

《心理学报》审稿意见与作者回应

题目：声旁语音信息对形声字加工的影响——来自眼动研究的证据

作者：迟慧 闫国利 许晓露 夏莹 崔磊 白学军

第一轮

审稿人 1 意见：本研究通过眼动技术考察声旁语音信息对形声字加工的影响。结果发现：当形声字声旁在左或声旁在上时，省略前面笔画的形声字的加工时间显著长于省略后面笔画的形声字的加工时间；当声旁在右或声旁在下时，省前与略后条件没有显著差异，即当略后的笔画中包含声旁时，亚词汇层面的声旁语音信息会影响到“笔画顺序效应”。研究结果支持了汉字语义通达的双通道模型。本文选题有意义，实验设计合理，数据处理方法正确，结论可靠，对深入认识汉字加工的特点有重要的启发意义。

意见 1：说明四个实验的被试是否为同一批被试。

回应：感谢审稿专家的意见，这一点在文中确实没有说明清楚，作者已经在修改稿中进一步说明（p9, p13, p16）。本研究中，实验一和实验二（只省略目标字探讨左右结构和上下结构形声字）是同一批被试 36 人，实验三和实验四（句子中全部汉字都进行笔画省略探讨左右结构和上下结构形声字）是另一批被试 36 人。

意见 2：说明问题回答正确率的结果。

回应：实验中会在一部分句子后面随机出现需要被试判断正误的问题句，目的是保证被试认真阅读句子。本研究的四个实验中，每个被试回答问题的正确率都超过了 86%，说明被试认真阅读并且正确理解了句意。这一点作者也会在修改稿中给予补充说明（p7, p11, p14, p17）。

意见 3：“除三个标准差”之外的数据修改为“删除平均数加减三个标准差”之外的数据。


回应：作者已在修改稿中修改了这一描述（p7, p11, p14, p17）。在以后行文中也会更加注意描述的准确性。

审稿人 2 意见：本研究采用眼动记录的方法，探讨了形声字中笔画缺省对形声字加工的影响，结果发现当省略的笔画中包含声旁时，加工受到了更大的干扰，作者认为这说明声旁的语音信息都对形声字的加工起到重要作用。我的具体意见如下。

意见 1：第一，本研究的创新性略显不足：作者在自检报告以及正文（如 p6）中反复强调：‘闫国利等人（2013）的实验中采用的是“单字命名任务”，任务本身存在校稿的语音要求和较低的语义通达要求’。命名任务确实会更凸显出语音在词汇通达中的作用，但是与芬兰语、德语等浅层正字法文字不同，中文作为一种深层正字法的文字，是否普遍存在命名任务中只通达语音而不通达语义的情况？大量中文阅读发展的研究将儿童正确命名的字（词）数量作为其语言发展的重要指标，而拼音文字的研究则不能采用这一指标，正是考虑到这种深层正字法的特征。也就是说，本实验的动机是否站得住脚，应需要理论上进一步的佐证。此外，即便承认命名任务的局限性，一个更直接的考察应该是词汇判断的任务而非眼动记录的手段。使用眼动记录的优势在于考察自然阅读中句子语境等因素的影响，希望作者加强这方面的讨论。

回应：感谢审稿人的建议，对于这个问题，作者的表述确实不够清楚，具体内容已经在修改

稿中进行修改（p4, p19），请审稿人进行查阅。

首先，作者的意思并不是“命名任务中只通达语音而不通达语义”，对于一般汉字命名任务确实是存在语义通达的。文中所说“任务本身存在校稿的语音要求和较低的语义通达要求”主要是针对实验特定的形声字而言。由于形声字包含表音的声旁，而被试的任务又是命名，所以被试会倾向于通过加工声旁来获取语音信息。特别是在闫国利等人（2013）实验中的形声字有笔画省略的操作，如命名“”，被试会更倾向于仅利用声旁来发音。总的来说，其实实验结果受实验任务本身特点的影响较大，在命名任务中，实验想探讨的“声旁的作用”被实验任务本身所提高了。所以实验结果更多的是说明在出声命名情况下存在这种语音的作用。而且，单字识别和自然的自然阅读在认知加工的特点上也存在一定的差异，因此，在单字识别情况下的得出的结论并不一定能够很好地说明在自然阅读中的情况。所以作者在前人研究基础上，想进一步探讨在不强调发音的自然阅读任务中声旁的作用。

