

《心理学报》审稿意见与作者回应

题目：聋人阅读中的副中央凹视觉注意增强效应——来自消失文本的证据

作者：刘璐，闫国利

第一轮

审稿人 1 意见：本论文用消失文本范式来考察聋人中央凹和副中央凹阅读加工的特点，其结果表明，聋人在阅读中存在副中央凹视觉注意增强的效应，但是，他们的中央凹加工却被削弱，作者猜测这可能是聋人阅读困难的原因。我觉得本文探讨了一个很有价值的问题，目前，对聋人这个特殊群体的阅读特点的研究并不多，从这个角度考虑，本文可以为大家了解聋人阅读的特点提供有价值的证据。但是，我觉得这一稿还不适合发表，我的问题主要有以下这些：

意见 1：当字或词呈现 40 毫秒后消失，读者可以获取多少信息，采用 40 毫秒或是 60 毫秒的根据是什么？作者认为在文本呈现 60 或 40 毫秒后消失，正常儿童和成人的阅读效率不受影响，那是否意味着文本消失条件下跟正常条件下，眼动指标没有差异？文本消失这个范式应该会带来视觉上的干扰，这个干扰效应是否需要考虑？

回应：非常感谢审稿专家的意见。

首先，在消失文本范式中，读者在字或词所呈现的 40 毫秒内主要是对字或词的视觉信息进行编码加工，即呈现时间并不是指大脑对视觉刺激进行加工所需要的时间，而是指读者对视觉信息的注视必须达到此呈现时间后，才能为大脑提供进一步的认知加工所需要的有效视觉信息 (Rayner, Liversedge, White, & Vergilino-Perez, 2003; Rayner, 2009; 闫国利, 巫金根, 胡晏雯, 白学军, 2010)。

其次，本实验中在以往大量关于中文阅读相关研究的基础上，设置文本呈现条件为 40 毫秒后消失，具体如下：在中央凹双字词呈现 40ms 时，儿童(10~12 岁)和成人读者(18~23 岁)的总阅读效率均未受到影响，即当中央凹词呈现 40ms 能够满足成人及小学生读者对文本信息的视觉编码(刘志方, 翁世华, 张锋, 2014; 熊建萍, 2014; 闫国利, 刘妮娜, 梁菲菲, 刘志方, 白学军, 2015; 张巧明, 2013); 当副中央凹双字词 N+1 呈现 40ms 时，对于五年级的小学生和成人大学生而言，其总阅读时间与正常阅读没有差异，即当副中央凹词 N+1 呈现 40ms 能够满足成人及小学生对文本信息的视觉编码；但对于小学三年级的小学生，文本呈现的时间要更长才能满足其文本信息的编码(刘志方, 王若逸, 张智君, 刘炜, 2016; 闫国利, 胡晏雯, 刘志方, 张智君, 2009; 张巧明, 2013)。

再者，在消失文本范式下，当文本短暂呈现条件下（如 40 毫秒后消失）读者的阅读效率不受影响，主要指读者的总体阅读效率（通常以总阅读时间为最重要的衡量指标，Blythe, Liversedge, Joseph, White, & Rayner, 2009; Blythe, Haikio, Bertam, Liversedge, & Hyönä 2011），而并非意味着在所有的眼动指标上与正常呈现条件无显著差异。实际上，消失文本的操纵会改变读者基于消失字或词的眼跳计划，为保证阅读的正常进行，读者会采取注视次数和注视时间之间的权衡策略，即在消失条件下的注视次数减少，而平均注视时间增长。而总阅读时间这一指标，是包含句子阅读过程中所有注视点持续时间以及眼跳持续时间之和，从整体上反映构成句子所有词汇的加工过程，因此作为读者总体阅读效率的衡量指标。

最后，关于消失文本范式，该范式与移动窗口范式、移动掩蔽范式、边界范式以及快速启动范式均属于“呈现随眼动变化技术 (eye-movement-contingent display-changes technology)” (Rayner, 1993; 1998)，即通过追踪读者眼睛注视的位置来引发屏幕上呈现的视觉刺激的变化，通常这种视觉刺激变化的用时很短（十几毫秒左右），因此即使读者能够意识到这种呈现的变化，但这种变化带来的影响通常被忽略不计 (闫国利, 巫金根, 胡晏雯, 白学军, 2010)。故在几乎所有应用这类技术的研究中，均不考虑视觉变化对读者阅读过程的影响。

意见 2：还需要澄清的是，刺激呈现 40 毫秒后消失，并不意味着读者在这 40 毫秒里对该字或词的加工结束了，眼睛还会停留在刺激呈现的位置，继续对先前出现的刺激进行加工。那种加工跟刺激没有消失，即正常条件下的加工有什么不同呢？这点作者没有解释，会影响

读者理解为什么文本消失范式是探讨副中央跟副中央凹加工的恰当的范式。

回应：非常感谢审稿专家的意见。确实如专家所提到的，原文中未详细解释消失文本条件下读者的阅读加工；因此为更好地阐释消失文本范式在本研究中的作用，在修改稿中增加相关的内容，具体如下：

用于探究这一问题最有效的方法是消失文本范式（disappearing text paradigm; Rayner, Liversedge, White, & Vergilino-Perez, 2003）：当读者注视一个词达到某特定时间（如 40ms）后，该词就会“消失”，即变成空白；当下一个注视开始时，原来消失的词又重新呈现，而当前注视词会在呈现相同的时间（40ms）后消失。尽管注视词在呈现 40ms 后“消失”，但并不意味着大脑对该词的加工已经完成，而是眼睛继续停留在注视词处进行认知加工；在注视词呈现的短暂时间内，读者需要快速地将该词的视觉信息转换为稳定视觉编码，进而保障随后的阅读加工过程（闫国利，巫金根，胡晏雯，白学军，2010；刘志方，张智君，潘运，全文，苏衡，2017）。换言之，通过操纵文本呈现的时间，可以考察读者阅读过程中词汇视觉编码所必须的时间，进而反映读者阅读中的视觉加工效率(Blythe, Liversedge, Joseph, White, & Rayner, 2009; Rayner, Liversedge, White, & Vergilino-Perez, 2003)。

