

# 《心理学报》审稿意见与作者回应

题目：汉语口语词汇产生的音韵编码单元：内隐启动范式的 ERP 研究

作者：王雪娇，张清芳

## 第一轮

### 审稿人 1 意见：

**意见 1：**前言部分对研究的合理性和必要性阐述不足。在该研究领域，已经有了一些 ERP 研究探讨汉语语音编码的时间进程，但是结果却不是非常一致（如是否存在音素效应，如存在其和音节效应的时间进程的差异）。因此作者希望采用内隐启动范式对这个问题进行研究。然而，作者使用内隐启动范式的理由是什么，在多大程度上以及在哪些关键问题上会比之前的研究范式更加有效地检测作者关心的问题，基本谈及较少。因此进行该研究的合理性和必要性阐述是不足的。建议作者在前言部分，从范式出发，进行一定的合理性阐述，并基于此对结果进行预测。

**回应：**感谢您的意见和建议，我们在修改稿中强调了本研究特色和贡献(见 P5~P6)。

基于研究现状的阐述，我们可以看到有关音节和音素在汉语口语产生中的作用及其发生阶段仍存在激烈争论。已有研究存在行为与 ERP 结果的不一致 (Wong & H.-C. Chen, 2008,2009; Wong et al., 2016)，以及 ERP 结果之间的不一致(Qu et al., 2012; Wang et al., 2017; Zhang & Damian, 2019)。在两个发现音素效应的研究中 (Qu et al., 2012; Yu et al., 2014)，均未将音素与音节效应进行比较，且存在被试的加工方式可能受到第二语言迁移的影响。已有 ERP 研究均采用了图画命名任务，包括了概念准备、词条选择、音韵编码、语音编码和发音等一系列加工过程，这些过程之间会产生相互影响，比如概念准备对词汇选择和音韵编码的影响。在 ERP 研究中，音节条件下的激活较强，而音素条件下的激活较弱，难以探测到 (Wang et al., 2017)，这提示研究者应该采用对音素效应更为敏感的实验任务。

在口语产生领域，研究者采用内隐启动范式对音韵编码进行了深入研究。在这一任务中，被试先学习几个词对，比如“海豹-夕阳，记忆-媳妇，皇帝-喜事，贝壳-细胞”。这四个单词对中的第二个单词第一个字的音节是相同的，称之为同 (homogeneous)条件。对照组单词对中的第二个单词在音韵上毫无关联，称之为异质 (heterogeneous)条件。在测试阶段呈现单词对中的第一个词(线索词如“海豹”，被试要说出对应的第二个词(目标词如“夕阳”)(张清芳，2008)。在内隐启动范式中，人们在目标词之间有音韵联系的情况下会产生准备效应，出现音韵促进效应。在印欧语系中，研究者采用内隐启动范式灵敏地探测到了单个音素的准备效应(Meyer, 1990, 1991; Damian & Bowers, 2003)，表明该范式对音韵编码单元是敏感的。此外，这一任务不包括概念准备阶段(Cholin, Schiller, & Level, 2004)，线索词呈现时被试仅需要回忆出相应的词汇，并对词汇进行音韵编码和语音编码，然后说出词语，任务中尽可能地避免了概念准备和词汇选择对音韵编码过程的影响。根据此分析，我们猜测内隐启动范式可能对音素效应会更为敏感。

在利用 ERP 技术发现早期音素效应(图画呈现后的 300 ms 之前)的研究中(Qu et al., 2012; Yu et al., 2014)，研究者只创造了音素相关与无关条件，这一研究设计不能比较音节与音素效应发现的时间进程。此外，这些研究中采用的皆为外语水平很高的双语被试，存在被试的加工方式可能受到第二语言迁移的影响。因此，本研究首次采用内隐启动范式和 ERP 技术相结合，而且我们选取了低二语水平且日常生活中均不接触英语的被试，以排除第二语

言对第一语言的影响，探索汉语普通话口语产生过程中音节效应与音素效应的时间进程。

**意见 2:** ERP 结果中对于 100~200ms 时间窗的效应的解释存在偏差。从作者使用的实验程序(视觉词汇呈现)、ERP 成分形态和分布看, 100~200ms 应该也是比较经典的 P2 成分。因为这是一个正成分, 因此越正其效应越大。基于此, 从图上以及作者的描述统计结果应该得知音节同质条件相比于异质条件诱发了更大的 P2 成分。而作者则对其解释则相反。建议进行修改。更重要的是, 解释早期出现的 P2 效应的原因何在。P2 往往跟注意相关, 而在心理语言学的研究中有研究者提出或许跟知觉层面的预期存在一定关系, 因此这似乎表明同质条件相比于异质条件诱发了更大的注意或者知觉层面的加工。这个效应是否是由于作者使用的特定的被试群体导致亦或者是由于这个特殊的范式导致(例如是否是由于被试年纪较大, 记忆能力退步, 导致在同质条件下出现了更大的干扰效应, 从而导致了更大的注意和知觉加工的卷入)。但是起码 P2 的结果一定程度上或许解释了为什么作者在行为学上没有看到任何差异的结果(同质条件 P2 更大, 异质条件晚期成分更大, 相互抵消)。建议作者对以往文献进行总结的基础上, 进行对比分析, 并在此基础上对 P2 进行解释, 更重要的是, 解释 P2 效应的干扰, 是否, 以及在多大程度上会对实验主要结果造成影响。

