

《心理学报》审稿意见与作者回应

题目：英汉双语者二语口语产生中音韵编码过程的同化与顺应机制

作者：辛昕 兰天一 张清芳

第一轮

审稿人 1 意见：

该文章从同化-顺应理论出发，尝试通过事件相关电位技术，考察英汉双语者二语口语产生过程中音韵编码过程中的同化与顺应机制，有理论意义。研究的设计及统计分析均有其特色，包括采用两组被试及不同任务条件的组合，以及时空微态分割分析。总体而言，研究逻辑较为清晰，方案设计合理。有以下意见或建议。

意见 1：前言 1.1 部分，第一句是“研究发现双语者两种语言正字法透明度(Orthographic Transparency)是影响双语者二语加工同化与顺应机制的重要因素”，但其后的两段内容都是讨论透明度不同的文字加工脑机制，没有谈如何影响同化-顺应机制。

回应：感谢审稿专家的宝贵意见。我们对 1.1 部分进行了全面的修改。在这一部分的第一段，我们指出双语者加工二语时表现出同化还是顺应机制，所比较的对象是两种语言作为母语时的认知神经机制。研究发现语言的正字法透明度影响了母语者语言加工的脑激活模式。在第二段，我们阐述了双语者加工二语的脑激活模式是与自己的母语相似还是与二语(作为母语时)类似，这与两种语言正字法透明度的相对高低有密切关系(见正文 P.1-2)。

意见 2：关于数据分析的基本逻辑，提一点不同想法与作者商榷。同化-顺应理论似乎应该以两种语言各自作为母语时，其加工机制存在差异为前提或基础。比如，以文章中的实验条件举例，首先汉汉组和英英组在比较中存在差异，在此基础上，才好讨论二语(英汉组)是以一语母语方式(英英组)加工呢还是以二语的母语方式(汉汉组)加工。文章中没有对汉汉组及英英组直接对比，所以无从知道两种语言作为母语的加工是否存在差异。

回应：感谢审稿专家的宝贵意见。我们非常赞同您提出的二语者加工二语的同化-顺应理论应该以两种语言各自作为母语时，其加工机制存在差异为前提或基础。本研究的实验设计遵循了此设计逻辑。已有研究一致地表明在口语词汇产生过程中，英语与汉语音韵编码单元是不同的，英语口语词汇产生中的音韵编码单元为音素(Dell, 1986; Shattuck-Hufnagel, 1979)，而汉语为音节(Chen, 2000; Feng, Yue, & Zhang, 2019)。我们在文章的前言详细地对此进行了阐述(见 1.2 口语词汇产生中音韵编码的单元)。

在行为结果中，针对汉汉组与英英组被试间两因素(组别和语音相关类型)混合线性模型分析，汉汉组和英英组的反应时存在显著差异，且两组都表现出音节相关效应，未出现音素

相关效应。我们认为出现这一结果的原因可能有：第一，实验设计中同时包括了音节相关、音素相关和无关，因为音节相关中语音重叠的程度大于音素相关，这在一定程度上可能掩盖了音素相关所引起的效应。同时，本研究发现英英组的反应时长于汉汉组，潜伏期变长也有可能使得较小的音素效应消失。第二，这可能是由于英汉双语者的语言环境是汉语所导致的。以往就有研究在行为结果上发现，第一语言的加工过程会受到第二语言学习迁移的影响(见讨论 P.15)。

尽管如此，本研究中 ERP 的结果显示，英英组在 200-250 ms 出现显著的音素效应，汉汉组在 250-300 ms 出现显著的音节效应，无论是语音效应类型还是时间窗口都具有显著差异。在微态分割分析上，汉汉组与英英组的微态对比分析发现，在音韵编码阶段，汉汉组与英英组具有完全不同的微态成分 2 和微态成分 3。ERP 和微态分割的结果一致表明汉汉组与英英组在音韵编码阶段存在差异。结合英汉组与汉汉组和英英组的对比，我们认为英汉双语者在加工第二语言汉语过程中采用类似于母语的认知神经模式，即同化机制。

