整体分析”，中脚注 2。

意见 5: 作者在讨论部分，对眼动指标的讨论很不清楚，导致很难清晰地认识到成人和儿童在眼动模式上的差异，表现在作者主要讨论平均注视时间和回视次数这两个眼动指标，以及分析二者之间的关系。然而平均注视时间和回视次数这两个眼动指标并不存在直接的关系。如果作者在讨论部分按照传统的分析方法，在首遍阅读时间，回视（跳进，跳出），再阅读时间，首遍跳读率（不仅涉及当前词，还包括前一个词以及后一个词的跳读情况，局部分析中缺少跳读率的数据结果）等眼动指标的结果之上，讨论呈现时间和眼动模式之间的关系可能结果会更好理解。

回应: 感谢审稿专家就讨论部分给出的具体意见和建议。参考您的建议，我们对讨论部分进行了较大幅度的修改。减少了对其中一些指标过多的讨论，而增加讨论了结果分析所涉及地

更多指标，期望能更好地说明呈现时间和眼动行为间的关系。首先对假设中最重要的两项指标：总阅读时间和回视进行分析，揭示了儿童和成人对文本呈现时间的需要。接着从相同和不同两个方面具体说明儿童和成人阅读消失文本时的眼动模式变化，具体结合早晚期注视时间，注视位置，跳读等多项指标共同说明。详见“4 讨论”部分蓝色修改内容。

对于您提出的修改建议，我们把握的不知是否准确。如果还不准确，请专家继续指教。

意见 6: 结论部分，儿童与成人相比，在消失文本条件下回视次数增加，表明儿童比成人读者获取文本视觉信息的速度慢。这一表述缺少理论依据。

22 页倒数第二段：“这很好的佐证了上述猜测，即读者采用减少首轮注视时间来补偿回视所需时间，从而保证总阅读时间不受影响。”这句话时序不对，为什么不是增加回视时间来补偿减少的首轮注视时间。

回应: 仔细推敲上述两项表达后，认为的确不太妥当。因此做出如下修改：

“与成人相比，消失文本操作对儿童读者的眼动行为造成更大影响。表现为儿童采用更多回视进行二次信息的提取，同时降低更多的再注视、增加更长的注视时间，以此完成正常的阅读加工。这表明，中文阅读中儿童读者比成人对文本视觉信息呈现时间的需要更长。”

“这很好佐证了上述猜测，即读者通过增加回视来补偿首轮阅读减少的时间。从而表现为总阅读时间没有受到显著影响。”

审稿人 2 意见: 本文采用消失文本范式，考察中文阅读中儿童(10~11 岁)与成人(18~21 岁)获取文本视觉信息速度的发展变化。论文整体上方法可行，结果较可信，结论对于推进阅读加工机制的研究有一定的贡献，但目前还存在如下几个问题：

意见 1: 前言部分关键性的理论铺垫还不足。尽管前言部分有提到该研究为眼动控制加工模型的建构提供支持，但并没有详细的论述，使得文章的理论意义或者理论创新上略显不够，建议在前言部分补充基本理论并在讨论中提升文章的理论价值。

回应: 感谢您提出的问题和建议。当前研究的目的是揭示中文儿童和成人对中央凹文本信息呈现的时间，比较儿童和成人在阅读文本信息呈现时间不完整时眼动行为的差异。研究儿童和成人读者眼动模式的差异能够为研究者提供新的视角去探讨眼动控制的发展机制以及眼动控制和阅读技能间的相互作用特点，最终有助于揭示阅读的认知加工机制。本研究作为一个阅读中的基本问题的探讨能够为阅读中眼动控制理论及模型的构建提供相应的基础数据，但由于与模型建构没有直接的关系，所以并未对此进行过多的叙述。

与本研究直接相关的理论是如何解释儿童和成人在眼动模式上存在的发展差异，Reichle et al. (2013)提出了如下三种假说：语言熟悉假说(linguistic proficiency hypothesis)，眼球运动协调假说(oculomotor-tuning hypothesis)以及语言熟悉和眼球协调运动相互作用假说(interactive-adaptation hypothesis)。就该理论内容我们在前言中补充了更详细的论述，并在讨论中也进行了理论的提升。具体见前言和讨论蓝色修改内容。

意见 2: 问题提出逻辑不够清楚明确，有待深化。比如在文章的第三段，为什么提出“那么这项阅读技能在儿童成为熟练读者过程中是如何发展变化的呢？进一步说，儿童和成人眼动行为的差异是否也能够通过获取视觉信息效率的差异来解释呢？”这两个研究问题，为什么研究儿童，为什么研究儿童与成人之间的差异，阐述不够详细清晰，两个研究问题之间的递进关系不明显。