**回应:** 非常感谢您所提供的解释, 这对于我们理解音节促进效应的机制有很大帮助。我们将此解释增加到讨论中相应的部分(见 P14)。

本研究发现无论是音节效应还是音素效应, 在波形上都表现为同质条件比异质条件更正, 这与其他 ERP 的研究发现一致(Qu et al., 2012; Yu et al., 2014; Wong et al., 2016; Wang et al., 2017; Verdonschot, Tokimoto, & Miyaoka, 2019), 反映了音节或音素相同时促进了口语产生过程, 表明音节或音素是音韵编码的加工单元。本研究中所发现的音节促进效应位于线索词出现后的 100~400 ms 之间, 这与经典的 P2 和 P3 波形所出现的时间窗口相似(Hackley, Woldorff, & Hillyard, 1990; Donchin & Coles, 1988)。研究表明 P2 成分与视觉特征的探测有关, 而且 P2 成分对正字法和音韵加工任务、语义分类任务和词汇分类任务中不同条件的设置比较敏感, 会出现波幅上的差异(Luck & Hillyard, 1994)。研究也发现 P3 的波幅大小与注意的程度有关(Gray & Burgess, 2004), 也可能与决策做出后的认知闭合(cognitive closure)(Desmedt, 1980; Verleger, 1998), 或与有意识地信息通达过程相关(Picton, 1992)。这一成分包括了两个子成分(P3a 和 P3b), P3a 分布在前额叶部位, 当有未预期的事件出现时引发这一成分, 反映了自动的注意调节机制; P3b 分布在顶枕区域, 与注意、工作记忆和复杂的认知加工(包括语言加工)相关。这提示与音节异质条件相比, 音节同质条件下可能诱发了被试更多的注意, 促进了被试的口语产生过程。关于注意如何调节口语产生中的音韵编码过程以及波形的变化, 则需要更进一步的研究。

#### 参考文献

- Desmedt, J. E. (1980). P300 in serial tasks: An essential post-decision closure mechanism. *Progressing in Brain Research*, 54, 682-686.
- Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of cognitive updating? *The Behavioral and Brain Sciences*, 11, 357-427.
- Gray, J. & Burgess, G. C. (2004). Personality differences in cognitive control? BAS, processing efficiency, and the prefrontal cortex. *Journal of Research in Personality*, 38, 35-36.
- Hackley, S. A., Woldorff, M., & Hillyard, S. A. (1990). Cross-modal selective attention effects on retinal, myogenic, brainstem and cerebral evoked potentials. *Psychophysiology*, 27, 195-208.
- Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, 31, 291-308.
- Picton, T. W. (1992). The P300 wave of the human event related potential. *Clinical Neurophysiology Lisse*, 9(1),

Verleger, R. (1988). Towards an integration of P3 research with cognitive neuroscience. *Behavioral and Brain Sciences*, 21, 150-154.

**意见 3:** 音素效应的解释需要更加谨慎。在这个研究中，作者在较晚的时间窗发现了较为微弱的音素效应，作者解释这个时间窗代表了音节分解音素的过程。然而之前的研究要么没有发现音素效应，要么发现音素效应在较早时间窗出现。此外，从理论上来说整体音节提取(和储存)的假设，也不一定需要进一步的音节分解的过程(而大量的结果证明似乎这个过程可能就是不存在的)。最后，这种较晚出现较为微弱的效应或许只是反映了一些跟任务相关的加工，而非实时的语音编码过程。因此建议作者对这个效应的解释要更加谨慎。

**回应:** 非常感谢您的宝贵意见。与音节效应相比，音素效应在口语词汇产生过程中确实比较微弱。例如，已有研究中在脑区之间比较音素相关和无关条件时并未进行统计校正(Qu et al., 2012; Yu et al., 2014; Wong et al., 2016; Wang et al., 2017)，我们尝试采用 *fdr* 校正对已有研究报告的 *p* 值进行校正后发现，基本上所有的 *p* 值在校正后都不显著( $p > 0.05$ )。本研究中的实验数据在统计校正后仍然显著，能在一定程度上避免这一问题，表明内隐启动范式下发现了显著的音素效应。

