

《心理学报》审稿意见与作者回应

题目：声调在汉语音节感知中的作用

作者：赵荣，王小娟，杨剑峰

第一轮

审稿人 1 意见：

本篇文章，采用 oddball 范式，利用两个行为实验探索了声调在汉语音节感知中的作用。结果发现，声调和声母变化的监测没有差异，两者都慢于韵母变化的监测；声调与声母或韵母的组合都促进了音节变化的监测。本文章的实验结果比较清晰，但是研究问题和创新之处需要进一步澄清。

意见 1：作者在引言部分的最后指出“综上所述，汉语声调可能具有独特的加工机制，在口语词汇识别中可能具有不同于声母和韵母的作用。然而，相对于音段信息(声母和韵母)，声调在汉语口语词汇识别中的作用究竟如何？它们又如何共同影响口语词汇的识别加工？目前对上述问题还缺少系统地研究。”。但是从这段话的描述中，我们尚不清楚，本研究的具体研究问题是什么？作者所说的加工机制，具体指的是什么？已有大量的研究考察了声调在词汇感知和语义加工中的作用，本研究的创新之处是什么？

在引言部分，作者较全面地综述了大量有关声调加工的研究，如，声调的范畴化加工、声调对词汇语义加工的影响、声调加工的大脑偏侧化、以儿童为被试的发展方面的研究。但是，这些内容比较分散，作者应该结合研究问题、按照一条主线把这些研究有机地组织在一起。最为重要的是，作者没有充分分析这些已有研究与本研究的研究问题之间的关系。即，已有研究中哪些问题还未澄清，本研究能提供什么新信息。

回应：此次修改对引言部分进行了比较大的修改和调整，主要突出了文章的思路和逻辑，并对前人的文献研究有针对性进行了分析和评述。分别与两个实验的研究目的结合起来，以突出本研究的两个重要发现：1)在口语感知中声调的单独作用弱于韵母而优于声母；2)声调依附于音段信息且共同影响汉语音节的感知。

另外，在讨论部分的修改中，我们也有意识地突出了本研究的创新之处。

意见 2：作者在引言的最后一段指出“因此，本研究采用被广泛应用的 Oddball 实验范式，通过两个行为实验，分别操纵偏差刺激与标准刺激变化的语音单元，考察汉语声调在口语词汇识别中的作用...”。为什么采用 oddball 范式的行为实验能考察本研究的问题？

回应：在修改的引言部分，说明了前人研究中声调与音段信息的相对作用不一致的原因可能是受语义信息的影响；而采用 oddball 实验范式可以排除主动任务以及语义因素的影响，更好地考察声调、声母和韵母在音节感知中的自动影响作用。在引言最后一段解释了选择该实验范式的两个理由。

意见 3：本研究主要有两方面的研究结果：1) 声调和声母变化的监测没有差异，两者都慢

于韵母变化的监测; 2) 声调与声母或韵母的组合都促进了音节变化的监测。对于第一条结果, 作者在讨论部分指出“...我们的数据结果与前人的 ERP 结果是一致的, Hu 等(2012)和 Li 等人(2014)都发现韵母违反诱发了 N400, 而声调违反更多诱发了相对晚期的 P600 效应, 表明声调可能比韵母的作用时间要晚一些。”对于第二条结果, 作者也指出“本研究一个有趣的发现是, 声调与音段信息的结合有效促进了对音节的识别加工。这与前人发现声调参与并影响了口语词汇的加工相一致(Zhou et al., 2004), 表明音调信息可能依附于音段信息之上起作用...”。那么, 本研究的新贡献是什么?