意见 3：表 3，消失文本范式举例。句子里每个词语都会在注视后消失还是只有目标词会消失？注视点的位置好像标记错了。

回应：非常感谢审稿专家的细致审阅。

在实验一使用的副中央凹消失文本范式，是在注视当前词 40 毫秒后，下一词（词 n+1）消失。本文中表 3 确实在注视点位置上标记错误，修改后的范式举例如下：

表 3 消失文本实验范式举例

注视	实验 1 副中央凹词 n+1 的消失	实验 2 中央凹词 n 的消失
当前注视开始	世界珍稀 植物 标本将于明天展出。	世界珍稀 植物 标本将于明天展出。
注视 40ms 之后	世界珍稀 植物 将于明天展出。	世界珍稀 标本将于明天展出。
下一注视开始	世界珍稀植物 标本 将于明天展出。	世界珍稀植物 标本 将于明天展出。
注视 40ms 后	世界珍稀植物 标本 明天展出。	世界珍稀植物 将于明天展出。

意见 4：数据整体分析里，再注视比率，跳读率，回视次数，应该是针对词语或是字进行分析的，不应该放在整体分析里。整体分析是把整个句子当作一个区域，可以算总阅读时间，总注视点个数，和平均注视时间，平均正向眼跳距离，和回视眼跳距离等。我觉得作者应该报告总注视点个数，和眼跳距离等指标，可以更好地衡量聋人是否采用冒险性的眼跳策略。建议作者考虑删掉跳读率，再注视率这些指标，或者加深对这些指标的讨论。

回应：非常感谢审稿专家的意见。

首先，正如专家所言，再注视比率、回视次数和跳读率通常是在对目标字或词进行局部分析时使用的。但是，在以往很多使用消失文本范式的研究中，也将再注视比率、回视次数以及跳读率(跳读次数)作为整体分析的眼动指标(Liversedge et al., 2004; Rayner, Liversedge, & White, 2006; Blythe, Liversedge, Joseph, White, & Rayner, 2009; Blythe, Haikio, Bertam, Liversedge, & Hyona, 2011; 刘志方，张智君，赵亚军，2011；刘志方，刘妮娜，全文，2014；闫国利，刘妮娜，梁菲菲，刘志方，白学军，2015；刘志方，王若逸，张智君，刘炜，2016)，并在整体分析中将每个词都作为目标词进行分析，可以更精细地分析被试对句子中词的平均认知加工情况，因此本研究在整体分析中亦选择使用了上述指标。

其次，专家建议使用诸如注视点个数、眼跳距离等指标以反映聋人是否采用冒险性的眼跳策略，这的确是非常值得去探究的问题。但实际上，正如在意见 1 的回复中，消失文本的操纵会改变读者基于消失字或词的眼跳计划，为保证阅读的正常进行，读者会采取注视次数和注视时间之间的权衡策略，即在消失条件下的注视次数减少，而平均注视时间增长。因此，这些指标的变化更多地可能是由于消失文本对眼动策略的影响，而难以确切地反映本研究所

关注的聋人读者副中央凹以及中央凹的加工效率问题。

另外，的确我们在文中未对再注视比率这一指标进行深入的讨论，而这一指标主要也是受到消失文本范式本身操纵的影响，因此在修改稿中删除对这一指标的分析。但我们保留了对跳读率以及回视次数的分析，这是由于在总讨论部分对这两项指标有进一步的解释，并且亦能反映聋人阅读中的副中央凹加工的特点：由于聋人将更多的注意资源分配于副中央凹处，因而能够更早地对副中央凹处的词汇信息进行加工，故当聋人读者在注视词 n 时，其副中央凹处的词 $n+1$ 得到更快速的加工，进而更可能做出朝向词 $n+2$ 的眼跳计划，使得对词 $n+1$ 的跳读率更高。但是，更高的跳读率可能带来读者在阅读理解上更大的风险；因此，聋人读者会采用回视的策略来弥补首轮加工中可能犯的错误（可能对词 n 的加工不够充分），以帮助其获得语义整合。

意见 5: 实验一发现有 *pof* 效应，表现为词频对目标词的注视时间有影响，但是作者并没有展开来说。只是在总讨论的地方比较笼统地提到 *pof* 效应，并没有结合本文的数据来讨论。

回应: 非常感谢审稿专家的意见。

需要说明的是，在本研究的实验 1 中，词频效应的影响是表现在消失词上，即词 $n+1$ 上。因此，在实验 1 中，我们并未考察是否存在 *POF* 效应（词频对注视词或词 n 的影响），原因主要在于（1）*POF* 效应在以往健听读者的阅读研究中尚不稳定，即是否存在 *POF* 效应以及是正向效应还是负向效应（胡笑羽，白学军，闫国利，2010）；（2）探究 *POF* 效应常用的范式为边界范式、词对阅读范式等（胡笑羽，白学军，闫国利，2010），暂无研究使用消失文本范式。而在总讨论中提及 *POF* 效应主要是从 *E-Z* 读者和 *SWIFT* 模型对比的角度，我们十分赞同专家提出的意见，这部分内容与本研究的主要目的关系并不密切，因此在修改稿中将其删除。

意见 6: 实验二的总阅读时间，*RC* 组在两种条件下的阅读时间差异不大，只有 82 毫秒，接近 *AC* 组，作者并没有对比聋人组讨论这个数据，这应该是一个很重要的对比。