本研究中发现的音素效应出现在线索词呈现后的 500~600 ms 之间，发生在音节效应之后。根据 O'Seaghdha 等(2010)提出的模型(见图 1)，我们对此音素效应尝试进行解释。在汉语口语产生过程中，音韵编码阶段音节被提取后会进一步分解为各个音素，在此之后各个音素的提取是平行进行的(Roelofs, 2015)或者是以从左至右的顺序提取的(J.-Y. Chen, O'Seaghdha, & T.-M. Chen, 2016)。同时人们也会提取相应的节律信息，形成单词形式的框架，将音素插入到相应的框架中去，进行语音编码为发音做好准备。音素效应的可能发生在音韵编码或者语音编码阶段。第一种可能性是音素效应发生在音韵编码阶段，在音节提取之后将其分割为各个音素，其仍处于音韵编码阶段。第二种可能性是这音素效应发生在语音编码阶段，Zhang 和 Damian(2019)在图画命名过程中发现了 500~600 ms 时间窗口内的音段效应，与图画命名元分析后的时间进程对比，研究者认为这一效应可能出现在语音编码阶段，反映了音段的启动作用。这一效应比较微弱，难以探测。本研究无法区分这两种解释，需要通过研究任务之间的对比(即时命名、延迟命名，以及延迟命名和发音抑制的结合)对这两种解释进行考察。

我们在对此进行了详细讨论，谨慎地提出了两种可能的解释(见 P12~P13)。

**意见 4:** 建议对于关键的概念添加对应的英文翻译(例如词素音位编码、音韵编码和语音编码，音素，音段等)。此外在论文中建议对音素和音段这两个概念的使用再仔细斟酌，目前有时候出现相互混淆的地方，建议统一概念的使用。

**回应:** 非常感谢您的宝贵意见。我们对您所指出的概念都添加了英文注释(见 P1)。对于音节、音素和音段之间的区分，我们在第一次出现时采用脚注的形式进行了解释(见 P1~P3)，并且在全文的使用中注意区分了音素和音段。

**意见 5:** 前言部分“不计声调有 400 个音节左右，计声调也只有 1200 个左右，而荷兰语中大约有 12000 个音节。”建议增加引用。

**回应:** 我们已增加相应的参考文献(见 P2)。

**意见 6:** 前言部分“合适单元假设对于印欧语言的单词形式编码过程的阐述是符合 WEAVER 模型的观点的。”建议增加对 WEAVER 模型的解释。

回应：关于 WEAVER 模型在正文增加了详细的解释(见 P4)。

意见 7：对图 3 中右侧图的 100~200ms 结果进行更新。建议使用平均波幅的绝对值作为 Y 轴。

回应：非常感谢您的宝贵意见。我们的主要发现是音节(音素)同质条件比异质条件引起的波幅更正，在 300~400ms 的左前、中前区域，音节同质引起的平均波幅为正，而音节异质引起的平均波幅为负，如果画图取了绝对值，则无法看出在所有脑区上的音节效应，因而结果呈现用原始波幅的数据更能表现音节(音素)效应的大小。因此，我们仍然保留了原图。

.....

**审稿人 2 意见：**

意见 1：鉴于 Zhang 和 Damian(2019)已考察了音节效应和音段效应发生的时间进程，本研究的创新点尚需明确。

回应：非常感谢您所提的问题。审稿人 1 的第一个问题与您的类似，我们修改了前言部分，在指出已有研究可能存在的问题基础上，强调了研究中所采用的实验任务和被试的选择是本研究的特色，具体的阐述见文章第 5~6 页，同时请您阅读对审稿人 1 第一个问题的回复。

意见 2：引言中提到“汉语口语词汇产生中单词形式编码单元是音节还是音素的争论尤为引人注目。”该提法有待商榷，语音编码单元不是非此即彼的，不是在音节与音素之间选其一。

回应：感谢您的意见。我们将此句修改为：“研究者对汉语口语词汇产生中单词形式编码过程中音节和音素的作用存在不同看法”(见 P2)。

意见 3：音节条件和音素条件的反应时差异较大，是否表明两种条件下的实验材料不匹配。如不匹配，比较效应发生的时间进程是否有效？

回应：在内隐启动范式中，被试先学习几个词对，比如“海豹-夕阳，记忆-媳妇，皇帝-喜事，贝壳-细胞”。这四个单词对中的第二个单词第一个字的音节是相同的，称之为同质条件。对照组单词对的第二个单词在音韵上毫无关联，称之为异质条件。正式实验中共使用了 32 对双字词，根据目标词之间的关系分为 4 套音节相同(但声调不同)和 4 套首音素相同的词汇，每套包括 4 对词汇，分别构成 4 套音节同质条件和 4 套音素同质条件。4 套异质条件组是从 4 套同质条件中分别选择一个词对构成的(具体的实验材料见附录 1)。这样所形成的音节同质条件和异质条件、音素同质和异质条件下的词对是相同的，唯一的不同是目标词所处的音韵线索不同。