采用眼动实验的主要目的就是要考察在自然阅读中的加工，而且眼动记录法相比命名或词汇判断任务，能更精确的记录被试对汉字的加工过程。

意见 2：阅读本文后，我发现文中最高频的一个短语是：与闫国利等人（2013）的研究结果一致。本研究自身独特的地方没有给读者留下深刻的印象。作者虽然指出，“本研究采用眼动技术，可以对形声字的加工进行较为精细的分析”，但是这句话基本流于形式：眼动分析究竟在什么方面比命名更“精细”，能比闫国利等人（2013）的研究回答哪些更多的问题，读者不得而知。

回应：同意审稿人的意见，文章相关的地方没有说清楚。本研究相比闫国利等人（2013）实验的创新之处：

第一，研究内容。如上所述，闫国利等人（2013）实验中，实验想探讨的“声旁的作用”有可能被实验任务本身所提高。本研究在此基础上，进一步考察了在不强调发音的自然阅读任务中声旁在形声字加工中的作用。

第二，研究方法。采用眼动记录技术来考察自然阅读是当前非常有效的探讨自然阅读的研究手段，这一问题在眼动阅读研究中已经有可靠证据。通过眼动记录法，可以在自然状态下检测到读者细微的眼动指标，可以揭示读者的早期加工和晚期加工，可以精确的记录到几毫秒的差异，比记录反应时（反应时包括被试的单纯加工时间和反应时间）的命名任务要更精确。

意见 3：作者一方面提出‘闫国利等人（2013）的实验中采用的是“单字命名任务”，任务本身存在较高的语音要求和较低的语义通达要求’（p6），另一方面又提出“本实验与闫国利等人（2013）的研究相比，……表明在自然句子阅读中，声旁的语音信息对形声字的识别加工仍起到重要作用，甚至可能起到更大的作用”（p10），到底作者认为在命名任务中还是在句子阅读任务中语音信息更重要？是否前后存在自相矛盾的说法？

回应：根据审稿人的意见，作者认为对于这一问题在文中说明的不够准确。本文四个实验的结果一致表明，当略后的笔画中有声旁时（声旁在右、下），笔画顺序效应受到干扰，即省前条件和略后条件下的各注视时间差异均出现逆转（略后大于省前），而且这种差异在四个实验中表现的稳定且一致，基本上都是差异非常显著的，即使有早期指标差异不显著，但是差异的逆转趋势也表现出来了（略后大于省前）。这一结果比闫国利等人（2013）的实验结果更稳定。在闫国利等人（2013）的实验一中（左右结构形声字），反应时的简单效应分析发现当声旁在右时，省前条件和略后条件差异不显著，但是趋势并没有逆转。所以，作者真正想表达的是本实验与闫国利等人（2013）的实验相比，在自然句子阅读中，声旁的语音信息对形声字的识别加工仍起到重要作用，而且结果比在命名实验中更加稳定，趋势更加一致明显。这个地方文中表述的不够准确，已经在修改稿中进行修改（p9）。另外，这一结果很

大部分的原因是由于不同的实验方法导致的。眼动记录法相比反应时记录法能更加精确地记录被试的加工时间。所以，这一结果也说明了眼动记录法的精确性，更能准确地揭示出结果。

意见 4: 第二，实验材料部分的介绍缺少重要细节：作者仅仅简单说“选取闫国利等人（2013）实验一中的左右结构形声字作为目标字放入正常的句子中”显然是远远不够的。实验材料是否进行了语言学属性上的匹配，进行了哪些匹配，应在文章中报告其平均数、标准差、统计检验结果等数据。一般来讲，至少要进行汉字频率、笔画数的匹配，此外，有些研究者还进行了可想象性、 获得年龄等维度的匹配。由于本研究的目的在于考察形声字的加工，应该对目标字的规则、一致性也加以严格匹配（关于规则、一致性的作用，作者在引言部分已有介绍）。

