

《心理学报》审稿意见与作者回应

题目：句子结构与控制动词类型对题元角色指派的影响

作者：李芳 李馨 张慢慢 白学军

第一轮

审稿人 1 意见：

该文作者采用眼动技术对句子理解过程中句子结构和控制动词类型的作用进行探讨，整个研究作者均尽力按照心理语言学研究标准来控制和匹配材料，严格实验流程，尝试在前人研究的基础上得出一些有益的结论。纵观全文，还有以下问题值得商榷。

意见 1：作者设置了两类句法结构，即居中结构和前置结构。那么两类结构在语料中的使用频率，读者的熟悉程度，句法引发的歧义，省略逗号所导致的加工难度等都是混淆作者实验设计结果的无关变量，虽然作者匹配了两类材料的通顺程度，但是，就汉语而言，语义加工可以在语法不够通顺的情况下顺利完成，因此，作者所讨论的实验结果很可能受到上述因素的干扰。

回应：非常感谢专家老师的宝贵意见。我们认同专家指出的两类结构在使用频率、熟悉程度、句法引发的歧义、省略逗号所导致的加工难度存在差异。前置结构的使用频率、熟悉程度低于居中结构，加工难度更大并带有句法歧义。但是，上述差异是自变量操纵的结果而不是由无关变量引起。

为了考察汉语读者对语序线索的依赖性，我们设置了居中结构（包含语序和格标记线索）和前置结构（仅包含格标记线索）两种句子结构。

前置结构（例如，“对老齐老徐懊悔说应该及早采取防护措施”）是省略了逗号的前置型“对”字结构（例如，“对老齐，老徐懊悔说应该及早采取防护措施”）。如果不省略逗号，前置型“对”字结构仍然同时存在语序和格标记线索，读者依据语序线索预期“老徐”是施事，并依据介词“对”的格标记线索预期“老齐”的语义角色。

省略了逗号之后，破坏了前置结构的语序线索，读者只能依据介词“对”的格标记线索预期“老齐老徐”的语义角色。

前置结构只保留格标记线索，从而与居中结构（例如，“老徐对老齐懊悔说应该及早采

取防护措施”）进行对比，来探讨汉语读者对语序线索的依赖性。

与此同时，破坏语序线索被使得前置结构的使用频率、熟悉度低于居中结构，且加工难度更大。也就是说，对自变量（句子结构）的操纵使得两种句子结构的使用频率、熟悉度和加工难度存在差异。

以往关于论元线索权重的研究，自变量操纵也会引起不同实验条件的句子在使用频率、熟悉度和加工难度等方面存在差异（Ferreira, 2003; Kuperberg et al., 2007; Thothathiri et al., 2018）。例如，Wang 等人（2009）探讨汉语语序与生命性线索（即名词论元是有生命的客体还是无生命的客体）的权重，共设置三种类型的首论元：（1）无生命-受事宾语（“小说理解了演员”）、（2）有生命-施事主语（“演员理解了小说”）和（3）无生命-施事主语（“小说教育了演员”）三种条件。显然，与条件（2）相比，条件（1）和（3）的句子不符合语用习惯，具有较低的使用频率、熟悉程度和较高的难度。

参考文献：

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, 47(2), 164–203.

Kuperberg, G. R., Kreher, D. A., Sitnikova, T., Caplan, D. N., & Holcomb, P. J. (2007). The role of animacy and thematic relationships in processing active English sentences: Evidence from event-related potentials. *Brain and Language*, 100(3), 223–237.

Thothathiri, M., Asaro, C. T., Hsu, N. S., & Novick, J. M. (2018). Who did what? A causal role for cognitive control in thematic role assignment during sentence comprehension. *Cognition*, 178, 162–177.

