

《心理学报》审稿意见与作者回应

题目：声调范畴感知模式及声调复杂度对音乐音高感知的跨领域影响

作者：李贤卓，肖容，梁丹丹

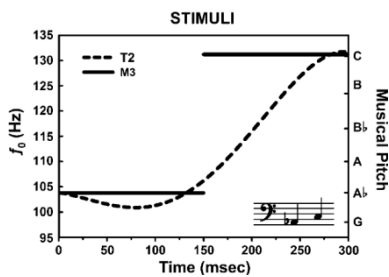
第一轮

审稿人 1 意见：本文创新地研究了声调感知与音乐音高感知的关系，被试选择得比较恰当，实验设计巧妙。但本文也存在一些问题。

意见 1：对于音乐刺激的描述只有 80 字，阅读后依然无法令读者全面地了解这些音乐刺激。

回应：感谢审稿专家的意见。初稿中只介绍了音乐刺激的生成方法，较为抽象。我们在修改稿中进行了举例说明并附上了频谱图（详见正文 2.2.2 节），这样可以为读者提供一个较为直观的了解。

以往研究为了追求语言和音乐刺激的生态效度，往往语音刺激采用真实词汇或负载声调的单音节，音乐刺激采用单个音符或音字符串。例如 Weidema 等(2016)，语音刺激使用了阴平-阳平和阴平-去声的声调连续统，但音乐刺激使用的单个音符：C4、E4、C5、G4。又如 Bidelman 等(2010)语音刺激是阳平调（T2），音乐刺激是一个从 A^b2 到 C3 的大三度音程（M3），下图为该实验的刺激：



这样做的好处是还原真实的声调和音乐音高，存在的问题是语音刺激和音乐刺激除起点和终点音高相同外，仍存在许多层面的差异，不仅内部音高不同，在共振峰等维度方面也存在差异。我们的研究设计尽可能地做到语音刺激和音乐刺激在这些变量上的统一，力求实验结果更加准确。

参考文献:

Bidelman, G. M., Gandour, J. T. & Krishnan, A. (2011). Cross-domain Effects of Music and Language Experience on the Representation of Pitch in the Human Auditory Brainstem. *Journal of Cognitive Neuroscience*, 23(2), 425–434.

Weidema, J. L., Roncaglia-Denissen, M. P., & Honing, H. (2016). Top-Down modulation on the Perception and categorization of identical pitch contours in speech and music. *Frontiers in Psychology*, 7, 1–11.

意见 2: 作者没有描述被试的音乐水平与天赋, 也没有调查被试中有没有失歌症患者。从文章呈现出来的音乐音高感知情况来看, 无法排除是否为其因素导致。

回应: 感谢审稿专家的意见。被试的音乐水平与天赋以及是否患有失歌症是可能影响实验结果的重要原因, 对此我们对被试进行了补测。音乐能力测试主要采用迷你音乐感知技能测试 (Mini Profile of Music Perception Skills, 简称 Mini-PROMS; Zentner & Strauss, 2017) (https://www.uibk.ac.at/psychologie/fachbereiche/pdd/personality_assessment/proms/take-the-test/mini-proms/), 该量表由旋律、音调、重音和节奏四个子测试组成, 耗时约 20 分钟。被试是否患有失歌症采用蒙特利尔失歌症评估量表(MBEA, Peretz et al., 2003), 该量表根据音乐加工模块理论发展而来, 分六个子量表, 分别是音阶、轮廓、音程、节奏、节拍和短时记忆, 耗时约 70-90 分钟。

由于初次实验时记录了被试多种联系方式, 我们经过努力与全部被试取得了联系。补测的两个测试全部线上视频连线完成, 主试及合作者全程指导被试完成, 两个量表初始版本为英文版, 主试及合作者分别口头指导被试填写, 或通过翻译工具翻译为中文、越南语和俄语, 并确保被试知晓测试内容。完成后付给被试酬金 50 元。

本实验在初始筛选被试时, 曾口头询问过被试是否有五音不全等现象, 因此, 补测没有发现音乐能力异常者或是失歌症患者 (具体得分见原始数据表)。对三组被试的成绩分别进行方差分析, 迷你音乐感知技能测试 $F(2,89)=0.621, p=0.54$, 蒙特利尔失歌症评估量表 $F(2,89)=1.509, p=0.23$, 三组差异均不显著。

参考文献:

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders. The montreal battery of evaluation of amusia. *Annals of the New York Academy of Sciences*, 999, 58–75.