回应: 非常感谢您的建议。在问题提出部分的确存在您提出的问题，为此我们对前言的结构

和内容做了较大的修改。从研究问题的引入，到大量成人研究中获得的结论和方法，到为什么要研究儿童和成人之间的差异，再通过“语言熟悉假说”为本研究问题的回答提供假设基础。之后通过比中文和拼音文字书写系统上的差异，以及分析拼音文字儿童读者研究的结果，做出了本研究的基本假设。详见“前言”蓝色修改内容。

意见 3: 本文只是简单介绍了消失文本范式，那么对于视觉信息获取的研究范式还有那些呢？是不是应该做适当的补充，体现本研究范式的优势呢？

回应: 感谢您提出的建议，我们在文章前言中补充了同类问题研究的其它研究范式，以及本研究选择消失文本范式的优势。具体见原文“前言”部分第二段标注的内容。

意见 4: 本研究只有一个实验。单纯依靠一个实验的结果来证明理论观点，似乎略显单薄。

回应: 您说的有道理，本研究的目的是揭示儿童和成人在阅读中对文本呈现时间基本需要上的差异。但由于这只是一个初步的研究，对影响文本呈现时间需要的其它可能变量还缺乏清晰的考察，例如，词频，词长，笔画数，部件数等。此外，本研究目前只能回答儿童和成人对当前注视词视觉信息呈现时间是否存在差异，而就附中央凹内非注视词（词 $n+1, n+2$, 以及 $n-1$ ）视觉信息呈现时间的需要等问题还无法给予回答，甚至在该问题上，成人读者的研究也还不够充分。因此，我们将在以后的实验中，对上述问题进行更具体深入的研究，从而丰富相关理论观点。

意见 5: 其他细节问题。

(1) 格式规范问题，比如在实验材料部分，统计符号 (M, SD) 应该斜体；再如，作者对数据进行了被试分析和项目分析，应明确注明。

(2) 语句通顺问题，比如“结果显示：所以消失文本条件下，虽然读者在首轮注视中的首次注视显著长于正常文本，”。

建议再朗读细看文章，对文章格式和细节进行润色。

回应: (1) 感谢您提出的不足，按照要求我们对格式部分进行了仔细校正，此外也增加了对两类分析的说明。具体见文章“2.3 实验材料”的第 1 段和“3 结果”的第 4 段中的蓝色修改内容。

(2) 非常感谢专家细致的阅读。按照建议对文章中语句表述不通顺和不规范的地方进行了修改。并多次通读全文，尽力避免类似文字表述不规范问题。

最后，再次感谢两位审稿专家对本文提出的宝贵意见和建议。采用消失文本范式考察中文阅读中读者获取文本视觉信息速度的发展变化，在国内尚属初步研究的阶段。因此，该研究在很多方面，可能还是存在许多不尽人意之处，在此，恳请审稿专家不吝赐教，我们将进一步认真修改。

第四轮

编委复审意见: 作者根据审稿意见对文章做了很好的修改，相关文献非常翔实，对研究问题 and 研究过程也有细致的思考。不过本研究内容上略显单薄，只用一个实验来描述儿童(10~11岁)与成人(18~21岁)阅读消失和正常文本时的眼动行为的差异，研究假设只是“儿童比成人可能需要更长的文本呈现时间才能获取到足够的视觉信息”。所以，希望作者能够对本研究的理论意义进一步提升一下。作者在前言中补充了，“如何解释儿童和成人在眼动模式上存在的发展差异，Reichle et al. (2013)提出了如下三种假说：语言熟悉假说(linguistic proficiency hypothesis)，眼球运动协调假说(oculomotor-tuning hypothesis)以及语言熟悉和眼球协调运

动相互作用假说（interactive-adaptation hypothesis）”。而在讨论部分，作者只是简单的使用语言熟悉假说解释了实验结果，那么本研究结果是不是可以确证语言熟悉假说更合理，其他两种假说不合理，希望作者在讨论部分能就此多做一些说明。从而更好的提升理论价值。修改后直接上交编委。

回应：感谢专家对本文已修改部分给予的肯定，同时非常感谢专家进一步提出的修改建议。按照您本次提出的理论价值提升的问题和建议，我们继续进行了修改。具体在前言和讨论中分别增加了相关的内容。所修改的内容主要是阐述本研究结果为什么倾向支持语言熟悉假说对儿童和成人眼动差异的解释，以及对不支持眼球协调运动假说给出了相应的结果佐证。具体见讨论的 P32 和前言的 P18、P21 的橙色修改文字。

