

《心理学报》审稿意见与作者回应

题目：身份信息与位置信息的加工进程及语境预测性的影响

作者：徐迺嘉；隋雪

第一轮

审稿人 1 意见：该文采用启动范式的反应时实验和句子阅读的眼动实验，探讨词素加工及词素位置加工在双字词识别中的作用。文章前言对领域进展的梳理较为清晰，实验设计比较有特色。有以下问题或建议：

意见 1：对研究的一个首要疑虑是，双字词的词素位置交换与拼音文字中单词内的字母位置交换是否具有对应关系？或是否具有可比性？如果可比性很弱，那么文章的导入及有关跨语言比较的一些讨论可能需要重新斟酌。在加工层次上，词素涉及到语义层面，字母加工则没有到这个层面。相比较而言，汉字内的部件位置交换，如“杏”与“呆”，与拼音文字中单词内的字母位置交换倒是更接近些。当然，由于不同文字系统的差异，二者也很难完全对应。

回应：感谢审稿人提出的意见。审稿人说的对，这个问题也是曾经困扰我们的问题。双字词的词素换位与拼音文字中的字母换位不具有完全对应关系。字母是没有意义的，汉字是有意义的。本研究想探讨汉语加工中位置与身份信息的加工情况，不能做到与拼音文字完全对应。我们考虑过笔画（stroke）换位、偏旁部首（radical）换位，但是都不可行。最后选择了词素换位。在英文中字母是构成单词的最小单位，而在汉语中，词素是构成词语的最小单位，从这个角度，汉语中的词素换位与英文中的字母换位有比较的可能。另外，审稿人说的对，部件位置交换与字母换位更接近。但是，汉字中部件换位，很多情况下是不能操作的，或者说，很多汉字的部件无法换位。从换位的可能性角度，拼音文字中的字母都可以换位，汉语中词素也都可以换位，具有一定可比性。在实际的书写、印刷以及计算机输入中，拼音文字会出现字母顺序错误的现象，汉字出现的也是词素顺序的错误，而很难出现笔画、部首顺序的错误。关于这个问题我们也痛苦了很久，最后，选择了词素换位与英文中的字母换位进行类似的实验，希望能回答位置信息与身份信息加工的不同。对这个问题，我们在文章中进行了认真的修改，和详细的说明。

意见 2：文章的研究问题没有清晰呈现。

回应：感谢审稿人的问题。我们对引言部分进行了认真的梳理，明确提出了研究问题：汉语词汇加工中是否存在换位效应，考察汉语词汇加工中身份信息和位置信息的加工过程，以及语境预测性对二者的影响。已经在文中进行了补充和修改。

意见 3：通常启动实验会设计一个无关对照组。但实验 1 的设计中，只有四种启动条件，没有设置单独的对照组，是有什么特别的考虑吗？

回应：审稿人说的对，通常来说启动实验确实会设计一个无关对照组，其他启动条件要与无关对照组比较，看启动效应的差异。在本研究中，我们要考察换位非词与替换非词之间的作用大小，我们把原词启动作为基线，看哪个更接近原词启动。字母换位效应，也是这样做的，是将换位词和替换词与原词（base word）进行比较。对于这部分内容，我们也进行了认真的修改和说明，在文中进行了补充。

意见 4: 有些重要概念需要进行界定。例如,什么是“身份信息”?摘要中,直接用“替换启动词”和“换位启动词”而没有适当说明,读者不一定理解其含义。

回应: 非常感谢审稿人的建议,我们进行了认真修改,在文章中的相应位置给出了相应的例子和说明。

.....