在内隐启动范式中，研究者一般不会去匹配线索词和目标词的词频和笔画数等，研究者主要关注的是不同音韵相关条件下，同质与异质相比是否存在显著差异。因为所用实验材料的不同，可能会出现音节条件和音素条件下反应时存在差异，但这并不影响对音节效应(音节同质-异质)和音素效应(音素同质-异质)的考察。大量的行为实验已经验证了这一实验逻辑的可靠性(Chen, Chen, & Dell, 2002; Cholin, Schiller & Level, 2004; Damian & Bowers, 2003; Meyer, 1990, 1991; 张清芳, 2008)。例如，Chen 等(2002)采用内隐启动范式，对汉语词汇产生过程进行研究，发现音节同质条件下的平均反应时在不同的音节和不同被试下分别为 658 ms(实验 2)，666 ms(实验 3)，662 ms(实验 4)，而音素同质条件下的反应时为 622 ms(实验 5)。张清芳(2008)采用同样的实验任务发现在音素同质平均反应时为 620 ms，而在音节同质的平均反应时为 633 ms。可以看到，平均反应时之间是存在较大差异的，且均表现为音节同质条件长于音素同质条件，与本实验的发现一致。研究者在做比较时并不关注不同材料之间的差异。这与本研究的发现一致，表明实验材料的选择确实会造成反应时上的差异，但

并不影响音节或音素效应的获得。

在实验中，为了匹配音节和音素条件中所用词对中线索词和目标词之间的语义联系度，我们招募了 21 名研究生，要求他们在 5 点量表上进行语义关联度的评定，1 代表完全无关，7 代表完全相关，结果发现音节材料( $M = 4.02, SD = 0.54$ )与音素材料( $M = 4.12, SD = 0.52$ )的语义关联度没有显著差异， $F(1,20)=1.97, p = 0.176$ 。这说明两种条件下的实验材料是相匹配的。基于内隐启动范式的逻辑，我们认为音节同质和异质的对比、音素同质和异质之间的对比是可行的。为了更清楚地说明这一点，我们修改了实验材料的介绍部分(见 P6)，增添了实验材料(见附录 1)，并增加了相应的讨论(见 P12~13)。

**意见 4:** 音节效应在反应时指标上不显著，很难解释。音节效应在汉语产生中是相对经典的、稳定的效应。

**回应:** 非常感谢您的宝贵意见。我们在讨论中探讨了可能影响音节效应的因素并尝试进行解释。第一，可能的原因之一是实验材料导致的，为了确认实验材料的选取是否合适。我们选择了 6 名大学本科生和研究生完成了相同的实验，结果表明，音节同质组的反应时显著快于音节异质组，而音素同质组与异质组之间无显著差异。实验结果与已有研究一致，这表明实验材料的选择是有效的。值得注意的是，已有的研究均选择了大学生或研究生作为被试，这些人日常生活中都会大量地接触英语材料，这在一定程度上表明二语水平可能会影响母语口语产生过程。事实上，本研究中的被试也表现出了音节同质组的反应时短于音节异质组，出现了微弱的促进效应。相比而言，研究生组对相同的实验材料和实验任务则出现了 20 ms 左右的促进效应，表明研究生组对于音节信息比本研究中的被试组更为敏感。第二，研究发现实验组产生目标词的平均潜伏期为 840 ms 左右，而材料测试组的平均潜伏期为 750 ms 左右，在内隐启动范式中，被试需要事先记忆线索词和目标词的联系，可能由于被试对词对的联系记忆不如材料测试组，当反应时延长时，被试在同质组和异质组产生的差异变得不显著。这一对比提示被试的教育程度以及日常与文字等材料接触的经验可能会影响讲话者的口语产生过程。关于二语水平和受教育程度对口语产生过程的影响机制需要进一步考察。

尽管行为结果未达到显著水平，但我们在 ERP 结果中发现了显著的音节启动效应和音素启动效应，表明虽然外显行为不显著，但即时的脑电指标反映出人们仍然先提取音节，再提取音素。已有汉语口语词汇产生的研究中，研究者也发现了行为指标无结果，而脑电指标有显著差异的模式。例如，Qu 等(2012)发表在 PNAS 的研究未发现音素相关和音素无关条件下反应时的差异，但是他们发现了脑电指标在特定的时间窗口内，音素相关和音素无关条件存在显著差异，因此他们认为在神经层面上，人们在汉语口语词汇产生过程中仍然表征了音素。

**意见 5:** 分析脑电数据时，行为学反应时小于 500 ms 的数据被删除，而脑电数据分析到 600ms 的时间窗口，如此，脑电结果是否包含一些反应时 500~600ms 的试次？建议分析脑电数据时，应删除反应时小于 600ms 的试次。