第五轮

主编终审意见：Xb13-392《中文读者词汇视觉信息获取速度的发展》一文从研究问题的份量上来看确实略显单薄：原因有三点：第一，研究只有一个初步性的验证实验，这样所得出的结果是否具有必然性值得怀疑；第二，本研究的被试和被试所阅读的材料较少也是一个问题，在稿件正文中，作者对被试和材料是这样描述的：在 75 个正式实验句中，每种呈现方式条件下各 15 个，在实验条件间按照拉丁方顺序进行轮组后形成 5 组实验材料，每名被试只阅读其中一组。”，但在回答专家提问时，作者对被试和材料是这样描述的：“本研究的被试内变量是 5 个水平，每个被试阅读 75 句正式实验句，即每个条件下 15 句。”到底哪一个的描述是正确的？？稿件正文中的表述感觉该研究中的呈现方式这一变量在实验设计上更像是被试间设计；第三，本研究的实验设计上是否存在一定的问题，即在呈现方式（控制条件、20ms 消失、40ms 消失、60ms 消失、80ms 消失）这一变量上，自变量的五个水平中，除控制条件外，其余 4 个自变量水平的取值是否合理，20ms 的变化范围能否有效地揭示自变量的变化范围？从研究结果来看，似乎结果不是很理想。鉴于以上原因，建议作者至少再补充一个实验，一是在新的实验中克服原有稿件中存在的问题，二是看两个实验结果是否具有一致性。

回应：非常感谢您对本研究提出的宝贵建议和意见。我们已经按照您的建议增加了一个新的实验，用以弥补实验一中存在的不足，同时验证原实验结果的稳定性。

实验二被试总数量增加到 70 名，其中儿童和成人各 35 名。按照您的建议，我们在实验二中改变了文本呈现条件的变化水平，使消失文本相邻呈现时间的间隔由原来的 20ms 增加为 40ms，时间范围由原来 20ms~80ms 扩大到 20ms~140ms。因此，实验二文本呈现条件的各水平分别为：20ms，60ms，100m，140ms 以及控制条件。这样设计的原因如下：首先，增加时间间隔是为了排除实验一间隔较小造成反应不灵敏的可能性。换句话说，自变量水平的变化不足导致因变量没有产生相应的变化。其次，将消失文本条件中最长的呈现时间增大到 140ms，是为了进一步考察满足儿童回视次数不受影响所必需的文本呈现时间，即儿童后期词汇整合以及高水平语言加工对文本呈现时间的需要。

按照您所提出的建议，当我们把实验二中消失文本间的间距增大后，所得到的结果一方面在整体趋势上与实验一保持基本一致，另一方面又在总阅读时间和回视两项重要指标上得到了新的发现。具体内容如下：

与实验一的结果相比较，实验二发现：

第一，从整体阅读水平上看，儿童和成人只在阅读 20ms 条件的消失文本时总阅读时间显著增加，而在大于 20ms 以上的消失条件上均与正常阅读无显著差异，这与实验一的结果一致。与实验一结果不同的是，实验二获得自变量间显著的交互作用，表现为儿童比成人在阅读 20ms 的消失文本时总阅读时间的增加程度更大。实验一的结论是，儿童需要更长的文

本呈现时间。因此，实验二的上述结果进一步支持了实验一结论。

第二，在回视行为上与实验一得出的结果趋势一致，表现为儿童的回视增加程度也显著大于成人。实验二的结果进一步显示，儿童的回视到 140ms 条件才开始不再增加，而成人在于 60ms 文本条件时便不再显著增加。这表明，满足儿童回视需要的文本呈现时间比成人更长。该结果再一次支持了实验一的结论。

第三，实验二也验证了读者在阅读消失文本时所表现出的与实验一一致的眼动变化策略，即首轮阅读时，读者的再注视较少，首次注视时间增加；首次登陆位置更靠近词的中心位置。这为总阅读时间和阅读理解率不受消失文本操作（除 20ms 条件）的干扰提供了保障。而上述眼动策略在国内外同类研究中也得到了证实。第四，在个别指标的事后检验或是简单效应分析中，也出现了与实验一不完全一致的结果，见“3.2.1 整体分析”中（3）平均注视时间和 3.2.2 局部分析中（2）凝视时间指标的结果部分，以及“3.3 讨论”的第三和第四段内容。然而以上结果并不影响总的研究结论。总之，我们认为实验二的结果有效地验证并进一步支持了实验一的结论。

此外，我们在原稿件中除了增加实验二的所有内容，还在如下几个方面做了修改（具体见原文相应部分的红色标注）：

1. 增加了实验一的讨论部分。详见文中“2.3 讨论”。
2. 简化了两个实验的结果部分的登陆位置分析，将其与其它局部分析合并在一起呈现。详见“2.2.2 局部分析”和“3.2.2 局部分析”。
3. 结合两项实验的结果对总讨论的内容进行了相应的修改，详见“4 讨论”中红色标注文字。
4. 关于您意见中的第二个问题，我们在稿件中的描述的确不清楚，很容易让读者将呈现方式理解为被试间变量，实际上该变量是被试内变量。为此我们在原文“2.1.3 实验材料”部分的第二段作如下修改：“将 75 个正式实验句随机分为 5 组，再按照拉丁方顺序使实验条件在各组材料之间进行轮组，最终形成五个实验材料的文件。每个文件包括相同的 75 个实验句，藉此保证所有的句子均接受 5 种实验处理，且在实验中，每名被试只阅读其中一个文件。”