审稿人 2 意见: 本论文通过两个实验对词素身份信息和位置信息加以改变探讨了词素信息的重要性,研究发现中文双字词识别过程中,在词语识别时位置信息的作用随时间的延长而降低,而在句子阅读中身份信息的作用更大,且预测性对位置信息的加工有促进作用。研究结果对于揭示汉语词素在汉语词汇加工中的作用机制具有重要的参考价值。经过修改之后,可以发表。论文尚有如下的问题,请作者予以考虑:

意见 1: 论文的题目似乎没有充分反映论文的工作内容,特别是实验二的内容,建议考虑论文题目的修改。

回应: 审稿人说的对,对于文章的题目我们考虑的不够全面,确实没有反应实验二的内容,之前是考虑到本研究的主要目的就是探究身份信息和位置信息的作用,忽略了对全文研究内容的概括性,我们已经将题目改为:“身份信息与位置信息加工的进程及语境预测性的影响”

意见 2: 研究这个问题的理论基础和理论意义是什么?作者应该在文中进一步加强这方面的阐述与说明。

回应: 感谢审稿人的建议。我们在文中进行了修改和说明。关于理论基础,借鉴了拼音文字的研究,拼音文字有许多模型已经考虑到了字母位置信息编码的灵活性,如空间编码模型(spatial-coding model)、重叠模型(overlap model)和贝叶斯读者模型(Bayesian reader model)等。而汉字的加工究竟与拼音文字的哪个模型相类似目前还没有统一的定论。因此,我们希望通过本研究探究汉字身份信息和位置信息的编码特点,为日后研究者提出中文加工模型提供实验数据。我们已经在文章中对对应位置对这部分的内容进行了补充。

意见 3: 英文研究字母的位置信息和身份信息的研究,已经有一些成果。拼音文字的换位效应,通常是以字母为单位的。而汉语的换位效应,更多的是用词素。这两种换位效应的机制一样吗?作者是如何考虑的?作者应该对此进行一定的阐述与说明。

回应: 感谢审稿人的意见,关于这点是我们考虑了汉语中的笔画、部首、词素与字母的对应关系,并参考了前人的研究,我们选择了词素的换位来对应拼音文字中字母的换位。“汉字笔画对应拼音文字字母”是不可行的,因为汉字笔画写法复杂,有八种基本笔画,而且这些笔画不是水平排列,跟英文字母排列完全不一样。换位后就导致无法识别;“汉字部首对应拼音文字字母”看起来是跟英文中最相似的,但中文中的部首存在上下结构和左右结构的情况,因此也无法在实验中实现控制;考虑到词素是组成汉语的最小单位,而字母是组成英文的最小单位;字母之间都可以平行换位,汉字之间也都可以平行换位,因此,我们选择了词素换位,来探讨身份信息和位置信息的加工。这个意见与审稿人一的意见 1 是类似的,因此我们已经在文中对对应位置用红色进行了补充说明和修改。

意见 4: 两个实验的逻辑关系是什么?为什么进行第二个实验,作者论证不够充分,请在文章中进一步予以说明。

回应: 感谢审稿人的问题,关于这个问题文中论述不清。关于两个实验的逻辑关系是这样的,

本实验的研究目的是探究在汉语中是否存在换位效应,这种换位效应是否会由于句子的语境预测性而发生改变。因此实验一探究的是在单独呈现的条件下的换位效应,而实验二探究的是语境预测性对于这种换位效应的影响。

在前言部分我们进行了修改和补充。

意见 5: 两个实验的正确率的标准为什么不一致? 一个 90%, 一个 80%。并给出理由。

回应: 感谢审稿人提出的问题,本研究中实验一是行为实验,任务是进行真假词判断,属于比较简单的任务,因此正确率标准的选择较高,为 90%,而实验二是阅读并对判断题进行回答,需要阅读并理解句子含义,任务难度加大,因此调低了正确率的标准,变为 80%。我们在文中的对应位置,对于正确率的设置进行了说明。

意见 6: 毫秒这个单位,通常后面不需要小数点。

回应: 感谢审稿人的耐心审稿,我们已经在对应位置按照审稿人的意见进行了修改,删除了以毫秒为单位的数据小数点后的数据。

意见 7: 行文中存在一些不通顺的句子,已经在文章中以标注的方式标出。

回应: 感谢审稿人,我们已经修改了不通顺的地方。

意见 8: 其他问题,已经在文章中以标注方式提出,请参考。

回应: 非常感谢审稿人提出的细节性问题,我们已经根据审稿人的意见,在文中进行了相应的修改。

第二轮

审稿人 3 意见: 该篇文章探讨了中文词汇加工中的位置编码问题,实验 1 用启动任务,实验二用句子阅读任务,虽然研究较有意义,但文章逻辑有些牵强,需要补充实验 3,在句子阅读条件下操作预视时间,来统一测量位置编码的时间进程以及预测性对时间进程的调控。

