

《心理学报》审稿意见与作者回应

题目：双语即时切换下非目标语言语音和语义的激活状态

作者：孟迎芳；林无忌；林静远；蔡超群

第一轮

审稿人 1 意见：

意见 1：本研究采用即时关联启动范式，探讨了非熟练中英双语者在目标语言（启动词）加工过程中非目标语言的激活层面。研究有一定的价值。但该文目前还存在以下问题：

第一，实验范式的创新性需进一步说明。先前多数研究都采用即时启动范式探讨双语者非目标语言的激活层面，但这些范式容易受到被试加工策略的影响，所以最新的研究中研究者开始使用基于内隐记忆的长时重复启动范式来研究双语者非目标语言的加工层面问题（如王瑞明等人，2011）。本研究又用回了即时启动范式，那么如何避免被试的加工策略的影响，研究者需要进一步的说明。

回应：非常感谢专家的意见，我们对前言部分进行了重新的梳理，针对以往研究中存在的问题提出了我们的思考，并在此思考基础上提出了本研究的实验设计。在实验中为了避免加工策略的影响，我们进行了如下设置：第一，采用了与翻译对等词有关联的其它词作为目标词；第二，采用了内隐关联启动范式，要求对启动词和目标词进行完全相同的非记忆任务，因启动词和目标词没有一一对应关系，并且与王瑞明等人的长时重复启动范式类似，被试也不知道启动刺激和目标刺激之间的联系，因此在一定程度上可减少被试加工策略的使用。对实验范式的具体说明请见前言中的第八段。

意见 2：研究结论的合理性需要进一步说明。为了探讨语言理解过程中非目标语言在词汇层面的激活状态，实验 1 要求被试对呈现的启动词和配对词进行语音判断。如果是中文词，按照声调进行归类；如果是英文词，按照单元音或多元音进行归类。这种语音加工任务也许干扰了非目标语言的语音加工，所以没有表现出语音效应。另外，该研究没有考察词汇层面其他加工是否产生“启动效应”，不能简单地将语音层面的实验结果推论到整个词汇层面，根据语音没有激活就认为整个词汇层面没有激活显得证据不足。特别是本研究根据阴性结果（语言相关和无关条件没有显著差异）来得出结论也是值得推敲的。

回应：确实如专家所说，我们只进行了语音层面上的实验，因此将本研究结果推论到整个词汇层面上有扩大化的嫌疑。但在总讨论部分我们也分析了前人在词汇其它层面的研究结果（如王瑞明等人，2011），这些结果也是间接支持本研究结果的，即如果只对目标语言进行词汇层面的加工，非目标语言的词汇层面是不会被激活。具体见总讨论的第四段。

针对专家的意见，我们也对文中词汇层面的描述进行了修改，尽可能避免推广扩大化。具体的更改见文中红色字体部分。

意见 3：实验材料和程序的严密性需要进一步说明。本研究实验材料的安排采用的是项目间比较，如何控制项目间的差异对实验结果的干扰，作者应进一步说明。比如英汉切换语义相关词对（donkey-蝌蚪）和英汉切换语义无关词对（eagle-浴室）在目标词上本身就有差异，如何保证目标词的差异是实验条件造成的，而非词汇本身差异造成的。作者为何不使用项目

内的比较，即使用同样的目标词，通过启动词的不同形成不同的实验条件。另外，L1-L2 转换中，语音相关和无关的差异体现为是否包括中间的元音，而 L2-L1 转换中，语音相关和无关的差异体现为发音是否完全相同，只是字型不同，两种转换条件明显不是完全对应的。作者在程序上还应该进一步明显因变量指标，明确项目间差异的控制，并补充项目检验的结果。另外，实验材料介绍时每种实验条件都应给出材料举例。

回应：本研究中主要通过相关组和无关组在行为或 ERP 上的差异来探讨非目标语言的激活状态。为了确保两组在材料上的平衡性，我们对中英文单词的长度，及语音上都进行了相应的平衡，同时，在文中也补充了两个实验各种条件下材料差异性的检验结果，结果均显示差异不显著，具体见两个实验的实验材料部分。因此我们认为各组材料在一定程度上是同质的。另外，我们也计算了两个实验中启动词的反应数据，如下表所示，然后把相同语言的启动词和两类配对词的反应时进行单一因素的方差分析，结果发现，在实验 1 中，不论是中文 [$F(2,34)= 1.754, p=0.188, \eta^2=0.094$] 还是英文 [$F(2,34)= 0.821, p=0.419, \eta^2=0.046$] 都没有发现反应时差异，说明对三组词的语音判断反应是一致的。在实验 2 中，中文的反应时差异显著 [$F(2,30)= 17.826, p<0.001, \eta^2=0.543$]，多重比较表明，差异主要表现在相关配对词和启动词 ($p=0.001$)，以及无关配对词 ($p<0.001$) 上，但启动词和无关配对词之间不存在反应时差异 ($p=1.000$)。在英文上的差异结果是一致的 [$F(2,30)= 36.142, p<0.001, \eta^2=0.707$ ；相关配对词 vs 启动词： $p<0.001$ ；相关配对词 vs 无关配对词： $p<0.001$ ；启动词 vs 无关配对词： $p=1.000$]。因此启动词与无关配对词上不存在反应差异可在一定程度上说明这些词组在性质上应该是同质的，而与相关配对词之间的差异是源于语义上的相关所导致的。

