

《心理学报》审稿意见与作者回应

题目：聋生“是……的”句句法意识研究

作者：张帆 李德高

第一轮

审稿人 1 意见：

意见 1：

1. 为什么选择这个句式，好像理由表述不充分。
2. 外国学生理解困难，但不一定说明聋人也理解困难。由于他们的母语不同。

回应：

我们在以下 2 个方面进行了修改和补充：

- (1) 在第三页倒数第三段倒数第三行中“然而”之前，增加了以下信息：“**其实，多数聋生因为在入学之后才开始学习和使用书面语，所以与健听人晚期双语者情况相近。所谓晚期双语者是指熟练掌握第一语言之后才开始学习使用第二语言的第二语言学习者，他们学习第二语言时不但会受到其第一语言经验的帮助，还会受到其第一语言经验的制约 (Saito, 2015)**”
- (2) 第四页第二段中“但迄今……在中国，”改为了“**但是，关于聋生书面语语法意识还有待深入开展研究(Traxler, Corina, Morford, Hafer, & Hoversten, 2014)。比如在中国，虽然有学者就聋生书面语语法困难现象进行了一些思考分析(许保生, 傅敏, 2014)，但是，通过认知心理学实验来回答下列这样问题的研究似乎还比较缺乏：**”
- (3) 第四页第二段最后一句“本研究……”起至第四页第四段调整和修改如下：

“是……的”句是现代汉语中一种常见句式，也是句法最为复杂的句式之一（熊仲儒, 2007）。在“是……的”句中，处于句尾的助词“的”与其前面的助词“是”构成语法框架，使框架内的内容在语用上得到强调。例如，在（1）“她爱护书籍”与（2）“她是爱护书籍的”这两句话中，除了（1）所表达的意思之外，（2）还强调，“爱护书籍”是“她”的一种习惯，或者，是“她”的一种品质。

然而，该句式不仅外国留学生学习起来特别困难（谢福, 2007），而且，也许因为书面语也是其第二语言的缘故，对于“是……的”句，聋生学习起来似乎也特别困难。比如说，聋生作文中经常出现把（2）写成“她是爱护书籍”或“她爱护书籍的”的现象。聋生写作中所犯这样的语法错误表明，一方面，他们在书面语学习实践中对“是……的”句式可能发展了一定的语法意识，另一方面，他们关于助词“是”和“的”语法框架意识似乎还不够明确。

聋生的读写技能是并行发展的（Albertini, Marschark, & Kincheloe, 2016），他们在“是……的”句阅读过程中也应该会有异常表现。**因此，本文拟通过阅读实验研究，就聋人大学生“是……的”句句法意识进行实验研究。**

意见 2：

3. 在后面的被试选择条件中，未体现出两类读者在语义知识方面的差异
- 4 本实验使用的被试，其听力缺失值是多少，请具体说明。
8. 来自某特殊教育职业学院的实验组和控制组被试参加了《瑞文标准推理测量》、阅读能力

测验。研究使用的余地能力测验是什么呢？应该具体说明。

9. 在表 2 中，列出耳聋年龄和学前语训效果的目的和依据是什么？两组被试在这两项上是否存在差异？

回应：

根据这 4 条意见，我们对 2.1 小节关于被试介绍的部分修改如下：

- (1) 来自某特殊教育职业学院的实验组和控制组被试参加了《瑞文标准推理测量》、阅读能力测验和背景信息问卷调查。耳聋不仅会影响聋人儿童（简称聋童）的语言发展，而且，还会伴随影响其智力发展（Shojaei, Jafari, & Gholami, 2016），因此，以聋人为对象的研究往往需要平衡被试的智力水平。关于阅读能力测量，我们借用了《聋人大学生汉语课程的开发》（张会文, 吕会华, 吴铃, 2009）附录中语文能力测试题的阅读理解部分，并假设在其他方面没有差异的条件下，若重度耳聋者的测试分数不低于轻度聋生的测量分数，则两种被试具备基本相同的词汇语义知识。

研究表明，聋童耳聋发生年龄越早，其口语语言能力发展受到制约的程度就越强；他们接受学前语言训练效果越好，入学后其书面语学习效率就可能越高（Shojaei, et al., 2016）；而且，如果聋童父母亲懂手语，那么聋童的手语能力就可能得到较好发展，而聋童的手语水平又会影其书面语学习（郑璇, 2004）。因此，背景信息问卷调查不仅包括性别还包括耳聋发生年龄、语训和父母亲是否使用手语等。

