

《心理科学进展》审稿意见与作者回应

题目：言语产生中音节频率效应的认知机制：跨语言视角

作者：潘家冰 张清芳

第一轮

非常感谢您审阅我们的稿件和提出宝贵建议。基于两位审稿专家的意见，修改稿中对文章的整体框架进行了调整，补充和详细说明了阐述不清楚的段落，增加了图 1，希望修改稿能清晰地阐明印欧语系和汉语口语产生中音节频率效应的理论假设、表现形式、发生阶段和加工机制等。

我们针对两位审稿专家的意见做了逐条回复，用蓝色标记了修改部分，请您和专家审阅。

审稿专家 1 意见：

从九十年代至本世纪初，国内外学者对音节进行了系统考察，试图通过音节相关效应（如音节频率）揭示言语产生中语音加工的认知机制。正如作者所述（“近 5 年考察言语产生过程中音节频率效应的研究本身数量较少”），在言语产生领域中，音节频率不是近年来的研究热点。虽然不是当前研究热点，但学界对该问题的全面综述、讨论确有需要，尤其是作者从跨语言角度论述了音节效应在跨语言之间的差异与一致性，具有重要的理论贡献，建议修改后发表。

现提出如下建议，供作者思考与修改。

1、关于不同音节位置上的音节频率，作者介绍了不同研究结果之间的差异，并介绍了前人对不同研究之间存在差异的原因（即刺激材料的词汇性），但该因素不足以解释现有的研究结果，例如英语双音节词命名的研究同时发现了首音节和第二音节的频率效应。鉴于此，作者提出“不同位置上的音节是否表现出频率效应，除了存在跨语言差异之外，还可能由于个体在命名真词和假词时依赖不同的语音检索策略。”该解释非常有趣，建议作者展开论述。例如，英语、荷兰语、西班牙语的研究结果不同，这三种语言之间的哪些特点可能造成了研究结果不同？另外，语音检索策略的个体差异如何影响音节位置频率效应？

回复：非常感谢您的积极评价和宝贵建议。三种语言之间的哪些特点差异导致了研究结果的不同？我们认为可能不同语言产生过程中音节边界的清晰度影响了发音计划广度。Cholin

等(2011)提出音节边界的清晰性会影响不同语言的发音计划广度。与英语相比,荷兰语和西班牙语的音节边界相对清晰,讲话者能够较早地检测到音节边界,音节化过程以及通达心理音节表的速度也较快。在快速命名任务中,荷兰语和西班牙语的讲话者倾向于采用音节大小的发音计划广度;而英语中音节边界相对模糊,检测到音节边界较为困难,因此讲话者倾向于采用跨音节边界的、多音节的发音计划广度。

审稿专家提出的语音检索策略的个体差异如何影响音节位置频率效应,可能是我们文中的表达不清楚导致的。原稿中想要表达的是人们加工真词和假词两类刺激时可能采用了不同的策略。研究发现命名真词时是在两个音节都完成语音编码后才开始发音,而命名假词时在完成首音节的语音编码后就开始发音,即对首音节进行发音输出的同时提取第二音节,因此在快速的命名任务中第二音节对频率操纵不敏感。Cholin 等人(2011)提出上述研究之间讲话者采取不同语音检索策略的原因可能在于实验材料的词汇性。在真词口语产生中,人们会在整个单词完成语音编码后才开始发音以正确地传达单词意义,而对于假词没有传达意义的要求,人们只需要发音即可。

有关个体差异对语音检索策略的影响,目前尚未有研究涉及。修改稿中我们指出不同位置上的音节是否表现出频率效应,既与讲话者在命名真词和假词时依赖不同的语音检索策略有关,又与发音计划广度的跨语言差异有关。

2、关于音节频率的发生阶段,作者介绍了印欧语系中音节频率两种可能的发生阶段,第一,音韵编码和语音编码交互界面,第二,语音编码阶段。作者总结“以往大量研究证据可以得出,印欧语系中的音节频率效应发生在语音编码阶段”。如何得出该结论,如何排除“音韵编码和语音编码交互界面”这一可能性,需要进一步明确论述。