意见 2：作者在前言中所讨论的所有语料均是基于对语法良好的分析和加工。但是，就作者提供的语料来看，主控动词和宾控动词例句中的“动词”并非句子中真正的动词。句 1 中，“懊悔”主要发挥修饰作用（状语功能），句子的主干是“老徐说采取防护措施”。而句 2 中，“提醒”发挥连动作用（动词功能），句子的主干是“老徐提醒说采取防护措施”。那么，作者所认为的“懊悔”是动词就不能成立，作者基于这一假设的设计以及后续的结果是否成立值得商榷。纵观作者的材料，类似的情况不少。

回应：非常感谢专家老师的宝贵意见。

本研究所选用的动词除了行为动词（例如，“筹备”等）和心理动词（例如，“钦佩”等），还包含言语动词。言语动词是指称语言性行为的动词，句型鉴别格式为“N₁+V+说/道（+N₂）”（王莹, 2005）。“懊悔”符合该句型鉴别格式，代入句型为：“老徐懊悔说应该及早

采取防护措施”，介词“对”介引与事角色之后，句子变为实验材料示例中的“老徐对老齐懊悔说应该及早采取防护措施”。“提醒”符合该句型鉴别格式，代入句型为：“老徐提醒说老齐应该及早采取防护措施”。介词“对”介引受事角色，句子变为实验材料示例中的“老徐对老齐提醒说应该及早采取防护措施”。因此，实验材料示例中的“懊悔”和“提醒”都是动词。本研究中的其他动词也都具有相应的语言学依据。

此外，我们采用评定任务，确认了这些动词在句中的词性。修改稿中增述了动词词性评定过程与评定结果：“为了确认两类控制动词在句中的词性，请不参与正式实验的 40 名本科生完成词性判断任务。评定材料分 4 个版本，每个版本有 10 人参与评定。每个版本共 80 个句子，包含 26 个填充句。填充句要求被试判断的词语均为名词。具体评定方式是：给出完整句子并设置词性判断问题（例如，请问“懊悔”在句中是什么词性？），要求被试填写两类控制动词的词性。以动词词性符合度（判断为动词的评定数与总评定数的比值）作为评价指标。如表 2 所示，四类实验句的动词词性符合度在 0.98~0.99 之间，表明两类控制动词在句中符合动词词性。”

基于以上分析可知，本研究所选用词语均为动词。

参考文献：

王莹. (2005). 现代汉语言语动词研究. *南开语言学刊*, (1), 76-82.

意见 3: 本研究的自变量为动词类型。但是对动词的选取使得两类材料的语法和语用并未完全匹配。在句法变换中，被字句或者把字句无论如何变换，都是基于句子主干的变换。句子完全可以独立。但是，作者所选用的两类动词，以“懊悔”和“提醒”为例，无论是“A 对 B 懊悔”还是“A 懊悔 B”均缺乏具体的语用环境，非常突兀。同时，“A 对 B 提醒”觉得很怪，而“A 提醒 B”则是正确表达。就以上分析，作者引入“对”线索是否合适，以及是否使相关动词保持原样值得商榷。

回应: 非常感谢专家老师的宝贵意见。

第一个问题，关于引入“对”线索是否合适。

两类控制动词置于“对”字结构之后，“A 对 B 懊悔”或者“A 对 B 提醒”显得有些突兀和奇怪，这种突兀和奇怪表现为通顺性的评定结果低于其他汉语阅读研究。本研究采用 7 点量表（4=介于不通顺和通顺之间，5=比较通顺，6=通顺，7=非常通顺）评定的实验材料平均通顺性为 4.80（两类居中结构的平均通顺性为 5.13）。检索近十年《心理学报》上发表的汉语

阅读研究中采用 7 点量表评定句子通顺性的结果（共 11 篇），平均通顺性在 5.75~6.48。

但是，引入“对”线索之后，两种控制动词条件下的通顺性没有差异，四种实验条件下的通顺性也没有差异，并且以通顺性为协变量进行线性混合模型分析也发现，通顺性不影响主要的实验结果。修改稿的脚注 2 概述了通顺性对结果的影响：“兴趣区 1、2、3 的所有眼动指标以及兴趣区 4 的第一遍阅读时间和回视路径时间上，通顺性的主效应均不显著， $t/s < 1.63$, $ps > 0.105$ 。兴趣区 4 的第二遍阅读时间和总回视入（出）次数上，通顺性的主效应（边缘）显著， $t/s > 1.89$, $ps < 0.060$ ，因此把通顺性作为自变量进行三因素的固定效应分析（正文报告的是该分析的结果），通顺性与其他变量的二阶和三阶交互作用不显著， $t/s < 1.47$, $ps > 0.145$ 。上述统计分析结果表明，通顺性不影响句子结构与控制动词类型的主效应和交互作用。”