结果如表 2 所示，每个实验中，两种被试在年龄、阅读能力和智力等方面没有显著性差异，但在听力损失程度方面有显著性差异。

- (2) 表 2 中补充了两类被试在各指标上差异与否的统计检验结果。

意见 3：

5. 作者指出：受口语实践影响，他们可能比重度聋生发展了更强的书面语句法意识。而重度聋生因缺乏口语实践经历所以对“是……的”句的语法意识较中轻度聋生弱。这样说的依据是什么？是否有文献支持？

回应：

第五页第一段修改如下：

……健听人交流。第二语言实践有助于双语者发展目标语语法技能（De Carli et al., 2015）。轻度聋生因为参与较多口语实践的缘故，所以肯定比重度聋生有更多的汉语语言实践经历，因此他们可能比重度聋生发展了更强的书面语句法意识。相反，重度聋生因缺乏口语实践经历所以对“是……的”句的语法意识较中轻度聋生弱。

意见 4：

6. 我们观察聋生的作文发现，其“是……的”句中缺少助词“是”的错误主要有 3 种类型（见表 1）。这种说法的依据是什么呢？

回应：

第五页第二段修改如下：

叶盼云和吴中伟（2008）把外国留学生容易发生错误的“是……的”句归为 3 种类型，即，助词“是”和“的”之间词构成的分别是形容词结构（如“平坦”）、动词结构（如“爱护书籍”）和带时间或地点状语的动词结构（如“在船上工作”）。我们观察发现，对应这 3 种类型“是……的”句，聋生在作文也均有缺少助词“是”的错误现象（见表 1）。因此，针对这 3 种“是……的”句，本研究拟进行 3 项实验来验证上述假设。如果 3 个实验取得一致结果，那么上述研究假设将会得到强有力的支持。

意见 5:

7. 虽然省略句有句法错误, 但被试读到句尾助词“的”时才可能会对这种错误有所觉察。为什么呢?

回应:

第五页第三段修改如下:

每个实验中, 被试需阅读完整的“是……的”句(简称完整句)和助词“是”省略的“是……的”句(简称省略句)。在逐词阅读任务中, 读者对每个词可能都会进行语法和语义上的信息提取、整合和预期(Luka & van Petten, 2014)。显然, 在句法信息整合方面, 当被试读到句尾助词“的”的时候, 他们在省略句中遇到的困难会大于他们在完整句遇到的困难。即, 被试对省略句的“收尾效应”会大于他们对完整句的“收尾效应”。

意见 6:

10. 不参加实验的聋生对句子中的词进行熟悉性评定(1 = 认识; 0 = 不认识)。进行熟悉度评定, 使用两点评定是否合适? 通常使用 5 点评定。

回应:

2.2 的第一行修改如下:

实验 1、2 和 3 的关键材料分别是 11、10 和 10 句完整句(见附录)和相应省略句。因为聋生在比较判断任务中倾向于使用是非判断策略, 他们似乎难以在多个选项之间做出比较(Marschark, Lang, & Albertini, 2002), 因此, 我们参照 Li, Zhang, and Zeng (2016), 请 20……

意见 7:

11. 实验组与控制组对句子理解题目应答的准确率均高于 85%。建议: 应该验组和控制组的正确率应该分别报告。并看看他们是否有差异。

回应:

第七页第二段第一行修改如下:

- (1) 原来的句子移动到了第七页第二段五行中“结果发现,”之后;
- (2) 在原来“实验组…个句子”替换为“**实验组(85.6% ± 4.9%)对句子理解题目应答的准确率显著低于控制组(89.6% ± 3.7%)对句子理解题目应答的准确率($t_{(107)} = 4.828$, $p < 0.05$, $d = 0.93$)。**”

意见 8:

12. 中几次提到晚期双语者, 晚期双语者的具体概念是什么? 本研究中的被试和晚期双语者有什么关系?