**回应:** 非常感谢您的宝贵意见。根据您的建议，我们在行为数据和脑电数据的分析中删除了反应时小于 600ms 的试次，约占总数据的 5.2%，并重新对波形进行了分析，结果与先前模式类似，具体的分析结果部分(见 P8~P11)。

**意见 6:** 在 500~600ms 的时间窗口内，未发现相关类型和相关条件的任何交互作用，做进一步的配对分析是否可行？

**回应:** 非常感谢您的宝贵意见。我们没有发现相关类型和相关条件的交互作用但仍然进行配对分析的理由是基于我们的研究假设，我们的关注重点在于分别在音节与音素条件下，比较

相关条件与无关条件引起的波幅差异是否显著，以此来探测音节与音素效应出现的时间进程。

最近，研究者认为即使在方差分析中没有发现交互作用，也是可以基于研究中所感兴趣的条件之间做配对分析的。例如,Howell(2010)在他的统计方法一书中指出：

"In the past, the general advice was that without a significant <ANOVA> group effect, individual comparisons were inappropriate. ... However, this is a case where the general advice is wrong. The logic behind most of our multiple comparison procedures does not require overall significance before making specific comparisons. ... requiring overall significance will actually change the FW <familywise error rate>, making the multiple comparisons tests unreasonably conservative. These tests were designed, and their significance levels established, without regard to the overall F."因此，我们在交互作用不显著的情况下做了进一步的条件之间的对比。

参考文献

Howell, D. C. (2010). *Statistical methods for psychology* (8th ed.). Belmont, CA: Cengage Wadsworth, pp. 372-373

**意见 7：**音节效应始于 100ms，相对较早。作者提出，这是因为“内隐启动任务中的产生过程不包括概念准备，是从词汇选择开始进行到发音。相比而言，内隐启动范式中的音韵编码过程会早于图画命名过程”。建议作者参考以往内隐启动范式中音节效应的时间进程，与本研究的发现做对照。

**回应：**感谢您的建议。遗憾的是，目前尚未有研究在内隐启动范式任务中采用 ERP 技术考察音节效应和音素效应的时间进程，因此无法进行直接比较。尽管如此，我们根据口语产生过程中图画命名任务和内隐启动任务中所包括的认知加工阶段不同，对较早的音节效应做了解释。

---

## 第二轮

**审稿人 1 意见：**作者回复了审稿意见，并做了相应修改。建议发表。

**回复：**衷心感谢审稿专家所提的意见及对修改稿的肯定。

**审稿人 2 意见：**本研究研究汉语普通话语言产生的基本单位，具有一定的理论意义。但是本研究存在的几个较大的问题，而作者在修改稿中并未很好地解决。鉴于解决该问题等于要重新执行实验(重新检查实验材料，重新招募被试)，本人建议可以直接退稿。如编辑部觉得可以再给修改机会，建议找第三位审稿人审议。以下是本人对该研究中的较大问题的总结，也供编辑部参考。

**意见 1：**由于实验设计的缺陷，致使最基本的效应未得到重复验证，很难让人相信作者研究中所衍生的其他结果。总的来说，内隐启动范式再语言产生研究中是比较经典的实验，使用该范式操纵音节变量，无论是在中文还是英文都较为稳定地发现了启动效应，然而作者在这个研究中则完全没有看到。这可能是作者实验设计不够严谨所致(异质性较高的被试样本，实验材料控制存在缺陷等)。因此这种基本效应都无法重复的研究，可能不一定适合在学报上发表。

**回复：**谢谢审稿专家的问题。根据实验设计，本研究中对所用实验材料进行了严格的删选，

测量并匹配了可能的混淆变量,保证目标词之间仅有音节相关或无关,音素相关或无关的变化。我们匹配了音节和音素条件中,同质组与异质组内线索词和目标词之间的语义联系度,以及音节和音素条件中同质组与异质组内目标词之间和线索词之间的语义关联度(具体指标见 P.7 的材料部分)。本实验的材料控制比已有研究更为严格,确保了选择的是在自变量水平上变化的实验材料。

从行为结果来看,我们未发现显著的音节促进效应。为了确认实验材料的选取是否合适,我们选择了 6 名未参加正式实验的研究生完成实验,结果发现 6 名学生都出现了音节启动效应,音节同质组( $M = 745\text{ms}$ ,  $SD = 153$ )比音节异质组( $M = 765\text{ms}$ ,  $SD = 157$ )的反应更快,  $t(5) = 3.10$ ,  $p = .022$ , 而未发现音素启动效应,  $p > 0.05$ 。从这一点来看,我们采用 6 名研究生就重复出了已有研究中的音节效应,表明实验材料的选择是有效的。尽管本研究中所发现的效应未达到显著水平,但能够观察到音节相关条件的反应时确实快于无关条件,呈现出微弱的促进效应。另一方面,已有有关言语产生的研究绝大多数选择了大学生或研究生作为被试,这些学生日常生活中都会大量地接触英语材料,这在一定程度上提示我们受教育程度和二语水平可能会影响母语口语产生过程。

