

汉语口语词汇产生中音素的作用阶段

朱嘉睿^{1*}, 赵黎明^{1,2,3#}

¹天津师范大学心理学部, 天津

²天津师范大学心理与行为研究院, 教育部人文社会科学重点研究基地, 天津

³学生心理发展与学习天津市高校社会科学实验室, 天津

收稿日期: 2022年2月25日; 录用日期: 2022年3月22日; 发布日期: 2022年3月30日

摘要

在汉语言产生中, 大多数研究聚焦于音节的发生阶段和作用大小, 认为其在音韵编码阶段起重要作用, 但目前关于音素的加工机制还不清楚, 对音素的作用阶段存在争议。本文将总结国内外相关研究, 提供音素可能作用于音韵编码的前期或后期, 或/和语音编码阶段的证据, 同时探讨音素在各阶段的效应大小和方向。最后基于研究现状, 从研究方法上, 如改进实验范式、结合其他技术手段等方面进一步探讨音素的作用机制, 为未来研究提供建议。对音素的探讨将有助于深化对汉语言产生的认识, 了解印欧语系语言与汉语在言语产生机制上的异同, 为建立汉语言产生模型提供基础。

关键词

口语词汇产生, 汉语, 音素

The Functional Stages of Phonemes in the Production of Spoken Chinese Words

Jiarui Zhu^{1*}, Liming Zhao^{1,2,3#}

¹Faculty of Psychology, Tianjin Normal University, Tianjin

²Key Research Base of Humanities and Social Sciences of the Ministry of Education, Academy of Psychology and Behavior, Tianjin Normal University, Tianjin

³Tianjin Social Science Laboratory of Students' Mental Development and Learning, Tianjin

Received: Feb. 25th, 2022; accepted: Mar. 22nd, 2022; published: Mar. 30th, 2022

*第一作者。

#通讯作者。

Abstract

In spoken Chinese production, investigations mainly focus on the occurrence stage and size of syllables, which are believed to play an important role in the phonological encoding stage. However, the processing mechanism of phonemes is still unclear, and the functional stages of phonemes are controversial. This article will summarize relevant research at home and abroad, provide evidence that phonemes may play a role in the early or late stages of phonological encoding, or/and phonetic encoding, and discuss the magnitude and direction of the effects of phonemes at each stage. Finally, based on the current research situation, from the perspective of methods, such as improving the experimental paradigm and combining other technical means, further explore the mechanism of phonemes, and provide suggestions for future research. The discussion of phonemes will help deepen the understanding of Chinese language production, understand the similarities and differences between Indo-European languages and Chinese in the speech production mechanism, and provide a basis for establishing a Chinese language production model.

Keywords

Spoken Word Production, Chinese, Phoneme

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

言语产生是将思想转化为言语表达的认知加工过程,涉及音韵表征的提取以及将其转换为发音命令,主要包括概念准备、词条选择、单词形式编码和发音阶段[1] [2]。根据 Levelt 等人[2]提出的言语产生模型,单词形式编码(Word-Form Encoding)是指从心理词典中提取语素,再进一步划分为音韵片段和节律框架,之后结合形成音韵音节并进行音节化的过程。单词形式编码又可进一步细分为词素音位编码(Morphophonological Encoding)、音韵编码(Phonological Encoding)和语音编码(Phonetic Encoding)三过程。在词素音位编码过程,讲话者根据所选择的词汇进一步选择相应的词素和句法特征;在音韵编码过程,根据词素选择音段和节律信息,并进行音节化,将音段和节律信息中的音节节点相联系;在语音编码过程,选择相应的音节程序节点为发音做好准备[1] [2] [3]。

近年来,研究者对单词形式编码进行研究,在这一阶段中音素和音节的作用,尤其是在音韵编码和语音编码过程中的作用受到了广泛关注[4] [5] [6] [7]。音素(phoneme)是根据语音的自然属性划分出来的最小语音单元,也是构成音节的最小单元或最小的语音片段,其按照发音特质可分为元音(Vowel,简称V)和辅音(Consonant,简称C)两大类[8]。在汉语中,普通话语音系统的音素可以分为声母和韵母两大类,韵母至少包括一个元音音素,声母大部分由辅音音素组成,例如,/hui2/,包含辅音音素/h/和两个元音音素/u/和/i/,两个元音音素构成一个韵母/ui/。在语音分析中,音节(syllable)被定义为每种语言都具有的一个单位,是辅音和元音组成的序列。所有语言都有CV和CVV音节结构,例如在汉语中/ha1/是一个CV结构的音节,其中/h/是辅音,/a/是元音,而/hai4/是一个CVV结构的音节,包括一个辅音/h/和两个元音/a/与/i/。

Dell 等人[1]和 Levelt 等人[2]提出的言语产生模型具有重要影响,两大模型均认为音素是音韵编码的

重要功能单元, 在音韵编码的前期阶段率先被提取, 之后结合相应的节律框架形成音节化的音韵词, 进行语音编码和发音。模型的不同之处在于, 前者认为音节内部的音素可以同时激活并插入音节框架, 而后者认为音素以序列递增的方式从左至右的插入音节框架。需要注意的是, 两大模型均是在英语与荷兰语等印欧语系语言的基础上建立的, 汉语作为一种非字母语言具有不同于印欧语系这种字母语言的特点, 具体而言, 字母语言需要在音韵编码阶段通过加工刺激的重音、音素等信息确定其音节结构, 音节边界模糊, 需要重新音节化; 而汉语的音节数量少, 音节边界清晰, 不存在重新音节化的现象, 这使得汉语母语者在音韵编码阶段可以直接提取刺激的音节信息[5]。因此, 印欧语系言语中对音素的解释未必适用于汉语, 在考察音素的作用阶段时必须结合汉语自身的特点进行研究。

