

《心理学报》审稿意见与作者回应

题目：二语语音辨别能力个体差异来源：来自 ERP 研究的证据

作者：范若琳 莫雷 徐贵平 钟伟芳 周莹 杨力

第一轮

审稿人 1 意见：

意见 1：从论文来看，三个实验应该是同时进行的，那么这三个实验在进行时的前后顺序怎样，是否会存在着一些非处理条件的其它效应，如练习效应，从而使两组在所有语音分辨上都表现出成绩差异显著。

回应：三个实验的确是一起做的，两组被试都采用拉丁方的方式平衡材料，每个实验中间间隔 5 分钟。

意见 2：从实验结果来看，PP 组只有在母语实验中诱发了可靠的 MMN，而在另外两个实验都没有发现 MMN，而 GP 组在三个实验中都诱发了可靠的 MMN，但作者对三个实验的结果都是单独进行分析，即使在讨论中也未有整合。

回应：审稿人提的意见很好，修改稿已在讨论部分加入了对本实验结果的整合性讨论，详见讨论部分 4.1 第 1 段

意见 3：作者针对 Díaz 等 2008 年实验中纯音的设计不易于得出相关的结论，从而在该实验中改为复合音，并在复合音中发现了 GP 组和 PP 组之间的差异，从而说明两组的差异来源于一般声音感知能力的差异。但如果实验中能补充一个纯音的设计，将会对该问题有更好的说明。

回应：的确如审稿专家的建议，如果补充一个纯音的设计，对于整个二语语音辨别差异的问题有更好的说明，只是本研究旨在从复合音出发，首先解决这种辨音差异是否真的与一般声音无关这一研究问题。我们的研究团队也认为 Díaz 等 2008 年实验中所采用纯音材料在设计上也有一些值得商榷的地方，正在计划开展后续的研究。

意见 4：实验二中分析了 LDN 成分，但该成分不论在前言还是讨论中都未加以说明。

回应：审稿人提的意见很好，已在讨论部分进行了补充说明，详见讨论部分 4.1 第 1 段。

意见 5：讨论中的 4.2 被试在声音加工过程上的差异分析，这点分析在论文中显得有点多余，且未对论文主题产生重要作用，建议删除。

回应：已按照专家的审稿建议删除了该部分。

审稿人 2 意见：

意见 1：在实验结论逻辑上存在问题。Díaz 等的研究比较并发现两组被试在简单纯音加工上不存在差异，从而得出 L2 语音区分能力高低只存在于语音感知方面，而一般声音感知没有问题。不同的是，本研究所用的一般声音材料是类元音复合音，在许多方面基本上保留了语音的信息（见作者自述“该声音保持了原语音的很多声音特征。”）。虽然研究发现了两组被试在这种材料上存在差异，也很难得出“二语语音辨别能力的个体差异...，也表现为一般声

音层面的辨音能力差异。”的结论，至少应该对所谓的“一般声音”进行明确的界定，但显然作者并没有对此进行适当的讨论，究竟实验 2 所采用的音有多大程度上更类似于语音？

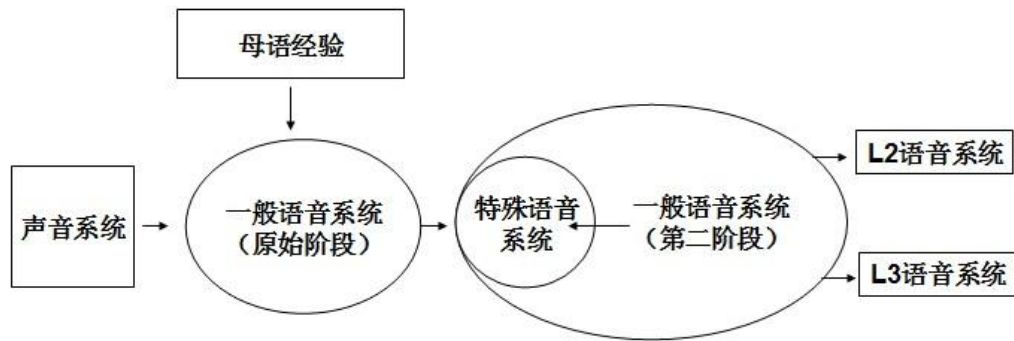
回应：审稿专家所提的问题非常重要，确实，如果本研究中采用的类元音实验材料和语音没有本质差异，那么本研究的结论逻辑就存在问题。我们的原稿对这个问题可能还没有表述清楚，现在已做了修改。本研究采用的声音材料是类元音复合音，将这种通过合成的类元音而成的复合音材料作为一般声音材料，是国际同类研究中比较常用的，例如，近年来对阅读障碍者(Mody et al. 1997; Molfese et al. 2004)，自闭症者(Čeponienė et al. 2003)以及正常被试(Jacobsen, Schröger, & Alter, 2004)的声音辨别研究中，均采用这种材料作为一般的声音材料。并且，本实验正式实验前，我们也通过预实验要求 28 名被试听该声音材料，然后判断所听到材料是否是语音。结果发现，28 名被试都判断为“否”，进一步检验了本研究中所采用的类元音并非语音而只是一般的声音材料。