读音判断任务中各类单词的平均反应时 (M±SE)

		L1-L2 切换组		L2-L1 切换组	
语音关联类型	中文启动词	英文配对词	英文启动词	中文配对词	
关联词对	1020 (155)	1196 (182)	1208 (159)	1054 (144)	
无关词对		1230 (200)		1056 (144)	

语义判断任务中各类单词的平均反应时 (M±SE)

		L1-L2 切换组		L2-L1 切换组	
语义关联类型	中文启动词	英文配对词	英文启动词	中文配对词	
关联词对	730 (24)	771 (34)	921 (30)	647 (25)	
无关词对		905 (32)		733 (27)	

(注：由于关联词对和无关词对中的目标词反应是相似的，因此在数据分析中将其合并。)

对于专家所提出的项目内比较，因为实验为被试内设计，担心同样的目标词出现多次会造成重复所带来的混淆效应，因此没有采用。

另外，根据专家的建议，在实验材料介绍时补充了每种实验条件的材料举例，具体见实验材料的红色部分。

意见 4：统计结果的呈现问题。统计结果中应该完整的呈现相关统计结果，并给出统计检验值，而不能只报告有差异的结果。另外，在进行 ERP 数据分析时，应该给出波幅的描述统计表，同时每个实验都要有不同条件的波形图。实验 1 的结果中，从图 1 看，L1-L2 中两种语音条件在波幅上似乎是有差异的，建议作者再仔细看一下。

回应：根据专家的意见，我们在文中补充了没有差异的统计结果，以及两个实验的 ERP 平均波幅表，具体见两个实验的结果部分，以红色字体显示。同时也添加了实验 2 的波形图，具体见图 2。

此外，我们对实验 1 的数据进行了再次分析，并仔细检查所有结果，发现在两个时间段都有配对词类型和脑区的交互作用，但进一步的简单效应检验并在任何脑区发现有两类配对词的差异，因此做出了无配对词类型差异的结论。

意见 5: 讨论还需进一步丰富。该研究沿袭了王瑞明（2011）等人的思路，从词汇和语义两个层面分析语言理解中非目标语言的激活情况。相较于之前的研究，该研究在实验任务、实验材料等方面进行了调整，得到的实验结果也存在一些差异。针对实验结果的差异性，作者进行了一定的解释和讨论，但尚不够全面。例如，该研究采用的被试群体是非英语专业的中英双语者，而之前研究采用了英语专业的中英双语者，语言背景和双语使用情况是否会影响实验结果。另外，本研究关注的是语言理解中的语音层面和语义层面，王瑞明等人（2011）关注的是语言理解中的字型层面和语义层面，不同的词汇层面是否也有差异。这些都需要进一步讨论。

回应: 非常感谢专家的意见，让我们对实验的结果进行了更深入的思考和分析，并对每个实验的讨论和总讨论部分进行了调整和完善。具体为每个实验的讨论部分主要针对该实验的结果进行一个小结，主要的理论解释放在总讨论部分，以使论文看起来更为协调一些。而在总讨论中，针对词汇层和语义层，尤其是词汇层与前人研究结果的差异，我们进行了详细的分析和解释。在词汇层，我们提出，前人研究中有发现词汇层面的激活一方面可能源于语义激活的干扰，另一方面可能源于词汇层面的被动激活，而不是自发激活，因此如果目标语言只是加工到词汇层面，非目标语言的词汇层面应该是不会自动激活的。在语义层，我们提出，目标语言语义加工过程中，非目标语言语义会被自动激活，但激活的程度存在着双语切换方向上的非对称性。而 ERP 上，不同切换方向的 N400 差异在脑区和时间进程上都存在差异，说明这种非对称性是源于双语在接通语义表征时的加工差异，这种差异与被试双语的熟练程度是有关联的。

审稿人 2 意见:

意见 1: 题目不清晰。

回应: 已将题目修改为“双语即时切换下非目标语言语音和语义的激活状态”，希望能更准确地表达本文的研究内容。

意见 2: 摘要不清晰

回应: 已按专家提出的具体要求修改了摘要，以使其突出本文的研究目的、研究方法、研究结果及结论。具体见摘要中红色部分。

意见 3: 前言脉络不清晰，尤其是本研究的创新点没能突出，对前人研究的分析不够深入。

回应: 根据专家的意见，我们对前言部分进行了认真的修改，添加了部分前人研究结果，并进行深入的分析，在此基础上提出本研究的思路，以突出本研究与前人研究不同的地方。具体见前言中的红色部分。