另外, Roelofs 等人[9]提出的言语产生模型认为, 在口语词汇产生的进程中, 在选择相应词汇后, 音韵编码阶段率先提取的是无声调音节和音节对应的音调信息, 之后音素信息才被提取。该模型认为音节和音素启动是串行的, 音素在音节之后进行加工, 音素在音韵编码的后期阶段起作用。Roelofs 等人的模型中对于音素的作用阶段的解释不同于 Dell 等人和 Levelt 等人的模型, 探讨哪一模型更符合汉语中言语产生的过程这一问题离不开对音素作用阶段这一问题的研究。对音素作用阶段的探讨将有助于修订汉语的言语产生模型, 不能宽泛地认为音素就是在音韵编码阶段起作用, 要进一步细化研究; 同时, 探讨音素的作用阶段这也将有助于验证合适单元假设[5]。合适单元假设认为音韵编码功能单元具有跨语言特异性, 音节或音素等语音单元在不同语言中所起的作用不同, 在汉语中, 音节是储存而非临场形成的, 且在音韵编码阶段率先提取, 之后分解为音素进行加工。对音素作用阶段的研究也可从侧面证明是否音节是汉语言语产生的合适单元。

在之前的研究中, Wong 和 Chen [10]发现音节内任意两个连续或不连续音素重叠时, 均会产生启动效应, 那么研究所发现的音韵编码前期阶段音节结构的启动效应是否不是因为音节结构的匹配, 而是因为音节内音韵内容(如音素)的匹配所导致的呢。Wong 等人[11]的实验中, CV 音节结构是潜在的形式准备成分, 对于该音节结构内的完全或部分重叠的音韵片段进行编码时, 毫无疑问, 这个重叠的音韵片段(如音素)也会产生一定的准备效应, 这就造成有关音节的研究中, 无法说明启动效应到底是由于音节结构还是音韵内容的重叠造成的。如果启动效应是由于音素这个音韵内容的重叠所造成的, 那么音素可能在音韵编码前期就开始起作用了。另外, 一部分研究表明在音韵编码过程中无声调音节与节律框架系列结合, 作为词汇水平下最先选择的加工单元, 音节被提取之后会进一步分解成音素/音段, 以平行或序列的方式提取[5] [9] [12], 这说明音素可能在音韵编码的后期, 甚至更晚的加工阶段, 如语音编码阶段发挥作用。

综上所述, 音素在口语词汇产生中发挥着重要作用, 但对音素的研究尚浅, 涉及音素的研究也大多是为了与音节作对比, 重点探讨音节和音素的相对时间早晚, 并且把侧重点放在了证明音节在音韵编码阶段起重要作用这个观点的探讨上。而对于音素效应的所在阶段、以及其在哪一阶段所起的相对重要作用(比如效应大小、效应方向)这方面的研究有所欠缺。根据以往的研究, 音素效应的作用阶段存在争议, 它可能在音韵编码或/和语音编码阶段起相对重要作用, 但具体在哪一阶段是没有定论的。因此, 本文从言语产生中音素的作用阶段入手, 深入探讨音素效应是在音韵编码还是语音编码阶段起重要作用, 对该问题的探讨有助于深化对汉语言语产生的认识, 完善现有的汉语言语产生模型, 揭示音素是否具有跨语言的特异性, 为未来有关音素和音节效应的相关研究提供参考。

2. 音素的作用阶段的相关研究

在汉语口语产生中, 音素效应的发生阶段难以确定, 音素在音韵编码或语音编码阶段发挥着特定的作用, 而合适单元假设只涉及音韵编码阶段, 而忽略了语音编码阶段。汉语口语产生中的音素激活是微弱的, 根据行为实验很难定位音素效应的发生阶段, 近年来有关音素效应的相关研究大多结合时间分辨

率较高的 ERP 技术进行探讨, 以此来确定音素效应的时间进程。其次, 也结合诸如 fMRI、TMS 等空间分辨率更高的技术来探讨言语产生各个阶段涉及的不同脑区激活, 从这个角度探讨音素效应的发生阶段。

2.1. 音素作用于音韵编码阶段的证据

2.1.1. 音素作用于音韵编码前期

在印欧语系语言中, 语误分析[13][14]的结果表明大多数的音韵错误涉及音素的插入、删除、替换和交换(例如, York library→lorck yibrary, reading list→leading list), 很少存在涉及整个音节的音韵错误(例如, napkin→kinnap); 反应时分析也表明音素更可能是口语产生的功能单元, 支持音素在音韵编码的前期起重要作用(内隐启动范式: Dutch: [15]; French: [16]; English: [17][18]; 形容词-名词命名任务: [19][20][21]; 图-词干扰范式: [22][23]; 掩蔽启动范式: [7][24][25])。