我们在修改稿进一步补充了关于对这个实验材料的介绍，见 1 前言部分倒数第二段、3.3.2 实验材料的介绍等。

从原理来说，一般声音是包含语音以及之外的所有声音，声音（除了语音外）在复杂程度上又分成两种声音，分别是纯音和复合声音。纯音是只有一种频率的最简单的声音，而复合音是由多个频率的声音复合而成，包括基音和泛音，人们在自然情境中听到的声音大部分都是复合音。本研究中采用的类元音材料之所以是区别于语音的一般声音材料，是因为该复合音材料是用元音的基音和泛音相同频率的正弦波叠加制作而成，又被称为“元音的类似物”。元音和复合音的主要不同在于其频谱包络的形状，元音频谱包络的形状主要由其共振峰频率和相应的带宽决定，复合音由于只采用了简单的正弦波叠加，而正弦波为等幅度且理论上的带宽为 0，其包络与正常语音相差很大，与语音明显不同。

意见 2：讨论部分的一些观点值得商榷。比如，“个体的一般语音系统首先是以个体的母语语音系统为基础发展起来的”，以及“...形成母语的语音系统，也就是一般语音系统的雏形”这个观点来自于何处？诡异的是，后续的讨论中引用了 Kuhl (2004) 的 NLNC 假说所持的观点恰恰与作者所说完全相反。该假说认为婴幼儿早期对语音的感知能力非常敏感，几乎能够区分任何语音，但是随着母语经验的输入，以及其他语音经验的缺失，从而渐渐地只对母语语音敏感，而对其他非母语的语音变得不再敏感。那么，到底如何去讨论母语语音系统和一般语音系统之间的关系，还需要更为周全的考虑。至少，目前的讨论（包括那个示意图）还存在明显的问题以及过于随意之嫌。另外，所使用的芬兰语中的两个元音与汉语或者粤语中的元音有多大程度上的相似性，文章也没有给出相关信息，虽然芬兰语和汉语分属不同语系，但未必这种跨语系的差异就一定大于同语系的加泰罗尼亚语和芬兰语。

回应：感谢审稿专家的宝贵意见，原稿关于一般语音系统与母语语音系统之间的关系的分析确实存在问题，修改稿对母语语音系统和一般语音系统的讨论进行了修改和补充，详见讨论部分 4.1 第 3 段。示意图修改如下：



另外,关于不同语系语音的问题,审稿人提出的问题也是非常重要的。我们认为,由于在 Diaz 等的研究中是西班牙母语的被试,所用的非母语材料是芬兰语语音对 (/o/-/ö/),由于西班牙语与芬兰语属于同一语系,因此,根据 Diaz 等的研究结果,只能得到个体在同语系语音之间具有辨音能力的一致性,还不能认为是一般语音系统的一致性。本研究用了汉语母语的被试,而采用了不同语系的芬兰语作为非母语语音,根据本研究结果,才能得出个体在不同语系语音之间也具有辨音能力的一致性的结论,或者说是在一般语音系统中具有辨音能力一致性的结论。至于两种语言之间的语音差异,可以从多种维度(包括物理维度)进行分析,很难认为相同语系的两种语言之间差异就会比不同语系的两种语言之间差异会小,因此,本研究不对这个问题进行分析,只是根据研究结果提出个体在不同语系的两种语言中也表现出辨音能力的一致性,从而对 Diaz 等的研究做出重要的补充。