意见 4: 存在格式不规范。

回应: 非常感谢专家的细致批改，对我们疏忽造成格式上的不规范表示歉意，并力求做到无格式上的差异问题。但可能有时会因电脑版本不同，个别地方会显示出不同的文字格式，对此造成的不规范还请专家多多谅解。

意见 5: 进行实验 1 和 2 的逻辑不够强。

回应: 非常感谢专家的意见,我们重新整理了两个实验之间的关系,并在前言的第七段和第八段添加了对两个实验设计的说明,两个实验希望从不同的加工层面来探讨非目标语言在语音和语义层面的激活状态,以期对非目标语言的激活状态有更全面的了解。

意见 6: 对结果的解释不够深入。

回应: 非常感谢专家的意见,让我们重新审视了我们对结果的分析 and 解释,并对与他人研究结果的差异进行了更深入的探讨,针对这些差异做出更为合理的解释,以期能够让读者更好地理解本研究的实验结果。具体见总讨论中的红色部分。

意见 7: 总讨论中,对本研究的结果与前人的结果进行分析,均欠缺很好的推论与证明。

回应: 与第 6 点的回答类似,我们对两个实验的讨论和总讨论都进行了修改。具体见文中的红色部分。

意见 8: 整篇文章,措辞太口语化,不符合学术文章的口吻。

回应: 已按专家的具体要求对一些描述进行了修改,使其更符合学术文章的口吻。

审稿人 3 意见: 该文通过两个实验探讨了非熟练中英双语者在目标语言(启动词)加工过程中非目标语言的激活层面及其对相关配对词(与非目标语言存在语音或语义关联的词)产生的影响,在语音判断任务中,非目标语言在语音层面上没有得到自动激活,而在语义判断任务中,非目标语言在语义层面上有自动激活,激活程度与语言的熟练程度有关。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: 实验 1 行为实验部分,应该报告配对词类型主效应及二因素交互作用的 F 值及 p 值,即使没有达到显著水平。ERP 部分,也应该报告没达到显著的检验相应的 F 值及 p 值,以帮助判断其效应的情况。另外,请检查 ERP 分析中 F 值自由度是否正确“切换条件×脑区的交互效应显著[F(1.503,27.050)=6.977]”。

回应: 已按专家的要求在文中补充了没有差异的统计结果。另外,ERP 分析中 F 值自由度给出的是当 Mauchly 球形检验不满足时使用 Greenhouse-Geisser 法校正后的自由度。

意见 2: 作者所言的熟练程度差别,似乎与母语/二语的差别有所混淆。母语和二语并非仅有熟练程度的差别,其表征结构本身可能就不同。因此用熟练程度来说明母语/二语的差别似乎不准确。

回应: 非常感谢专家的意见,确实本实验中英语与汉语并非仅有熟练程度的差别,其表征结构本身也不同,英语是拼音文字,而汉语是表意文字。因此在前言和结论解释中添加了中英文结构差异的说明,以及切换方向上非对称性差异的解释。

第二轮

审稿人意见: 作者根据审稿意见进行了认真修改,建议在讨论部分再适当扩充一下非目标语言和非加工语言两种不同操作可能造成的结果差异,以及为何有的实验中词汇层面有激活,有的实验中词汇层面没有激活,对原因做进一步的说明,从而更好的提升本研究的理论价值。

回应：非常感谢专家对本文修改内容的肯定，针对专家建议中提出的两点，我们在讨论部分都进行了适度的扩充，“为何有的实验中词汇层面有激活，有的实验中词汇层面没有激活，对原因做进一步的说明”具体见总讨论第三段中的蓝色部分；“非目标语言和非加工语言两种不同操作可能造成的结果差异”具体见总讨论第四段中的蓝色部分。

第三轮

审稿人意见：作者经过几轮修改，目前文章逻辑清楚，数据可靠，可以接受。但文章经过多轮修改后显得过于冗长，目前正文部分已一万三千余字，语言略显啰嗦，请作者进行删减，尤其前言和讨论部分，尽量言简意赅，总分与主体结果关系不大的理论背景可从略，但同时仍需将论点阐述清楚，将全文删减至 8000-9000 字左右。

回应：已按编委专家意见将正文重新梳理，最后字数为 9387 字。其中重点删减部分为前言和讨论。修改较大的部分如下：

1、前言部分简化了理论背景的描述，把重点放在非目标语言激活层面相关研究的描述以及述评上，并在此基础上提出本文的研究内容。

2、结果部分去除了两个实验的波幅描述统计表。考虑到两个实验都有相应的波形图可供读者对应统计结果，去除波幅描述统计表不会影响读者对论文的理解。

3、去除了每个实验后面的小讨论部分，直接在结果后总结每个实验的结论，使论文看起来更简洁。

4、总讨论部分对“非目标语言和非加工语言两种不同操作可能造成的结果差异”这一段的解释进行了重新梳理，重点突出非加工语言和非目标语言之间的差异所带来的混淆。

5、总讨论部分对“语义层面激活的非对称性解释”进行了重新梳理，将原来的两段合为一段，使之更为简洁。

6、总讨论部分将最后一段修改为本文所获得的主要结论。