而在汉语中, 研究者对于音韵编码过程中音节和音素的作用存在争议。一部分研究认为音节是音韵编码的功能单元, 在音韵编码的前期阶段起主要作用[12][26][27]。但 Verdonschot 等人[28]发现, 音节不是汉语普通话初始音韵建构中必不可少的组成部分, 音节和亚音节都是起作用的, 一些粤语研究也证明了音素等亚音节在汉语口语词汇产生中可以作为功能单元被提取, 这表明音素也可能在音韵编码的前期起作用(形式启动任务: [11]; 掩蔽启动范式: [28][29]; 图-词干扰范式: [10][30][31])。有关汉语普通话的电生理研究也揭示了音素重叠和不重叠条件之间的差异。例如, Qu 等人[21]采用有颜色图片的命名任务, 要求被试使用“形容词+名词”的短语形式对有颜色图片进行出声命名, 结果在图片呈现后的 200~300 ms 内发现了显著的音素重叠效应, 结合 Indefrey [32]中关于音韵编码阶段的时间进程的估计(275~455 ms), 证明音素在音韵编码的前期阶段起重要作用。同样地, Yu 等人[33]采用首音素重叠范式结合图片命名任务, 发现无论音素所处的词内位置如何, 都在音韵编码的前期阶段起重要作用, 支持音素在汉语中有重要的功能参与。另外, Qu 等人[34]使用内隐启动范式结合 ERP 技术, 操纵首音素和首音节重叠条件, 要求被试对同源组和异源组的图片对命名(例如, 同源组: 蘑菇, /mo2gu1/, mushroom-拇指, /mu3zhi3/, thumb-蜜蜂/mi4feng1/, bee-马桶, /ma3tong3/, toilet; 异源组: 书包, /shu1bao1/, schoolbag-插头, /cha1tou2/, plug-开关, /kai1guan1/, switch-蜜蜂, /mi4feng1/, bee)。在传统的内隐启动范式中, 被试需要先学习几个词对, 待被试熟悉所有词对后进入测验阶段, 即呈现单词对中的第一个词(线索词), 要求被试在呈现线索词的同时说出对应的第二个词(目标词)。每组词对中的目标词之间存在音韵相关为同源条件, 目标词之间不存在音韵相关为异源条件。Qu 等人对这一范式进行改进, 采用图片命名代替词汇命名, 排除了注意对启动效应的影响[35], 结果发现音节和音素效应具有相似的时间进程, 因此研究者认为音节和音素在音韵编码的前期阶段共同作用。

值得注意的是, 有关音素启动效应的研究中行为结果很少显著, 和 ERP 结果显示出不对称性[36][37]。Chen 等人[12]认为干扰项和目标项之间单一的首音素重叠可能无法引出可在行为数据中观察到的足够的激活[21][33]。涉及到多个音素重叠条件的研究在行为结果上发现了显著的音素重叠效应[10][11][27][28][29][30][31]。因此我们推测, 单个音素确实被激活了, 但这种激活不够强, 不足以引起在较不敏感的指标(如命名反应时)中可以被注意到的变化。其次, Qu 等人[38]提出自我监测机制可能会抵消音韵编码过程中产生的基于音素/音段的微弱的促进作用。最初可以通过重复提取相同的音素来促进命名过程, 但是随后由于增加的语误概率, 导致认知工作量增加, 从而阻碍了命名过程, 导致行为指标上未发现相应的变化。

2.1.2. 音素作用于音韵编码后期

前文可知, 汉语和印欧语系语言口语产生过程是不同的, 汉语母语者在音韵编码阶段可以率先提取

词的音节信息, 这样更加经济。结合 O'Seaghdha 等人[5]提出的合适单元假设, 我们推测音素在音韵编码阶段起作用可能不是作为功能单元, 音素效应可能反映的是两个具有相同起始部分(如首辅音)的词的共同激活, 是通过音节扩散到对应音素的间接激活[9] [39] [40]。另外, 我们观察到的较早的音素启动效应也可能是源于“词首效应”。词首效应是指词首这个特定位置上的字母在言语产生中的特殊作用, 首音素比词内其他位置的音素更容易产生语误[1] [41]。

Cai 等人[37]使用掩蔽启动范式探讨音节和音素效应的相对时间进程, 在该范式中要求被试对目标图片命名, 启动词的呈现在目标图片呈现之前, 持续时间短暂, 并且启动词呈现前后伴有掩蔽图形, 被试对启动词的加工处于无意识水平。结果显示音素启动晚于音节启动效应, 出现在图片呈现后 400~440 ms, 且相比于音素不重叠条件, 在音素重叠条件下明显缩短了与音韵编码阶段相关联的 Map7 的持续时间, 起始潜伏期的分析也表明了汉语口语产生中音节(284 ms)和首音素(396 ms)的先后时间进程。因此, 研究者认为音节作为加工单元先于音素被提取, 音素在音韵编码的后期阶段起促进作用。同样, Qu 等人[34]使用内隐启动范式, 发现音素效应存在于两个不同的时间窗(230~290 ms 和 350~400 ms), 均表现为同源组的平均波形小于异源组的平均波形, 两个音素效应的时间窗均大致符合 Indefrey [32]元分析中音韵编码阶段的时间窗范围(200~455 ms)。