意见 3：被试信息有些过于简略，特别是英汉双语者。只有一句话“匹配英语母语被试的汉语水平，通过汉语水平考试(HSK)5 级及以上”。这儿的“匹配”是谁和谁匹配呢？另外，他们的汉语学习经验如何？学了多久？是否有在中国学习的经历？除了汉语水平考试，是否有其他衡量其汉语水平(如听说读写)的测试或指标等？总之，有些太简略。

回应：感谢审稿专家的宝贵意见。我们在 2.1 部分补充了被试的相关信息：“英汉双语被试 18 人(13 名男性，5 名女性，平均年龄 22.9 岁)，两组被试在年龄($t=1.347, p=0.187$)上无显著差异，均为在校大学生。英语母语被试为外籍留学生，是晚期英汉双语者，学习汉语至少三年以上，平均学习汉语年龄 18 岁开始，均具有一年以上的中国留学经历，通过汉语水平考试(HSK)5 级及以上”(见 P.6)。

意见 4：行为结果部分：1)“汉汉组和英汉组对比中语音相关类型和组别在被试分析中交互作用显著， $F_1(2, 34)=0.413, p=0.024, \eta^2=0.007$ ”。请检查 F 值是否有误，若为 0.413，不可能显著吧。2)“汉汉组音节相关条件下的反应时边缘显著快于无关条件， $t(17)=-8.067, p<0.001$ ”。边缘显著？比较困惑。3)“音素相关条件反应时和无关条件显著差异， $t(14)=-0.481, p=0.636$ ”，到底显著不显著？

回应：感谢审稿专家的指正。在修改稿中，我们采用混合线性模型分析了行为数据，详细的数据分析方法和结果见“3.1 行为数据”部分。

意见 5：从行为结果(图 1)看，汉汉组、英汉组和英英组的模式其实非常相似，而且汉汉和英英尽管反应时的绝对值有差异，但三种实验条件(音节相关、音素相关与语音无关)之间的模式都非常相似，而且汉汉组和英英组的相似性似乎更高些。因此，但从交互作用的显著与否来说是顺应机制还是同化机制，似乎说服力不强。

回应：感谢审稿专家的宝贵意见。我们同意您的看法，行为数据的结果确实不能明确地说明英汉双语者汉语加工的认知机制。我们修改了文中有关的表达，对于行为数据的结果做了深入讨论(见讨论 P.15)。关于英汉双语者汉语加工的同化机制，在讨论中我们主要是基于 ERP 和时间空间微态分析的结果得出了有关英汉双语者汉语加工的认知神经机制。这一问题与前面的第二个问题密切相关，请审稿专家同时参阅对问题 2 的回复。

意见 6：ERP 结果，表 1 中的统计结果也应该进行多重比较校正。

回应：感谢审稿专家的指正。已对表 1(现表 2)的统计结果进行多重比较校正，采用的是目前较为通用的 *fdr* 校正方法(见表 2)。

意见 7：确定显著和边缘显著的标准比较含糊，如“英英组 200-250 ms 时间窗口，后半球中右音素效应显著， $F(1, 17)=8.809$ ， $p = 0.054$ ， $\eta^2 = 0.341$ ； $F(1, 17)=5.756$ ， $p = 0.058$ ， $\eta^2 = 0.253$ 。前半球左中音素效应显著， $F(1, 17)=4.886$ ， $p = 0.058$ ， $\eta^2 = 0.223$ ； $F(1, 17)=4.497$ ， $p = 0.058$ ， $\eta^2 = 0.209$ 。”这里都称为显著。

回应：感谢审稿专家的指正。我们根据统计学上的显著性检验指标将行为和微态分割结果中 $p > 0.05$ 定义为不显著， $p < 0.05$ 定义为显著， $p < 0.001$ 定义为极其显著；在 ERP 结果中， p 值校正前小于 0.05 校正后大于 0.05 小于 0.1，定义为边缘显著。我们修改了文中表述不一致的地方(见 P.12)。

意见 8：汉汉与英汉的对比是组间比较，英英和英汉的比较是组内比较，这种比较本身是否会影响结果呢？比如，组间比较会不可避免地受被试变量的影响，而组内比较则不受此影响。建议在讨论中对此略作说明。