意见 1: 实验 1 的基线存在问题,转置效应的意义在于转置之后相对于控制条件存在加工优势,作者用原词作为基线进行比较,只能得出转置和替换同样干扰词汇加工,而转置本身的两个汉字与原词身份相同的优势无法体现,无法说明转置的优势,与传统转置研究存在差异。而且,在 97ms 下,换位非词条件反应时与两种替换条件下的差异 $>.05$,并没有给出详细的 p 值,从趋势看,非词与原词以及两种替换条件趋势相同。

回应: 接受审稿人的意见。我们补充了实验条件,加上了无关启动条件,以无关启动条件作为基线更为合理。重新进行实验及采集数据,并且,结果报告部分,给出了详细的 P 值。

意见 2: 实验 1 的启动时间并没有给出很好的说明,因为在 97ms 到 314ms 之间存在较大的区间,传统上的启动时间更短就可以出现启动效应,作者设置的启动时间的范围过大,但是分区较少。

回应: 接受审稿人的意见。我们重新设计了启动时间,进行了三点区分。80ms、150ms、300ms。重新进行了实验,采集数据。并在文章中进行了修改。

意见 3: 理论基础和理论意义仍然不清楚。总讨论中,虽然加了一部分,作者提到:根据本

实验结果，汉语的加工模型可能类似于英文中的双重途径加工模型（Whitney, Bertrand, & Grainger, 2012）。但是，作者并没有深入讨论为什么以及如何类似。

回应：感谢审稿人的意见。认真研究了本实验的结果与相关理论之间的关系。在引言中加入了本研究涉及的相关理论，并在讨论中认真分析了本研究对相关理论的贡献。

意见 4：预测性对眼跳指标可能会造成重要影响，作者可以尝试分析加入眼跳落点位置的指标，看是否预测性对其存在调节，在不同替换和转置条件下是否存在差异。

回应：感谢审稿人的意见。我们在第二、三个实验中，选择了在兴趣区内的首次注视的注视次数(first run fixation count, FRFC)和总注视次数(fixation count, FC)，来反应注视位置的差异。审稿人的意见很好，我们在未来研究中会多设置兴趣区，进行注视位置的详细考察。

意见 5：几个小问题：1，重复测量的效应量是 partial eta square，不是作者写出的 eta square。2，所有的 p 值需要给出具体数字，统一精确到小数点后三位，依据 APA 的规范。3，常规的句子阅读，需要评定句子通顺性。此外，高低预测性词汇需要报告其词频、笔画数、字频等，确保这些无关变量不存在差异，否则会显著影响阅读时间。

回应：感谢审稿人的提示。问题 1、2，已经在文章中对对应位置进行了修改，对于 p 值小于 0.01 的，我们报告为 $p < 0.01$ ，0.01-0.05 之间的，我们精确报告到小数点后三位，大于 0.05 的，我们报告为 $p > 0.05$ 。问题 3 有关通顺性问题，由 28 名不参加实验的被试对句子通顺性进行 5 点评分，1 为不通顺，5 为通顺，句子的通顺性平均分为 4.32。已经在对应位置进行补充。对于词汇的词频、字频、笔画信息，我们也在实验一中进行了补充。在实验二和实验三中由于我们是对同一个目标词编制高-低预测性的句子框架，因此目标词本身的属性在实验设计中已经被控制。我们已经在对应位置进行了补充说明。

第三轮

审稿人 3 意见：这篇文章加入实验三，质量有所提升，但依然存在一些问题，需要进一步修改。

意见 1：作者采用的替换条件并非真实的汉字，如何和英文中采用真实的字母替换做类比？这种非字从视觉上会形成极大的突出性，对正常阅读和词汇识别的影响与真实汉字构成的非词有何不同？