回应: 非常感谢审稿专家的意见。

的确如专家所说，在实验 2 中，我们发现健听读者（阅读能力匹配组和生理年龄匹配组）在两种文本呈现条件下的总阅读时间均无显著差异，表现读者在短暂的中央凹词 n 呈现时间内获取了足够的文本视觉信息；同时我们也发现了，聋人读者（聋生组）在消失文本条件下的总阅读时间显著长于正常文本条件下的总阅读时间，这表明 40 毫秒的文本呈现时间，聋人读者未能快速将中央凹词 n 的视觉信息转换为稳定视觉编码，从而影响其总体阅读时间。可能在实验 2 的讨论中，我们对这一重要结果的论述较少，修改稿中我们对此进行补充。

意见 7: 局部和整体数据分析表格的格式，标准差是放在括号里还是用加减号表示，应该统一。图 1 建议高低频放一起比较，分三组，更容易看出词频效应。

回应: 非常感谢审稿专家的细致审阅和建议。已在修改稿中已经将数据表结构统一，并将图 1 重新更改。具体请见修改稿表 5、表 7 和图 1。

意见 8: 虽然实验一的总阅读时间没有发现聋人组的阅读时间受消失文本范式的影响，实际上，消失文本和正常阅读范式下的差异有 119 毫秒，而 *AC* 组的差异只有 18 毫秒。局部分析上也有类似的问题，数据的模式并不如作者描述的那么清楚。数据报告的时候，建议同时解释数据模式的意义，而不仅仅是报告有没有主效应跟交互作用。

回应: 非常感谢审稿专家的意见。

由于另一位审稿专家建议使用 *Lineal mixed effects model & logist mixed effects model* 分析数据，因此数据结果部分稍作改动，具体请见修改稿。

意见 9: 在讨论里，有些指标是被忽略的，例如实验一重点报告句子总阅读时间，可以考虑删掉不重要的指标，例如局部分析里的首次注视时间。

回应: 非常感谢审稿专家的意见。

正如专家所说，本研究两项实验均重点报告了总阅读时间这一指标，并以此展开讨论和推论。而局部指标上，由于消失文本范式的操纵会产生读者眼动策略的改变，因此反映更为

早期词汇加工阶段的首次注视时间，其数据模式通常更容易受到影响，从而难以清晰地说明读者的词汇通达过程。因此，我们采纳专家的意见，删除对首次注视时间的分析，主要考虑凝视时间以及总注视时间两项局部分析指标。

意见 10: 作者参考了不少文献，但是感觉太泛，最相关的概念和理论没得到展开。例如“注意增强效应”在我看来还是解释得不够，眼动阅读的串行和并行模型也像是在中间跳出来，副中央凹中央凹效应也是在讨论里跳出来的感觉。如果是检验串并行模型，建议开篇写清楚两种理论对中央凹和副中央凹加工的不同预测，或者抛开这两个理论，单纯从注意分配或者加工效益方面考虑中央凹跟副中央凹的加工。具体的修改作者可以自己考虑，目的是为了文章更统一精炼。

回应: 非常感谢审稿专家的这个建设性意见。

本研究的主要目的还是在于，从视觉注意分配的角度，探究聋人阅读过程中的副中央凹以及中央凹加工效率的问题；而对于聋人读者的眼动控制模型，更多是作为理论支撑，且未直接检验串并行模型问题。这可能在一定程度上使得文章整体的重点不够突出，因此，我们考虑删除关于眼动控制模型部分的论述，并对聋人视觉注意的特点及其对阅读的影响做进一步阐述，具体请见修改稿。

最后，再次感谢审稿专家的审阅与建议，我们已对您的建议逐一进行了认真修改，如果还有其他方面的问题，请专家进一步批评指正，我们将一步仔细修改。

审稿人 2 意见: 该研究尝试提出“聋人阅读中的副中央凹视觉注意增强效应”，并采用消失文本范式，设计两个眼动实验考察这一效应对聋人读者副中央凹及中央凹处词汇信息视觉编码效率的影响。实验一数据模式发现，在词 $n+1$ 短暂呈现时，聋生快速编码副中央凹处文本信息的加工效率与其生理年龄匹配组相近，高于阅读能力匹配组；但在词 n 短暂呈现时，聋生快速编码中央凹处文本信息的加工效率落后于健听控制组，其词汇通达和总阅读效率均受阻（这是该论文摘要原文，审稿人对此持保留意见）。研究者基于上述两个实验，认为上述研究证实了聋人阅读中的副中央凹视觉注意增强效应，这可能是导致其阅读困难的原因之一。审稿人认为该研究选题具有较高创新性，仔细阅读了相关文献，推敲了研究者的实验设计，假设的逻辑起点，数据分析，认为仍存在一些问题需要进一步明确。

意见 1: 第二段中，“有研究认为，由于早期听觉损伤，聋人视觉注意资源的空间分布发生了改变——从中央视野转移到了中央以外视野，使得聋人中央以外视野处视觉注意资源分配增多（Bavelier, Dye & Hauser, 2006; Dye, Hauser & Bavelier, 2009; Dye & Bavelier, 2010; 项明强, 胡耿丹, 2010; 王庭蔚, 杨娟, 2013）。聋人这种视觉注意再分配的特点具有环境适应性：尽管其听力受损，但可以通过在中央以外视野视觉注意资源分配增多对来自更广范围的信息输入进行功能性补偿。”这里存在一个问题，即聋人视觉注意资源分配增多以及由此而引起的视觉认知的变化是对所有认知功能的泛化，还是与视觉-听觉 convergence 的功能的变化。这种变化与高级认知功能的变化存在什么样的关系？请研究者进一步仔细阅读 Bavelier 等 Trends in Cognitive Sciences 的论文，做准确和细致阐述。

回应: 非常感谢审稿专家的意见。

专家建议对聋人视觉注意资源分配的变化做细致阐述，通过进一步阅读 Bavelier, Dye 和 Hauser (2006) 的文章，我们对这一问题的补充说明具体如下，并请见修改稿。