回应: 对于这一问题，感谢审稿人提出，这确实是作者在文章撰写中忽视的一个问题，因为在闫国利等人（2013）的实验中已经了属性匹配，所以在本文书写中没有把属性统计信息放入文中。事实上，对于实验材料的选取和编排的操作都很严格，只是由于写文章的经验不足，以后会进一步加强，感谢审稿人的指导。具体内容已经在修改稿中补充（p5, p10），请审稿人检查。本实验匹配了笔画数、部件数、字频、规则性、一致性、声旁家族总字数以及家族内与声旁读音一致字数，差异都是不显著的（ $p>0.05$ ）。具体数据如下：

上下结构形声字属性统计检验结果：

	笔画数	部件数	频率	规则性	一致性	家族总字数	声旁一致
声旁在上	11.43 (3.33)	3.03 (0.85)	32.85 (45.83)	0.40 (0.50)	0.27 (0.45)	10.83 (8.16)	5.57 (3.53)
声旁在下	11.73 (3.68)	2.73 (0.74)	34.67 (50.81)	0.33 (0.48)	0.33 (0.48)	11.2 (10.47)	6.00 (5.69)

左右结构形声字属性统计检验结果：

	笔画数	部件数	频率	规则性	一致性	家族总字数	声旁一致
声旁在左	11.2(2.92)	2.63(0.61)	52.31(64.19)	0.23(0.43)	0.03(0.18)	14.4(8.36)	6.67(4.59)
声旁在右	9.6(2.93)	2.6(0.77)	64.6(118.61)	0.33(0.48)	0.13(0.35)	11.1(7.87)	4.93(3.90)

注：声旁一致指家族内与声旁读音一致字数

一致性：一致——1；不一致——0 规则性：规则——1；半规则——0

关于规则性控制的问题，本实验的材料确实不全是规则字。由于各种属性及句子编排上的要求（成对目标字嵌套在同一句式、合理性、预测性等），实验材料的选择很有限。作者是在保证控制重要无关变量的基础上，最大限度进行了材料的匹配。所以实验材料中基本上是一半的规则字，一半的半规则字。但在各条件下都进行了匹配（如上表所示），所以不存在不同实验条件间的差异。一般在做汉字亚词汇的研究中（Ding, Peng, & Taft, 2004），由于要保证重要无关变量的控制，实验材料的选择上都存在很大难度，只能尽可能匹配。

意见 5: 实验 1 中目标字的可预期性如何？作者提到有进行预测性的检测，但并未报告是否匹配已经统计检验结构，看似这句话说到一半就没有了，不知何故。

回应: 感谢审稿人细心的审阅，这一问题应该是作者上传到编辑部的过程中出现了一些失误。作者的原始稿件中是有完整叙述的（同实验 2 描述，p10）“另外选取 30 名大学生被试对目标字的预测性进行了检测，呈现实验句中目标字前的部分，要求被试补充句子，实验设置的目标字预测值为零。”作者已在修改稿中补充完整（p6）。对于出现的这一问题作者深感歉意，以后在文章校对及上传稿件的时候会更加注意。

意见 6: 数据筛选标准不明确：“删除三个标准差之外的数据”这一标准引起的具体标准是多少？即，长于多少毫秒和短于多少毫秒的注视点被排除？此外，删除数据的具体过程是什么，是对于每个指标分别删除极端数据，还是说某个目标词的一个指标超标就将这个目标词

排除在分析之外？

回应：这一问题是在作者在文中阐述的不够详细，已经在修改稿中进行了补充（p7）。数据删除的具体标准解释：

第一，删除被试阅读中少于四个注视点的实验句，一般这种情况都是由于被试没有读完句子就不小心按了翻页键的原因（所占比例很少），表明被试没有通达整个句子，所以需进行删除。具体删除过程：在实验后对每个被试阅读的实验句进行逐一查看，哪个句子上的注视点少于4个，就把这个被试所阅读的这个句子删除，不进入统计分析。

第二，删除平均数三个标准差之外的数据，这一操作是对每个指标分别进行的。具体删除过程：对每个指标分别计算其平均数和标准差，只删除每个指标平均数三个标准差之外的数值，并不删去目标字。