意见 3: 前面也提到,文中并没有明确给出所采用材料具体信息,请用图示展示所用刺激的频率、能量、共振峰等信息,并详细阐述类元音和元音的异同,以及如何制作。引用文献或资料详细论述类元音和元音在物理属性以及感知觉属性上的特点。

回应: 审稿人提的这个问题。详细内容已在问题 1 中回答了。现在再进一步作补充:这两种材料最大的区别就来自于不同的制作原理,本实验采用的语音材料主要是用 Praat 语音合成软件制作的,而复合音则是采用了简单的正弦波(与相应语音 F0-F4 的频率相同的正弦波)叠加,其包络与正常语音相差很大,因此在听觉上会有明显的不同。

意见 4: 希望能够看到平均后的 ERP 波形(而不仅是差异波),以及多个电极情况。

回应: 已在修改稿中增加了电极点 F3、F4、Fz 和 LM 的原始波形图。图 1 实验 1a(母语语音实验),图 3 实验 1b(第三种语言语音实验),图 5 实验 2(复合音实验)。

意见 5: 文中的一些术语或概念容易混淆,比如摘要里“一般复合音”、“声音条件下”、“一般语音层面”、“一般声音层面”,以及前言里的“一般语音辨别能力”和“特殊语音辨别能力”;前言和讨论里提到“完全没接触过的三语(L3)”,没有接触过还叫 L3?

回应: 如审稿专家所言,原文在这些术语或概念的界定上不够清晰,修改稿中进一步明确(详见前言部分第 6 和第 7 段):一般语音辨别能力(language-general)是指普遍的语音加工能力,不受具体语言的影响,而特殊语音辨别能力(language-specific)则指与具体语言相关的语音辨别能力(Minagawa-Kawai, Mori, Naoi, & Kojima, 2006)。一般声音系统是包含语音以及声音在内的整个普遍心理声音的处理系统。而声音(除了语音在外)在复杂程度上又分成两种声音,分别是纯音和复合声音。我们的研究用复合音的辨别能力为指标研究被试的一般声

音系统。此外，修改稿将“完全没接触过的三语（L3）”改为更准确的“第三种语言”。

意见 6：一些说法上有待商榷。比如，“二语语音辨别的个体差异是源于一般语音辨别能力还是源于特殊语音辨别能力”，由于未能回答因果关系，是不是只说差异存在于哪个方面更合适些？

回应：如审稿专家所言，本研究没有回答因果关系，用“差异存在于哪个方面”很合理，前人的相关研究采用了“差异来源”的说法，他们认为一般声音加工系统是包含语音和声音的一个普遍的心理声音加工系统，所以假设，如果个体的二语语音能力只是依赖个体的语音加工机制，那么个体的能力差异就只会发生在语音材料上。相反，如果个体的辨音能力在声音材料上也有差异，就可以说明个体之间不同的二语语音能力是由于个体的一般声音加工机制所引起(D áz, Baus, Escera, Costa, & Sebasti án-Gall és, 2008)，因此，在综合考虑下，本研究沿用了 D áz 等人（2008）的研究逻辑而采用了这种说法。

意见 7：A'分数转换，虽然引用了文献，也应该简单介绍下如何转换。

回应：答：A'分数的转换是一种基于信号检测论的统计方法，根据被试完成任务的击中率 $P(S | s)$ 以及虚报率 $P(S | n)$ 查表所得，该表在文献 (McNicol, 1972) 的附录部分。计算公式如下：

$$P(\bar{A}) = \frac{1}{2} \left(\frac{x^3 + x^2(y-2) - x(y+3)(y-1) - y(y-1)^2}{2x(1-y)} \right)$$

式如下：

$$x = P(S | s) \text{ and } y = P(S | n).$$

意见 8：语言上还需要进一步润色，包括英文摘要部分。比如标题“The origin of different perception in the sounds of a second”，显然不清楚。

回应：修改稿更加注意语言的润色问题，英文标题修改为“Electrophysiological evidences for general phoneme and sound discrimination reveal the origin of individual differences in the sounds perception of a second language”。