回应：感谢审稿专家的宝贵意见。这一问题使得我们深入地思考本研究实验设计的逻辑和数据处理方法。我们在讨论中对组间和组内比较对结果可能产生的影响进行了说明。组间比较时汉汉组和英汉组所用材料相同，而组内比较时英英组和英汉组为同一组被试，所用材料不同(汉语和英语)。针对被试和项目，组间比较中包括了被试间变异，而组内比较中包括了项目间变异。作为随机变量的被试因素和项目因素都有可能对结果产生影响。为了综合考虑这两个随机变量，我们在修改稿中采用了混合线模型分析了行为数据(见“3.1 行为数据”)，主要的分析结果与之前方差分析的结果类似。

本研究的实验设计遵循了已有对于双语者二语加工同化与顺应机制的研究逻辑，与已有研究设计类似(例如 Tian et al., 2019; Timmer & Shiller, 2012; Sun et al., 2015)。Sun 等(2015)利用 fMRI 技术考察英汉双语者和汉英双语者在阅读母语与二语词汇过程中的脑区激活差异。其中将英汉双语者进行二语汉语词汇判断任务组(英汉组)与英汉双语者进行英语词汇判断任务(英英组)和汉英双语者进行汉语词汇判断组(汉汉组)分别进行组内和组间对比。结果表明，与英英组相比，英汉组更多的激活右侧额下回区域，该脑区的激活与汉语加工相关。

也就是说，英汉双语者在词汇识别任务中利用了与母语加工类似的顺应机制，这一结果与 Liu 和 Cao(2016)元分析中的结果相一致，因此我们认为该机制受到组内或组间比较的影响较小。尽管如此，我们认为由于双语者的两种语言之间存在相互影响的可能，在考察二语加工同化与顺应的机制时，更为理想的情况是三组被试分别为汉语单语组、英语单语组和英汉双语组，比较英汉组与其他两组的加工机制是否相同(见讨论 P. 19)。

.....

审稿人 2 意见：

意见 1：文中提到的音段(segments)实际上是次音节单元(sub-syllabic units)。音段(segments)这一概念的一般用法可参考 Eric Bakovi? (2014) Phonemes, segments and features, *Language, Cognition and Neuroscience*, 29:1, 21-23, DOI: 10.1080/01690965.2013.848992

回应：感谢审稿专家的宝贵意见。我们根据您的建议修改了有关音段的阐述。具体内容为：“音段(segment)在汉语中指的是亚音节语音表征单元(subsegmental units)(Bakovi & Eric., 2014; Wong & H. -C. Chen, 2008, 2009)。例如，对于一个 CVG(G 表示 glide)结构的汉语音节 /qing/, 可以划分为两个音段，首 CV 结构的/qi/和尾音段 VG 结构的/ing/”(见脚注 1)。

意见 2：文中对之前文献的总结描述及引用有不准确之处，请仔细检查。例如，Wang, Wong & Chen, 2017 中的被试是在香港求学的内地普通话被试，并非香港当地的粤语被试；Wang 等(2017)的 ERP 研究并非利用图-图干扰范式(应为启动范式)，而且没有出现在参考文献列表中。另写作上有待斟酌或改进之处，已在原文高亮标示，请仔细查看全文。

回应：感谢审稿专家的指正。原文中提到“Wang, Wong & Chen, 2017 中的被试是在香港求学的内地普通话被试，并非香港当地的粤语被试”，此处引用的是文献为“Wang, J. , Wong, W. K. , & Chen, H. C. . (2017). Time course of syllabic and sub-syllabic processing in mandarin word production evidence from the picture-word interference paradigm. *Psychonomic Bulletin & Review*, 3, 1-6.”。该研究选取的被试为来自大陆的汉语普通话被试，其中部分具有粤语经验，已在文中修正表述。Wong(2012)等选取的粤语被试是来自香港的大学生，具有较高的英语水平(p.14)。

文中所引用的 Wang 等(2017)的 ERP 研究指是 Wang 等 2017 年发表在 *Scientific Reports* 期刊上的研究。这篇文中报告的实验方法为图-图干扰范式和延迟命名任务的结合。具体文献信息为：“Wang, J., Wong, A. W.-K., Wang, S. P., & Chen, H.-C. (2017). Primary phonological planning units in spoken word production are language-specific: Evidence from an ERP study. *Scientific Reports*, 7, 5815–5822.”。相应的参考文献信息已增加至列表。