回应: 本研究的贡献在于使用反应时技术首次在行为层面揭示了声调的作用。两条实验结果虽然在前人文献里已经有所体现, 但都是间接的证据, 如 Zhou 等人(2004)的研究探讨的问题是声调在口语词汇的语义通达中的作用, 与本研究探讨音节感知不同。此次修改, 在讨论和结论部分明确提出了本研究是在行为层面首次提供实证数据。本研究的另一个贡献在于, 提示了探讨音段与超音段信息的整合机制应该成为汉语音节感知加工研究的新问题。

审稿人 2 意见:

该文通过两个实验探讨了汉语声调在音节感知中的作用, 发现单独检测声调变化和检测声母变化的时间没有差异, 都比检测韵母变化的时间长。但声调与声母、韵母的同时变化则显著促进了被试对偏差刺激的觉察。文章实验设计严谨, 分析得当, 结果可信。有以下问题或建议。

意见 1: 作者的关键结论是: “声调的单独作用相对有限, 而与声母和韵母结合的共同作用对汉语音节识别的影响更大”。不知这个“相对有限”该如何理解和界定, 容易让人误解为声调单独不起作用。从实验 1 的结果看, 反应时上声调条件与声母条件相似, 但错误率上声调条件要优于声母条件。所以说一个合理的结论应该是“声调单独的作用比不上韵母, 但优于声母”。结论的另外一个要点, “声母和韵母结合的共同作用影响更大”, 在讨论中没有对“结合能产生更大作用”的原因进行展开论述。

回应: 1)此次修改对声调“相对有限”的作用更加明确化了, 在摘要以及结论部分都明确提出“声调的作用弱于韵母而优于声母”, 这一结论不仅来源于实验 1 的结果, 还来源于实验 2 的结果, 在两个实验的讨论部分都已经明确提出。

2)在讨论的第 5 段文字, 对声调与韵母结合产生的更大影响作用进行了讨论分析。

意见 2: 引言的最后一段, 即实验 1 的前一段, 应该主要是提出问题, 陈述研究假设、研究逻辑及关键的变量操作等。这儿先把研究结果描述出来, 有些不太“常规”。

回应: 接受审稿人的意见, 对本段文字进行了重写。主要介绍了两个研究的实验思路和逻辑。

意见 3: 鉴于本研究主要关注声调的作用, 实验 2 中的“声母+韵母”条件的设置意义不清楚。希望作者能予以说明。

回应: 在实验 2 的“设计和材料部分”增加了对于设置声母+韵母条件的说明, 并在实验 2 的讨论部分增加了针对这个条件的结果讨论。

意见 4: 实验 2 和实验 1 的性别比例差别很大,让人担心这个变量会影响两个实验中的结果模式。作者应该在讨论中给予适当讨论和解释。

回应: 我们也意识到了这个问题:实验 1 的 28 名被试中仅有 4 名男生,在做实验 2 时我们注意到了这个问题,31 名被试有接近一半的被试是男生(14 名)。在修改稿的“总讨论”最后一段,我们对性别可能的影响以及本文实验 1 的不足进行了讨论。

第二轮

审稿人 1 意见:

该研究采用 oddball 范式考察了声调在汉语音节感知中的作用。作者针对前面两位审稿人的意见和建议做了一些修改,但文章仍然存在如下问题。

意见 1: 实验目的:“实验结果将首次的行为层面揭示声调在汉语音节感知中的作用机制,为揭示口语词汇识别中音段与超音段信息的作用机制提供证据。”尽管作者认为自己的研究是“首次”,但事实上已经有一些行为研究考察了音段与超音段信息的认知机制,如 Tong 等(2008)的研究。作者在文章中却忽略了这些文献。

Tong, Y., Francis, A. L. & Gandour, J. T. (2008). Processing dependencies between segmental and suprasegmental features in Mandarin Chinese. *Language and Cognitive Processes*, 23, 698–708.

回应:“首次”的确让人感觉语气比较强烈,我们分析这里存在的问题还是表述的精确性问题。虽然前人已有研究探讨了音段或超音段信息的不同作用机制,但两种信息的共同作用机制无论在行为和神经层面都是比较新的研究问题,还缺乏相应的研究证据。因此,本文是在这个问题上的“首次”探讨。

此次修改我们对这个问题进行两方面的修改:

1) 去掉“首次”的表述,而且对探讨的问题进行了限制,精确为“探讨音段和超音段信息的共同作用机制”。(修改见前言最后 1 段以及讨论第 1 段)

2) 接受审稿人的建议,再次梳理文献,并对 Tong 等人(2008)的研究进行了综述和讨论,进而说明了本研究相对于前人的重要意义:系统探讨声调与音段信息的相对作用,以及多种信息的整合作用。(修改见前言倒数第 2 段)

这里很感谢审稿人的意见,通过与前人文献的更多对比,让本文更加准确地定位了研究所探讨的问题及其重要性。

意见 2: 作者在前言中重点总结了采用 ERP 技术对该问题的研究,作者是如何考虑行为结果和 ERP 结果之间的联系?行为结果与 ERP 结果如何对应?