对聋人而言，听觉通道信息输入的缺失使其视觉通道的信息获取尤为重要。由于大脑可塑性的特点，聋人缺乏听觉刺激经验会促进视觉认知的增强；但这种视觉认知的增强是具有选择性，聋人仅表现出对边缘视野处或运动刺激的视觉加工增强（Lork & Song, 1991; Stevens & Neville, 2006; Buckley, Codina, Bhardwaj & Pascalis, 2010）。有研究认为，由于视-听跨通道整合（auditory-visual convergence）的补偿机制，多感觉联合皮层以及早期视觉皮层在听觉输入丧失后发生重组，表现出对视觉通道，尤其是边缘视野处刺激更大的敏感性（Bavelier & Neville, 2002; Bavelier, et al., 2001）；但由于空间注意资源是有限的，聋人的视觉注意资源在空间上进行再分配——从中央视野转移到了中央以外视野，在中央以外视野处视觉注意资源分配增多，以对来自更广范围的视觉信息输入进行功能性补偿（Bavelier,

Dye & Hauser, 2006; Dye, Hauser & Bavelier, 2009; Dye & Bavelier, 2010; 项明强, 胡耿丹, 2010; 王庭蔚, 杨娟, 2013)。

意见 2: “其次, 在副中央凹获取信息的类型方面: 有研究发现, 相比于健听读者, 中国聋人读者能够在更早期的加工中(反映在首次注视时间上)就从副中央凹预加工中获得语义预视效应(Yan, Pan, B danger, & Shu, 2014) 和手语表征预视效应(Pan, Shu, Wang, & Yan, 2015), 这表明聋人读者在副中央凹处的加工效率足够高, 使其能够更早地获得对预视信息的词汇通达。”这里论述不准确。研究者认为副中央凹获取信息类型不同, 但是又说加工效率足够高。此外, 正常人也可以从副中央凹预加工中获得语义预视效应。 所以, 这里只有可能的早晚差异, 而没有类型差异。

回应: 非常感谢审稿专家的意见。

原文在此处的表述确实不准确, 因此在修改稿中将“类型”更改为“加工阶段”。具体修改请见修改稿。

意见 3: “既然聋人读者在阅读过程中表现出副中央凹处获取信息的范围更广、加工效率更高, 那为何聋人读者的阅读水平却远落后于同龄健听读者呢?”这是研究者拟回答的问题, 并基于此提出了可能的解释, 即“聋人阅读中的副中央凹视觉注意增强效应”。 这里存在一些逻辑推论问题:(1) 阅读能力是人类的一种高级能力。视觉系统完成视觉信息输入, 从枕叶皮层向顶叶、额叶、颞叶皮层传递信息, 完成信息解码、在编码和整合。这里还存在自上而下信息的驱动和交互, 各种外在环境的信息的交互等等。最终, 人们表现出不同的阅读行为。所以, 阅读水平低与副中央凹视觉注意增强效应没有必然冲突, 也不构成必然关系。(2) 从后续两个眼动实验看, 聋人阅读句子的正确率与阅读能力匹配组、生理年龄匹配组之间的差异不显著, 甚至实验一中高于阅读年龄匹配组。 那么如何理解这两者之间的矛盾呢? 请研究者仔细思考上述问题, 进一步完善问题提出的合理性, 论述的逻辑严密性。

回应: 非常感谢审稿专家的意见。

首先, 正如专家所提到的, 阅读能力是一项复杂的高级认知能力, 阅读过程中有多个加工阶段和加工策略的相互作用。原文中, 将聋人总体的阅读能力与其副中央凹视觉注意增强进行比较, 确实欠妥当。实际上, 本研究最主要的目的在于考察聋人视觉注意资源分配的特点对其阅读加工的影响, 关注的焦点在于副中央凹以及中央凹的加工效率问题。因此, 修改稿中我们不再以“聋人的副中央凹视觉注意增强效应”来解决聋人阅读水平落后或困难原因的问题, 而集中阐述聋人视觉注意资源分配对其副中央凹以及中央凹加工的影响。

其次, 关于三组被试的阅读正确率差异不显著的问题, 由于本研究采用的实验句非常简单、较短且语义简单的句子(对句子难度和通顺性的评定能够支持这点), 读者在阅读时非常容易理解(经过聋校教师的评估, 这些句子对于聋生是易于理解的), 因此三组被试的阅读正确率均能够保证较高的水平, 表现出在正确率上没有显著差异; 而在实验一中, 聋生的阅读正确率略高于阅读年龄匹配组, 可能是由于副中央凹词 $n+1$ 的短暂呈现对阅读年龄匹配组造成的影响, 但在统计上三组被试的阅读正确率并无显著差异。

意见 4: “最终为汉语阅读中序列加工与平行加工之争提供支持。”请研究者删除该论述, 因为该研究没有直接解决这个问题。

回应: 非常感谢审稿专家的意见。

正如专家所说, 本研究并未直接检验或解决序列加工与平行加工之争, 之所以提及眼动控制模型, 主要是作为聋人阅读加工的理论支撑。我们采纳专家的建议, 删除相关论述。

意见 5: 报告智力测查工具与水平的量化数据。

回应: 非常感谢审稿专家的意见。

十分抱歉的是, 在本次实验过程中, 我们未对被试的智力水平进行统一的测量。这主要是由于目前尚没有能够广泛应用于聋人被试的智力测验, 并且在已有的智力测验(如韦氏智力测验、瑞文推理测验等)中, 缺乏特殊群体被试的常模, 因此难以获知聋人被试的智力水平。但本实验中的聋生被试, 在入学前均经过聋人学校的智力测量(《中国比内测验》), 其智力水平正常; 同时聋生被试在进行阅读能力匹配时, 还进行过三项阅读能力测验(正字法