Zentner, M. , & Strauss, H. (2017). Assessing musical ability quickly and objectively: development and validation of the short-proms and the mini-proms. *Annals of the New York Academy of Sciences*, 1400, 33–45.

意见 3：关于“接受过器乐训练”这一条，请作者详细描述一下接受过哪些器乐的训练，因为不同乐器对音高感知的要求不同。

回应：感谢审稿专家的意见。在实验中，我们记录了受过业余音乐训练者的时长、受训乐器种类等信息，以受训时长作为筛选标准，超过 3 年的被试并未纳入试验。这一标准来源于 Cooper & Wang (2012)、Bidelman & Moreno(2013)、Chang, Hedberg & Wang (2016)、Weidema 等(2016)等人研究：非音乐家指未接受过专业院校的音乐训练或接受过业余音乐训练且时长不多于 3 年者，且在过去的 5 年里，他们都没有接受过任何正式的音乐培训。

本研究中，有音乐训练经验的非音乐家受训时长为 1 周–3 年，具体如下表所示：

被试组别	被试编号	学习内容	学习时长
越南	2	钢琴	11 个月
越南	6	钢琴	1 个星期
越南	8	声乐	5 个月
越南	13	中国古筝	1 个月
中国	3	钢琴	2 年
中国	16	钢琴	3 年
中国	28	二胡	1 年
中国	29	钢琴	3 年

的确，如审稿专家指出的那样，不同乐器对音高感知的要求不同，钢琴对于音乐精度的影响可能优于古筝。但我们并未将受过业余音乐训练的乐器种类作为指标，主要基于以下考虑：第一，本研究中的被试接受音乐训练的时长较短；第二，从实验结果看，这一类人群的表现与未受过音乐训练的被试趋于一致；第三，以往研究也未对接受过一些业余音乐训练的被试进行进一步区分。未来的研究可以就这一问题招募更多被试，展开进一步讨论。

参考文献：

- Bidelman, G. M., Hutka, S., & Moreno, S. (2013). Tone Language Speakers and Musicians Share Enhanced Perceptual and Cognitive Abilities for Musical Pitch: Evidence for Bidirectionality between the Domains of Language and Music. *PLoS ONE*, 8(4), e60676..
- Chang, D., Hedberg, N., & Wang, Y. (2016). Effects of musical and linguistic experience on categorization of lexical and melodic tones. *The Journal of the Acoustical Society of America*, 139(5), 2432–2447.
- Cooper, A., & Wang, Y. (2012). The influence of linguistic and musical experience on Cantonese word learning. *The Journal of the Acoustical Society of America*, 131(6), 4756–4769.

Weidema, J. L., Roncaglia-Denissen, M. P., & Honing, H. (2016). Top-Down modulation on the Perception and categorization of identical pitch contours in speech and music. *Frontiers in Psychology*, 7, 1–11.

意见 4: 关于“汉语母语者”被试, 我们知道, 很多北方方言的声调和普通话有一定差异, 这是否会影响声调感知?