第二轮

审稿人 1 意见：

意见 1：一般声音分为纯音和复合音，D áz 等人（2008）的研究表明 GP 和 PP 组在纯音上无差异，提出二语语音掌握的个体差异仅源于语音加工系统差异，而本研究表明 GP 和 PP 组在复合音上有差异，提出二语语音掌握的个体差异源于一般声音加工系统差异。那么这两种不同的结果作者如何解释。由于本实验是在 D áz 等人（2008）的基础上提出的，因此这点解释是极其重要的。另一方面，从作者对复合音的说明来看，其复合音的设计与语音的区别并不明显。如“而语音是由多个频率的声音复合而成的复合音”，而作者在文章只提到“在听觉上会有明显不同”，建议应把作者对复合音的前期测试结果放在正文中。

回应：审稿人提的意见很好，作者在文章中的讲述的确不够清楚。已经在文章讨论部分做出了修改。在 D áz（2008）的语音研究中，也考察了被试的声音辨别能力，但是他们选择用纯音材料来考察被试的一般声音加工系统，并且发现 L2 辨音水平不同的被试对于纯音的敏感性是没有显著差异。然而一般声音除了纯音还有复合音，因此纯音没有差异不能说明两组被试对于声音的辨别能力就是没有差异的。并且本研究还认为纯音是一种最简单的声音，在自然状态不常听见，不是考察被试一般声音辨别能力最佳的代表，所以就选择用复合音，一种生态效度更高的声音材料，来考察两组被试的一般声音加工能力。结果发现，在复合音实

验中，两组被试在复合音感知能力上同样存在普遍差异，GP 感知声音的能力要比 PP 强。因此，作者认为二语语音辨别能力的个体差异不仅可以追溯到一般语音层面的辨音能力差异，而且可以追溯到一般声音层面复合音的辨音能力差异。另外，已经按审稿人的意见将复合音的前期测试结果放在正文中了，见原文 2.2 实验材料。

意见 2: 如前审稿专家所述，两个实验是同时进行，且设计上完全相同，只是材料不同，是否应整合在一起说明，一方面可使文章更简洁，另一方面，可进行不同材料的 MMN 比较，这对于论文目的的说明更有说服力，且作者在文章中也应对实验的顺序进行说明，这可能会影响到实验结果。

回应: 已经按审稿人提的意见将两个实验合并。见原文 2 方法、3 结果 等部分。实验的顺序的说明详见 2.3 实验程序部分。

意见 3: 在讨论中，作者在理论解释的基础上提出“母语辨音能力为个体的一般语音系统的形成和发展奠定了基础，而这种一般语音系统会进一步影响个体学习其他语言”，这一结论并不合理，因实验中并不能确定几种能力之间的因果关系，并且作者前面又提到“个体的母语语音是从原始一般语音系统中分化并壮大的那一部分语音系统”，这两种的论述是相互矛盾的，也可以说明是一般语音系统的能力差异为母语以及其它语言的学习奠定基础。因此本文中的讨论有对结果扩大化之嫌。

回应: 已经按审稿人的意见大幅修改了讨论。见原文 4 讨论部分。

意见 4: 图 7 应为图 4，该图为个体语音辨别差异来源图，但并不能很清晰地体现差异究竟来源于何处，更类似于个体语音系统的发展过程。

回应: 审稿人的意见很好，该图不能很好地反映文中观点，修改稿已删去改图。见原文 4 讨论部分。

审稿人 2 意见:

意见 1: 最重要的一点，讨论部分逻辑不清楚，甚至有点儿混乱，应该重新组织。另外，那个图建议去掉，没有给人太多的信息，而且未必准确。我建议讨论部分能否分这样两个问题讨论：一个是讨论二语语音能力的差异来源；一个讨论一般声音感知与语音加工之间的关系。

回应: 非常感谢审稿人的宝贵意见，已经按审稿人的意见重新组织了讨论部分。见原文 4 讨论

意见 2: 讨论部分关于早期 MMN 和晚期 MMN（或者 LDN）的讨论还有待商榷，作者说早期 MMN 是浅加工，晚期 MMN 是深加工。是不是换种说法更合适？比如早期 MMN 反映的是声音的自动化感知，而晚期 MMN 更多与受认知控制的声音加工。可能作者还是需要结合相关的文献来讨论。