意思、阅读流畅性以及阅读理解), 其测验成绩与小学五年级学生相当, 这也从侧面印证了聋生的智力水平。这类问题我们会在以后的研究中注意这个问题, 恳请专家的谅解。

意见 6: 报告词频的差异性统计检验结果; 报告笔画数、词汇语义家族大小, 以及差异性统计检验结果; 补充调查句子合理性; 报告难度和通顺性的差异性统计检验结果。

回应: 非常感谢审稿专家的意见。已按照专家的意见, 对实验材料部分数据进行了补充和完整, 具体请见修改稿。

另外, 根据专家的意见, 我们对目标词的词汇语义家族大小进行了统计, 具体方法如下: 通过检索《现代汉语词典》(第六版), 对每个目标词的语义解释的项目数量进行了统计, 具体数据请见修改稿。

十分抱歉的是, 关于专家建议补充调查句子合理性, 我们非常赞成这一建议。由于目前中小学已经进入假期, 所以短时间内难以完成该项补充调查, 但我们会尽快完成, 恳请专家的谅解。

意见 7: 数据处理方法需要改进。仔细阅读实验一数据, 在很多眼动指标上, 多个自变量的主效应或者交互效应在被试分析上显著, 在项目分析上不显著, 或者相反模式。F1、F2 的数据分析在语言认知研究中已经落伍, 更重要的是该研究中存在很多两种统计结果不一致。这妨碍了我们对结果的正确判断。所以, 建议研究者采用 *Lineal mixed effects model & logist mixed effects model* 分析数据。这是当前研究需要认真和着重解决的关键问题。

回应: 非常感谢审稿专家的意见。

修改稿中的数据分析, 已采用 *lme4* 构建线性混合模型在 R 中分析各项指标(Bates, Maechler, & Bolker, 2011; Barr, Levy, Scheepers, & Tily, 2013), 具体请见修改稿。

意见 8: 与实验一类似的问题请解决, 含数据处理方法的改进。

回应: 非常感谢审稿专家的意见。

修改稿中的数据分析, 已采用 *lme4* 构建线性混合模型在 R 中分析各项指标(Bates, Maechler, & Bolker, 2011; Barr, Levy, Scheepers, & Tily, 2013), 具体请见修改稿。

意见 9: 实验二选用了与实验一相同的被试, 请报告两个实验间隔时间。

回应: 非常感谢审稿专家的意见。

两项实验采用了相同的被试, 两个实验的间隔时间为 1 周。已在实验 2 的“被试”部分进行报告。

意见 10: 实验二局部分分析中, 数据模式令人疑惑。以表 7“凝视时间”为例, RC 和 AC 组在正常阅读条件下的阅读时间竟然普遍高于 40ms 消失文本。DS 组在正常条件下对低频词的时间高于 40ms 消失文本的低频词约 32ms。为什么? 更需要注意的是, 在 40ms 消失条件和正常条件下, DS 聋生组的凝视时间普遍低于对应的 RC 组和 AC 组。这表明在正确率没有差异的前提下, 聋生组的加工效率更高, 哪怕中央凹只呈现 40ms。这种数据模式表明, 聋生组的中央凹视觉注意不但没有减弱, 还得到增强。如何理解上述问题?

回应: 非常感谢审稿专家的意见。

首先, 对于被试在正常呈现条件下对目标词的凝视时间更长, 这与以往多篇文献中的结果是一致的 (Blythe, Liversedge, Joseph, White, & Rayner, 2009; Blythe, Haikio, Bertam, Liversedge, & Hyona, 2011; 刘妮娜, 2015)。正如前面提到的, 词 n 消失文本的操纵会改变读者基于消失文本的眼跳计划, 迫使读者的注视迅速离开当前词, 因此读者的再注视比率会显著减少; 并为保证阅读的正常进行, 读者需要延长注视时间, 本研究中我们发现目标词的总注视时间确实在消失条件下显著长于正常文本呈现条件。但凝视时间是指读者第一遍加工时, 在目标词上多次注视的时间之和, 它与再注视比率之间存在正相关; 消失条件下的再注视比率减少, 凝视时间也表现出减少。同时, 读者取消原本眼跳计划、缩短凝视时间是获取充足的时间进行回视, 通过增加回视次数, 读者获得对文本更多的视觉编码时间, 从而保证了对句子的阅读理解。因此, 在整个数据模式下, 我们发现了读者再注视比率的减少、凝视时间减少, 但回视次数增加、总注视时间增加。

另外,关于聋人读者在两种文本条件下的凝视时间均短于健听读者(RC和AC组),不一定表明聋生对中央凹局部词汇的加工效率更高。原因如下:对于局部词汇通达的有效判断指标是目标词的词频效应,而在消失条件下,聋生组在凝视时间上的词频效应并不显著,这说明聋生组在消失条件下的词汇通达受到阻碍;凝视时间的减少可能更多是由于目标词的消失而迫使读者注视离开所导致。并且,综合整体指标上的数据结果,我们发现聋生在中央凹短暂呈现的条件下,其总阅读时间明显增加,这表明其阅读加工受到词n呈现时间不足的影响,而健听读者则与之相反。所以,从整个实验结果推论,仍支持聋生中央凹加工效率的下降。

最后,再次感谢审稿专家的审阅与建议,我们已对您的建议逐一进行了认真修改,如果还有其他方面的问题,请专家进一步批评指正。

第二轮

审稿人1意见:跟第一稿相比,我觉得修改稿的易读性高了很多,尤其是在写作上,逻辑更通顺,对研究问题的阐述更清晰。基于这篇文章探讨了一个很有价值的问题,我倾向于接受它发表在心理学报上。但是,这一版也还存在很多问题,我在下面一一列举,供作者参考修改。