回应: 感谢审稿专家的意见。诚然如专家所言, 北方方言的声调和普通话有一定差异, 一些方言的差异甚至很大。鉴于客观条件, 为了排除方言的影响, 我们在招募被试时在北方方言母语者中优先选择普通话母语者以及日常使用语言为普通话的被试, 此外, 本研究的被试全部来自南京师范大学, 普通话已通过二级甲等。实验时告知被试, 听到的语音部分是普通话而非方言, 确保被试激活普通话的声调表征。

意见 5: pratt 是什么软件? 请前后保持一致, 包括大小写。

回应: 感谢审稿专家的意见, 这是我们的疏忽, 这里指的是语音学软件 Praat, 主要用于对数字化的语音信号进行分析、标注、处理及合成。修改稿中已经统一为 Praat。

意见 6: 请提交打包后的实验刺激和实验程序。作者在“《心理学报》论文自检报告”中提供的原始数据网盘是打不开的, 麻烦一并提供。

回应: 感谢审稿专家的意见。我们随修改稿附上了 (1)原始数据 (2)实验 Eprime 程序 (包含实验刺激) (3)蒙特利尔失歌症评估量表 (另一个迷你音乐感知技能测试量表 Mini-PROMS 见: https://www.uibk.ac.at/psychologie/fachbereiche/pdd/personality_assessment/proms/take-the-test/mini-proms/)。

.....

审稿人 2 意见:

意见 1: 专业术语不规范。引言第一段“如汉语普通话中, 有阴平、阳平、上升和去声四个声调”中“上升”应为“上声”; 文中“声调母语者”、“非声调母语者”应为“声调语言母语者”、“非声调语言母语者”等。

回应: 感谢审稿专家的意见。我们在修改稿中已改为“上声”, 全文中的“声调母语者”、“非声调母语者”都改为了“声调语言母语者”、“非声调语言母语者”。

意见 2：引言部分论述条理不清，问题不突出，略显混乱。

回应：感谢审稿专家的意见。我们对引言部分相关段落进行了改写并增加了节标题，进一步强调了前人研究的不足与本研究在这些问题上的探索。节标题如下：1.1 语言音高和音乐音高加工的“模块论”与“共享论”；1.2 声调范畴感知对音乐音高感知的影响；1.3 母语声调复杂度差异对跨领域音高感知的影响。

修改之后的引言内部逻辑如下：

首先介绍音高等关键术语，在 1.1 小节，主要介绍语言和音乐音高加工两种矛盾的观点——“模块论”与“共享论”，指出前人从语言音高到音乐音高迁移这一角度的研究多支持了“共享论”，但这些研究并未考虑到语言音高感知的领域特异性，即本研究需要讨论的两个问题——范畴感知和母语声调复杂度的影响，因此前人得出的结论可能存在问题。

接着，在 1.2 和 1.3 节，分别从范畴感知和母语声调复杂度对音乐音高感知的影响角度，回顾了前人的相关文献。在 1.2 节，主要提出声调范畴感知将一系列不同的刺激感知为有限的范畴，而音乐音高感知则需要精确感知每个刺激的区别，这是两种相反的模式，如果存在迁移，声调范畴感知模式可能会抑制音乐音高感知，而前人文献并未明确讨论过这一问题且实验设计存在缺陷。在 1.3 节，根据已有语音学研究“声调系统复杂的语言使用者有更精细的语言音高感知能力”这一结论，认为这一能力可能迁移到音乐音高感知中。

最后提出研究问题以及假设。

意见 3：2.2.1 中语音刺激 yī 用国际音标标写更加规范。

回应：感谢审稿专家的意见。我们已将所有的拼音改为国际音标标写。其中，yī 改为[i¹]，yí 改为[i⁴]。

意见 4：语音刺激连续统为什么前 100ms 没有采用平稳段，请阐述原因。

回应：感谢专家的意见。经过对前人文献的考察，我们认为设置平稳段并非必要条件。

首先，设置平稳段的理由值得商榷。例如，Wang (1976)是汉语声调感知研究中最先设置平稳段的研究，但并未说明为何设置平稳段。Chang 等(2016)的研究设置了平稳段，理由是阳平(第二声)和去声(第四声)在上升或下降之前都分别包含了一个平缓的 F0 初始阶段，因此，设置平稳段目的在于模拟真实语音。但是，众所周知，阳平(第二声)在上升前的初始阶段并非“平稳段”，而是“下降段”(沈炯, 1985; 吴宗济和林茂灿, 1989; Shen & Lin, 1991; Moore & Jongman, 1997; 石锋和王萍, 2006)，设置“平稳段”实际上并不能模拟真实语音。