第二个问题，关于引入“对”线索是否使相关动词保持原样。如回答专家老师的问题 2 所述，依据语言学知识和客观的评定结果（动词词性符合度在 0.98~0.99 之间）可以确认句中动词的词性，表明引入“对”线索之后动词仍保持原样。

因此，可以引入“对”线索考察语序线索的作用以及论元线索与控制动词论元表征一致性。

意见 4: 作者采取较为精细的兴趣区分析，但是眼动校准为何未采取 9 点校准？每个 BLOCK 都有 54 个实验句子，那么设置 18 个填充材料是否太少？

回应: 非常感谢专家老师的宝贵意见。

第一个问题，关于眼动校准类型的选择。本研究使用的实验材料均为单行句子，水平 3 点校准（左中、中间中、右中）即可满足本实验的需求，这不仅可以简化校准过程，而且还可以提高校准数据的精确度，与以往使用单行句子阅读的研究一致（例如，白学军 等, 2015; 王永胜 等, 2018; Juhasz et al., 2006; Li, Liu, & Rayner, 2015; Liu et al., 2020; Zang et al., 2019; Zhang et al., 2019）。9 点校准一般适应于实验材料呈现为全屏的情况（例如，篇章阅读）。为了便于理解校准点的选用，我们在“2.5 实验程序”增述了选用三点校准的一些解释：“实验材料均为单行句子，因此采用既简便又具有足够精度的水平三点校准”。

第二个问题，关于填充材料的数量是否过少。

在以往实验材料为特殊句型的研究中，例如，Ye 等人（2006）的研究借助“把”字结构考察汉语句子理解过程中的句法和语义加工，特殊句型与一般句型的比例为 5:1。还有些研究，如 Ye 等人（2007）以及 Wang 等人（2013）考察“把”字句和“被”字句的句法和语义加

工时, 实验句和填充句均为特殊句型的句子。在阅读的眼动研究中, 实验句与填充句的比例通常为 3:1~4:1 (白学军 等, 2015; 关宜韞 等, 2019; 杨帆 等, 2020; 王永胜 等, 2018; 钟伟芳 等, 2012; Slattery et al., 2011)。本研究实验句与填充句的比例为 3:1。对比以往研究, 本研究的填充材料数量尚可。

参考文献:

- 白学军, 王永胜, 郭志英, 高晓雷, 闫国利. (2015). 汉语阅读中词 $n+2$ 的预视对高频词 $n+1$ 加工影响的眼动研究. *心理学报*, 47(2), 143–156.
- 关宜韞, 宋悉妮, 郑玉玮, 张颖靓, 崔磊. (2019). 中文词间词和词内词预视加工的差异:词间阴影的作用. *心理学报* 51(9), 969–981.
- 杨帆, 隋雪, 李雨桐. (2020). 中文阅读中长距离回视引导机制的眼动研究. *心理学报*, 52(8), 921–932.
- 王永胜, 赵冰洁, 陈茗静, 李馨, 闫国利, 白学军. (2018). 中央凹加工负荷与副中央凹信息在汉语阅读眼跳目标选择中的作用. *心理学报*, 50(12), 336–1345.
- 钟伟芳, 莫雷, 金花, 徐贵平. (2012). 句子阅读理解过程中句意的建构时间. *心理学报*, 44(6), 735–744.
- Juhasz, B. J., Liversedge, S. P., White, S. J., & Rayner, K. (2006). Binocular coordination of the eyes during reading: Word frequency and case alternation affect fixation duration but not fixation disparity. *Quarterly Journal of Experimental Psychology*, 59(9), 1614–1625.
- Li, X., Liu, P., & Rayner. (2015). Saccade target selection in Chinese reading. *Psychonomic Bulletin & Review*, 22(2), 524–530.
- Liu, W., Inhoff, A. W., & Li, X. (2020). Attention shifting during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception and Performance*, 46(9), 979–990.
- Rayner, K., Schotter, E. R., & Drieghe, D. (2014). Lack of semantic parafoveal preview benefit in reading revisited. *Psychonomic Bulletin & Review*, 21(4), 1067–1072.
- Slattery, T. J., Angele, B., & Rayner, K. (2011). Eye movements and display change detection during reading. *Journal of Experimental Psychology: Human Perception and Performance*, 37(6), 1924–1938.
- Wang, S., Mo, D., Xiang, M., Xu, R., Chen, H. (2013). The time course of semantic and syntactic processing in reading Chinese: Evidence from ERPs. *Language and Cognitive Processes*, 28(4), 577–596.
- Ye, Z., Luo, Y.-J., Friederici, A. D., & Zhou, X. (2006). Semantic and syntactic processing in Chinese sentence comprehension: Evidence from event-related potentials. *Brain Research*, 1071, 186–196.
- Ye, Z., Zhan, W., & Zhou, X. (2007). The semantic processing of syntactic structure in sentence comprehension:

An ERP study. *Brain Research*, 1142, 135–145.

Zang, C., Du, H., Bai, X., Yan, G., & Liversedge, S. P. (2019). Word skipping in Chinese reading: The role of high-frequency preview and syntactic felicity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(4), 603–620.

Zhang, M., Liversedge, S. P., Bai, X., Yan, G., & Zang, C. (2019). The influence of foveal lexical processing load on parafoveal preview and saccadic targeting during Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance*, 45(6), 812–825.

意见 5: 两类句子的通顺性不同，句子结构的主效应显著，是否暗含了前置结构不常用、不熟悉，省略句号导致的加工难度？

回应: 非常感谢专家老师宝贵意见。

首先，尽管两种结构的通顺性存在差异，但是以通顺性为协变量的分析结果显示，通顺性并不影响句子结构的主效应(请见脚注 2: “兴趣区 1、2、3 的所有眼动指标以及兴趣区 4 的第一遍阅读时间和回视路径时间上，通顺性的主效应均不显著， $|t/s| < 1.63$ ， $ps > 0.105$ 。兴趣区 4 的第二遍阅读时间和总回视入（出）次数上，通顺性的主效应（边缘）显著， $|t/s| > 1.89$ ， $ps < 0.060$ ，因此把通顺性作为自变量进行三因素的固定效应分析（正文报告的是该分析的结果），通顺性与其他变量的二阶和三阶交互作用不显著， $|t/s| < 1.47$ ， $ps > 0.145$ 。上述统计分析结果表明，通顺性不影响句子结构与控制动词类型的主效应和交互作用。”)。因此，句子结构的主效应并不是由于二者的通顺性差异所致，而是由于两种句子结构的语序线索差异所致。

其次，关于前置结构不常用、不熟悉、省略句号导致的加工困难。如回复专家老师的问题 1 所述，前置结构与居中结构在使用频率、熟悉度和加工难度上的差异恰恰是由对自变量（即句子结构）的操纵引起的。前置结构破坏语序线索而仅保留格标记线索，导致前置结构的使用频率和熟悉度低于居中结构（同时包含语序和格标记线索），且加工难度更大。

意见 6: 在表格呈现中，星号应呈现为 T 结果的上标。在表 9 的结果中，P 值为 0.08，作者均标记为边缘显著，是否合适？

回应: 非常感谢专家老师的宝贵意见。

第一个问题，关于表格中星号的位置。为了精简结果，修改稿去掉了各兴趣区固定效应分析结果的庞大表格，而采用在正文中呈现的方式，具体请见结果部分的修改。

第二个问题,关于原表 9 的 $p=0.08$ 标记为边缘显著是否合适。Pritschet, Powell 和 Horne (2016) 统计分析了标记 p 值边缘显著的 459 项实证研究发现, 92.6% 的研究者把 0.05-0.10 范围内的 p 值标记为边缘显著, 甚至有研究者把 $p=0.18$ 标记为边缘显著。并且, 研究者们 (Greenland et al., 2016; Wasserstein & Lazar, 2016) 认为, 边缘显著时相对较大的 p 值并不是支持虚无假设的证据, 而是意味着拒绝虚无假设的证据稍弱。因此, $p=0.08$ 可以标记为边缘显著, 并且可以作为支持研究假设的证据。此外, 本研究的结论并不唯独基于该结果, 而是结合其他多项统计显著的结果进行推断。结合兴趣区 4 句子结构与控制动词类型的交互作用, 该结果支持汉语读者依赖语序线索 (见“讨论 4.1 论元线索权重”)。本研究还发现前置结构的名词 1 区域、名词 2 区域和动词区域的阅读时间长于居中结构 (多个指标在 $p < 0.001$ 水平显著)。结合显著的句子结构主效应, 我们得以推断读者依赖语序线索 (见“讨论 4.1 论元线索权重”)。