回应: 本文的问题来源于认知神经科学研究的进展,从而综述了目前已有的行为与认知神经科学研究证据。但是引言的撰写可能存在不足,误导了读者将 ERP 与行为实验技术对立或区分,为了在内容上消除这种误导,此次修改:

1) 修改前言各段、各部分内容的逻辑性和条理性,避免了将行为与 ERP 技术的对立。

2) 增加了更多的行为实验证据,以认知机制为主导,消除 ERP 与行为研究的界限。

3)更清晰地说明本研究所探讨的问题在行为和 ERP 研究中都没有相应的证据支持,基于本文的行为结果,今后的 ERP 研究如果能够验证或揭示更多的证据,将对音段与超音段信息的相互作用机制进行更加深入地探讨。我们在讨论中提出了对未来 ERP 研究的展望。

意见 3: 根据作者的陈述“声调在汉语音节感知中的作用,及其如何与音段信息共同作用来影响音节感知还缺乏系统的探讨”。在音节感知的研究中,已经有很多这方面的成果出现,希望作者能够全面检索文献,并在前言中总结已有的研究成果,在此基础上提出本研究的目 的及贡献。

回应: 感谢审稿人的建议,此次修改,我们在引言第 2 段和第 7 段都增加了相应的文献综述,第 7 段专门探讨了汉语音节感知中的声调与音段信息的相互作用。

意见 4: Meng 等的研究发现“声母和韵母偏差刺激在 400ms 之后诱发了较大的负波,而声调偏差刺激并没有出现此现象,说明声调具有与声母、韵母不同的加工机制”。对此有疑问:第一,与什么条件相比,声母和韵母的偏差刺激在 400ms 之后诱发了较大的负波,而声母偏差刺激未出现负波?第二,有无负波的差异是如何说明声调与声母、韵母具有不同的加工机制的? 第三,儿童这一特别的被试是否会对结果产生影响?

回应: 针对上述疑问,我们进行了如下修改:

1)Meng et al.(2005)的研究中的偏差刺激都是相对标准刺激来说的,已经在修改版中进行了详细说明;另外,这里的“400ms 之后”表述容易误导为 N400,经过对原文的数据分析发现,使用“250ms 之后”更为准确。

2)有无负波仅能说明声调是否与其它两种音段信息的加工机制有区别,遗憾的是,原文没有关注这方面的对比,缺少对这方面机制的详细解释,我们也无法给出更多解释。

3)感谢审稿人的“第三”条意见,儿童被试的声调与音段信息感知能力发展不均衡的确可能是影响实验结果的因素,需要进一步得到成人实验的证据支持。修改稿中已经反映出来,成为本段综述相关实验研究的逻辑。

意见 5: Lee 等(2012)的结果与 Meng 等存在显著不同,作者是如何考虑这两个研究结果之间的差异的?

回应: Lee 等(2012)与 Tong 等(2014)的结果是相似,但与 Meng 等(2005)的结果模式不同。可能是被试年龄段的差异(修改版已经提及)、母语环境(北京 vs 香港)等因素的影响。但是, Meng et al(2005)的研究没有直接考察声调与声母、韵母的相对关系,可以通过文献里的数据图看到结果趋势,但无法进行数量化比较。就很难说明两类研究结果到底是否存在差异,所以没有深入讨论,只说明这些研究共同表明了声调与音段信息的作用不同,这也正是本段内容拟表达的中心意思。

意见 6: Li 等(2010)的脑区研究结果又与 Lee 等的不同,作者是如何考虑的?