回应:

我们做了三处修改:

- (1) 在第三页倒数第三段倒数第三行中“然而”之前, 增加了以下信息: “**其实, 多数聋生因为在入学之后才开始学习和使用书面语, 所以与健听人晚期双语者情况相近。所谓晚期双语者是指熟练掌握第一语言之后才开始学习使用第二语言的第二语言学习者, 他们学习第二语言时不但会受到其第一语言经验的帮助, 还会受到其第一语言经验的制约(Saito, 2015)**”
- (2) 在第四页第二段第四行中“……怎样的表现?”之后增加了以下内容: “**毫无疑问, 对这些问题的回答不仅会对聋人语文教育教学提供借鉴, 而且, 对于揭示晚期双语者第二语言语法技能发展困难(Schwartz & Sprouse, 1996; Clahsen & Felser, 2006)**

认知本质有启发性意义。”

- (3) 在讨论部分，突出讨论重度聋生语法意识较弱对聋生双语教育和双语者第二语言语法技能习得的关系

审稿人 2 意见：

意见 1：

1. 文章写得挺清楚的，但理论意义有待加强，并存在一些方法学上的问题。接下来的修改重点讨论本研究的理论意义。例如：严重聋生对“是…的”句的句法意识比中等程度聋生差，这个结果从理论上讲反映了什么？为什么会出现这种情况？对于语言心理学、认知心理学或教育心理学有什么理论贡献？另外因为实验材料和程序的问题，降低了本研究结果的可信度，建议作者重做实验，并大修论文。

回应：

我们做了两处修改：

- (1) 在第四页第二段第四行中“……怎样的表现？”之后增加了以下内容：“**毫无疑问，对这些问题的回答不仅会对聋人语文教育教学提供借鉴，而且，对于揭示晚期双语者第二语言语法技能发展困难（Schwartz & Sprouse, 1996; Clahsen & Felser, 2006）认知本质有启发性意义。**”
- (2) 讨论部分做了重新调整，突出了对主要实验结果及其理论意义的解释。

意见 2：

2. 实验材料和程序问题（严重）。三个实验的实验材料分别是“11、10 和 10 句完整句（见附录）和相应省略句”，每个省略句与完整句是完全对应的（如，“道路是平坦的。”和“道路平坦的。”）。然后在实际实验中，每个被试每个完整句和对应的省略句都看了（也就是说，被试既看了“道路是平坦的。”，也看了“道路平坦的。”）。（文中没有直接描述这一点，但可以推测出来，如文中指出“实验 1、2 和 3 中的填充句数量是相应关键句的 3 倍，分别为 66、60 和 60 句。”，说明被试看了 22、20、20 个实验句，即所有的完整句和相应的省略句）。这将给实验结果带来严重的影响。（挺可惜的，可以看出作者阅读了相关的心理语言学文献，很多细节也注意到了，但是可能没有受过正规的心理语言学研究训练，建议作者重做实验，如果需要，可以通过学报编辑，我帮忙看一下实验程序）

3. 阅读理解题都出在填充句上，这样不好，如果作者重新做实验，阅读理解题最好是实验句也有。

回应：

我们愿意接受审稿人关于重做实验的建议。

但是，因为学生都已经放假回家，所以在这段时间我们没有办法重新采集数据。

不过，在讨论部分的结尾处，我们增加了以下内容：

虽然本研究中同一组被试同时阅读了完整的“是……的”句和省略了助词“是”的错误的“是……的”句，但是，不同实验重复取得同样的结果。我们相信，如果对同样的设计进行 ERP 实验，那么，对实验组在完整句和省略句句尾助词“的”的信息加工会主要测量到 N400，对控制组省略句句尾助词“的”的信息加工会主要测量到 P600。

意见 3：

4. 结果分析方案：作者“在完整句和省略句中，分别对他们在句尾助词“的”和其之前词的 RT 之间进行 t 检验”。这个分析对作者的研究实践有意义，但对本文意义不大，可以去掉。

首先句尾组词“的”和之前的名词词性、位置、信息量等各方面都不一样，是不可比较的。另外这个分析与主题关系不大。这样的实验设计主要是比较两个被试组在完整句和省略句中对应的词上的加工是否有差异，这样的比较才能尽可能地排除无关因素的干扰。

回应：

我们删除了相关部分，对结果陈述重新做了调整。

意见 3：

5. 实验结果描述较为混乱。实验 1：(1) 刚刚提到了，结果的第一段就不要了。(2) 当 $p < .05$ 时，写 $p < .05$ 就可以。(3) 关于第二段的简单效应分析：首先，边缘显著是不太稳定的效应，在这上面做进一步的分析和讨论要谨慎。其次，如果做简单效应分析，应该在一个因素的两个水平上分别检验另一个因素的效应，重点是比较在这个因素的两个水平上另一个因素的效应有没有差别，而不是要挑出所有差异显著的结果报告。实验 2、3 类似的问题也一起改掉（重要）

回应：

根据审稿人意见，我们进行了 3 个方面的修改：

- (1) 结果的第一段删掉了。
- (2) 所有 $p < .05$ 的地方全部改写成了 $p < .05$ 。
- (3) 实验 1 中删除了对边缘显著交互作用的效应分析。