其次, 平稳段对人们感知声调的作用不大。在声谱上, F0 线条是声调的标志。一些声调 F0 曲线在起始和末尾时, 会出现弯头和弯尾的音段。“这些弯头和弯尾的音段在人们声调感知过程中, 并没有成为语言音高信息, 这多半与这些弯头、弯尾的时长太短有关系, 而中部的调型段时长最长, 对人们的声调感知所起的作用也最大, 是人们感知声调的有效时长”(郭锦桴, 1993)。

第三, 我们再次回溯了前人相关的研究, 部分研究设置了平稳段, 如 Wang (1976)、Chang 等(2016)、Zheng 等(2012)、Zheng 等(2014), 这些研究中, 平稳段的持续时间也不统一。更多的研究并未设置平稳段, 如 Xi 等(2010)、Zhang 等(2012)、Zhao & Kuhl (2015)、Weidema 等(2016)、Yu 等(2019)等。

参考文献:

- Chang, D., Hedberg, N., & Wang, Y. (2016). Effects of musical and linguistic experience on categorization of lexical and melodic tones. *The Journal of the Acoustical Society of America*, 139(5), 2432–2447.
- Moore, C. B., & Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *The Journal of the Acoustical Society of America*, 102(3), 1864–1877.
- Peng, G., Zheng, H. Y., Gong, T., Yang, R., Kong, J., & Wang, W. S. (2010). The influence of language experience on categorical perception of pitch contours. *Journal of Phonetics*, 38(4), 616–624.
- Shen, S., & Lin, M. (1991). A perceptual study of Mandarin tones 2 and tone 3. *Language and Speech*, 34(2), 145–156.
- Wang, W. S. (1976). Language change. *Annals of the New York Academy of Sciences*, 280(1), 61–72.
- Xi, J., Zhang, L., Shu, H., Zhang, Y., & Li, P. (2010). Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience*, 170(1), 223–231.
- Yu, K., Li, L., Chen, Y., Zhou, Y., Wang, R., Zhang, Y., & Li, P. (2019). Effects of native language experience on Mandarin lexical tone processing in proficient second language learners. *Psychophysiology*, 56(11), 1–20.
- Zhang, L., Xi, J., Wu, H., Shu, H., & Li, P. (2012). Electrophysiological evidence of categorical perception of Chinese lexical tones in attentive condition. *NeuroReport*, 23(1), 35–39.
- Zhao, T. C., & Kuhl, P. K. (2015). Effect of musical experience on learning lexical tone categories. *Journal of the Acoustical Society of America*, 137(3), 1452–1463.
- Zheng, H. Y., Minett, J. W., Peng, G., & Wang, W. S. Y. (2012). The impact of tone systems on the categorical perception of lexical tones: An event-related potentials study. *Language and Cognitive Processes*, 27(2), 184–209.

Zheng, H. Y., Peng, G., Chen, J. Y., Zhang, C., Minett, J. W., & Wang, W. S. Y. (2014). The influence of tone inventory on ERP without focal attention: A cross-language study. *Computational and Mathematical Methods in Medicine*, Article ID 961563.

Weidema, J. L., Roncaglia-Denissen, M. P., & Honing, H. (2016). Top-Down modulation on the Perception and categorization of identical pitch contours in speech and music. *Frontiers in Psychology*, 7, 1–11.

郭锦桴(1993). *汉语声调语调阐要与探索*. 北京语言学院出版社.

沈炯(1985). 北京话声调的音域和语调, 载林焘、王理嘉编. *北京语音实验录*, 北京大学出版社.

石锋, 王萍(2006). 北京话单字音声调的统计分析, *中国语文*, (1), 33–40.