回复:原稿中关于“音韵编码和语音编码交互界面”的表达不清楚,我们想要表达的是印欧语系中的音节频率效应可能发生在音韵编码或(和)语音编码阶段。Laganaro 和 Alario(2006)的研究以及 Bürki 等人(2015)的研究明确得到了音节频率效应发生在语音编码阶段,但 Croot 等人(2017)的研究结果并不能排除音节频率效应发生在音韵编码阶段的可能性。

关于印欧语系中音节频率效应的发生阶段,修改稿中明确指出对于音节频率所表现出来的促进效应,已有研究对其发生在音韵编码还是语音编码阶段存在争论(见修改稿 P.5)。多数研究证据支持印欧语系中的音节频率效应发生在语音编码阶段,但目前也并不能排除音节频率效应发生在音韵编码阶段的可能性。音节频率的促进效应可能出现在音韵编码或(和)语音编码阶段(见修改稿 P.6)。

3、建议在最后一节着明确音节以及音节频率效应的跨语言差异（如发生阶段不同、音节效率表现等等），这将利于读者加工并记忆该综述的亮点。

回复：感谢审稿专家的建议。修改稿在展望部分进行了比较和总结。音节频率效应的跨语言差异主要体现在表现形式和发生阶段两个方面。在表现形式上，印欧语系语言中发现了音节频率促进效应、抑制效应或无任何效应，而汉语中则是以音节频率的促进效应为主。音节频率效应的方向与所操纵的频率指标(类别频率或象征频率)和任务密切相关。在发生阶段上，尽管目前无论是印欧语系还是汉语的研究对于音节频率效应的发生阶段均存在争议，但已有理论和研究证据多数支持印欧语系口语词汇产生中的音节频率效应发生在语音编码阶段，而汉语口语词汇产生中的音节频率效应可能发生在音韵编码和语音编码阶段。音节频率效应的跨语言差异可能与不同语言系统的特性有关，为音节在字母语言和汉语口语产生过程的不同加工阶段发挥作用提供了新的视角（见修改稿 P.13）。

4、一些细节：

a. 建议阐释相关的专业术语，以便增加相关内容的可理解性，如形式编码、后词汇编码水平、有发音抑制的延时命名任务、单纯的延时命名任务、音素、音段（图 1）等。

回复：感谢您的宝贵建议。我们通读了全文，对文中所涉及到的概念、任务等在相关段落中进行了详细描述。

（1）音节准备效应包含了单词形式编码的所有加工阶段(即词素音位编码、音韵编码和语音编码)，根据相加因素法的原理，两个因素之间存在交互作用表明二者可能作用于序列加工过程中相同的阶段，因此研究者认为该结果表明心理音节表可能位于整词通达之后的语音编码阶段。（2）延时命名任务并不要求被试在看到图片后立刻命名，而是只有当屏幕上出现提示符时才开始出声命名，在发音前被试已经完成了对目标词的音韵编码和语音编码，当提示线索出现时执行已经准备好的发音运动程序即可，因此该任务的命名潜伏期仅反映了发音阶段的加工。（3）有发音抑制的延时命名任务则要求被试在图片出现后一直出声重复某个特定语音(例如/da/)，当看到提示符时立即停止并完成图片命名。由于重复发音将占用讲话者的发音环路(articulatory loop, Baddeley et al., 1984)并阻断对目标词的语音编码，因此提示符出现后所测量的命名潜伏期反映了语音编码和发音两个阶段的加工。（4）音素(phoneme)是能够区分意义的最小语音单元，分为元音和辅音(Clark & Yallop, 1995)。（5）汉语的音韵编码过程中讲话者首先提取音节，再进一步将音节分解为音段信息(即音素或大