回应：感谢审稿专家的问题。在这几年的研究中，我们一直在思考替换条件的设置。英文换位研究中，采用真实的字母，目的是让单词中某个字母的身份发生改变。然后比较身份信息错误与位置信息错误之间的差异。英文单词中的字母本身基本没有语义（a，代表一个，除外），字母替换一般不会有语义激活，所以使用真实的字母替换不会带来非常大的语义干扰。我们最初的设置，也使用真实的汉字进行替换，后来发现，每个汉字都有自己的意义，当被试看到汉字替换形成的假词时，会产生很多因为替换汉字引起的加工。这时的加工已经超出了某个位置上身份信息错误带来的影响，增加了新的身份信息导致的加工。所以，我们在后来的实验中，选择了一个类似于汉字的符号进行替换，而且，尽量减少视觉上的突出性，又不会像真的汉字替换那样导致语义激活。

意见 2：作者前面介绍中文中的相关研究都已经提到，然而，文中漏掉了近期一些重要的中文位置编码研究。譬如“Gu, J., Li, X., & Liversedge, S.P. (2015). Character order processing in

Chinese reading. *Journal of Experimental Psychology: Human Perception & Performance*.41(1), 127-137”“Gu, J., & Li, X. (2015). The effects of character transposition within and across words in Chinese reading. *Attention, Perception, & Psychophysics*, 77, 272-281”“滑慧敏, 顾俊娟, 林楠, & 李兴珊. (2017). 视觉词汇识别中的字符位置编码. *心理科学进展*, 25(6), 1-7. doi:10.3724/SP.J.1042.2017.001”。这些研究和综述对相关理论以及中文中可能的位置编码都有所讨论, 作者是否关注过这些近期的研究。

回应: 感谢审稿专家的提醒, 我们下载了这三篇文献, 已经在文章中加入了这三篇文献。Gu, J., Li, X. (2015). The effects of character transposition within and across words in Chinese reading. *Attention, Perception & Psychophysics*, 77: 272–281; Gu, J., Li, X., Liversedge, S. P. (2015). Character Order Processing in Chinese Reading. *Journal of Experimental Psychology: Human Perception and performance*, 41(1): 127-137; Hua, H., Gu, J., Lin, N., Li, X. (2017). Letter/ character position encoding in visual word recognition. *Advances in Psychological science*, 25(7): 1132-1138.

[滑慧敏, 顾俊娟, 林楠, 李兴珊.(2017). 视觉词汇识别中的字符位置编码. *心理科学进展*, 25 (7): 1132-1138]第一篇文献 Gu, J., Li, X. (2015)四字词的使用, 很成功, 尤其边界范式的使用。一般认为注视广度为注视点右侧 2 个汉字, 作者用了 4 个字, 成功的证明了词汇边界对汉字换位加工的影响。第二篇文献 Gu, J., Li, X., Liversedge, S. P. (2015).用词汇判定任务和阅读任务, 比较了非相关条件、换位条件和本词条件。发现字序编码在阅读的早期阶段, 而位置编码没有那么严格。第三篇文献很好的总结了国内外的研究及相关理论。

意见 3: 实验二中, 作者选择性的报告了 FFD 和 DT, 这与传统方法不符, 一般考察预视效应最基本的指标包括 FFD 和 GD, 为什么作者没有报告 gaze duration。

回应: 非常感谢审稿专家提出来。我们在实验二中没有使用边界范式, 是直接阅读的条件, 这里面选用的 DT 是 IA_dwell_time, 就是在兴趣区的总的注视时间, 没有添加 gaze duration 指标是我们的疏忽。我们已经在文章中增加了 gaze duration 的指标, 与大部分论文保持一致的用法。

意见 4: 实验二中兴趣区即目标词注视时间的凝视时间原词达到 1362ms, 这对于两字词来说极为罕见, 是否因为作者的操作方法导致被试很容易看到 boundary change, 对于知觉到变化的这组数据是否剔除。而且作者的 dwell time 指代的是否是 gaze duration 或者 total time, 如果是前者, 就更需要说明, 与传统研究发现有很大出入, 参考 Gu et al., 2015.