吴宗济, 林茂灿(编) (1989). *实验语音学概要*. 高等教育出版社.

意见 5: 在正式实验前是否进行了预实验?

回应: 感谢审稿专家的意见。在本研究开始之前, 我们曾经对于连续统刺激的个数选择、控制组母语的选择以及实验任务的选择进行过三项研究, 作为本研究的预实验。

第一项是《汉语母语者与非声调语言母语者的范畴感知研究》。实验材料采用阴平-阳平连续统, 但只有 7 个刺激, 按照与本研究完全相同的方法统计数据。研究发现汉语普通话母语者产生范畴感知现象, 而非声调语言母语的汉语初级学习者是连续感知。**该研究对本研究的启示:** 第一, 该阴平-阳平声调连续统可以作为正式实验材料, 但 7 个刺激所体现的音高精度信息可能过少, 如果将音高精细度纳入实验变量, 可以尝试提高连续统中的刺激数量, 因此, 本研究中, 连续统被等分为 9 个刺激。第二, 确定了本研究中, 非声调语言母语控制组采用俄语母语者。预实验非声调语言母语者被试共 15 名 (7 男 8 女), 分别来自乌兹别克斯坦 (8 人)、塔吉克斯坦 (4 人)、吉尔吉斯斯坦 (3 人), 被试全部为双语者, 另一语言为俄语。汉语学习时长为 2 个月。他们的感知结果都不属于范畴感知, 但略有差异。我们认为差异主要是个体的汉语水平不同造成的, 而非母语。为证明这一点, 我们分别将被试汉语课期末考试成绩 (代表语言水平) 与其范畴感知区分任务中的区分率进行简单线性回归分析。结果显示, 语言成绩可以解释区分结果 56% 的变化, 这说明成绩差异, 而非母语因素是被试区分结果不同的主要影响因素。已有研究 (张林军, 2010) 也有类似结论。所以, 预实验的初级组被试虽然来自不同国家, 但都能熟练使用俄语, 这说明俄语并不能促进普通话声调范畴感知。因此, 正式实验中, 我们选择了母语为俄语的被试作为非声调语言控制组。预实验中之所以没有直接测量俄语组被试, 是由于客观原因, 即俄语组被试人数较少, 仅能满足一次正式实

验。预实验中发现汉语学习时长为 2 个月的非声调语言被试都不能表现出范畴感知，正式实验中我们遵循更加严格的标准，被试学习汉语的时长不超过 1 个月。

第二项实验是《不同母语对于汉语普通话声调范畴感知的影响》，研究材料同上一项研究，被试为汉语普通话母语者和老挝语母语者。研究发现：老挝语母语者感知汉语普通话也呈现范畴感知模式，但辨认曲线的斜率、范畴边界的宽度、区分峰位置与汉语普通话母语者具有显著差异，主要是受母语迁移影响，导致老挝被试感知普通话阴平调域扩大，一定程度上阻碍二语声调感知。**该研究对本研究的启示：**如果将母语声调复杂度作为考察变量，必须确保刺激材料在被试的声调母语中都能够找到相同或相似声调。该实验刺激采用的是汉语普通话声调，但在老挝语中不存在类似的对应声调，未能做到两种语言声调匹配。鉴于以上考虑，正式实验我们采用了越南语母语者作为另一种声调母语被试。原因有二：第一，越南语标准语有六个声调调类，比普通话更为复杂，见正文表 1。第二，汉语普通话阴平与越南语横声调值相同，汉语普通话阳平与越南语锐声调值相同，这使得实验结果具有可比性。

第三项实验是《母语声调复杂度对范畴感知的影响》，刺激与正式实验相同，任务为区分测试，被试为汉语普通话母语者和越南语母语者各 15 人。结果发现，汉语母语者和越南语母语者在语言和音乐刺激上都表现出了范畴感知，这表明声调范畴感知机制能够向音乐领域迁移。证明实验材料可行。

参考文献:

张林军. (2010). 母语经验对留学生汉语声调范畴化知觉的影响. 华文教学与研究(2), 15–20.