意见 7：第三，实验设计和结果的解释存在的问题：本研究和闫国利等人（2013）的研究均报告了所谓“声旁位置的主效应”，当形声字声旁在左比在右的字识别更快，错误率更低，注视时间更短。但是这种效应究竟是不是能够完全归因于声旁的位置？仔细查看作者的材料举例（图1），可以发现材料中存在形声字的规则性偏差：声旁在左的“剧”为半规则字，因为其声旁“居”的发音与目标字声母韵母均相同，只存在声调差异；而“版”字虽然也算作半规则字，但是属于不同类型，即整字与声旁的韵母相同。王晓怡，马力飞，于薇，谭向杰，张兆琪，翁旭初（2004）的脑成像研究发现，同音同调（规则字）和同音异调半规则字的脑激活模式比较接近，而异音同调的半规则字则与异音异调（不规则字）的脑激活模式更接近。这说明与声调的作用相比，读音的作用更明显。根据王晓怡等人（2004）的研究结果，我们有理由推断本研究中声旁在左的目标字相比更规则，更容易加工；因此本研究以及闫国利等人（2013）的研究结果是否仅仅反映了规则性效应，而并非声旁位置的效应？同理，前述形声字的规则性偏差也可以本研究观测到的交互作用：表1中，声旁在左时破坏声旁（省前条件）造成的影响（与不省略条件比）较大，为 $301-251=50$ （首次注视）和 $335-293=42$ （凝视），而声旁在右时破坏声旁的影响较小，为 $266-261=5$ （首次注视）和 $354-333=21$ （凝视）。与前述同理，破坏一个较规则（能提供更多语音信息）的声旁导致的影响肯定应该大于破坏一个较不规则（能提供较少语音信息）的声旁导致的影响。在实验2中也存在类似的问题，图2中发现，声旁在上的目标字“袋”是规则字，而声旁在下的目标字“箱”是半规则字。同理，破坏一个规则字声旁的影响 $[314-258=56$ （首次注视）和 $369-297=72$ （凝视）]大于破坏半规则字声旁的影响 $[294-254=40$ （首次注视）和 $378-310=68$ （凝视）]。综上所述，同时结合我在实验材料部分的提出问题，目前的实验结果究竟应归因于规则性效应还是声旁位置效应，不能得出明确结论。作者如果在本实验中没有控制目标字的规则、一致性，应重新选取材料并进行实验。

回应：同上对于实验材料控制问题的回答，本研究的材料均进行了严格的匹配，并不存在规则性偏差。所以本实验结果是可以归因于声旁位置的。再次对审稿人表示歉意，由于作者的疏忽，给审稿人的审阅增加了麻烦。

意见 8：关于实验3和实验4的动机，作者指出，“目标字在整句中突显出来，进而会吸引读者更多的注意和注视”（p14），这个说法固然有道理，但是如果作者对不同条件间删去的笔画进行了较好的匹配，在吸引注意力这一问题上应不存在条件间的差异，因此实验1和实验2的结果是可比较的（但如前述，由于实验设计上可能存在的缺陷，作者的解释不一定合理）。但是，实验3和实验4则比实验1和实验2存在更大的问题：如图3所示，无省略、省前条件、省后条件不仅应用在关键字上，还应用到了其他非关键字上，这导致读者在加工

三种条件下目标词之前就存在了加工上的差异,因此在目标字区段观测到注视点的差异不能单纯归因于目标字本身,导致三种条件不可比较。

回应: (1) 对于审稿人提到的实验 3 和实验 4 的动机问题,确实在匹配后不会影响最后各条件之间实验结果的差异。但是“pop out 效应”确实是存在的,而且目标字上的“pop out 效应”会影响被试在整句阅读过程中的眼跳计划,而这种影响不通过实验手段是无法准确预期的,所以想验证“pop out 效应”是否产生的影响,就有必要再进行一个消除“pop out 效应”的实验来进一步验证所得到的实验结果。

(2) 对于审稿人提到的实验 3 和实验 4 的操作会使“目标词之前存在了加工上的差异”这一问题确实有道理,也是作者在实验时反复考虑过的问题。作者考虑到为了严格控制 pop-out 效应的影响,就必须保证目标字前的字和目标字的形式是一致的,而这就必然带来不同条件间“目标词之前加工上的差异”。如果为了保证一致,让三种条件都采用“省前”的形式,那对于“不省”和“略后”两种条件就又会出现了 pop-out 效应;如果三种条件均采用“略后”的形式,同理,“不省”和“省前”两种条件会出现 pop-out 效应。所以作者考虑通过结果进一步分析。实验结果证明,实验 3 和 4 的结果跟实验 1 和 2 (仅有目标字省略) 的结果一致,这也就说明“目标词之前加工上的差异”没有对目标字的加工造成很大的影响。总之,如果仅省略目标字(目标字突出),可能出现 pop-out 效应;如果省略整句所有汉字(目标字不突出),目标字前的差异可能导致结果的差异。但是事实上,两种条件的结果出现了一致性,也就是说实验结果的趋势没有显著地受到“pop-out 效应”的影响,也没有显著地受到“目标词之前加工上的差异”的影响。所以,虽然存在这一问题,但是实验结果是可以解释清楚并可以说明问题的。