于音素而小于音节的音韵组合单元)和超音段信息(即声调), 然后完成语音编码和发音过程(见图 2)。

b. 有发音抑制的延时命名任务、单纯的延时命名任务有何不同, 如何用音节频率效应在两种任务上不同的结果做出相应的结论。其逻辑需阐明。

回复: 感谢审稿专家的建议。Laganaro 和 Alario(2006)采用即时命名任务、延时命名任务以及有发音抑制的延时命名任务来考察法语口语词汇产生中音节频率效应的发生阶段。在三种任务中, 即时命名任务包含从概念化到发音的全部言语产生过程; 延时命名任务并不要求被试在看到图片后立刻命名, 而是只有当屏幕上出现提示符时才开始出声命名, 在发音前被试已经完成了对目标词的音韵编码和语音编码, 当提示线索出现时执行已经准备好的发音运动程序即可, 因此该任务的命名潜伏期仅反映了发音阶段的加工; 有发音抑制的延时命名任务则要求被试在图片出现后一直出声重复某个特定语音(例如/da/), 当看到提示符时立即停止并完成图片命名。由于重复发音将占用讲话者的发音环路(articulatory loop, Baddeley et al., 1984)并阻断对目标词的语音编码, 因此提示符出现后所测量的命名潜伏期反映了语音编码和发音两个阶段的加工。研究者在即时命名任务以及有发音抑制的延时命名任务中均观察到了音节频率效应, 而在延时命名任务中未观察到。即时命名和有发音抑制的延时命名任务中均包括语音编码过程, 而延时命名任务中未包括, 任务之间的对比结果表明音节频率效应发生在语音编码阶段(见修改稿 P.5)。

c “研究者指出尽管可能在统计上存在二类错误”, 需阐明。

回复: 修改稿中删除了这一难以理解的描述, 指出可能是效应比较微弱而未能在有发音抑制的延时命名任务中检测到(见修改稿 P.5)

d. "尽管有研究发现音素在汉语口语产生的音韵编码阶段也会激活(Qu et al., 2012, 2020; Yu et al., 2014), 但由于汉语音素加工方面的证据相对较少且存在争议(屈青青 等, 2018), 因此已有研究主要证实了音节在汉语口语词汇产生中的重要作用。"分句之间的逻辑不清楚。

回复: 我们将此句修改为: “尽管有研究发现音素在汉语口语产生的音韵编码阶段也会激活(Qu et al., 2012, 2020; Yu et al., 2014), 但目前有关汉语音素加工方面的证据相对较少且存在争议(屈青青 等, 2018), 已有研究主要证实了音节在汉语音韵编码中的重要作用。”

e. 类别频率是指目标音节所拥有的全部音节邻近项的数量，与语音邻居大小(phonological neighborhood size)有何区别？

回复：语音邻居大小(phonological neighborhood size)指目标词所拥有的音韵邻近项的数量。与音节邻近项不同，它是指与目标词只存在一个音素差异的单词，这种差异包括音素的替换、添加或删除，例如“fat”、“kit”、“cab”、“at”、“scat”和“cats”都被认为是“cat”的音韵邻近项(Hameau et al., 2021)。

音节邻近项在字母语言中是指相同的位置上共享同一个音节的所有词，例如西班牙语词 casa、caro、caja 和 cama 等互为音节邻近项，它们共享首音节 ca。

Hameau, S., Biedermann, B., Robidoux, S., & Nickels, L. (2021). Effects of phonological neighbourhood density and frequency in picture naming. *Journal of Memory and Language*, 120, 104248.

f. 请确定文中是否有对图 1 的介绍。

回复：感谢审稿专家的提醒。修改稿中补充了一张新的图 1，原稿中的图 1 调整为图 2。修改稿中在“4.1 音节在汉语口语词汇产生中的作用”部分介绍了图 2（见修改稿 P.11）。图 2 主要描述了汉语单音节词汇的口语产生过程，在音韵编码过程中讲话者会首先提取音节，再进一步将音节分解为音段信息(即音素或大于音素而小于音节的音韵组合单元)和超音段信息(即声调)，然后完成语音编码和发音过程(见图 2)。

.....
审稿专家 2 意见：

本综述聚焦口语词汇产生过程中音节频率效应，重点从其表现形式（促进或抑制）和发生阶段（音韵编码和语音编码）阐述了音节频率效应的跨语言差异。本文聚焦的问题新颖，有较强的科学意义。作者应更好地组织和表达重要观点，从而让读者可以更清晰和容易地理解其观点。下面是我的一些具体建议，希望可以帮到作者精进本文。

1、“在早期词汇通达阶段，高音节频率词由于拥有更多的音节邻近项，能够激活更大的初始候选词群，且该词群更有可能包含高频词，这些激活的词汇节点之间相互竞争并干扰目标词的通达，故表现为抑制效应；而在晚期语音输出阶段，由于高频音节的发音运动程序能够更快地从心理音节表中提取出来，使讲话者在构建高音节频率词的语音输出时更加容易，故表现为促进效应。”这个是本文的最主要假设之一，除了在之后引用的一些间接证据外，下