需要注意的是, 该研究在音韵编码后期阶段发现的是抑制而非促进作用, 音素重叠的抑制效应可能反映的是音韵编码内部对音素的自我监测。自我监测涉及内部关于音韵编码的监测环路和外部对于自行生成的口语的环路[42]。Qu 等人[21]的研究中, 在 200~300 ms 的靠前时间窗内, 音素重叠相比于不重叠条件, 在脑后区表现出更大的正波, 但在 300~400 ms 的靠后时间窗, 在脑前区则表现出更小的正波。他们认为靠前时间窗内的 ERP 波形反映的是音素重叠在音韵编码阶段的促进作用, 表示两个单词在重叠条件下存在相互的形式启动; 而靠后时间窗内的 ERP 波形反映的是音韵编码阶段内部的自我监测。内部的自我监测的时间进程大约是刺激呈现后的 355 ms, 这与该研究中 300~400 ms 时间窗内发现的音素效应保持一致。自我监测的一个基本机制是在言语产生的计划过程中, 讲话者判断是否合适的音素被提取以及正确排列[1]。因此在 Qu 等人的研究中, 讲话者口语产生形容词 - 名词短语, 相比于音素无关条件, 在音素相关条件下检测系统是在较高的负载下去防止产生言语错误, 从而导致了更小的正波。随后, Yu 等人[33]的研究也发现了类似结果, 靠后时间窗内的音素效应被解释为较强的言语自我监控所引起的抑制效应, 这也与之前较早的研究保持一致[43], 即为了减少语误, 内部自我监测引起神经抑制, 证明了音韵编码之后出现的类似的更小的正向波形与对音素的自我监测有关。

综上, 音素在音韵编码阶段的作用不容忽视, 但是对其具体是在前期作为功能单元起促进作用, 还是在后期起促进或抑制作用, 这方面的研究有待进一步的讨论。为了区分后期音素提取的解释与自我监测的解释, 我们可以使用更精细的空间分辨率的技术进行探讨, 例如 fMRI [44]、TMS [45]。

2.2. 音素作用于语音编码阶段的证据

有研究表明音素效应可能出现在语音编码阶段, 而不是反映音韵编码阶段对音素的自我监测, 两种猜测中音素效应均表现为抑制作用, 但效应发生的时间进程不同。对命名反应时的分析中, 我们发现了轻微的音素重叠抑制效应[12] [27] [46]。Zhang & Damian [27]采用掩蔽启动范式, 发现在图片呈现后 500~600 ms 这个较晚的时间窗内音素重叠和音素不重叠条件下的平均波形存在差异, 音素效应的时间进程大致与 Indefrey 和 Levelt [42]以及 Indefrey [32]估计的语音编码阶段的时间进程(455~600 ms)是相同的。张清芳和王雪娇[47]使用内隐启动范式, 也发现音素效应出现在线索词呈现后 500~600 ms 的时间窗内, 发生在音节效应之后, 但同源组的平均波形相比于异源组的平均波形更正。他们推测这种音素重叠的抑制效应可能是源于音韵编码阶段对音素的自我监测, 或者音素在语音编码起作用。前一种解释不符合之

前研究中我们对于自我监测时音素重叠条件下(同源组)更小的正向波形的观点,因此我们有理由排除自我监测的影响;对于后一种解释,较晚时间窗内的音素抑制效应可能发生在语音编码阶段,有一定的可能性。

一些研究者提出了前运动和运动执行过程分离的假设[48][49]。他们认为词汇表征的准备和执行过程是独立的。音韵编码阶段是对词汇音韵内容的提取,属于词汇表征的准备过程,而语音编码和发音过程属于对运动执行过程的准备。岳源和张清芳[50]使用图-词干扰范式,与即时命名、延迟命名以及延迟命名和发音抑制任务相结合,比较三个任务中是否存在音节和音段效应,并据此推测这些效应的发生阶段。结果发现,音节和音段的促进效应发生在音韵编码阶段,而音节和音段的抑制效应发生在语音编码或发音阶段。且效果量(Cohen d)的分析表明,在语音编码和发音阶段,音段相关的效果量 Cohen d 值强于音节相关的效果量,说明在音韵编码阶段,音节的作用大于音段;在语音编码和发音阶段,音段的作用大于音节。他们的结果支持了该假设。那么作为最小音段的音素(两者同为亚音节),是否也和音段类似,在音韵编码起微弱的促进,在语音编码起抑制作用并且相比于音节更加重要,是否同样支持该假设,这些需要我们进一步探讨。综上所述,靠后时间窗内的音素效应表现为抑制作用,可能反映的不是音韵编码阶段的自我监测,而是音素在语音编码阶段起作用。

3. 总结

在汉语口语词汇产生中,音素起作用的认知机制尚有疑虑,目前已知的是音素效应起相对重要作用的阶段可能是音韵编码的前期或后期阶段,也可能是语音编码阶段,但音素效应具体在哪一阶段起相对重要作用,以及在各阶段的效应大小和方向问题仍需要进一步探讨。已有的研究倾向于认为音素在音韵编码的后期或是语音编码阶段起重要作用,在音韵编码后期表现为促进或抑制作用,语音编码阶段则表现为抑制作用,具体是两阶段的哪一阶段还需进一步研究。但这一结论已支持合适单元假设,一定程度上符合 Roelofs 等人[9]的观点:音素具有跨语言特异性,在汉语中倾向于在言语产生的后期起重要作用,与印欧语系中建立的言语产生模型不同,在汉语中我们认为音韵编码阶段最先提取的是音节信息,之后才会激活相应的音素信息进行加工。目前的研究结论将有助于我们建立和完善适合于汉语的言语产生模型,对之后的研究提供一定的启示。

二语的熟练度、使用的实验材料或实验范式的差异均会导致实验结果出现差异,且破坏研究间的可比性。目前已有研究中对音素作用阶段进行研究的实验范式也不够敏感,未来还需寻找对音素更敏感有效的实验范式和技术进行研究,需进一步控制实验材料和实验范式,以此来增加研究间的可比性。

另外,虽然传统的 ERP 波形分析在时间进程上帮助我们区分了音节和音素的加工顺序,但 ERP 的波幅大小和方向在研究中表现出不一致,这使我们很难进行进一步的解释,这种实验间波幅的差异是源于实验范式的不同还是加工机制的不同仍需探讨。音素相关效应是否支持前运动过程和运动执行过程分离的观点也需结合相应的实验任务进行考量。