第二轮

审稿人 2 意见:我曾为贵刊审阅了《声旁语音信息对形声字加工的影响——来自眼动研究的证据》(编号 xb13-373)一文的第一版。整体来说本文的修改稿水平有所提升,作者比较全面、中肯地回答了我初审时提出的一系列问题:作者对本文与以往相关文章的讨论有所加强,对实验材料的细节进行了合理的报告。针对修改稿我仍然有一些问题希望与作者进行进一步的探讨。

意见 1:针对问题“1-1”的回答:我基本认同作者探讨的独立字词识别和自然阅读的差异。但是本文作者指出“由于形声字包含表音的声旁,而被试的任务又是命名,所以被试会倾向于通过加工声旁来获取语音信息。”我对这一说法存疑。研究指出,汉字中只有 30%左右的字可以从亚词汇水平得到正确的读音(即所谓规则字),因此换言之,亚词汇(声旁)语音信息对于汉字整体语音的贡献是非常有限、不可靠的。本研究的受测对象为成人被试,他们通过大量学习应当已经掌握这样的统计规律,因此,为何被试要“倾向于通过加工声旁来获取语音信息”?同时,作者举出的例子(去掉部分形旁的“袋”)有一定的偏向性,本例中声旁完整保留,是读者可利用的主要信息;自然,对于这样的声旁完整的字,读者自然会倾向于利用声旁发音。但这样的例子不一定具有代表性:当读者看到声旁不完整的字时(如图 1 中第二行的关键字),如何“倾向于仅利用声旁发音”?

回应:首先,审稿专家提出的第一个问题“为何被试知道声旁语音不可靠的规律,但仍‘倾向于通过加工声旁来获取语音信息’?”,对于成人被试来说,确实已经形成了一套完整的汉字提取加工策略,并不会仅仅简单地依据声旁来确定汉字的读音,但是在其特定的实验任务(命名)要求下会让被试产生这种倾向。由于命名任务要求被试尽快反应进行发音,在舒华,张厚粲(1987)的研究中也提到,这迫使被试只能更多地运用声旁类似推理策略。尤其在其实验中存在笔画缺失的汉字,被试想要识别就必须根据已有的笔画信息进行补全,而命

名任务仅要求被试进行发音，所以当声旁保留时，被试能有效利用的就是声旁所提供的语音信息。例如，当被试看到义旁笔画省略的汉字时（“屈”），尽管不确定自己的判断是否是完全正确的，也会根据声旁进行发音。所以，作者认为闫国利等人（2013）的研究可能存在“实验任务使被试倾向通过加工声旁来获取语音信息”这一问题。

审稿专家提出的第二个问题“当读者看到声旁不完整的字时（如图 1 中第二行的关键字），如何‘倾向于仅利用声旁发音’？”这一问题应该是作者在文中说明的还不够清楚，作者想表达的正如专家所说，当省略义旁时，被试会倾向利用声旁发音；而省略声旁时，则不能利用声旁信息来发音。这也正是作者对于闫国利等人（2013）的实验结果的质疑及本实验的出发点之一，其命名实验发现“当声旁在右或声旁在下时（略后的笔画中包含声旁），省前与略后条件没有显著差异，甚至出现差异的逆转。”而这种逆转的结果很可能是由于实验任务（命名）导致的，不管省略位置如何，均表现为省略声旁不容易发音，而省略义旁容易发音。对于这一问题作者已在修改稿中进行进一步修改（p4），如果仍存在没有说明清楚的地方，还请审稿专家进一步指正。