参考文献:

- Pritschet, L., Powell, D., & Horne, Z. (2016). Marginally significant effects as evidence for hypotheses: Changing attitudes over four decades. *Psychological Science*, 27(7), 1036–1342.
- Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European Journal of Epidemiology*, 31, 337–350.
- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA’s statement on p-values: Context, process, and purpose. *The American Statistician*, 70, 129–133.

意见 7: 讨论中 4.1 作者完全未展开? 而且句子也很难理解。

回应: 感谢专家老师的宝贵意见。我们对该部分的表述作了较大修改, 修改稿“4.1 论元线索权重”基于两项主要研究结果 (前置结构加工负荷大于居中结构; 前置结构比居中结构的线索整合发生在更晚的句子成分) 验证预期 1, 通过与英语 (同属于固定语序的语言) 和德、意、日、土耳其语 (语序自由的语言) 的对比, 讨论扩展论元依存模型所主张的论元线索权重的跨语言差异性, 并从竞争模型的角度解释论元线索权重跨语言差异的内在机制。

.....

审稿人 2 意见：

该研究采用眼动技术考察了句子结构和控制动词类型对题元角色指派的影响，并验证了论元线索权重具有语言特异性，题元角色再分析具有跨语言普遍性。整个研究具有一定的创新性，设计严谨，写作规范，语言简练。综合来看不失为一个优秀的研究，但有以下几个问题需要注意。

意见 1：研究的结果呈现部分过于繁琐，需进行适当删减以突出重点。

回应：非常感谢专家老师的宝贵建议。修改稿“3.2 各兴趣区的分析结果”部分作了如下删减和调整：（1）删除各兴趣区固定效应分析结果的庞大表格，改为在正文中报告相关的统计数值；（2）正文不再逐个兴趣区报告通顺性的主效应，改为在脚注 2 概述通顺性对实验结果的影响：“兴趣区 1、2、3 的所有眼动指标以及兴趣区 4 的第一遍阅读时间和回视路径时间上，通顺性的主效应均不显著， $t/s < 1.63$ ， $ps > 0.105$ 。兴趣区 4 的第二遍阅读时间和总回视入（出）次数上，通顺性的主效应（边缘）显著， $t/s > 1.89$ ， $ps < 0.060$ ，因此把通顺性作为自变量进行三因素的固定效应分析（正文报告的是该分析的结果），通顺性与其他变量的二阶和三阶交互作用不显著， $t/s < 1.47$ ， $ps > 0.145$ 。上述统计分析结果表明，通顺性不影响句子结构与控制动词类型的主效应和交互作用。”（3）不以眼动指标分别介绍结果，而是围绕兴趣区，按照句子结构的主效应、控制动词类型的主效应以及二者的交互作用的形式分别呈现结果。

意见 2：文中认为“由于自定步速阅读和反应时记录的方法限制，无法确定该差异究竟源于动词论元表征的差异还是源于语序线索与动词论元表征不一致效应”，而本研究采用整屏呈现实验材料，请详细说明是如何确定这一差异。

回应：非常感谢专家老师的宝贵意见。动词论元表征属于词汇语义加工过程，反映在早期眼动指标。语序线索与动词论元表征不一致效应发生于句子整合加工过程，反映在晚期眼动指标。本研究主要通过早期眼动指标（第一遍阅读时间）和晚期眼动指标（第二遍阅读时间等）区分这两种差异来源。修改稿的表述为：“但是由于自定步速阅读和反应时记录的方法限制，无法确定该差异究竟源于动词论元表征的差异还是源于语序线索与动词论元表征不一致效应。前者属于词汇加工过程，后者属于句子整合加工过程，其出现时间晚于前者。采用眼动技术不仅可以记录读者自然阅读过程，而且可以通过早期眼动指标（例如，第一遍阅读时间）和晚期眼动指标（例如，第二遍阅读时间）分别揭示词汇加工过程和句子整合过程，考察晚