4. 展望

目前,研究无法区分音素效应是发生在音韵编码阶段的后期还是语音编码阶段。前文提到,岳源和张清芳[50]使用图-词干扰范式,通过研究任务之间的对比表明音段在运动执行过程中可能起了相对重要的作用。该研究使用图-词干扰范式的原因是该范式对音段更加敏感,但如果要探测音素效应,使用图-词干扰范式或许不是最好的选择。且该研究的实验 2 和 3 中效应量 Cohen d 值都较小,这一结果更加佐证了我们的观点,即图-词干扰范式与延迟命名和发音抑制的结合任务在程序上相对复杂,对音节或音段不够敏感,更不用说比它们更小的语音单元-音素了。因此,未来需要寻找对音素效应更敏感的范

式来进行进一步探讨。比如掩蔽启动范式, 可能更适合对音素的研究。选择掩蔽启动范式的两点原因: 首先, 在目标图片之前仅短暂地(大约 50 ms)呈现了一个掩蔽启动词, 讲话者难以感知这个词, 因此来自掩蔽启动任务的证据是在不受策略干扰的情况下自动加工音韵的; 其次, 它排除了其他范式(如图-词干扰任务、音素准备任务或音素重复任务)中难以排除的自我监控的重要干扰因素[12] [26]。未来的研究可以采用多个研究任务之间的对比, 结合其他对音素更敏感的言语产生范式, 如掩蔽启动范式或重复启动范式, 来探测出音素所处的阶段和作用。

由前文可知, 行为研究中的音素效应是微弱的, 难以测量的, 且不够敏感, 不能对言语产生各阶段的时间进程进行较为细致的分析。其指标之一的“命名反应时”是指发音阶段前的图片命名的时间长度, 包括概念准备, 词汇选择和单词形式编码阶段, 反映了相关条件和无关条件下图片命名的总差异。以往的研究多采用的视觉干扰词或启动词, 未发现亚音节的相关效应[51]。但 Meyer 和 Schriefers [23]的实验中采用听觉呈现干扰词的 PWI 任务, 却在 SOA = 0、+150 ms 时发现了亚音节效应。另外, 关于粤语的 PWI 任务的研究中发现, 视觉呈现干扰词的音韵相关效应(尾音段效应)不是很强, 但听觉呈现干扰词时相关效应较强, 即感觉通道对音韵相关效应存在影响[10] [30] [31]。因此, 之后对音素的研究中可使用听觉呈现干扰词或启动词, 结合对所研究对象更敏感的实验范式, 来验证反应时指标上是否可能更敏感。

近年来, 言语产生的相关研究开始结合 ERP 技术, 对词汇产生的加工过程进行探讨。ERP 技术具有高时间分辨率和一定的空间分辨率, 可以弥补行为指标上的局限, 对言语产生各阶段的时间进程更加敏感, 相关条件和无关条件之间的比较反映的是在特定时间间隔, 特定阶段之间的差异, 比如音韵编码阶段。因此在确定音素效应的所处阶段时, 可结合 ERP 技术, 将其与 Indefrey [32]的元分析中各阶段时间进程节点进行对比, 以此来确定效应所处阶段。但是在进行时间窗的比较时, Indefrey 的时间进程元分析只有一定的参考意义。因为它是有一定局限的: 首先, Indefrey 的元分析大多基于印欧语系语言的研究建立, 用于与汉语的口语词汇产生中的音素效应进行对比, 可能存在一定偏差。其次, 元分析是基于图片命名任务, 因此在与使用其他范式进行研究时所发现的音素时间窗对比时, 存在一定的差异。例如, Zhang & Damian [27]中使用掩蔽启动范式, 发现音素相关条件下的命名反应时是 690 ms, 相比于 Indefrey 和 Levelt [42]图片命名任务确定的命名反应时(大约 600 ms)有很大不同, 所以这样就造成在对音素效应的较后时间窗进行分析时, 不能确定其时间进程是在音韵编码后期阶段还是语音编码阶段的时间范围内。另外, 与图画命名范式相比, 内隐启动范式中言语产生过程不包括概念准备(直接从工作记忆中提取), 音韵编码阶段相对较早。最近的研究已经证实, 音韵编码发生在不同任务的大约 275~400 ms 的时间窗内, 通过比较出声的图片命名和不出声的观看相同图片时的 ERP 差异, 确定图片呈现后 275~400 ms 的时间窗内属于音韵编码阶段[52]; 在图-词干扰任务中, 音韵相关性的促进作用也发生在 250~400 ms 的相似时间窗内[53]。这些发现与一些出声的口语词汇产生过程中的音韵效应的研究一致[51] [54] [55], 但与 Indefrey 元分析中确定音韵编码阶段的时间进程(大约为 275~455 ms)存在一定出入。因此, 利用 ERP 研究中音素效应发生的时间窗口来推测音素作用的加工阶段是有局限性并且需要谨慎解释。