在写作上有待斟酌和修改之处，我们根据您的标记出来的地方，逐一地进行了修改，具体的修改我们放到了回复的最后。

意见 3：三组中配对比较的结果显示，每组都产生显著的音节效应以及不显著的音素/次音

节效应。汉汉组和英汉组对比中语音相关类型和组别的交互作用只在被试分析中显著，在项目分析中不显著。因此行为数据的结果并不倾向于支持作者所说的同化机制。

回应：感谢审稿专家的宝贵意见。我们同意您的看法，行为数据的结果确实不能明确地说明英汉双语者汉语加工的认知机制。我们修改了文中有关的表达，对于行为数据的结果做了深入讨论(见讨论 P.15)。关于英汉双语者汉语加工的同化机制，在讨论中我们主要是基于 ERP 和时间空间微态分析的结果得出了有关英汉双语者汉语加工的认知神经机制。这一问题与第一位审稿专家的第二个问题密切相关，请审稿专家同时参阅对该问题的回复。

意见 4：英汉双语被试的行为数据结果和 ERP 结果不一致。行为数据中看到音节效应而无音素效应，ERP 结果则相反。另 ERP 结果中很多效应仅达到边缘显著，降低了实验结果的可信度，无法有力支持同化机制的结论。

回应：感谢审稿专家的宝贵意见。英汉双语被试的行为结果与 ERP 结果确实存在不一致之处：英汉组在行为上在对母语英语的加工中表现出音节效应，而脑电结果表现出音素效应。反应时是最终输出的结果反映，会受到一系列加工过程的影响，而脑电指标反映了实时的加工。基于已有研究的结果和本研究的分析，我们倾向于认为英汉双语者在英语的加工中，其神经层次上仍然保留了对英语音素的敏感性。而行为上未表现出音素效应的原因比较复杂，比如同时呈现的语音相关条件，以及作为二语的汉语对母语英语的影响，需要在下一步的研究中进行考察(见讨论 P. 18)。

关于 ERP 结果，英英组与英汉组相同时间窗口的音素效应，其中报告的效应在 FDR 校正前均为显著，得到校正后报告为边缘显著。我们在文章中补充分析了方差分析的效果量大小(见 ERP 分析 P.12)，根据方差分析效应量 Cohen's f 的评价标准，0.10 为的低效应，0.25 为中等效应，0.40 为高的效应(Cohen, 1988) (见脚注 2)。结果发现，所有条件上的效应量值 $Cohen's f \geq 0.51$ ，即具有较高的效应量，表明自变量引起统计学上的显著性是可靠的。进一步结合时空微态分割结果，我们认为在音韵编码阶段英汉组的脑微态成分更类似于英英组而非汉汉组，支持了同化机制。

意见 5：作者将英英组无显著音素效应解释为二语学习对一语加工过程的影响。这是否间接地反映了顺应机制的存在？如果这一解释成立，那么此英英组并不是英汉组的合理比较对象。应该考虑加入一组英语母语被试(新英英组)，排除汉语学习对英语加工可能产生的影响，再对英汉组和新英英组进行被试间比较。

回应：感谢审稿专家的宝贵意见。有关英汉双语者加工二语的同化于顺应机制研究，针对的是二语的加工策略采用了母语还是二语的加工机制，并未涉及到二语对一语的影响。如果同化指的是用相似的策略加工两种语言，那么我们认为母语受到二语的影响表现出的是同化机制。这一问题比较复杂，且不是我们研究的主要目的，我们对此未做深入讨论。

此外，我们同意您的看法，由于双语者的两种语言之间存在相互影响的可能，在考察二

语加工同化与顺应的机制时，更为理想的情况是三组被试分别为汉语单语组、英语单语组和英汉双语组，比较英汉组与其他两组的加工机制是否相同。我们在文中增加了对此问题的思考和讨论(见 P.15)。

意见 6: 加入新英英组的另一个考虑是目前英汉组和汉汉组为被试间比较，而英汉组和英英组为被试内比较。被试间比较所呈现的更大差异性可能源自个体差异，而非语言加工的差异。将英汉组和新英英组同样进行被试间比较则更加公平，使结果更加可信。