意见 6：识别实验为什么采用 ABX 任务，不采用 AXB 形式？

回应：正如审稿专家提示的那样，AXB 任务更好。AXB 任务的优势在于刺激 A 和刺激 B 与目标刺激 X 的距离相等，对于被试来说，A 与 B 有相同的记忆负担 (Strange & Shafer, 2008)。而 ABX 中，目标刺激 X 与 AB 距离不等，很可能出现刺激 A 由于距离目标刺激远而被遗忘的可能性。

本研究中，我们认为 ABX 任务与 AXB 任务差异不大，是基于以下两点考虑。其一，汉语二语者被试有不超过一个月的汉语学习经历，知晓阴平和阳平的区别，因此遗忘对他们的影响较小。其二，声调范畴感知研究传统上多采用 ABX 任务，如 Wang (1976), Francis(2003), Wayland & Guion(2004), Zheng(2010), Ong 等(2016)。

但不可否认的是，AXB 是一个更好的选择，我们在今后的研究中应该采用。

参考文献:

- Francis, A. L. , Ciocca, V. , & Ng, B. K. C. (2003). On the (non)categorical perception of lexical tones. *Perception & Psychophysics*, 65(7), 1029-1044.
- Ong, J. H., Burnham, D., Stevens, C. J., & Escudero, P. (2016). Native learners show cross-domain transfer after distributional learning: The case of lexical and musical pitch. *Frontiers in Psychology*, 7, 1–10.
- Strange, W. , & Shafer, V.L. (2008). Speech perception in second language learners: The re-education of selective perception. In Hansen-Edwards, J.G. , & Zampini, M.L. (Eds.) *Phonology and Second Language Acquisition* (pp. 153-185). Amsterdam: John Benjamins Publishing Company.
- Wang, W. S. (1976). Language change. *Annals of the New York Academy of Sciences*, 280(1), 61–72.
- Wayland, R. P., & Guion, S. G. (2004). Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Language Learning*, 54(4), 681–712.
- Zheng, H. Y. (2010). *Categorical perception of lexical tones: Behavioral and psychophysiological study*(Unpublished doctoral dissertation). The Chinese University of Hong Kong.

意见 7：图表不规范，如表 2 标题与内容分属两页；图 2、图 3 图例位置不统一等。

回应：感谢审稿专家的意见。我们已经进行了调整，将分属两页的图表调至一页，统一了图 2、图 3 (修改稿是图 4、图 5)的图例位置。

第二轮

审稿人 1 意见：作者修改认真，疫情期间找被试已不是易事，还进行了补测。对审稿意见理解基本到位。对提出的问题都进行了回应和相应的修改。但目前还存在以下问题：

意见 1：实验性的论文应该注意可重复性，也就是说，读者读罢可以把实验重复出来。建议把每个刺激，用表或图的形式罗列，读者一目了然每个音的音程。

回应：感谢审稿专家的意见。实验中的音乐刺激是在语音刺激的基础上，通过 Praat 软件中的音高提取功能，提取每个刺激的音高轮廓合成的。因此，这些音乐刺激与语音刺激在音高轮廓、振幅和持续时间上完全相同。之前的表述过于笼统，我们在修改稿中，完善了相关说明与图表。图 1 是 9 个语音刺激/音乐刺激的示意图，表 2 对每个语音刺激与音乐刺激的起始点频率和终点频率等参数进行了统一说明，并添加了相应语音刺激的调类与音乐刺激的音名。同时以刺激 9 为例，给出了语音刺激和音乐刺激的波形图和频谱图(图 2)。

意见 2：作者所列的乐器有一些的演奏需要辨别音准(弦乐)，有一些不需要(键盘乐等)，对音高的感知要求不同。“这一类人群的表现与未受过音乐训练的被试趋于一致”具体的数据支撑