回应: 非常感谢审稿人的宝贵意见，已经按审稿人的意见结合相关文献进行了修改。见原文 4 讨论部分。

意见 3: 数据统计分析问题。首先，文章三个实验的方差分析的“被试类型”和“偏侧性”主效应检验的自由度都是错误的（其他请作者自行检查），所以不得不让人怀疑统计结果是否可信。第二，三个实验均没有发现任何与电极位置的交互作用，但是从图上差异波情况来看，至少 LM1 和 LM2 的模式是与其他电极的模式是相反的，不知道统计数据究竟是什么样的？

回应：已经按审稿人提的意见检查了统计部分，修改见原文 3 结果。第二，关于 LM1 和 LM2 的波幅模式与其他电极的模式是相反，是听觉 MMN 的一个特点 (Silvia, Naseem, Barbara, & April, 2010)，因此，在统计中是将 LM 位置的波幅值按绝对值不变、极性反转为与其他电极相同的负值来进行的，同样的做法见 (D áz, Baus, Escera, Costa, & Sebastián-Gallés, 2008)。

意见 4：虽然作者对英文摘要进行了修改，但还是有比较明显的问题，比如还是标题！

回应：已经按审稿人的意见修改了英文摘要。见原文英文摘要。

第三轮

审稿人 1 意见：

意见 1：在前一稿基础上，作者进行了比较大的修改，已经比原来有了比较大的提升。但是作者关于实验采用复合音的原因的阐述和讨论仍并不令人满意，在一些重要的方面并没有进行讨论。这篇文章的靶文章 (Diaz 等 2008) 使用了简单纯音，但它是有非常明确的目的，而且对声音的频率和时长等都有操作，以此来看两组人在这些简单的声音属性加工上是否存在差异，这是在以往研究中被普遍采用和接纳的。但本文的作者并没有提到这些考虑因素，作者强调“复合音作为一种生态效度更高的声音材料”，那么，我们为什么需要考察生态效度更高的声音材料？它与简单声音和语音之间的主要差别究竟在哪些维度上？以及这些差异可能是怎样影响人的声音感知加工的？作者在这些问题上的讨论仍有待深入。比如，前人关于复合音的研究 (Nääänen 等 1993 等) 均没有被提及。作者在将前人的研究结果结合起来讨论方面也有待加强，比如 Diaz 等使用简单纯音发现两组被试没有差异，而本研究使用复合音发现有差异，结合其他相关研究发现，共同能说明什么问题？

回应：谢谢专家宝贵的意见，原文稿在关于实验采用复合音的原因的阐述和讨论方面确实不够清晰，在认真思考并参看了专家提供的文献以及相关文献之后，我们认真对原文相关部分做了修改，首先在第一部分“问题与目的”中，对本研究的立题依据做出了修改，解释本研究为什么要专门考察复合音。修改内容如下（见修改稿第4页涂蓝部分）：

第二个问题是，根据 D áz 等人的研究结果，是否可以确定个体二语语音加工的差异真的与一般声音的加工能力无关？D áz 等人的研究为了探讨个体语音辨别加工能力的差异根源，首先对语音辨别能力不同的个体的纯音辨别能力进行考察，这是正确的，因为，纯音是最基本的音，要考察个体语音辨别差异根源，首先应该从最基本的层面即纯音层面开始。研究表明，语音辨别能力不同的个体，在纯音辨别方面没有差异，因此，个体语音辨别的差异并不是源于纯音层面辨别能力的差异，这个研究结果是有意义的。然而，本研究认为，确定了个体语音辨别差异并不是来源于纯音辨别的差异，不等于就能确定语音辨别的差异与一般声音的辨别能力无关，因为，声音有两个层面，一个是基本的纯音层面，第二个层面是复合音层面，因此，在 Diaz 等人 2008 年研究的基础上，应该进一步探讨不同语音加工能力的个体是否在复合音辨别能力也存在差异，才能确定地回答个体二语语音加工的差异是否与一般声音的加工能力无关，才能对个体语音辨别差异的来源做出确定性的结论。