回应: 感谢审稿专家的宝贵意见。本研究中英汉组和英英组是同一组被试，但所用实验材料不同，而英汉组和汉汉组为不同组被试，所用实验材料相同。在行为数据分析中我们采用混合线性模型同时纳入了被试随机变量和项目随机变量，在考虑被试差异和项目差异影响的基础上考察了语言和语音相关类型对英汉组、英英组和汉汉组口语词汇产生过程的影响。混合线性模型的分析结果与先前方差分析的结果类似。

需要注意的是，本研究的实验设计遵循了已有对于双语者二语加工同化与顺应机制的研究逻辑，与已有研究设计类似(例如 Tian et al., 2019; Timmer & Shiller, 2012; Sun et al., 2015)。Sun 等(2015)利用 fMRI 技术考察英汉双语者和汉英双语者在阅读母语与二语词汇过程中的脑区激活差异。其中将英汉双语者进行二语汉语词汇判断任务组(英汉组)与英汉双语者进行英语词汇判断任务(英英组)和汉英双语者进行汉语词汇判断组(汉汉组)分别进行组内和组间对比。结果表明，与英英组相比，英汉组更多的激活右侧额下回区域，该脑区的激活与汉语加工相关。也就是说，英汉双语者在词汇识别任务中利用了与母语加工类似的顺应机制，这一结果与 Liu 和 Cao(2016)元分析中的结果相一致，因此我们认为该机制受到组内或组间比较的影响较小。我们在文中增加了对上述问题的讨论(见 P.19)。

其他问题: 文中标亮部分已修改，具体如下:

- (1) 修改参考文献引用格式为“Nakayama, Kinoshita, & Verdonschot, 2016”。
- (2) 将“从词汇产生过程的研究很少”修改为“从词汇产生过程角度进行探索的研究很少”。
- (3) 将“英语较为一致的形音匹配规则”修改为“英语的形音匹配规则较为一致”。
- (4) 将“音素和音节相同的干扰词不会连续呈现”修改为“各个条件下相同的干扰词不会连续呈现”。
- (5) 将“共 54 幅”，修改为“共 50 幅”。
- (6) 将文中“语音类型”修改为“语音相关类型”。
- (7) 将 ERP 结果部分校正后 $0.1 > p > 0.05$ 修改为边缘显著。核对 F 值无误，经校正后导致 p 值较为接近。
- (8) 将“图 3 所示是音节相关、音素相关和语音无关条件下 P4 电极点上的波形”修改为“图 2 所示是音节相关、音素相关和语音无关条件下 Cz 电极点上所有被试的平均波形”。
- (9) 将“英汉组的音素效应出现在 200-300 ms 之间，位于左侧半球前部区域，英英组的音素效应出现在 200-250 ms 之间，位于左侧半球前部区域，这与英语中发现的音素效应出现的

时间一致(Damian & Martin, 1999; Malouf & Kinoshita, 2007; Schiller, 2008)”文献引用修改为“Timmer & Schiller, 2012; Timmer, Ganushchak, Ceusters, & Schiller, 2014”。

(10) 将“这一微态成分在汉汉组中为出现”修改为“这一微态成分在汉汉组中未出现”。

第二轮

审稿人 1 意见

意见 1: 修改稿有明显提高, 并较好回应了之前提出的问题。有一个建议供作者参考: 图 2 目前呈现的是所有被试的平均结果。鉴于文章聚焦于三组条件的对比, 这种平均结果提供的有效信息有限。建议分为三种条件呈现, 读者可以直观看到不同条件的模式。

回应: 感谢审稿专家对修改稿的肯定和修改意见。文中图 2 包括了三个图形, 从左至右分别表示了汉汉组(左)、英英组(中)和英汉组(右)在不同时间窗口内三种条件下在 Cz 电极点上的平均波形图, 以及相应的时间窗口内的地形分布图。在图注中明确说明: 波形图中的阴影部分表示的是统计达到显著的效应 ($p < .05$), 分别为汉汉组音节效应(250-300 ms) (左)、英英组音素效应(200-250 ms) (中)和英汉组音素效应(250-300 ms) (右)。原文表述不当造成了误解, 已进行了修改, 做了详细说明(见正文 P.13)。