意见1:实验1中,聋人组跟生理年龄匹配组在40ms消失跟正常呈现条件比较没有显著差异,作者解释是他们能够在40ms消失的条件下就从副中央上获取了正常呈现条件下的信息。这个结果跟Rayner, Liversedge, & White, (2006)的是不一样的。作者需要解释两个研究结果不同的可能原因。而且,还存在另外一个可能,聋人读者不管是在正常还是40ms消失条件下,都无法对副中央上的词进行加工,所以两个条件也没有差异。作者在前言假设部分需要阐述清楚,什么样的数据模式支持研究假设。相同的数据模式(即40ms消失跟正常呈现差异不显著)可能反映两种不同的加工,2.3的讨论要再修改。

回应:非常谢谢审稿专家的肯定以及新提出的宝贵意见。

首先,确实如专家所说,Rayner, Liversedge和White(2006)的研究发现,读者在词n+1呈现60ms后消失条件下,总阅读时间受到显著影响,认为词n+1的短暂呈现对读者的总体阅读加工造成了影响;这与关于汉语阅读的相关研究结果不一致,成年的中文读者(大学生)在词n+1呈现40ms后消失条件下的总阅读时间未受到影响(刘志方,张智军,赵亚军,2011;刘妮娜,2015),并且小学五年级学生亦有同样的表现(张巧明,2013);本研究结果与上述汉语阅读研究的结果一致。与Rayner等(2006)研究结果不同的原因可能与汉语文本的特性有关:在拼音文字中(如英语),由于词长较长以及空格的原因,词n+1距离注视点(中央凹区域)较远,而随着文本距离中央凹区域越远,人的认知系统对其加工速度也越低,因而对于拼音文字的读者而言,词n+1呈现60ms难以满足读者快速地从文本中获取所需的全部视觉信息并进行编码;而汉字是一种空间密度较高的书写文字,词与词之间不存在空格,因此,汉语文本中的词n+1相对于拼音文字而言,可能更靠近于注视点,距离中央凹区域更近,因而汉语读者能够在词n+1短暂呈现的条件下完成视觉信息的获取及编码。

第二,专家的意见中提出了另一种可能的数据解释:“聋人读者不管是在正常还是40ms消失条件下,都无法对副中央上的词进行加工,所以两个条件也没有差异。”这点在理论上确实是可能存在的,但是在句子阅读过程中,读者仅依靠中央凹区域处的文本信息并不能高效地完成阅读加工,副中央凹区域处的信息对阅读加工也至关重要(Hyönä 2011);当副中央凹区域处的信息被掩蔽时,读者阅读效率将受到严重的影响(Rayner, 1998; 2009)。正如关于聋人知觉广度的研究发现,相比于健听读者,聋人读者的阅读知觉广度更大(B danger, Slattery, Mayberry, & Rayner, 2012),中国聋人大学生的右侧阅读知觉广度约为2~3个字(乔静芝,张兰兰,闫国利,2011)。因此,无论是正常呈现还是40ms消失的阅读条件下,聋人读者必然会对阅读知觉广度范围内的副中央凹词n+1进行加工,以保证完成阅读任务。

综上所述,我们参考专家的建议,对前言中假设部分以及2.3的讨论做出了相应修改,具体请详见修改稿。

意见 2: 实验 1 局部分析, 目标词在正常阅读条件下, 词频效应不显著。词频效应是很稳定的效应, 在许多研究中被重复, 为什么在本研究中没有呢?

回应: 非常感谢审稿专家的疑问。

正如专家所说, 词频效应是阅读中较为稳定的效应; 尽管在实验 1 正常文本呈现条件下凝视时间这一指标上词频效应不显著, 但在总注视时间上的差异是显著的 ($b=44.82$, $SE=16.51$, $t=2.46$), 即该实验中的词频效应在晚期指标上是存在的, 并非未发现词频效应; 并且, 以往研究也表明词频效应在语言加工的后期阶段仍很明显 (Yang & McConkie, 1999)。

意见 3: 文章一个重大的修改是重新分析了数据。我很赞同审稿专家 2 的意见用线性混合模型。但是, 作者在汇报线性模型时候, 很多信息没有提供。例如, 模型是怎么定义的, 对于不同组的被试, 组间的比较是建立了 contrasts 还是以某组为基线, 其它组跟它比较呢? 要检查主效应、交互作用和简单效应是否需要建立两个模型? 模型是如何定义的, 是否包含 random intercepts 和 random slopes 呢, 作者需要说明。特别是, 作者没有提到不同因素之间的交互作用, 是因为不显著, 还是没有包含到模型里呢?

回应: 非常感谢审稿专家的意见。上一稿中关于线性混合模型的相关描述确实太简略, 在二轮修改中, 我们对此进行了补充, 具体如下并详见修改稿中“2.1.5 统计方法与处理”:

对每一项眼动指标, 我们建立了以被试和项目因素为交叉随机效应 (crossed random effects), 被试类型、文本呈现条件、词频为固定变量 (fixed factors) 的模型; 并在模型内进行了两次连续差异对比 (successive difference contrasts, Venable & Ripley, 2002) 以比较聋生组与阅读能力匹配组以及聋生组与生理年龄匹配组之间的差异, 由此产生两个三阶交互作用 (DS vs RC * 文本呈现方式 * 词频, DS vs AC * 文本呈现方式 * 词频)、两个与文本呈现方式因素的二阶交互作用以及两个与词频因素的二阶交互作用; 模型首先包含全部的随机效应 (随机截距 random intercepts 和随机斜率 random slopes), 之后, 通过似然比检验 (likelihood ratio tests) 对模型进行比较, 并逐渐简化至包含被试和项目的随机截距。