此外, 目前的结合 ERP 技术的研究大多是从音素的时间进程上进行考虑, 以此来确定音素效应的发生阶段, 很少有人从音素的空间特性上进行研究。Indefrey [32]评估了言语产生的大量神经影像研究, 发现左侧额下回(IFG)和左侧额中回(MFG)在音韵编码阶段常被激活, 中央前回、左丘脑和小脑在语音编码和发音阶段常被激活, 而双侧颞上回(STG)、扣带皮层等脑区的激活涉及自我监测, 这表明音韵编码和语音编码涉及不同的加工, 引起不同的脑区激活, 我们有可能从空间特性上区分言语产生各阶段的加工进程; Siok 等人[56]研究了书写产生中汉语的音素与音节加工, 发现左中额叶皮层与音节处理有关, 而左下前额回与音素加工有关, 因此, 从这个角度也可探讨在口语产生中音素效应的发生阶段, 除了使用 ERP 技术, 还可使用空间分辨率更高的 fMRI 或者 fNIRS 进行研究, 结合时间和空间特性同时对音素效应进

行考察。

二语的熟练度对于音韵的加工也存在影响。Verdonschot 等人[29]的研究中,以高熟练度的英-汉双语者为被试,表现出行为上显著的首音素启动效应。而 Verdonschot 等人[28]的研究中,使用与 Verdonschot 等人[29]的实验中相同的中文材料,但以英语水平较低的汉语母语者为被试,结果未发现首音素启动效应。这说明有关英语的音韵加工对汉语的目标音韵的加工施加了影响,以至于在汉语中检测到了首音素的启动效应。因此,在汉语研究中研究音素效应,有必要控制被试的二语的熟练度水平,在进行实验时选择英语水平低的汉语母语者,排除二语的音韵编码单元可能对汉语口语词汇产生过程的影响。

基金项目

天津市哲学社会科学规划项目(TJJX21-002)。

参考文献

- [1] Dell, G.S. (1984) Representation of Serial Order in Speech: Evidence from the Repeated Phoneme Effect in Speech Errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **10**, 222-233. <https://doi.org/10.1037/0278-7393.10.2.222>
- [2] Levelt, W.J.M., Roelofs, A. and Meyer, A.S. (1999) A Theory of Lexical Access in Speech Production. *Behavioral and Brain Sciences*, **22**, 1-38. <https://doi.org/10.1017/S0140525X99001776>
- [3] Roelofs, A. (1997) The WEAVER Model of Word-form Encoding in Speech Production. *Cognition*, **64**, 249-284. [https://doi.org/10.1016/S0010-0277\(97\)00027-9](https://doi.org/10.1016/S0010-0277(97)00027-9)
- [4] Chen, J.-Y., Lin, W.-C. and Ferrand, L. (2003) Masked Priming of the Syllable in Mandarin Chinese Speech Production. *Chinese Journal of Psychology*, **45**, 107-120.
- [5] O'Seaghdha, P.G., Chen, J.Y. and Chen, T.-M. (2010) Proximate Units in Word Production: Phonological Encoding Begins with Syllables in Mandarin Chinese but with Segments in English. *Cognition*, **115**, 282-302. <https://doi.org/10.1016/j.cognition.2010.01.001>
- [6] Schiller, N.O. (2004) The Onset Effect in Word Naming. *Journal of Memory and Language*, **50**, 477-490. <https://doi.org/10.1016/j.jml.2004.02.004>
- [7] Schiller, N.O. (2008) The Masked Onset Priming Effect in Picture Naming. *Cognition*, **106**, 952-962. <https://doi.org/10.1016/j.cognition.2007.03.007>
- [8] 邵静敏. 现代汉语通论[M]. 上海: 上海教育出版社, 2001.
- [9] Roelofs, A. (2015) Modeling of Phonological Encoding in Spoken Word Production: from Germanic Languages to Mandarin Chinese and Japanese. *Japanese Psychological Research*, **57**, 22-37. <https://doi.org/10.1111/jpr.12050>
- [10] Wong, A.W. and Chen, H.-C. (2015) Processing Segmental and Prosodic Information in Spoken Word Planning: Further Evidence from Cantonese Chinese. *Japanese Psychological Research*, **57**, 69-80. <https://doi.org/10.1111/jpr.12054>
- [11] Wong, A.W., Huang, J. and Chen, H.-C. (2012) Phonological Units in Spoken Word Production: Insights from Cantonese. *PLoS ONE*, **7**, Article ID: e48776. <https://doi.org/10.1371/journal.pone.0048776>
- [12] Chen, J.-Y., O'Seaghdha, P.G. and Chen, T.-M. (2016) The Primacy of Abstract Syllables in Chinese Word Production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **42**, 825-836. <https://doi.org/10.1037/a0039911>
- [13] Dell, G.S. (1986) A Spreading-Activation Theory of Retrieval in Sentence Production. *Psychological Review*, **93**, 283-321. <https://doi.org/10.1037/0033-295X.93.3.283>
- [14] Shattuck-Hufnagel, S. (1979) Speech Errors as Evidence for a Serial-Ordering Mechanism in Sentence Production. In: Cooper, W.E. and Walker, E.C.T., Eds., *Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett*, Erlbaum, Hillsdale, 295-342.
- [15] Meyer, A.S. (1991) The Time Course of Phonological Encoding in Language Production: Phonological Encoding inside a Syllable. *Journal of Memory and Language*, **30**, 69-89.
- [16] Alario, F.X., Perre, L., Castel, C. and Ziegler, J.C. (2007) The Role of Orthography in Speech Production Revisited. *Cognition*, **102**, 464-475. <https://doi.org/10.1016/j.cognition.2006.02.002>
- [17] Damian, M.F. and Bowers, J.S. (2003) Effects of Orthography on Speech Production in A Form-Preparation Paradigm. *Journal of Memory and Language*, **49**, 119-132. [https://doi.org/10.1016/S0749-596X\(03\)00008-1](https://doi.org/10.1016/S0749-596X(03)00008-1)