遵照审稿人意见，对各实验结果中的交互作用重新进行了简单相应分析，即，在一个因素的两个水平上分别比较了另一个因素不同水平之间的差别显著性。

意见 3：

6. 另外，实验 1 中，按照作者报告的统计显著性检验结果（“实验组在省略句中的 RT 显著长于他们在完整句中的 RT (t 项目(10) = 2.658, $p = 0.024$, $d = 1.68$)；省略句中，实验组的 RT 显著长于控制组的 RT (t 项目(10) = 3.271, $p = 0.008$, $d = 2.07$)。”），似乎实验组在两种不同句子中的 RT 不同，而控制组则差不多，这与结果图提供的信息矛盾（按照图上来看，实验组有显著差异，控制组则差不多）。请作者提供原始数据以做进一步分析。（重要）

回应：

是的，在这里我们表述欠清楚。

“实验组..... 2.658, $p = 0.024$, $d = 1.68$ $p = 0.008$, $d = 2.07$ ”是关于句尾助词“的”之前的“词 1”上边缘交互性作用的报告。如我们对前一条意见的回答，这一段已经删除。

图 1 中，在“的”上的 RT，控制组在正确句和错误句之间有差异，实验组则差不多。对此，我们补充了完全的简单效应分析结果。

意见 4：

7. 讨论部分：第一段的“结果发现”应当报告全部结果。另外，目前报告的这部分结果是不建议做的分析，所以可以删除。

回应：

遵照审稿人意见，我们重写了实验结果小结：

.....对 RT 分数进行方差分析表明，(1) 实验 2 中，实验组对句尾助词“的”的 RT 显著长于控制组对句尾助词“的”的 RT。(2) 实验 1 和实验 3 中的结果相似：实验组对句尾助词“的”的 RT 在完整句和省略句之间没有显著性差异，然而，控制组在完整句中对句尾助词“的”的 RT 显著短于他们在省略句中句尾助词“的”的 RT；在完整句中，实验组对句尾助词“的”

的 RT 显著长于控制组对句尾助词“的”的 RT，然而，在省略句中，实验组对句尾助词“的”的 RT 和控制组对实验组对句尾助词“的”的 RT 没有显著性差异。

意见 5:

8. 重度和中轻度聋生在句子阅读上的相似性：这一部分的讨论可以简略一点，本研究重点是下一部分。

回应:

我们压缩了两种被试间句子阅读相似性讨论，突出了其差异性讨论：

- (1) 删除了小标题；
- (2) 把原来相似性缩减至一个自然段
- (3) 扩展了关于“因为语言差异，晚期双语者对第二语言中某些句式可能难以发展语法意识”的讨论。

意见 6:

9. 结论部分：“本研究带给人们的启示是，因为语言差异，晚期双语者对第二语言中某些句式可能难以发展语法意识”，这个类比是可行的，但是需要更详细的阐述。在引言和讨论部分的逻辑要连贯，且作为一个启示，放在结论中似乎不太合适。

回应:

我们进行了两方面修改：

- (1) 在结论中删除了“因为语言差异.....语法意识”；
- (2) 关于“因为语言差异，晚期双语者对第二语言中某些句式可能难以发展语法意识”，我们在讨论中扩展了这方面的讨论。

意见 7:

10. 最后，考虑到《心理学报》注重研究的理论贡献，而不仅仅是实际应用价值，本文的理论意义有待加强

回应:

我们做了两处修改：

- (1) 在第四页第二段第四行中“.....怎样的表现？”之后增加了以下内容：“**毫无疑问，对这些问题的回答不仅会对聋人语文教育教学提供借鉴，而且，对于揭示晚期双语者第二语言语法技能发展困难（Schwartz & Sprouse, 1996; Clahsen & Felser, 2006）认知本质有启发性意义。**”
- (2) 讨论部分做了重新调整，突出了对主要实验结果及其理论意义的解释。

第二轮

审稿人 1 意见:

意见 1:

1.选题具有一定创新性，且具有一定的理论意义。但是，论文对于理论意义部分，应该再进行加强和丰富。

回应:

我们改写了讨论部分的第二小节，加强的理论意义上的讨论。

意见 2:

论文中,以后不少不太符合心理学论文写作规范的地方,我已经采用标注的方式标出,供参考。

此外,我也同意第二位审稿人的意见以上意见仅供参考。

回应:

我们对行文又进行了润色,并请专家帮我们进行了校对性阅读。

审稿人 2 意见:

意见 1:

1、为何需要重做实验:

在第一轮审稿意见中,我指出,实验中每名被试都看了每套实验材料的所有两个版本,即被试既看了“道路是平坦的。”也看了“道路平坦的。”这会污染实验结果(见上轮审稿意见第 2 点)。作者回复“不同实验重复取得同样的结果”,因此认为影响不大,这个逻辑是不合理的。原因有两点:(1)并非不同实验重复取得同样结果,实验 2 的结果与实验 1、3 是不一样的。(2)即使不同实验取得相同结果,也不能说明实验设计上的缺陷没有影响,很有可能,实验缺陷以一种系统性的方式影响实验结果。建议参考《实验心理学》中关于实验设计的章节,特别注意如何控制混淆变量、无关因素等。

具体来讲,在心理语言学研究中,被试看同一套材料的两个版本,会导致若干方面的问题。首先,最简单的,是被试先看的句子会对后看的句子产生影响,考虑到重复效应,可能导致后看的句子的阅读更容易;然而考虑到“是…的”句式的强调作用,后看的句子受到的影响可能因其是正常句还是异常句而有所不同。另外,请作者思考:“为什么实验要加入填充材料?”当然这个问题的答案作者在文稿中已经回答了,

“为避免被试意识到句型结构特点或启用阅读策略等(Rüschmeyer, Zysset, & Friederici, 2006),实验 1、2 和 3 中的填充句数量是相应关键句的 3 倍,分别为 66、60 和 60 句。”可是,当被试既能读到“道路是平坦的”,又能读到“道路平坦的”,难道还意识不到句型结构特点吗?所以作者调控无关变量的操作(加入填充材料等)是失败的。被试意识到句型结构特点后,对其阅读过程是有影响的。一个可能的影响,请作者查阅关于“实验者期望效应”的心理学书籍或文献。实际上,心理语言学研究中,每个被试只能看同一套语言材料的一个版本。这在心理学研究中叫做项目内设计,请作者学习这方面的知识。

所以,实验设计是存在缺陷的,该缺陷可能污染实验结果,导致文中的观点和结论错误。这是必须重做实验的原因。

回应:

我们重新做了实验,对论文进行了相应修改。

意见 2:

2、除上述情况外,实验设计还存在其它小缺陷,在重做实验时可一起改正,数据和结果将更为可靠:

(1)请参考第一轮审稿中的第 3 点。

(2)每个实验只有 11 套或者 10 套材料,严格来说有点少,建议增加到每种条件 15-20 个材料,也就是每个实验 30-40 套实验材料。

(3)如果能采用眼动技术进行实验,结果会更可靠,可考虑。

回应:

(1)全部实验材料(包括关键句和填充句)均要求被试做一个阅读理解题。

- (2) 三个实验中，实验材料均为 36 套。但是由于技术上原因，实验 2 中一个关键完整句的数据记录不全，因此，相应省略句的数据也没有参与统计分析。
- (3) 目前，我们还不能在聋人学校进行眼动实验测量。

意见 3:

- (1) 要说明如何处理和分析数据，譬如说明是对哪个词做的统计显著性检验
- (2) 应该做两种分析：被试分析和项目分析，作者只做了项目分析。
- (3) RT 应统一写成反应时，作者同时使用了反应时和 RT 这两个词。
- (4) 作者在讨论部分写到：“至于为什么实验 2 中被试对完整句和省略句句尾助词“的” RT 的相对差异和他们在另外两个实验中的情况有所不同，还需要进一步研究。”请试着进行解释和讨论，而不是不加思考，用“需要进一步研究”来解决。
- (5) 一个补充：p 值也可以视实际情况写 $p<.01$, $p<.001$ 等不同显著性水平。第一轮审稿中，我的意思是，当达到显著水平时，不需要报告 p 值的具体数值，只要标明其显著性即可。当然，现在这样写也可以。

回应:

- (1) 在报告数据分析之前和结果讨论部分，明确说明了数据分析方法：“**对被试在句尾助词“的”的反应时数据进行 2（被试）× 2（句型）两因素混合测量方差被试分析和项目分析**”。
- (2) 我们报告了被试分析和项目分析结果。
- (3) RT 应统一写成了反应时。
- (4) 和上次相比，实验 2 的结果更加丰富一些了。我们也着重讨论了实验 2 与其他两个实验的差异。
- (5) 所有 p 值均只标明了其显著性水平。