意见 4: 一般 t 值大于 2 才被视为效应显著。

回应: 非常感谢审稿专家的意见。已参考专家的建议进行了修改, 详见“2.1.5 统计方法与处理”。

意见 5: 结果部分还有不少复制粘贴过来的错误。例如 2.2.2 整体分析里, (2) 跳读率的指标里出现“进一步比较三组被试在两种文本呈现条件下阅读时间上的差异”, 阅读时间上的差异这个表述出现在好几个指标里。还有 2.2.3 局部分析, 进行 $3*2*2$ 的重复测量方差分析。该问题在实验 2 也存在。

回应: 非常感谢审稿专家的细致审阅。实在抱歉, 文中确实存在诸如此类的错误, 二轮修改稿中已对此类问题进行了更改。

意见 6: “用于探究这一问题最有效的方法是消失文本范式”, 这个表述是不准确的, 移动掩蔽范式也可以用来探讨该问题, 究竟哪种范式更好, 我们没有比较, 而且这个也不是本文的重点。作者可以改一下表述, 不要用“最有效的方法”, 而是其中之一的方法。

回应: 非常感谢审稿专家的建议。诚如专家所说, 使用“最有效的方法”这种表述确实不妥, 因此在二轮修改稿中已修改为“有效的方法之一”。

意见 7: 类似的, 在解释呈现 40ms 后消失, “读者需要快速地将该词的视觉信息转换为稳定视觉编码”, 我觉得用“稳定的视觉编码”是不太准确的, 因为在不同的加工情况下, 词语呈现 40ms 能达到什么样的加工水平不是一概而论, 我参考了 Rayner, Liversedge, White, & Vergilino-Perez. (2003) 相关的文章, 他们是这样解释的, “A striking result from this paradigm is that reading proceeds quite normally if the reader gets to see the text for 50 – 60 ms before it disappears or the mask appears. This does not mean that words can be identified in 50 – 60 ms, but it does mean that readers are able to encode all of the visual information that they need from the text in this time frame. 供作者参考。

回应: 非常感谢审稿专家的指导和建议。我们在参考 Rayner et al.(2003) 研究中的相关表述之

后，在二轮修改稿中将“读者需要快速地将该词的视觉信息转换为稳定视觉编码”修改为“读者需要快速地从文本中获取所需的全部视觉信息并进行编码”。

意见 8：前言第三段，“中国聋人读者能够在更早期的加工中（反映在首次注视时间上）就从副中央凹预视加工中获得语义预视效应”，这个表述是否准确？以大学生读者为被试，也能在首次注视时间上看到语义预视效应（Yan et al., 2009）。前言部分最后一段，“词频...并在语义加工的早期阶段产生影响”。这个表述我不是很理解，“语义加工的早期阶段”通过什么指标反映？以及在讨论部分第四段“能够在更早期地从副中央凹预加工中获得语义预视效应”，我觉得作者应该小心用这样的表述。

回应：非常感谢审稿专家的意见。

首先，关于专家对原文中“中国聋人读者能够在更早期的加工中（反映在首次注视时间上）就从副中央凹预视加工中获得语义预视效应”这一表述的疑问，我们再次研读了原文献 Yan, M., Pan, J., Banger, N. N., & Shu, H. (2014). Chinese deaf readers have early access to parafoveal semantics.在其研究中，确实采用了“parafoveal semantic information is obtained earlier among the deaf readers”这类表述，原因在于其数据结果发现，聋人读者在首次注视时间上表现出了显著的早期语义预视效应，而健听读者的语义预视效应则不显著；但在相对晚期的眼动指标上（凝视时间、总阅读时间），健听读者表现出了显著的语义预视效应。因而，Yan 等人认为，相比于聋人读者，健听读者的语义加工可能稍晚一些。而专家提到的 Yan 等人（2009）研究中发现，成年健听读者能够在首次注视时间上即表现出语义预视效应，两项研究结果存在差异的原因可能在于：①两项研究中健听读者的阅读水平有差异，由于 Yan 等人（2014）研究中采用了与聋人读者（中学生，平均年龄为 18.6 岁）阅读水平匹配的健听读者（小学生，平均年龄为 10.7 岁），因而其阅读水平可能略低于以往研究中的成年读者；②两项研究中所采用实验材料的句子框架及难度可能与以往研究不同，因而两项研究的结果难以进行直接比较。原文中，我们可能未将 Yan 等人研究中被试匹配的差异阐述清楚，因而在修改稿中将这点特意说明，具体请详见修改稿。

其次，十分抱歉的是，专家所提出的第二点“词频...并在语义加工的早期阶段产生影响”中“语义加工”为笔误，此处应为“语言加工”；语言加工的早期阶段在眼动研究中主要通过单一注视时间、首次注视时间等指标来反映。再次向专家致歉，并感谢专家的细致审阅。

意见 9：聋人视觉注意资源的空间分布发生改变“这是作者引用来的观点，但是为什么这个改变不是增强中央凹的加工，而是副中央凹的加工呢？正如作者文中说的，中央凹是视敏度最好，加工效率最高的地方，那这种补偿加工为什么不是发生在中央凹加工上呢？希望作者可以提供多一点这方面的信息。

回应：非常感谢审稿专家的意见。聋人视觉注意资源的空间分布是与其环境适应性相关的，即聋人在生活环境中缺乏来自更广知觉范围的听觉线索，例如健听者能够听到呼啸而来的汽车声或咯吱作响的开门声，但聋人无法通过听觉感知这些信息，需要借助其他感官通道，特别是视觉通道（Dye, Hauser, & Bavelier, 2008），因而聋人会将更多的注意资源分配到副中央凹及边缘视野以获取更大范围的信息。尽管中央凹处的视敏度更大，大脑对中央凹处信息的加工效率更高，但中央凹的视野范围较小（2°视角左右），难以满足聋人通过视觉加工进行功能性补偿的需求。