回应: 非常感谢审稿专家的意见, 在实验二中的 dwell time 指的是 total time, 也就是包含所有注视的总注视时间, 使用这个指标是为了探究目标词位置信息和身份信息的变化, 对整体的阅读的影响。我们已经在文章中加入了 gaze duration 的指标, 这个指标与以往研究发现的 gaze duration 也有一定的差异, 这可能是由于本实验二中, 句子中的目标词区域直接呈现换位和替换后的双字词, 导致注视时间增加。

意见 5: 作者需要讨论对比实验三和实验一, 在孤立词和阅读中, 位置和身份加工的时间进程的异同。

回应: 感谢审稿专家的建议, 我们已经在实验三的讨论部分增加了基于实验一和实验三的结果, 对比单独呈现词语和句子阅读中, 位置信息和身份信息加工时间进程的相关讨论。

意见 6: 几个小问题: 1, 摘要中“整句及目标词的相关眼动指标值都低”表达不清。2, 作者既然将原词作为基线, 在图 1 作图标注显著性时, 应当所有比较都与原图进行, 而不是不同

SOA 下切换了基线。3，作者在 3.2 中先介绍了整体指标，然后介绍局部指标，而结果呈现时位置顺序颠倒，需要调整。

回应：感谢审稿专家的提醒，关于问题 1，我们已经参考了国内外眼动研究中，摘要部分对于眼动指标的表述，并进行了相应的修改。问题 2 中审稿专家说的对，是我们图做的有问题，我们已经以原词为基线，对图一进行了修改。针对问题 3 我们已经在对应位置调整了文章中指标呈现的顺序。并且将数据呈现部分的相关顺序都进行了统一。

第四轮

审稿人 3 意见：经过作者几轮修订，文章质量得到提升，建议发表。希望作者进一步规范写作，严格按照 APA 格式修订格式。几个细节问题：

意见 1：摘要中，第一次提到两种替换，读者不清楚是什么，建议写成首字和尾字替换。

回应：我们已经对文章格式进行了认真的修改。感谢审稿人的提醒，我们已经对摘要部分进行了修改，修改为“首字和尾字替换”。

意见 2：参考文献滑惠敏，顾俊娟，林楠，李兴珊，2017，第一作者的名字错误，滑慧敏。

回应：非常感谢审稿人耐心审稿，我们已经对参考文献部分整体进行了认真的检查和修改。

意见 3：新加入的几片参考文献，引用不符合 APA 格式，建议修正。譬如 Gu, Li, 2015, 应修订为 Gu & Li, 2015。

回应：感谢审稿人的细心。我们已经进行了相应的修改。

意见 4：虽然作者统计信息的 p 值做了修订，但依然有些不符合 APA 格式，建议修订。

回应：非常感谢审稿人的意见，我们已经将全部 p 值按照 APA 的格式进行了修订。

意见 5：换位非词条件与首字替换词条件差异也显著($p = 0.41$)，这句话与后面统计指标不符。

回应：感谢审稿人提出的意见，这部分在写作的时候出现了笔误，p 值应该为 0.041，已经在文章中进行了修改。

意见 6：，摘要中“高预测性条件下，换位条件和替换条件差异不显著，而低预测性条件下，换位条件与替换条件差异显著”与结果描述不符。

回应：审稿人说的对，我们已经将摘要部分比对结果和讨论进行了认真的修改。

第五轮

主编意见：文章经过多轮修改，增加了一个实验，研究逻辑清晰，写作也有了很大提高。建议进一步规范写作，修后发表。具体建议如下：

意见 1: 图 1 的图例是不同颜色, 黑白版很难区分, 请修改图例。

回应: 非常感谢主编提出的意见, 我们已经对文中的图 1 及图例部分进行了修改。

意见 2: 表 4、表 5、表 6 应该为三线表。

回应: 感谢主编的耐心审稿, 我们已经对表格的格式进行了修改, 并对文章剩余部分进行了检查和更改。

意见 3: 与评审人建议 p 给出具体数值, 请参考修改。

回应: 感谢主编提出的宝贵意见, 我们已经根据 APA 格式, 将 p 值小于 0.0001 的写为 $p < 0.001$, p 值小于 0.01 的写为 $p < 0.01$, p 值在 0.01 和 0.05 之间的写出了具体数值, p 值在 0.05 以上的写为 $p > 0.05$ 。如果多种条件的 p 值情况一致则写为了 ps。文中对应修改的部分已经用红色字体标出。