面两篇文章直接操控两种不同层级的频率，会作为更直接的证据来支持次观点【Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological science*, 9(4), 325-329.; Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of memory and language*, 40(3), 374-408.】另如果按这个假设汉语的效应在早期，则预测结果应该是频率效应在中文中应该是更偏向抑制？

回复：感谢审稿专家的建议。审稿专家推荐的文献中主要操纵了音位配列概率和相似性邻域密度，结果发现在出声朗读单词的任务中，音位配列概率(probabilistic phonotactics, 即音段和音段序列的频率)促进了亚词汇水平的加工，而相似性邻域密度(similarity-neighborhood density, 即音韵邻近项累计词频)产生了抑制效应，这可能是由词汇水平上音韵邻近项(即由一个音素替换、添加或删除得到的词)之间竞争导致的(Vitevitch & Luce, 1998; 1999)。因此，在言语产生过程中与音韵信息相关的效应中，促进效应可能反映了词汇和亚词汇水平上由音韵表征本身特性导致的加工优势，而抑制效应则源于词汇水平上激活词条之间的竞争（见修改稿 P.4）。

原稿中关于言语产生早期、晚期阶段的表述不够清楚，修改稿中不再区分“早期”和“晚期”，而是明确地指出效应可能发生的阶段（例如词汇选择、音韵编码或语音编码）。汉语音节频率效应表现为促进效应还是抑制效应，取决于词汇选择阶段的抑制效应与音韵编码/语音编码阶段促进效应的相对强度大小。目前尚未明确音节频率对汉语词汇选择阶段产生影响的程度如何，已有汉语图画命名任务中音节频率效应均表现为促进效应。修改稿中我们尝试提出了一个有关汉语口语词汇产生中音节作用机制的模型(图 1b)，希望能清晰地表达汉语口语产生中音节起作用的认知机制。

2、“语音编码的双通路模型 (dual-route model, Varley & Whiteside, 2001; Whiteside & Varley, 1998)也指出存在两种通路将音韵音节转化为语音音节：高频音节通过直接通路从心理音节表中进行提取，低频音节通过间接通路一个音段接一个音段地在线组装而成(同见 Cholin, 2008; Cholin & Levelt, 2009; Hagoort et al., 1999)”。文中相应位置时提到这个模型的目的是什么？这个模型（或是假设）和文中之前提到的音节频率效应在早、晚阶段的关系是什么？应更好的把这个语音编码的双通路和之前相关的理论叙述衔接，更好的融合到全文和观点中。

回复：感谢审稿专家的建议。介绍双通路模型是因为研究者对于口语产生中音节频率促进效应的发生机制存在争论，目前存在心理音节表理论、混合模型和双通路模型三种理论模型。其中，混合模型和双通路模型是对心理音节表理论的补充和发展，同样预测讲话者对高音节频率词的命名快于低频音节词，表现为促进效应，但促进效应出现的认知机制是不同的，即高频音节从心理音节表中直接提取，而低频音节通过在线计算或组装而来。根据审稿专家的建议，修改稿中阐明了这三个理论模型之间的关系以及模型发展的脉络（见修改稿 P.7-8）。

上述三个理论模型解释的是音韵编码和语音编码阶段中音节频率促进效应的认知机制。修改稿中不再强调早、晚期阶段，而是直接指明抑制或促进效应发生的具体阶段。此外，修改稿中调整了文章结构，在介绍理论观点之前首先阐明了音节频率效应的表现形式和发生阶段，然后在此基础上进一步阐述了有关音节频率促进效应的相关理论模型。

3、本文开始主要阐述语音频率效应的处理过程及理论【早或晚阶段、促进或抑制】，但中间插了一部分讨论影响音节频率效应的因素，实验和实际操控。最好把理论和实际因素更加清晰的分开表达，并明确为什么要在讨论理论中加一段实际因素的讨论，这样就可以从这两个方面更有机的探讨音节频率效应。