意见 2：针对问题“3-1”的回答：我认为作者针对此问题的答复可能没有十分恰当地回答我先前提到的问题。我在初审中提到的王晓怡等人（2005）的这项研究表明（在上次的审稿中我错写为 2004，本研究发表于《中国医学影像技术》），同样属于“不规则”这类的字，可以进一步划分为“同音异调半规则字”、“异音同调半规则”和“完全不规则”这三类。王晓怡等人的这篇近 10 年前的脑成像研究清楚地说明这三类不规则字的加工模式有显著差异，同时，在国际刊物上发表的相关文章也探讨过半规则字的问题（如 Chen, Shu, Wu, & Anderson, 2003; Shu, Chen, Anderson, Wu, & Xuan, 2003 等）。因此，回到本文上，我认为仅仅将材料中的字区分为规则、不规则两个水平（表 1），对于 2014 年的研究来说，显然是有些粗糙的。我建议作者应该将规则性进一步更细致地加以区分。

回应：感谢审稿专家提供的宝贵文献资源，作者会继续加强相关文献的搜集和整理，也会加强实验材料控制的严谨度。

作者细心研读了审稿专家提供的这篇文章《阅读半规则汉字时的脑激活模式：声调的作用》，此文是从汉字声旁与整字的读音及声调是否一致的角度，将汉字分为完全规则字（同音同调字）、半规则字（同音异调字和异音同调字）及完全不规则字（异音异调字），主要考察“声调”的作用，研究结果发现了不同类别半规则字加工的差异。作者在进行材料选择过程中也考虑到了半规则字的进一步分类问题，根据前人研究文献（陈俊和张积家 2005；舒华，毕雪梅，武宁宁，2003），本研究选择材料依据的半规则字划分为：声旁与整字同音不同调的字，声母不同韵母相同的字，韵母不同声母相同的字。具体统计结果如下：

左右结构形声字规则性分类个数统计：

规则字		半规则字		
		同音不同调	声母不同	韵母不同
声旁在左	7	3	18	2
声旁在右	10	3	12	5

上下结构形声字规则性分类个数统计：

规则字		半规则字		
		同音不同调	声母不同	韵母不同
声旁在上	12	7	9	2
声旁在下	10	8	11	1

根据如上所述进行准确的划分，本研究实验材料采用的是规则字和半规则字，并没有不规则字。这也是作者在第一次修改稿中没有表述准确的问题，还请专家见谅！对于王晓怡等

人（2005）文章中“同音不同调”的半规则字，本研究在左右结构和上下结构形声字的不同条件下均没有显著差异，故本研究结果不会受到声调的影响。另外，其文中提到的“异音同调字”，作者理解的是“声旁的声母韵母与整字均不相同但发音相同的字”。如果作者理解正确的话，本研究的实验材料并没有异音同调字。

本研究最初进行材料选取时，确实只控制了规则字与半规则字的匹配，并没有对半规则字的几种子分类进行严格匹配。这确实是本研究的不足之处，感谢审稿专家再次的耐心指正，让作者准确清晰地明确了这一问题。在今后的研究中，作者一定会更加注意材料的控制问题。

意见 3：针对问题“3-1”的回答：我认为作者的回答比较实事求是，我也理解实验 3 和实验 4 的动机。但是，我希望作者至少可以在文章中讨论实验 3 和实验 4 的操作会带来的问题（即，目标词之前的加工存在条件间的差异）；对审稿人已经指出且作者本人已经了解并承认的问题在文中避而不谈，似乎有所欠缺。

回应：对于这一问题确实应该在文中进行明确的分析和解释，以便使读者更好地理解。作者已在文中讨论部分增加了对问题的表述(p20)。

第三轮

审稿人 2 意见：我曾为贵刊审阅了《声旁语音信息对形声字加工的影响——来自眼动研究的证据》（编号 xb13-373）一文的第一版和第二版实验。我认为经过两次修改，作者比较全面地回答了审稿人提出的问题，对材料的描述比较到位，对于实验的创新性和不足之处也进行了适当的讨论。

意见 1：我在这一轮的修改意见是，希望作者考虑将第三页提供的规则性分类个数统计信息在正文中提供，并在正文中讨论此处提到的相关研究，以便使本文的读者能更快捷地把握本研究所采用的实验材料。除此之外，我认为本文是一篇合格学术论文。

回应：在本次的修改稿中，作者已经将“规则性分类个数统计信息”表格放入文中相应位置（p5, p6, p11）并进行了相关研究的论述。