期句子整合阶段的语序线索与动词论元表征不一致效应。”

此外，讨论部分“4.2 论元线索与控制动词论元表征的线索整合”也进行了相应调整。我们从早期眼动指标和晚期眼动指标两个方面分别讨论控制动词类型的主效应，探讨早期眼动指标与两类控制动词论元表征差异的关系、晚期眼动指标与语序线索和动词论元表征不一致效应的关系。

意见 3: 由于“对”字结构的格标记线索不够明确，因此对于结论中，汉语题元角色指派更依赖语序线索而非格标记线索这一论述需谨慎。

回应: 非常感谢专家老师的宝贵建议，原文表述确实不够严谨。修改稿的表述为：“在本研究条件下发现：（1）即使存在介词“对”的格标记线索，汉语题元角色指派仍然依赖语序线索，论元线索权重具有一定的语言特异性；（2）汉语论元线索与控制动词论元表征不一致产生更大加工负荷，题元角色再分析具有一定的跨语言普遍性。研究结果支持扩展论元依存模型。”

.....

审稿人 3 意见:

源于 20 世纪 70 年代的句法—语义接口问题，是当前语言学研究的热点议题之一。在这一背景下，该文通过分析汉语语序和格标记线索权重以及控制动词论元表征与论元线索一致性对题元角色指派的影响，探讨句法—语义接口问题的选题，具有一定的创新性。相较于以往研究，该文最明显的创新性在于分析了句法—语义接口中“语序和格标记线索”的权重问题。但是，作为一篇期刊论文，还存在如下不足。

意见 1: 实验材料: 文中提及“参照 Zhai (2011) 的实验材料，选取主语控制动词与宾语控制动词各 54 个”。主语和宾语控制，实为动词的语义指向。对于不同的动词，在不同的句子结构中，语义指向的程度存在较大差异，比如话题（徐晓东等 2013; 2017）。因此，该文作者需要补充说明选取材料的原则和方法。如果能补充语义指向（就是该文作者所说的主语和宾语控制的测评数据），就更加科学了。（徐晓东等（2017）. 汉语话题回指如何受动词语义关系约束：来自脑电研究的证据，外语教学与研究，49（3），323-334. 徐晓东等（2013). 话题结构和动词语义对代词回指的影响：一项基于语言产生和语言理解任务的实证研究，现代外语, 36(4), 331-339.)

回应：非常感谢专家老师的宝贵建议。首先，我们在引言部分介绍主语控制动词和宾语控制动词时，引入语义指向并在脚注 1 对语义指向的概念进行解释。正文修改后为：“主语控制动词指主语是后续动作的施事（动词语义指向¹主语），其论元表征为施事在前；宾语控制动词指宾语是后续动作的施事（动词语义指向宾语），其论元表征为施事在后”，脚注 1 为：“语义指向，指句中某成分在语义上跟哪个成分直接相关（徐晓东 等, 2013; 徐晓东 等, 2017）。比如：人们倾向于从主语控制动词“懊悔”的主语方面说明懊悔的原因，而倾向于从宾语控制动词“提醒”的宾语方面说明被提醒的事宜”。其次，我们根据专家老师建议补评了两类控制动词的语义指向，并在“2.3 实验材料”增述了语义指向的评定过程与评定结果：“为了确认两类控制动词的语义指向（徐晓东 等, 2013; 徐晓东 等, 2017），请不参与正式实验的 40 名本科生完成施事判断任务。评定材料分 4 个版本，每个版本有 10 人参与评定。具体评定方式是：给出完整句子并设置施事判断问题（“谁应该及早采取防护措施？”），要求参与评定者判断施事角色是 NP1（“老徐”）还是 NP2（“老齐”）。如果包含主语控制动词的句子被判断为 NP1 是施事或者包含宾语控制动词的句子被判断为 NP2 是施事，则符合控制动词的语义指向。以语义指向符合度（符合语义指向的评定数与总评定数的比值）作为动词语义指向的评价指标。如表 2 所示，四类实验句的动词语义指向符合度在 0.89~0.94 之间，动词语义指向非常符合该动词的控制信息。”同时，我们还对四类实验句的语义指向符合度进行统计检验：“对动词词性符合度、语义指向符合度进行两因素重复测量方差分析，各主效应和交互作用均不显著， $F_s < 2.55$, $p_s > 0.11$ 。”，各实验条件下的动词语义指向度不存在显著差异。