- [18] Jacobs, C.L. and Dell, G.S. (2014) “Hotdog”, Not “Hot” “Dog”: The Phonological Planning of Compound Words. *Language, Cognition and Neuroscience*, **29**, 512-523. <https://doi.org/10.1080/23273798.2014.892144>
- [19] Damian, M.F. and Dumay, N. (2007) Time Pressure and Phonological Advance Planning in Spoken Production. *Journal of Memory and Language*, **57**, 195-209. <https://doi.org/10.1016/j.jml.2006.11.001>
- [20] Damian, M.F. and Dumay, N. (2009) Exploring Phonological Encoding Through Repeated Segments. *Language and Cognitive Processes*, **24**, 685-712. <https://doi.org/10.1080/01690960802351260>
- [21] Qu, Q.Q., Damian, M.F. & Kazanina, N. (2012) Sound-sized Segments Are Significant for Mandarin Speakers. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, 14265-14270. <https://doi.org/10.1073/pnas.1200632109>
- [22] Damian, M.F. and Martin, R.C. (1999) Semantic and Phonological Codes Interact in Single Word Production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **25**, 345-361. <https://doi.org/10.1037/0278-7393.25.2.345>
- [23] Meyer, A.S. and Schriefers, H. (1991) Phonological Facilitation in Picture-word Interference Experiments: Effects of Stimulus Onset Asynchrony and Types of Interfering Stimuli. *Journal of Experimental Psychology: Learning Memory, and Cognition*, **17**, 1146-1160. <https://doi.org/10.1037/0278-7393.17.6.1146>
- [24] Forster, K.I. and Davis, C. (1991) The Density Constraint on Form-priming in the Naming Task: Interference Effects from A Masked Prime. *Journal of Memory and Language*, **30**, 1-25. [https://doi.org/10.1016/0749-596X\(91\)90008-8](https://doi.org/10.1016/0749-596X(91)90008-8)
- [25] Kinoshita, S. and Woollams, A. (2002) The Masked Onset Priming Effect in Naming: Computation of Phonology or Speech Planning. *Memory and Cognition*, **30**, 237-245. <https://doi.org/10.3758/BF03195284>
- [26] You, W., Zhang, Q. and Verdonschot, R.G. (2012) Masked Syllable Priming Effects in Word and Picture Naming in Chinese. *PLoS ONE*, **7**, Article ID: e46595. <https://doi.org/10.1371/journal.pone.0046595>
- [27] Zhang, Q. and Damian, M.F. (2019) Syllables Constitute Proximate Units for Mandarin Speakers: Electrophysiological Evidence from A Masked Priming Task. *Psychophysiology*, **56**, Article ID: e13317. <https://doi.org/10.1111/psyp.13317>
- [28] Verdonschot, R.G., Lai, J., Chen, F., Tamaoka, K. and Schiller, N.O. (2015) Constructing Initial Phonology in Mandarin Chinese: Syllabic or Subsyllabic? A Masked Priming Investigation. *Japanese Psychological Research*, **57**, 61-68. <https://doi.org/10.1111/jpr.12064>
- [29] Verdonschot, R.G., Nakayama, M., Zhang, Q.F., Tamaoka, K. and Schiller, N.O. (2013) The Proximate Phonological Unit of Chinese-English Bilinguals: Proficiency Matters. *PLoS ONE*, **8**, Article ID: e61454. <https://doi.org/10.1371/journal.pone.0061454>
- [30] Wong, A.W. and Chen, H.-C. (2008) Processing Segmental and Prosodic Information in Cantonese Word Production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **34**, 1172-1190. <https://doi.org/10.1037/a0013000>
- [31] Wong, A.W. and Chen, H.-C. (2009) What Are Effective Phonological Units in Cantonese Spoken Word Planning? *Psychonomic Bulletin & Review*, **16**, 888-892. <https://doi.org/10.3758/PBR.16.5.888>
- [32] Indefrey, P. (2011) The Spatial and Temporal Signatures of Word Production Components: A Critical Update. *Frontiers in Psychology*, **2**, Article No. 255. <https://doi.org/10.3389/fpsyg.2011.00255>
- [33] Yu, M., Mo, C. and Mo, L. (2014) The Role of Phoneme in Mandarin Chinese Production: Evidence from ERPs. *PLoS One*, **9**, Article ID: e106486. <https://doi.org/10.1371/journal.pone.0106486>
- [34] Qu, Q.Q., Feng, C., Hou, F. and Damian, M.F. (2020) Syllables and Phonemes as Planning Units in Mandarin Chinese Spoken Word Production: Evidence from ERPs. *Neuropsychologia*, **146**, Article ID: 107559. <https://doi.org/10.1016/j.neuropsychologia.2020.107559>
- [35] O’Seaghdha, P.G. and Frazer, A.K. (2014) The Exception Does Not Rule: Attention Constrains Form Preparation in Word Production. *Journal of Experimental Psychology Learning Memory and Cognition*, **40**, 797-810. <https://doi.org/10.1037/a0035576>
- [36] Chen, T.-M. and Chen, J.-Y. (2013) The Syllable as the Proximate Unit in Mandarin Chinese Word Production: An Intrinsic Or Accidental Property of the Production System? *Psychonomic Bulletin & Review*, **20**, 154-162. <https://doi.org/10.3758/s13423-012-0326-7>
- [37] Cai, X., Yin, Y. and Zhang, Q. (2020) The Roles of Syllables and Phonemes during Phonological Encoding in Chinese Spoken Word Production: A Topographic ERP Study. *Neuropsychologia*, **140**, Article ID: 107382. <https://doi.org/10.1016/j.neuropsychologia.2020.107382>
- [38] Qu, Q.Q., Damian, M.F. and Kazanina, N. (2013) Reply to O’Seaghdha *et al.*: Primary Phonological Planning Units in Chinese Are Phonemically Specified. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, Article No. E4. <https://doi.org/10.1073/pnas.1217601110>
- [39] O’Seaghdha, P.G., Chen, J.-Y. and Chen, T.-M. (2013) Close but Not Proximate: the Significant of Phonological Seg-