尽管行为反应时上未发现显著的音节效应,但本研究的 ERP 结果发现在内隐启动范式中,讲话者对音节的提取先于音素。这一模式重复出了已有研究(Zhang & Damian, 2019)的模式。而且无论是音节效应还是音素效应,在波形上都表现为同质条件比异质条件更正,这与其他 ERP 的研究发现一致(Qu et al., 2012; Yu et al., 2014; Wong et al., 2016; Wang et al., 2017; Verdonchot, Tokimoto, & Miyaoka, 2019),反映了音节或音素相同时促进了口语产生过程,表明音节或音素是音韵编码的加工单元。根据 ERP 的结果,我们认为所发现的结果是可靠的,符合口语词汇产生过程的特点。在文章中,我们重点讨论了实验材料是否可靠、以及 ERP 结果与已有研究的相同点和不同点,并强调了本研究的特点和贡献。

**意见 2:** 实验结果较为意外,且较难解释。作者在音节操纵的实验中发现了之前该领域较少发现的 P2 效应,而且是启动相比于非启动有更大的 P2 效应。由于 P2 效应与音节编码之间关系并不密切(起码之前的相关研究基本很少报告),再者该效应与知觉、注意等心理过程密切相关,因此这个结果一定程度上似乎与作者的实验设计的缺陷(由于涉及的缺陷导致了额外的变量的涉入,从而导致基本的行为学效应消失)相互印证。而更加重要的是,作者该研究中最重要发现:音素效应,首先出现的时间窗较晚很难说清是否是由于实验任务导致,其次效应较弱,因此基于该结果无法对现有理论争议提供实质性的推动。

**回复:** 感谢审稿专家的问题及对口语产生中音节效应机制所提供的解释。本研究的实验设计与已有研究一致,而且实验材料的选择比已有研究更为严格,排除了混淆变量的影响(见对问题 1 的回复),所以我们认为实验设计是合理的。

从结果上来看,我们发现的音节效应所出现的时间窗口是刺激呈现后的 100-400 ms 之间,表现为音节相关条件比无关条件引起的波幅更正,这与已有研究中所发现的音节效应引发的 ERP 波形走向一致(J. Wang, Wong, S. Wang, & H.-C. Chen, 2017; Zhang & Damian, 2019)。同时早期 100~200ms 的音节效应与经典的 P2 和 P3 波形所出现的时间窗口相似(Hackley, Woldorff, & Hillyard, 1990; Donchin & Coles, 1988)。研究表明 P2 成分与视觉特征的探测有关,而且 P2 成分对正字法和音韵加工任务、语义分类任务和词汇分类任务中不同条件的设置比较敏感,会出现波幅上的差异(Luck & Hillyard, 1994)。研究也发现 P3 的波幅大小与注意的程度有关(Gray & Burgess, 2004),也可能与决策做出后的认知闭合(cognitive closure)(Desmedt, 1980; Verleger, 1998)或与有意识的信息通达过程相关(Picton, 1992)。这一成分包括了两个子成分(P3a 和 P3b),P3a 分布在前额叶部位,当有未预期的事件出现时引发这一成分,反映了自动的注意调节机制;P3b 分布在顶枕区域,与注意、工作记忆和复杂的认

知加工(包括语言加工)相关。这提示与音节异质条件相比,音节同质条件下可能诱发了被试更多的注意,促进了被试的口语产生过程。关于注意如何调节口语产生中的音韵编码过程以及波形的变化,则需要更进一步的研究。

音素效应出现在线索词呈现后的 500~600 ms 之间,发生在音节效应之后。根据 O'Seaghdha 等(2010)提出的模型(见文中图 1),在汉语口语产生过程中的音韵编码阶段,音节被提取后会进一步分解为各个音素,在此之后各个音素的提取是平行进行的(Roelofs, 2015)或者是以从左至右的顺序提取的(J.-Y. Chen et al., 2016)。同时人们也会提取相应的节律信息,形成单词形式的框架,将音素插入到相应的框架中去,进行语音编码为发音做好准备。音素效应可能发生在音韵编码或者语音编码阶段。第一种可能性是音素效应发生在音韵编码阶段,在音节提取之后将其分割为各个音素,仍处于音韵编码阶段。第二种可能性是音素效应发生在语音编码阶段,Zhang 和 Damian(2019)在图画命名过程中发现了 500~600 ms 时间窗口内的音段效应,与图画命名元分析的时间进程对比,他们认为这一效应可能出现在语音编码阶段,反映了音段的启动作用。本研究中发现的音素效应与已有研究一致,可以在汉语口语词汇产生的模型框架中得到合理的解释,同时提示研究者音素效应是较微弱的,探测到这一效应需要采取敏感的实验任务。