我们采用的具体评定方法与徐晓东等人（2013, 2017）的方法有所不同，但都是基于完整句意评定动词语义指向。徐老师采用句子补全任务，要求参与者用前句的 NP1 或者 NP2 开头补全逗号后的句子，该评定方法是基于逗号前句的完整句意评定动词的语义指向。本研究采用施事判断任务，要求被试判断整句的施事，因为读者需要基于整句信息才能准确识别句中控制动词的语义指向。采用施事判断任务可以准确、直接地考察句中控制动词的语义指向符合度。

意见 2：该文的预期是两个，而发现和结论都是三条。其实，该文在前言中的预期，可以修改为三条，与文中其他地方的描述保持一致。

回应：非常感谢专家老师的宝贵建议。为了使预期、结果和结论部分更好地相对应，我们经过认真思考，把引言部分的预期修改为：“扩展论元依存模型认为相对固定的语序使得读者

非常依赖语序线索，那么可以预期：（1）前置结构语序线索的缺乏使得其动词前区域的加工负荷大于居中结构。扩展论元依存模型认为论元线索与控制动词论元表征不一致时的题元角色再分析产生更大的加工负荷，那么可以预期：（2）控制动词论元表征与论元线索进行整合时，宾语控制动词条件的加工负荷大于主语控制动词条件，且反映在晚期加工阶段”。预期（1）对应讨论“4.1 论元线索权重”和结论（1）；预期（2）对应讨论“4.2 论元线索与控制动词论元表征的线索整合”和结论（2）。另外，讨论和结论部分的描述也已与其保持一致。

意见 3：眼动指标的选择：不同学者持有不同的观点。如果作者没有使用有些指标，特别是不同类型但可能与该文研究相关的指标，如：瞳孔大小（pupil size），也需要在文中提及未入选的理由。这样，显示作者在指标选择上的全面性。

回应：非常感谢专家老师的宝贵建议。本研究旨在探究汉语读者对语序线索的依赖性，以及控制动词论元表征与论元线索一致性对题元角色指派的影响。前人研究（Li et al., 2015; Gattei et al., 2017）发现，论元线索以及控制动词论元表征与论元线索一致性效应通常在时间维度的眼动指标更加灵敏。前者既可以反映在早期眼动指标，也可以反映在晚期眼动指标。后者涉及句子整合加工，主要反映在晚期眼动指标（早期眼动指标反映的是词汇加工过程中的控制动词论元表征差异）。瞳孔大小可以很好地反映记忆加工以及情绪、兴趣、态度等的变化（杨晓梦 等, 2020; 于洋 等, 2020），也可以反映整体的语言认知加工负荷（闫国利 等, 2013），但是在区分词汇加工和句子整合加工方面的灵敏性比较低，因而不适用于考察控制动词论元表征与论元线索一致性效应。修改稿“3.1 数据处理及眼动指标”关于眼动指标的选用：“参照同类研究（Li et al., 2015; Gattei et al., 2017），采用时间维度的不同眼动指标：……”。

参考文献：

- 闫国利, 熊建萍, 臧传丽, 余莉莉, 崔磊, 白学军. (2013). 阅读研究中的主要眼动指标评述. *心理科学进展*, 21(4), 589–605.
- 杨晓梦, 王福兴, 王燕青, 赵婷婷, 高春颖, 胡祥恩. (2020). 瞳孔是心灵的窗口吗? ——瞳孔在心理学研究中的应用及测量. *心理科学进展*, 28(7), 1029–1041.
- 于洋, 姜英杰, 王永胜, 于明阳. (2020). 瞳孔变化在记忆加工中的生物标记作用. *心理科学进展*, 28(3), 416–425.