意见 10：缺参考文献，例如闫国利，巫金根，胡晏雯，白学军，2010；刘志方，张智君，潘运，全文，苏衡，2017，在“参考文献”中没有。请作者仔细检查是否还存在类似的情况。

回应：非常感谢审稿专家的细致审阅。实在抱歉，文章中参考文献确实存在诸多问题，在二轮修改稿中已对参考文献进行了仔细的核查。具体请见修改稿的参考文献部分。

意见 11：Buckley, Codina, Bhardwaj&Pascalis ,2010 请核对该文献是否准确，并补充在“参考文献”中

回应：非常感谢审稿专家的细致审阅。实在抱歉，文章中参考文献确实存在诸多问题，在二轮修改稿中已对参考文献进行了仔细的核查。具体请见修改稿的参考文献部分。

意见 12: 逗号的使用很混乱, 有在同一索引中中英文逗号混用, 如: 刘志方, 王若逸, 张智君, 刘炜, 2016; 闫国利, 胡晏雯, 刘志方, 张智君, 2009; 张巧明, 2013; 或是英文索引人名后不加空格 (如: Bavelier, Dye & Hauser, 2006; Dye, Hauser & Bavelier, 2009; Dye & Bavelier, 2010), 中文索引人名后加空格的。不一一列举, 请作者自行检查修改。

回应: 非常感谢审稿专家的细致审阅。实在抱歉, 文章中参考文献确实存在诸多问题, 在二轮修改稿中已对参考文献进行了仔细的核查。具体请见修改稿的参考文献部分。

最后, 再次感谢审稿专家的审阅与建议, 我们已对您的建议逐一进行了认真修改, 如果还有其他方面的问题, 请专家进一步批评指正, 我们将进一步进行修改。

.....

审稿人 2 意见: 论文作者对审稿人提出的问题做了仔细回答, 同时也采用基于 R 软件的 LME 模型详细分析了数据, 论文质量有了显著改善。尽管数据分析的方法仍有待提升, 基本模式可以接受。仔细审阅论文后, 审稿人认为仍有两个问题需要解决。

意见 1: 补充调查句子合理性。

回应: 非常感谢审稿专家的建议。二轮修改稿中, 我们补充了对实验材料合理性的评定。具体做法如下: 另请不参与眼动实验的小学五年级学生 20 人、高中一年级学生 20 人分别对两项实验句子的合理性进行 5 点量表评定 (非常不合理为 1 分 ~ 非常合理为 5 分), 小学生合理性评分为 3.70 ± 0.37 , 中学生合理性评分为 3.65 ± 0.34 , 句子合理性差异不显著 ($t = 1.49$, $p > .05$)。

意见 2: 跳读率、回视率应该采用 logic mixed-effect model。当前, 采用 linear mixed-effect model 处理数据存在问题。

回应: 非常感谢审稿专家的建议。一轮修改稿中, 我们采纳专家的建议, 使用 R 软件的 LMM 模型对数据重新进行了分析; 其中, 在对跳读率这一指标进行分析时, 尽管采用的是 linear mixed-effect model, 但我们根据跳读率的分布检验 (qq-norm) 将跳读率进行了对数转换 (log-transform) 后符合 LMM 的基本假设, 进而将对数转换后的跳读率作为因变量进行分析。这在以往研究中有类似做法 (Kliegl, Masson, & Richter, 2010; Yan, Pan, B dänger, & Shu, 2014; Pan, Shu, Wang, & Yan, 2015), 因此, 本轮修改稿中采用的是上述做法, 并在数据分析中进行了简要说明。

最后, 再次感谢审稿专家的审阅与建议, 我们已对您的建议逐一进行了修改, 如果还有其他方面的问题, 请专家进一步批评指正, 我们将认真进行修改。

第三轮

审稿人 2 意见: 研究者对审稿人提出的问题做出了回应, 审稿人基本满意。但是, 关于跳读率, 审稿人仍然认为正确的做法是 logist mixed-effect model。目前, 研究者将跳读这种二项分布的数据转换成连续性数据, 同时通过 log 转换, 使之线性。这里对数据进行了两次修正, 尽管符合线性要求, 但是仍不符合科学的统计思维。本来就是二项分布数据, logist mixed-effect model 就可以很好解决问题, 为什么要两次转换呢? 同时, 转换后是否更好呢? 在跳读率上没有研究来证实。但是, 我可以提供一个发在在 Journal of Memory and Language 上的论文, 研究者讨论了正确与否 (二项分布) 和正确率, 哪一种更好? 统计计算后, 发现基于二项分布的正确与否是更好的数据源。这对于跳读率的数据处理是一个借鉴。所以, 期待研究者加强多广义线性模型的学习。

回应: 非常谢谢审稿专家的宝贵意见。

首先, 需要致歉的是, 作者在广义线性模型的掌握方面确实存在一定的欠缺, 今后会加强相关的学习, 专家提供的相关文献也十分值得拜读。

我们查阅了大量眼动研究文献, 以往研究中鲜有使用“logist mixed-effect model”处理跳读率或回视率等数据。专家的指导很有借鉴意义, 作者之后会继续研究这一模型在处理二项分布数据中的优势。由于以往研究 (Kliegl, Masson, & Richter, 2010; Yan, Pan, B dänger, & Shu, 2014; Pan, Shu, Wang, & Yan, 2015; Cop, Dirix, Assche, Drieghe, Duyck, 2017; Chen et al., 2017)

中也有采用与本研究类似的做法，因此，与专家商榷，我们能否仍沿用国外大部分专家的做法，继续使用 **linear mixed-effect model**，将跳读率进行了对数转换（**log-transform**）后进行分析。希望能够得到外审专家的谅解。与此同时，我们会进一步认真阅读您提供的文献，学习您建议的新统计方法。

再次感谢审稿专家的审阅与建议，敬请编委专家进一步批评指正，如果还有其他方面的问题，我们将认真进行修改。