纯音的辨别之所以与复合音的辨别不同，主要原因有三点：

第一，纯音与复合音在物理属性上是有明显差异。纯音是最简单的声音，具有单一频率的正弦波，它只有频率、振幅与模式三个维度的特征；而复合音则是由频率不同、振幅不同和相位不同的正弦波叠加形成的声音，由基音和泛音构成。泛音的多寡及泛音间的相对强度决定了声音的特定音色，音色由三个成分构成：时间成分，频谱成分以及两者的交互作用

(Tervaniemi, Winkler, & Näätänen, 1998; Toivainen, Tervaniemi, Louhivuori, Saher, Huotilainen, & Näätänen, 1999), 纯音是没有音色的, 音色是复合音才具有的声音属性。因此, 对复合音的辨别, 不仅包括了对声音的频率、振幅与模式的辨别, 而且更为重要的是包括了对不同音色的辨别。因此, 研究具有音色特征的复合音是很有必要的。第二, 纯音与复合音感知加工的脑机制也是不同的。有研究发现加工纯音和复合音的相同频率物理变化是在听力皮质的不同区域, 相比于纯音, 复合音的加工发生在更中间, 更深的位置(Alho, Tervaniemi, Huotilainen, Lavikainen, Tiitinen, Ilmoniemi, Knuutila & Näätänen, 1996)。这可以证明纯音加工和复合音加工是不同的脑机制。因此, 应该要分别进行研究。第三, 语音是一种特殊的复合音, 同样具有音色这一物理特征的, 有可能语音辨别能力的差异实际上就是一般复合音辨别能力差异在语音领域的表现。基于以上三点, 必须在 Díaz 等人(2008)研究的基础上进一步探讨二语辨音能力不同的个体对一般复合音(除语音外)的辨别状况, 才能确定地回答“二语语音辨别的个体差异是否与一般声音的加工能力无关”这个问题。

另外, 按照审稿人的意见, 我们在第四部分“讨论”中, 结合了 Diaz 等纯音的结果进行讨论, 见修改稿 21 页涂蓝部分。

意见 2: 另外, 现有讨论部分“这一过程正好可以反映出 GP 最终在语音上的知觉差异恰恰在过程上也是源于对其声音特征差异感受再过渡到语音感受上的。由此可以得到, 被试 L2 语音感知能力的差异可能来源于个体对于一般声音加工能力的差异。也就是说, 个体对于一般声音某些特征的加工能力会对个体加工 L2 语音产生影响。甚至可以说, 被试对于一般复合音的辨别力可以预测个体在新语音学习上的表现。”以及最后一段中的“但是据此可以设想, 也许不同语言阅读障碍儿童对母语语音辨别困难, 其实是他们在语音辨别困难并进一步在声音辨别困难的表现。”存在无依据臆想之嫌!

回应: 已经按审稿人的意见删改了以上论述。见原文 4 讨论部分。

意见 3: 再次建议作者找英语写作熟练者修改摘要, 尽量避免有一天被老外检索到的时候被笑话。文章并没有做相关分析, 英文摘要中的 *positively correlated with ...* 的说法欠妥。

回应: 谢谢专家的指正! 已经按审稿人的意见找英语写作熟练者修改了英文摘要。见原文英文摘要。

意见 4: 图 7 应为图 4, 该图为个体语音辨别差异来源图, 但并不能很清晰地体现差异究竟来源于何处, 更类似于个体语音系统的发展过程。

回应: 审稿人的意见很好, 该图不能很好地反映文中观点, 修改稿已删去改图。见原文 4 讨论部分。

第四轮

对专家以及编委审稿意见的修改:

意见: 较前一稿, 作者进行了比较大的修改, 讨论更为恰当合理。只是在摘要部分的最后一句话, 还是应该对“一般声音的辨别能力”进行合适地限定为“复合音”。另外, 建议对图进行适当编辑, 现在每张图的各个元素是分散而没有合成一体的, 从而造成了现实效果不佳, 比如图 4 的图标显示不全等。

回应: 谢谢专家宝贵的意见, 已经修改了摘要, 将“一般声音的辨别能力”限定为了“复合音”。另外, 已经按专家要求将图进行编辑, 修改了图标不全等等问题, 并将每张图的各个元素合

成一体。