意见 4: 结论部分需要增加该文研究的不足和对未来研究的展望。

回应: 非常感谢专家老师的宝贵建议,修改稿结论部分已增述本研究的不足以及对未来研究的展望:“本研究发现汉语题元角色指派非常依赖语序线索,这是基于介词“对”的格标记线索不明确时汉语母语读者的论元线索权重。如前所述,未来研究还需探索格标记线索明确时的汉语论元线索权重,对比在不同线索有效性的格标记条件下的语序与格标记线索权重,以探究语序线索依赖在汉语论元线索权重中的普适性。另外,汉语二语学习者的语序与格标记线索权重是否以及如何受母语语言特性的影响也是值得探究的问题。Bates 和 MacWhinney (1987) 认为初学者对第二语言的题元角色指派仍然依据母语的论元线索权重,但熟练者则转向依据第二语言的论元线索权重。该观点已得到来自汉语等不同语言的证据支持 (Harrington, 1987; McDonald & Heilenman, 1991; Hernandez et al., 1994; Su, 2001; Morett & MacWhinney, 2013), 但这些研究主要考察的是语序和生命性线索(即,论元是有生命客体还是无生命客体)的权重。未来研究还需探究不同汉语习得水平的二语学习者是否表现出语序线索依赖性以及其如何受母语论元线索权重影响。”

此外,我们再次通读全文进行自查,做出了如下修改:

第一,“引言”关于介词“对”以及“对”字结构的论述,改用更有权威性的参考文献以及更严谨的语言表述,修改为:“介词“对”跟随的名词论元在句中承担受事角色或者与事角色(即施事动作的参与者)(孟晓琦,贾秀英,2009)。……(2)前置型结构,“对”字结构位于句首,例如“对老齐,老徐懊悔说应该及早采取防护措施”。居中结构同时包含语序线索和格标记线索;省略了逗号的前置型结构(即“对老齐老徐懊悔说应该及早采取防护措施”,称前置结构)缺乏语序线索而仅包含格标记线索。借助“对”字结构可以探究汉语读者对语序线索的依赖性……。”

第二,“2.1 被试”关于估算样本量的表述修改为:“采用 G*Power 3.1 软件 (Faul et al., 2009) 估算样本量,设置 $f=0.25$ (中等效应量,见 Cohen, 1988), $\alpha=0.05$, Power=0.80, 计算得到样本量为 24 人。”

第三,“4 讨论”增加“下面分别从论元线索权重、论元线索与控制动词论元表征的线索整合两个方面进行讨论。”

第二轮

审稿人 2 意见：

通过对审稿意见的回复和修改，稿件质量得到很大提升，基本同意发表，只是摘要中有些句子不够通顺，如“宾语控制动词条件加工负荷大于主语控制动词条件，且该更大加工负荷在前置结构条件下出现在更晚的句子成分。”后半句的表达还需斟酌。未来研究展望建议放在讨论的最后，而非结论部分。

回应：诚挚感谢专家老师对稿件的充分认可和宝贵建议！

第一个问题，摘要中不通顺的句子，我们修改为“宾语控制动词条件加工负荷大于主语控制动词条件，且与居中结构相比，前置结构时该更大加工负荷出现在更晚的句子成分。”

第二个问题，未来研究展望在文中的位置，我们按照专家老师的建议改放至讨论部分的最后一段，并标记为红色字体。

此外，我们再次通读全文进行自查，把一些语句修改为更加准确、严谨的表述，并标记为橙色字体。

第三轮

编委意见：请作者根据审稿专家的意见进一步修改文字后再审。

回应：诚挚感谢编委专家的宝贵意见。我们认真研读全文并组织多轮讨论，对文章作出如下修改：（1）在第二段，增述对专业术语的解释；（2）在第四段、第六段和第七段补充若干语句，增强逻辑衔接；（3）对不通顺、不准确、不清晰、不简洁的文字表述加以修改。具体修改见正文玫红色字体部分。如有未尽之处，敬请专家老师批评指正。

第四轮

编委意见：本文经过修改，质量有了较大的提高，建议刊用。

主编意见：同意外审和编委意见，建议录用。