回应: Li 等(2010)的脑区研究结果说明声调可能使用了与音段信息不同的神经回路,而前面的三个 ERP 研究说明声调在特定认知加工中的作用与音段信息不同。这里看到的这两类研

究共同揭示了声调与音段信息的不同作用(已在修改版中加强这种观点)。但是,要将 ERP 成份与脑区激活进行对比分析还没有足够的实验证据,因此,修改时没有进行深入地对比分析。

意见 7:“Gandour 等人(2003)要求被试判断连续呈现的第三个音节与第一个音节的声母、韵母和声调是否相同,结果发现判断声母不同的反应时最短,判断声调与韵母不同的反应时没有差异,结果支持声母的作用更大”。声调与韵母不同的反应时没有差异,为什么却得到结论“声母的作用更大”?

回应:在上一版中存在语言表述问题,此次修改了表述,消除了误导(见前言第 4 段)。

意见 8:“无论是单字词的语音违反(Zhao, Guo, Zhou, & Shu, 2011),还是句子的尾音节违反(Brown-Schmidt & Canseco-Gonzalez, 2004),声调、声母与韵母违反诱发的 N400 成分之间没有差异,表明三者的作用相同”。已有大量研究表明,声调、声母与韵母在一个音节的构成中扮演了不同的角色,其作用也是不同的。作者仅仅基于声调、声母与韵母违反诱发的 N400 成分之间没有差异,即推论出三者的作用相同。这样的结论有点草率,也与已有研究不符。

回应:我们承认,这里的表述过于武断,容易使读者产生误解。此次修改,我们将这一段与下一段合并,指出声调与音段信息在 N400 上没有表现出差异,但随后的研究发现声调的作用表现在 P600 上,比音段信息更能影响到对词汇语义的再分析,从而指明声调的作用区别于音段信息。

意见 9:“Zhao 等人(2010)的研究也发现整体音节违反比单独语音成分的违反要诱发更早的 N400 反应。也就是说,声调与音段信息的共同作用要强于单独的声调或音段信息,但是,这种更强的作用可能有不同的解释,一种可能是声调依附于音段信息,它与音段信息的结合会起更大的作用;另一种可能是两种信息的累加要强于单一信息的作用”。诱发出更早的 N400 反应能得到的结论是:被试对声调与音段信息共同违反的觉察要早于仅有音段信息或仅有声调信息的违反。作者是依据什么逻辑得到“声调与音段信息的共同作用要强于单独的声调或音段信息”?作者对此的两种解释有什么本质区别?

回应:这里的表述的确存在两方面的问题,一方面是两种解释的文字表述容易让人误解是针对 Zhao 等人(2010)的研究;另一方面是解释的比较强势。经过修改:

- 1)删除这两种可能的解释;
- 2)仅提出上述研究还不足以揭示出“声调在整体音节感知中的作用”。

意见 10:“实验结果将首次 in 行为层面揭示声调在汉语音节感知中的作用机制,为揭示口语词汇识别中音段与超音段信息的作用机制提供证据”。与 ERP 或 fMRI 技术相比,本研究中提到的行为层面的研究能比 ERP 或脑成像的研究多提供哪些信息?新的贡献是什么?

回应:此次修改:

- 1)已经去掉了“首次”的表述,从而消除了行为与神经证据的对立。
- 2)无论是行为还是神经方面的研究都没有揭示出声调与音段信息如何共同起作用。我们的研究并非是在给 ERP 或 fMRI 技术的结果补充行为证据,而是从认知神经科学研究的角度

度提出了认知行为层面的研究问题，同时可以为未来的认知神经科学研究提供方向。这一点已经在引言和讨论的修改部分中体现出来。

意见 11: 在汉语口语产生研究领域,有一些有关声调、声母和韵母作用的研究(如 Chen,Chen& Dell,2002;Zhang & Zhu,2010 等),作者应该对此进行检索,并综述其研究结果,考察汉语加工中音段与超音段信息的不同作用。

回应: 接受审稿人的建议,我们在前言第 5 段增加了关于汉语口语产生方面的研究综述,从口语产生的角度来表明音段和超音段信息具有不同的作用机制。

意见 12: 方法:本研究试图“有效考察声母、韵母和声调在音节感知中的自动化作用机制”,但在实验中要求被试有意识地去判断三类偏差刺激与标准刺激的是否相同,这样的实验方法能否达到考察自动化作用机制的目的?