是？

回应：感谢审稿专家的意见。正如审稿专家所言，键盘乐器(如钢琴)，调好音就可以直接演奏，音准是固定的，而弦乐(如二胡)是自由定音的乐器，不像键盘乐器和多数弹拨乐器那样有固定的品位，对于演奏者的音准能力提出了更高的要求。实证研究也发现，从可辨阈角度看，大、小提琴和铜管专业音乐家对 4 音分音差已有分辨能力，木管、打击乐和指挥专业音乐家是 6 音分，中提琴、钢琴、和弹拨专业音乐家是 8 音分(韩宝强, 1992)。可见，不同乐器的确对于专业音乐家的音准能力起到了不同程度的塑造作用，弦乐专业的音乐家对于音高的辨识能力更强。

本研究中的 60 名非音乐家被试中仅包含 8 名受过业余音乐训练的被试，他们受音乐训练时长为 1 周-3 年(非音乐家接受过业余音乐训练时长不多于 3 年这一标准源于已有文献，见第一轮审稿意见第一位专家意见 3 的回应)，除 1 名被试接受过声乐训练外，5 人接受的是钢琴训练，仅 2 人接受的是弦乐训练。我们将这 8 人语言和音乐的识别和区分指标分别与其所在组的均值和标准差进行对比，如下表所示。可以发现，绝大多数指标都在 1 个标准差之内，只有极个别指标(黄色)在两个标准差之内。而接受弦乐训练的业余被试的表现并未全部好于接受业余键盘乐训练的被试。据此我们认为受过业余音乐训练的被试与非音乐家表现一致。我们在第一次回应中提及，“这一类人群的表现与未受过音乐训练的被试趋于一致”，这个表述并不精确，“这类人群”主要是指本实验中 8 名受过业余音乐训练的被试。

表 1 受过业余音乐训练的被试识别任务指标与组均值的对比

被试组别	被试编号	学习内容	学习时长	语言斜率	语言范畴边界宽度	音乐斜率	音乐范畴边界宽度
越南	2	钢琴	11 个月	-2.19	1.00	-1.82	1.21
越南	6	钢琴	1 周	-2.02	1.09	-1.89	1.16
越南	8	声乐	5 个月	-1.43	1.54	-2.20	1.00
越南	13	中国古筝	1 个月	-1.94	1.13	-2.03	1.08
越南组均值	-	-	-	-1.83±0.44	1.95±1.73	-2.20±0.58	1.06±0.27
中国	3	钢琴	2 年	-2.00	1.10	-2.05	1.07
中国	16	钢琴	3 年	-3.54	0.62	-3.58	0.62
中国	28	二胡	1 年	-2.56	0.86	-1.95	1.13
中国	29	钢琴	3 年	-2.26	0.97	-2.95	0.74
中国组均值	-	-	-	-2.3±0.70	1.0±0.38	-2.43±0.54	0.94±0.17

表 2 受过业余音乐训练的被试区分任务指标与组均值的对比

被试组别	被试编号	学习内容	学习时长	范畴内语言	范畴间语言	范畴内音乐	范畴间音乐
越南	2	钢琴	11 个月	0.50	0.60	0.53	0.68
越南	6	钢琴	1 周	0.50	0.55	0.45	0.64
越南	8	声乐	5 个月	0.55	0.63	0.49	0.59
越南	13	中国古筝	1 个月	0.43	0.60	0.53	0.73
越南组均值	-	-	-	0.56 ± 0.08	0.61 ± 0.08	0.51 ± 0.05	0.59 ± 0.10
中国	3	钢琴	2 年	0.50	0.60	0.55	0.50
中国	16	钢琴	3 年	0.50	0.55	0.48	0.79
中国	28	二胡	1 年	0.55	0.63	0.53	0.60
中国	29	钢琴	3 年	0.43	0.60	0.60	0.83
中国组均值	-	-	-	0.57 ± 0.06	0.64 ± 0.13	0.54 ± 0.06	0.69 ± 0.10

参考文献:

韩宝强(1992). 音乐家的音准感——与律学有关的听觉心理研究. *中国音乐学*(3), 5-19.