回复：感谢审稿专家的建议，在修改稿中我们调整了文章的整体结构，分开阐述了影响因素和理论观点。在“2 音节频率的测量：类别频率和象征频率”中阐述了音节频率的两种测量方式；针对印欧语系语言中的音节频率效应，我们在“3.1 音节频率效应的表现形式和发生阶段”中阐述了音节频率效应中有关抑制效应和促进效应的认知机制，并介绍了可能影响音节频率效应的因素，比如音节位置等对印欧语系言语产生中音节频率效应的影响；在“3.2 音节频率促进效应的认知机制：理论观点”中阐述针对音节频率促进效应的三种理论模型。

4、后来的‘音节频率效应的发生阶段’在之前 2 中已经提过了，显得重复。可以把两个部分更好的整合，减少冗余，更清晰的在一处表达讨论。

回复：感谢审稿专家的建议。修改稿分别针对印欧语系（见 3.1）和汉语（见 4.2），将音节频率效应的表现形式和发生阶段整合在一处阐述。

5、“对于汉语母语者来说，将音节存储起来并在音韵编码阶段直接提取是更为高效的加工方式(蒋宇宸等，2020；张清芳，王雪娇，2020)，心理音节表理论也尤为适合解释汉语口语产生中的音节加工过程(Levelt e al., 1999)”，这与文中之前提出的汉语在早期阶段的理论假

设的关系是什么？

回复：感谢审稿专家的问题。原稿中提到的汉语效应发生在“早期”阶段即为音韵编码阶段，修改稿中我们删除了有关“早期”和“晚期”的描述，在讨论具体效应时直接描述其可能发生在哪个具体的阶段。此外，“心理音节表”这一概念主要是针对印欧语系语言的研究提出来的，汉语中在语音编码阶段是否存在“心理音节表”尚需研究证实。修改稿中删除了这一部分。

6、“合适编码单元(proximate units)是指词素信息被激活之后最先选择的音韵加工单元，印欧语系中的合适编码单元是音素，汉语中则是音节。”这个假设又给汉语和印欧语系中音节频率效应提供了另一种可能的原因（加上之前的早、晚阶段区别，和心理音节表的可能贡献）。作者应该把所综述的所有可能假设，系统的综合和对比，这样可以给读者更清晰的理解。比如这段标蓝表示作者自己的观点的话：“由于汉语单字词可作为一个表征意义的独立语素而存在(邹丽娟, 舒华, 2013)，因此在词汇选择过程中受到音节邻近项的干扰可能相对较弱。此外，不同于字母语言中的双音节词或多音节词，汉语单字词一旦通达了词汇水平的表征，就可以顺利提取出相应音节的发音运动程序并进一步完成整词的发音，而不存在首音节表征的激活反馈到词汇水平并激活大量首音节邻近项的情况，故汉语音节频率效应受到早期词汇通达阶段的影响可能相对较小。”这个观点是如何受到文中之前提到的早晚阶段、心理音节表，及合适编码单元三种假设的影响和启发的呢？作者可以更直接、明确的表达出来。

回复：感谢审稿专家的建议。修改稿中新增了图 1a 和 1b，系统地综合和对比印欧语系和汉语关于音节频率效应的理论假设、发生阶段和加工机制等。根据已有研究发现，我们提出了一个有关汉语口语词汇产生中音节作用机制的模型（见图 1b），详细描述了汉语音节频率效应的发生阶段、加工机制和表现形式（见修改稿 P.13）。

原稿中蓝色标记的段落目的是阐述词汇选择过程中音节邻近项之间存在激活和竞争机制，基于汉语不同于字母语言的特点，预测汉语音节频率效应受到词汇选择阶段中邻近项之间相互抑制的影响可能相对较小，由此解释汉语音节频率的促进效应。修改稿中我们根据已有汉语研究中音节频率的促进和抑制效应，尝试提出了可能的认知机制。修改稿中删除了这一段。