回应: 这是对初稿修改不彻底导致的问题,我们在文章的初稿写作中曾经涉及了自动化加工的问题,正如审稿人指出的一样,我们在修改稿中感觉不合适就去掉与自动化加工相关的表述。此次修改经过通篇检查,已经确定去掉了所有与自动化相关的表述。

意见 13: 实验 1 结果:作者认为本研究的实验设计可以避免“语义”在音节感知中的作用,但在实验 1 的讨论中却一直提到“汉语口语词汇识别”,而非“感知”。口语词汇识别必然包含了语义的加工。该研究的目的是要考察“词汇识别”还是“音节感知”?

回应: 这里的确存在术语表达的不精确问题,此次修改通篇统一了术语,仅在介绍前人文献中关于口语词汇识别的研究中出现“口语词汇识别”术语,与本研究相关的都局限于“汉语音节感知”。

意见 14: 如何解释实验 1 所得到的对声母变化的觉察要长于韵母,这与已有的研究结果矛盾。

回应: 对此矛盾结果的讨论,总讨论第 2 段虽有提及,但并没有深入讨论。此次修改在实验 1 的讨论部分(第 2 段)增加了对矛盾结果的专门讨论:一方面指出实验任务的差异可能导致的结果差异;另一方面也指出,声母与韵母的相对重要性仍然是研究者争论的问题,还需要更多的实验证据进行深入探讨。

意见 15: 总讨论:针对前言中提到的问题进行讨论,与已有的行为结果做比较。现在的总讨论主要借用了脑成像的实验结果,由于行为层面与脑层面结果对应的间接性,用脑成像层面的结果对本研究中所发现的行为结果进行解释是不妥当的。

回应: 本文在总讨论倒数第二段提及了部分脑成像研究的成果,主要有两方面的目的:

- 1)借助 fMRI 的最新研究进展说明本研究所探讨问题的前沿性和重要性;
- 2)同时也反映了本文的行为实验结果对于今后 ERP 或 fMRI 实验的促进作用。

此次对这段文字的表述进行了仔细修改,使上述两个目的更加明确。

意见 16: 行文中有一些小问题:

例如,Bonferroni—》 Bonferroni

“不如韵母变化的作用却优于声母”,这一句子与上下文读起来不太通顺。

回应: 非常感谢您的细心审阅, 上述地方已经修改。

另外, 此次个修改共增加了 13 篇参考文献, 已经在返修稿更新了相应的参考文献列表, 以及投稿说明中的文献比例。

第三轮

编委复审意见:

意见 1: 统计分析缺乏效应量的数据。

回应: 感谢您的建议, 这是我们的疏忽, 误读了投稿指南的“建议给出”效应量, 就没有提供这方面的数据。此次修改, 我们添加了相应统计分析的效应量数据。

意见 2: 在实验 2 中, 声调加声母、声调加韵母、声调加声母和韵母条件差异不显著, 显然有违逻辑, 如果真的如此, 作者整个研究结果就十分令人怀疑。作者的天花板效应的解释显然说不过去。是研究技术不敏感还是别的原因, 建议作者用其他方法与技术检验这种可能。

回应: 实验 2 的结果发现上述三种复合条件的差异不显著, 也出乎我们的意料:

首先, 天花板效应肯定不能解释全部, 我们对此部分讨论的不足, 导致可能的误解, 此次我们修改了这部分讨论内容, 消除了这种误解。

其次, 我们承认, 可能需要其他的方法与技术来检验声调与音段信息的整合作用。事实上, 之前我们曾经做过一个被动的实验任务, 结果发现所有复合条件(CV, CT, VT, CVT)之间都没有差异。而采用主动任务时, 实验 2 发现了 CV 条件与 VT、CVT 条件之间的差异显著, 因此本文只报告了主动任务的实验 2 结果。从而表明, 信息的复合作用是存在的, 但更加深入地研究可能需要借助 ERP 等先进技术手段实现, 修改稿对此进行了讨论, 同时也指出了今后 ERP 研究技术的方向。