- ments in Speaking Depends on Their Functional Engagement. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, Article No. E3. <https://doi.org/10.1073/pnas.1217032110>
- [40] O'Seaghdha, P.G. (2015) Across the Great Divide: Proximate Units at the Lexical-phonological Interface. *Japanese Psychological Research*, **57**, 4-21. <https://doi.org/10.1111/jpr.12074>
- [41] MacKay, D.G. (1972) The Structure of Words and Syllables: Evidence from Errors in Speech. *Cognitive Psychology*, **3**, 210-227. [https://doi.org/10.1016/0010-0285\(72\)90004-7](https://doi.org/10.1016/0010-0285(72)90004-7)
- [42] Indefrey, P. and Levelt, W.J.M. (2004) The Spatial and Temporal Signatures of Word Production Components. *Cognition*, **92**, 101-144. <https://doi.org/10.1016/j.cognition.2002.06.001>
- [43] Maess, B., Friederici, A.D., Damian, M.F., Meyer A.S. and Levelt, W.J.M. (2002) Semantic Category Interference in Overt Picture Naming: An MEG Study. *Journal of Cognitive Neuroscience*, **14**, 455-462. <https://doi.org/10.1162/089892902317361967>
- [44] Yu, M., Mo, C., Li, Y. and Mo, L. (2015) Distinct Representations of Syllables and Phonemes in Chinese Production: Evidence from fMRI Adaptation. *Neuropsychologia*, **77**, 253-259. <https://doi.org/10.1016/j.neuropsychologia.2015.08.027>
- [45] Schuhmann, T., Schiller, N.O., Goebel, R. and Sack, A.T. (2012) Speaking of Which: Dissecting the Neurocognitive Network of Language Production in Picture Naming. *Cerebral Cortex*, **22**, 701-709. <https://doi.org/10.1093/cercor/bhr155>
- [46] 蒋宇宸, 蔡笑, 张清芳. θ 频段(4-8 Hz)的活动反映了汉语口语产生中音节信息的加工[J]. *心理学报*, 2020, 52(10): 1199-1211.
- [47] 张清芳, 王雪娇. 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[J]. *心理学报*, 2020, 52(4): 414-425.
- [48] Hickok, G. and Poeppel, D. (2007) The Cortical Organization of Speech Processing. *Nature Reviews Neuroscience*, **8**, 393-402. <https://doi.org/10.1038/nrn2113>
- [49] Eickhoff, S.B., Heim, S., Zilles, K. and Amunts, K. (2009) A Systems Perspective on the Effective Connectivity of Overt Speech Production. *NeuroImage*, **47**, S39-S41. [https://doi.org/10.1016/S1053-8119\(09\)70789-X](https://doi.org/10.1016/S1053-8119(09)70789-X)
- [50] 岳源, 张清芳. 汉语口语产生中音节和音段的促进和抑制效应[J]. *心理学报*, 2015, 47(3): 319-328.
- [51] Wang, J., Wong, A.W., Wang, S.P. and Chen, H.-C. (2017) Primary Phonological Planning Units in Spoken Word Production Are Language-specific: Evidence from An ERP Study. *Scientific Reports*, **7**, Article No. 5815. <https://doi.org/10.1038/s41598-017-06186-z>
- [52] Eulitz, C., Hauk, O. and Cohen, R. (2000) Electroencephalographic Activity over Temporal Brain Areas during Phonological Encoding in Picture Naming. *Clinical Neurophysiology*, **111**, 2088-2097. [https://doi.org/10.1016/S1388-2457\(00\)00441-7](https://doi.org/10.1016/S1388-2457(00)00441-7)
- [53] Dell'Acqua, R., Sessa, P., Peressotti, F., Mulatti, C., Navarrete, E. and Grainger, J. (2010) ErP Evidence for Ultra-fast Semantic Processing in the Picture-word Interference Paradigm. *Frontiers in Psychology*, **1**, Article No. 177. <https://doi.org/10.3389/fpsyg.2010.00177>
- [54] Wong, A.W., Wang, J., Ng, T.Y. and Chen, H.-C. (2016) Syllabic Encoding during Overt Speech Production in Cantonese: Evidence from Temporal Brain Responses. *Brain Research*, **1648**, 101-109. <https://doi.org/10.1016/j.brainres.2016.07.032>
- [55] Zhu, X., Damian, M.F. and Zhang, Q. (2015) Seriality of Semantic and Phonological Processes during Overt Speech in Mandarin as Revealed by Event-Related Brain Potentials. *Brain and Language*, **144**, 16-25. <https://doi.org/10.1016/j.bandl.2015.03.007>
- [56] Siok, W.T., Jin, Z., Fletcher, P. and Tan, L.H. (2003) Distinct Brain Regions Associated with Syllable and Phoneme. *Human Brain Mapping*, **18**, 201-207. <https://doi.org/10.1002/hbm.10094>