审稿人 2 意见

意见 1: 文中提到的音段 (segments) 实际上是次音节单元 (sub-syllabic units)。音段 (segments) 这一概念的一般用法可参考 Eric Bakovi? (2014) Phonemes, segments and features, *Language, Cognition and Neuroscience*, 29:1, 21-23, DOI:10.1080/01690965.2013.848992。文中对此问题的解释为: “音段(segment)在汉语中指的是亚音节语音表征单元(subsegmental units)(Bakovi & Eric., 2014; Wong & H.-C. Chen, 2008, 2009)。”其中“亚音节语音表征单元”的英文翻译是“subsegmental units”吗? Bakovi & Eric., 2014: “Features are thus subsegmental units of phonological representation.” 此处引用的三篇文献真的支持文中 segment 的用法吗? 如有的话, 请在回复中摘出文献中的原文。

回应: 感谢审稿专家的宝贵意见, 我们根据您的建议核对了关于音段表述的文献引用。

Bakovi 和 Eric(2014)认为音段是语音表征的中心单元, 与辅音和元音的组合相对应。原文表述为: “what phonologists call phonemes, but rather what we call segments (or, as H. more properly calls these on p. 15, “individual speech sounds”): a central unit of representation corresponding — though sometimes only roughly — to the union of what we call consonants and vowels” (P.21, 正文 6-11 行)。

Wong 和 Chen(2008)认为音段是一种亚音节成分, 在音素组成数目上少于音节。原文表述为: “we examined how lexical tone, subsyllabic components (onset and rhyme), and segmental

syllable, as well as their different combinations, are processed in Cantonese monosyllabic word production” (P.1183, General Discussion, 1-4 行)。

Wong 和 Chen(2009)在认为音段是起始音素辅音和元音组合, 或元音和辅音组合的韵。原文表述为: “Furthermore, the null effect observed in the VT-related condition, together with the facilitation effects obtained in the CV-related and the VC-related conditions, indicates that the lexical tone is represented or processed differently from segmental units” (P.891, General Discussion, 18-22 行)。

根据上述表述, 我们修改了有关音段的阐述。具体修改内容为: “音段(segmental units)在汉语中指的是亚音节语音表征单元(sub-syllabic units) (Bakovi & Eric., 2014; Wong & H.-C. Chen, 2008, 2009)”(见脚注 1)。

意见 2: “Wang, Wong, ~~WK~~, Wang 和 Chen(2017)利用图图干扰范式结合延迟图画命名任务”中所说的图图干扰范式并不恰当, 因为被试对先后呈现 prime 和 target 图片均需进行命名, 并不是只命名 target 图片。

回应: 感谢审稿专家的宝贵意见。正如您所言, Wang, Wong, Wang 和 Chen (2017)的研究中在关键试次中先后呈现两幅图片, 两幅图片存在音节相关、音段相关或无关关系。要求被试在看到第一幅图片和第二幅图片(目标图)时均先不要命名, 当提示线索出现时才对图片进行口语命名(即延迟命名任务)。研究在目标图画呈现后的 200-400 ms 与 400-600 ms 发现了音节相关效应, 却未发现任何音段相关效应。我们在修改稿中详细地阐述了这一任务, 并做了相应的修改(见正文 P.5)。

意见 3: “也就是说, 英汉双语者在词汇识别任务中利用了与母语加工类似的顺应机制, 这一结果与 Liu 和 Cao(2016)元分析中的结果相一致, 因此我们认为该机制受到组内或组间比较的影响较小。”一句中“利用了与母语加工类似的顺应机制”存在歧义, 请修改。

回应: 感谢审稿专家的宝贵意见。我们根据您的建议修改了有关表述。具体内容修改为“英汉双语者在词汇识别任务中采用了与母语类似的加工方式, 即表现出顺应机制”(见正文 P.19)。

第三轮

审稿人 2 意见

意见 1: “英汉双语者在词汇识别任务中采用了与母语类似的加工方式, 即表现出顺应机制”中“与母语类似”是否应改为“与二语类似”? 待此处明确之后, 推荐发表。

回应: 感谢审稿专家对修改稿的肯定和修改意见。我们根据您的建议核对并修改了有关表述。具体内容修改为“英汉双语者在词汇识别任务中采用了与二语类似的加工方式, 即表现出顺应机制”(见正文 P.19)。