7、“音节的神经表征既存在跨语言差异，也具有跨语言间的一致性。一方面，针对汉语的

研究发现双侧颞上回(superior temporal gyrus, STG, Yu et al., 2015; Zhang et al., 2016)以及左侧额中回(Siok et al., 2003; Tan et al., 2005)参与音节加工。作为一种符号语言, 汉语的正字法表征可以直接映射到音节上(Tan et al., 2005; Yang & Tan, 2020; Yu et al., 2015)且具有词素音节的特点, 每个音节可对应于独立的词素(Burchfield & Bradlow, 2014; Long et al., 2022), 汉字阅读过程不存在字母语言中从字母到音素的形音对应规则(Siok et al., 2003, 2004; Tan et al., 2005)。鉴于汉字的符号特点和单音节性质, 有研究者指出 STG 可能参与了汉字命名过程中将正字法表征转换为音节表征的过程(Lee et al., 2004); 类似地, 左侧额中皮层(或左背侧额叶系统)也与汉字的视觉空间分析以及从正字法到音节的映射有关(见 Siok et al., 2003; Tan et al., 2005)。另一方面, 基于印欧语系的研究也证实了 STG 在音节加工中的重要性(Bhaya-Grossman & Chang, 2022; Oganian & Chang, 2019; Scott, 2019), 有研究发现 STG 对包含多音节的语音序列高度敏感而对辅音簇不敏感(McGettigan et al., 2011), 关于 STG 时间敏感性的研究也表明神经活动对语音的追踪出现在单词/音节而非音素的时间尺度上, 表明 STG 对音节序列结构更敏感(Keitel et al., 2018)。因此, 有研究者指出 STG 的语音表征中更为重要的是由音素组成的、在音节水平上进行表征的语音序列, 而非单一的音素(Scott, 2019)。”这一段均是言语感知方面(perception)的研究, 其与所讨论的言语产生有什么关系? 需要直接及明确的指出。

回复: 感谢审稿专家的建议。修改稿中从神经机制的层面指出了言语感知和言语产生之间的联系。在神经机制层面, 有关音节的研究多数集中在言语感知方面, 对词汇产生中音节编码的神经机制研究较少。在语言加工中, 言语感知和言语产生之间存在密切联系, 两类过程之间会产生相互影响(Pickering & Garrod, 2013)。研究发现两类加工过程中会使用相同的加工单元, 例如, 研究者发现汉语的词汇理解和词汇产生过程中基本的加工单元是音节(Lin et al., 2022; Sereno & Lee, 2015; Yang & Chen, 2022), 其神经表征可能是类似的。根据已有研究发现, STG 可能是负责音节加工的特别脑区, 具有跨语言的一致性。然而, 根据音节在汉语和印欧语系语言的口语词汇产生中不同的认知机制, 这一脑区在两类语言产生过程中所起的作用可能完全不同。目前考察言语产生中音节加工神经机制的研究较少, 研究者要重点关注音节频率效应是如何体现音节神经表征的跨语言差异性和普遍性的, 为言语产生中音节频率效应的跨语言比较提供新的视角(见修改稿 P.14-15)。

8、本文包括了众多理论、假设、阶段、效应表现形式、多语言对比等, 信息多且乱, 需要一到两张图, 更清晰和有效的表达本文的主要观点。

回复：感谢审稿专家的建议。修改稿中补充了“图 1 (a) 印欧语系和 (b) 汉语口语词汇产生过程中的音节频率效应”，系统对比了印欧语系和汉语关于音节频率效应的理论假设、表现形式、发生阶段和加工机制等。

第二轮

审稿专家 1 意见：

作者按照审稿意见进行了修改，建议发表。

审稿专家 2 意见：

十分感谢作者的努力，修改稿已全部回答了我的问题

编委 1 复审意见：

二位审稿人非常专业，提出的问题很有水平。作者对问题作了认真考虑回复，并对文章作了大幅修改，使文章质量有很大提升。同意发表。

编委 2 复审意见：

有两个建议

1. 象征频率(token frequency) 这个翻译我觉得一定不合适，没有“象征”的意思，简单百度了一下，一般是例频(token frequency)和型频(type frequency)，实例频率(token frequency)，另一类叫类型频率(type frequency)。

回复：感谢您的宝贵建议。修改稿中采用了“实例频率(token frequency)”和“类型频率(type frequency)”的表述。

2. 整体篇幅挺长的，部分可以缩减，比如最后很长一段是讨论语言感知，和文章无关建议删去。“在神经机制层面，有关音节的研究多数集中在言语感知方面.....”

回复：感谢您的宝贵建议。修改稿已将展望部分的最后一段有关言语感知方面的内容删去。

.....

主编意见：根据编委和审稿专家的意见，建议发表。