我们也注意到,口语词汇产生领域的 ERP 研究一般都采用实验条件与基线条件对比的设计,比较在各个时间窗口下两类条件是否达到了显著差异来推测各个变量起作用的时间进程。目前,相关的研究较少,随着研究的积累,期待研究者能够发现特定的加工过程与特定的成分有密切联系。针对于某一确定认知涵义的成分,会更有利于跨研究之间的比较。通过讨论可以看到本研究的 ERP 结果与已有研究模式相似,并且在汉语口语词汇产生的理论框架内可以得到合理的解释。

**审稿人 3 意见:** 汉语口语词汇产生中音节和音段的作用是一个重要的研究议题。本研究将内隐启动范式与 ERP 技术相结合,并且针对二语水平低的汉语母语者,有一定的创新性和理论贡献。但是本文需对研究结果的解释作进一步讨论。

**意见 1:** 内隐启动任务的产生过程中是否包含概念准备值得讨论。在 Cholin, Schiller & Level (2004)的研究中,每组目标词都来自同一个词根,例如 *leiden* ([to] lead), *leider* (leader), *leidend* (leading), *leidde* (led)都来自词根 *leid-*。在这种情况下,被试在准备阶段不仅能提前开始音韵编码,还可以提前进行概念准备。但是大多数内隐启动任务中(包括本研究),每组目标词包含不同的概念,被试无法在准备阶段预测哪一个概念将会出现。因此,在刺激出现后,被试应该仍需进行概念准备。

**回复:** 感谢您的宝贵意见。研究者采用内隐启动任务,选择音韵相关或无关词进行对比,主要的目的是为了考察音韵编码过程的加工机制。从实验程序的介绍可以看出,在每组测试中,被试仅需要针对四个线索词说出目标词,且在测试之前被试已经多次学习了这四个词对,能在线索词呈现时快速说出目标词。研究者认为这四个词在测试过程中都是从工作记忆中提取出来的,不涉及到概念的加工(Meyer, 1990, 1991; J.-Y. Chen et al., 2002)。相对于图片命名过程来说,内隐启动任务中不涉及到口语词汇产生中的概念准备阶段。原文的表述不太清楚,我们对文中相关的句子进行了修改(见 P.6)。

**意见 2:** Wong, Wang, Ng 和 Chen (2016)只考察了音节效应,并没有考察音段效应。

**回复:** 感谢您指出的问题。Wong 等(2016)的研究中设置了两种类型的音节相关,没有涉及到音段的相关。我们已经修改了文中的错误。



**意见 3:** 建议对每组材料中四对词语之间的语义关联度进行评分。例如, 音节异质组第 3 组中, 屁股-尾巴和剃须-胡子, 烧水-锅炉和食物-饭馆可能有较高关联度。如果同质组和异质组的语义关联度不匹配, 可能对结果的解释产生影响。

**回复:** 非常感谢您的宝贵意见。本实验的核心逻辑是比较分别在音节、音素材料中, 同质与异质组的反应结果是否有差异, 因此匹配同质组与异质组中词对的语义关联度是关键。为此, 我们从音节同质条件下的 4 套词汇中均选择 1 个词对, 形成 4 个目标词无任何语音联系的词对, 共组成 4 套音节异质条件。用同样的方法形成 4 套音素异质条件(实验材料见附录 1)。这样所形成的音节同质条件和异质条件、音素同质和异质条件下的词对是相同的, 唯一的不同是目标词的音韵线索不同。在后续的数据分析中, 我们比较的是整体音节异质组与音节同质组反应之间的差异, 因此从整体上来说, 同质与异质组之间词对的语义关联度是一样的。

尽管如此, 从不同的同质组和异质组来说, 它们所形成的小组中目标词和线索词之间的语义关联度可能是不同的, 以往的研究没有严格控制这一点。根据您的建议, 为了匹配音节和音素条件中同质组与异质组内目标词之间的语义关联度, 我们招募了 20 个研究生, 要求他们在 5 点量表上对两个目标词之间(例如“危机-围墙”、“危机-卫星”或“围墙-尾巴”等词对)的语义关联度评定, 1 代表完全无关, 5 代表完全相关, 结果发现在音节材料中, 两组目标词之间的语义关联度都很低, 且同质组( $M = 1.34, SD = 0.29$ )与异质组( $M = 1.32, SD = 0.35$ )的语义关联度没有显著差异,  $F(1, 19) = 0.177, p = 0.862$ 。在音素材料中, 同质组( $M = 1.28, SD = 0.29$ )与异质组( $M = 1.27, SD = 0.32$ )的语义关联度没有显著差异,  $F(1, 19) = 0.18, p = 0.79$ 。同时我们也对同质组与异质组中线索词之间的语义关联度进行了测量, 结果显示, 在音节材料中, 两组线索词之间的语义关联度都很低, 且同质组( $M = 1.41, SD = 0.36$ )与异质组( $M = 1.33, SD = 0.33$ )的语义关联度没有显著差异,  $F(1, 19) = 1.71, p = 0.103$ 。在音素材料中, 同质组( $M = 1.24, SD = 0.24$ )与异质组( $M = 1.31, SD = 0.27$ )的语义关联度没有显著差异,  $F(1, 19) = 1.32, p = 0.23$ 。上述结果表明同质组和异质组的语义关联度是匹配的。