意见 3: 有没有证据可以证明方言背景不影响(普通话声调感知)?

回应: 感谢专家提出这一尖锐的问题。我们重新对文献进行了梳理, 并未发现相关证据。恰恰相反, 一些对比普通话使用者和汉语方言使用者的研究表明, 方言使用者会影响普通话声调感知, 如 Huang(2004)、Peng 等(2010)。具体来说, 根据 Peng 等(2010)对于粤语使用者和普通话使用者对普通话声调连续统的感知研究, 粤语使用者仅会影响范畴感知的程度, 但不会改变范畴感知这一声调感知模式。

我们在筛选被试时尽可能降低方言的影响, 但我们认为, 即便有方言的影响, 也不影响本实验的结果。主要理由如下:

第一, 本研究的研究问题一是讨论汉语母语者对于普通话的范畴感知模式能否迁移, 因此, 只要得出汉语母语者对普通话声调范畴存在范畴, 就已经达到目的, 方言的影响并非一个关键变量。本实验的结果表明, 即便存在潜在的方言影响, 汉语母语者对普通话阴平-阳平连续统的感知也是范畴感知, 这与已有研究一致。前人的研究中(如 Chang 等, 2016; 王韞佳, 2010), 汉语母语者被试也未区分或未报告方言, 但仍得出他们对普通话阴平-阳平连续统的感知是范畴感知。

第二, 本实验的研究问题二是对比汉语母语者与越南语母语者对声调感知精度的影响。考虑到可能的方言影响, 我们选取的被试方言声调的丰富程度不高于越南语标准方言的六个声调。此外, 我们在选取被试时, 选取了北方方言母语者或是普通话母语者。他们日常使用语言为普通话, 并且普通话水平达到二级甲等或以上。北方方言母语者具体方言情况见表

3. 我们尽可能保证方言与普通话声调调值有重合。例如邯郸方言的四个声调调值分别为 214、55、35、51(秦学武, 2006), 与普通话的 55、35、214、51 在调值上相同, 只是音名不同。通过以上这些手段尽可能将方言的影响降至最低。

表 3 北方方言母语被试的方言及母语声调调类情况

被试序号	方言	方言声调调类
1	山东菏泽	四声
2	河北石家庄	四声
3	江苏南京	四声+入声*
4	江苏南京	四声+入声
5	河北秦皇岛	四声
7	江苏徐州	四声
9	山东枣庄	四声
11	辽宁抚顺	四声
14	河北石家庄	四声
17	安徽芜湖	五声
18	江苏南京	四声+入声
19	河北邯郸	四声
21	河北承德	四声
22	江苏淮安	四声+入声
23	内蒙古通辽	四声
24	江苏淮安	四声+入声
25	山东济南	四声
26	吉林吉林	四声
27	山西大同	四声+入声
30	山东淄博	四声

*入声短而急促, 在被试的方言中体现为喉塞音韵尾[ʔ], 不会影响阴平-阳平连续统的知觉。

参考文献:

- Chang, D., Hedberg, N., & Wang, Y. (2016). Effects of musical and linguistic experience on categorization of lexical and melodic tones. *Journal of the Acoustical Society of America*, 139(5), 2432–2447.
- Huang, T.(2004). *Language-specificity in auditory perception of Chinese tones*. Doctoral dissertation, The Ohio State University.
- Peng, G., Zheng, H. Y., Gong, T., Yang, R., Kong, J., & Wang, W. S. (2010). The influence of language experience on categorical perception of pitch contours. *Journal of Phonetics*, 38(4), 616–624.
- 秦学武(2006). *普通话教程*. 北京: 中国农业出版社.
- 王韞佳, 李美京(2010). 调型和调阶对阳平和上声知觉的作用. *心理学报* 42(9), 899–908.