另外, 如果要突出 T 的作用, 一种可能的办法是减少实验条件, 只考查 C 与 T, V 与 T, CV 与 CVT 的差异, 但是, 这样的实验就很难说明 C、V 与 T 三种信息的整合作用。另一种减少实验条件办法是只对比 CT、VT 与 CVT 三者的差异。但是, 将这些孤立实验结合在一起时, 将导致另一种可能的解释, 即构建的实验条件多少导致了被试判断标准的变化, 并非真实反映了复合条件内部的加工机制差异。而探讨被试的判断标准是否受实验条件多少的影响超出了本文的探讨范围。

因此, 从经验上来说, 我们的确应该发现 CT,VT 和 CVT 之间的差异。但是很遗憾目前的行为实验结果仅能揭示出有限的作用, 还不能深入地揭示出声调与音段信息的整合机制。这也为今后语言认知神经科学的研究提供了方向(已经在讨论部分指出)。

意见 3: 文章的行文不够简洁, 标点符号的运用问题也多, 不要用“我们的”之类的第一人称表述。

回应: 我们再次对全文进行了校正阅读, 通篇不再出现第一人称的表述, 还修正了多处文字表述以及标点符号等错误。

第四轮

主编终审意见：

研究脉络清晰，统计合理，建议对以下几点进一步思考完善。

意见 1：建议进一步说明文章的创新性贡献。

回应：感谢您的建议。本文的最大贡献在于弥补了对音节感知中声调作用的认知不足，在认知行为层面提供了声调与音段信息共同影响音节感知的直接证据。本研究为探讨音段信息与超音段信息共同作用的认知神经机制提供了基础，并指明继续深入地研究将完善和发展口语词汇识别的认知和神经生理模型。为了加强本文的研究意义，此次做了以下修改：

1) 摘要：在开头部分指明了现有的汉语声调研究集中于口语词汇的语义通达阶段，在相对早期的音节感知阶段，声调与音段信息的共同作用机制还缺乏系统地研究；在结尾部分进一步指出了本研究的意义，以及对认知神经科学研究的潜在作用。

2) 引言：第 2 段明确指明探讨音节感知中音段与超音段信息共同作用的理论意义。最后 1 段指出现有研究对声调在音节感知中作用的认识不足，明确成为本研究的出发点。

3) 讨论：在第 1 段最后增加了对本研究的意义说明；第 2 段通过对前人研究的分析来明确指出本研究提供了“直接的证据”。倒数第 2 段结合认知神经科学的研究进展，指明探讨音段与超音段信息共同作用的意义，从而指明了“今后研究的方向”。

4) 结论部分，重申引言第 1 段指出的口语词汇识别模型的不足，进一步指明探讨音段与超音段信息共同作用的重要理论意义。

意见 2：第一段的概念介绍中，着重介绍了口语词汇识别的模型，没有提及音节感知，二者之间是否为从属关系或者其他？并不明确，音节感知是本文研究的重点，建议明确概念，以免混淆。

回应：此次修改增加了引言第 2 段，指明音节感知是口语词汇识别相对早期阶段。

意见 3：实验一结果讨论中认为“声调比声母的作用强”，但这二者在反应时和正确率上均无显著差异，请确认。其他小的问题：1) 前言最后一段第三行：有必要在排除语义加工的影响考察声调与音段信息在音节感知中的作用机制。读起来不通顺。2) 实验程序中的“run”和“block”应改为中文。3) 文中部分 t, p 等统计量应斜写。

回应：这里的确存在过度解释的问题。提出“声调比声母的作用更强”的结论只能从实验二的结果“声+韵与调+韵之间存在显著差异”得到。实验一虽然在正确率上声调与韵母相当，并优于声母，但统计结果不够显著，在实验一讨论部分就提出这个结论的确是不合适的。

此次修改进一步精练了实验结果，集中在讨论两个主要结果：即单独的声调作用弱于韵母，而声调与音段信息的结合作用明显。而对“声调比声母的作用略强”的结果不再重点提及，仅在实验二的讨论部分有所提及。

非常感谢您的仔细阅读，此次修改校正了上述三个细节问题，同时还修改了几处文字表达问题。