相应的材料评定我们增加到了方法 2.2 材料部分(见 P.7)。

**意见 4:** 正文第 12 页“未参加正式实验的研究生(单位名称 XXX)”; 正文第 13 页“这一结果与已有采用掩蔽图画-词汇干扰实验范式的 ERP 结果一致(Zhang & Damian, 2019)”, 应改为“掩蔽启动范式”。

**回复:** 谢谢您的指正。在投稿中, 根据《心理学报》的匿名要求, 我们没有明确地标出是哪一个单位。如果稿件能被接受, 我们会在此添加参加者的单位名称。另外, 我们已经修改了关于“掩蔽启动范式”的错误(见 P.13)。

---

### 第三轮

**审稿人 3 意见:** 无进一步意见。

**编委意见:** 该文经多轮评审, 审稿人对论文提出许多尖锐的问题及建设性的建议, 作者已经根据审稿人提出的问题进行修改或回答, 两审稿人均认为可以发表, 一审稿人认为创新性略显不足。我自己也仔细阅读了文章, 觉得整个研究修改后质量有较大改善, 同时考虑到口语产生领域的研究, 尤其是汉语的研究仍然十分缺乏, 综合多方面因素, 我认为这一文章可以考虑接受。

**主编修改意见:** 本文经过几轮修改基本达到发表水准, 发表前只需对一些细节问题加以修改, 具体如下:

**意见 1:** 引言部分第一段末尾“音韵编码的加工单元一直是言语产生研究的争论焦点之一”后面需要给出代表性参考文献。

**回复:** 感谢您的宝贵意见。已根据您的要求在“音韵编码的加工单元一直是言语产生研究的争论焦点之一”后面加入了印欧语系和汉语中研究言语产生的音韵编码加工单元的代表性文献(见 P.1~2)。

**意见 2:** 引言部分第二段末尾“Dell(1986)认为音韵编码的单元可以是音节或音素，音节内部的音素可以同时被激活并插入音节框架；而 Levelt 等(1999)认为音韵编码的单元是音素，音素以序列的方式从左至右被插入音节框架。”这里与上面描述为重复语句，应当简洁总结”为重复语句，作者应当简洁总结。

**回复:** 非常感谢您的宝贵意见。这句话并非重复语句，意在表明两名研究者的观点有所不同，可能因为表述不清楚造成了误解，现将这句话改为“Dell(1986)认为音韵编码的单元可以是音节或音素，而 Levelt 等(1999)认为音韵编码的单元只是音素。两种观点的不同在于，前者认为音节内部的音素可以同时被激活并插入音节框架，而后者认为音素以序列的方式从左至右被插入音节框架。”(见 P.2)。

**意见 3:** 引言部分第三段前半部分将 Dell 和语误分析的研究作为“音素是音韵编码的加工单元”的支撑证据，这与第二段的文献综述似乎矛盾，根据第二段的表述，Dell(1986)认为音韵编码的单元可以是音节或音素，而 Levelt 等(1999)认为音韵编码的单元是音素。

**回复:** 非常感谢您的宝贵意见。Dell(1986)认为音韵编码的加工单元是音节或音素，即他认为二者都是口语产生过程中音韵编码阶段的加工单元，音节和音素都在此阶段起作用；而 Levelt 等(1999)认为音韵编码的单元只是音素，否认了音节的作用。因此用 Dell(1986)发现的音素替换现象的语误研究作为“音素是音韵编码的加工单元”的支撑证据与第二段的表述并不矛盾。

**意见 4:** 文中 2.5 程序部分的第 4 行“屏幕上呈现刚才被试学过的词对中的第一个词(线索词)”屏幕上呈现线索词的位置应该报告。

**回复:** 非常感谢您的宝贵意见。我们已按要求做了修改（见 P.8）。

**意见 5:** 文中 2.5 程序部分的最后一句“如果被试出现错误，就再进行一次学习，然后测试，直至被试准确熟练地记住全部词对。”有歧义。是一出错就对该词对进行再一次学习？还是完成全部测试后，对所有错误的词对进行再一次学习？请说明

**回复:** 非常感谢您的宝贵意见。正确的程序是完成全部测试后，对所有错误的词对再一次学习，然后再进行一次测试，直到被试准确熟练地记住全部词对。我们已在原文中进行了补充和修改（见 P.8）。