科学通报 2016 年 第 61 卷 第 7 期: 682 ~ 690
《中国科学》杂志社
SCIENCE CHINA PRESS
Science125个科学前沿问题系列解读(III)
What are the evolutionary roots of language and music?

语言与音乐进化的起源

蒋存梅[①†*], 张清芳[②③†*]

① 上海师范大学音乐学院, 上海 200234;
② 中国人民大学心理学系, 北京 100872;
③ 中国科学院心理研究所, 北京 100101
† 同等贡献
* 联系人, E-mail: cunmeijiang@126.com; qingfang.zhang@ruc.edu.cn

2015-10-14 收稿, 2015-11-23 修回, 2015-11-25 接受, 2016-01-05 网络版发表
国家自然科学基金(31470972, 31170977, 31471074)和中央高校基本科研业务费专项基金(14NXLF12)资助

摘要 自达尔文始, 人们就开始探究语言和音乐的进化起源. 迄今为止, 这个问题尚未定论. *Science*在创刊125周年之际, 将其列为未来最具挑战的科学问题之一. 从进化角度说, 自然选择应满足3个标准: 人类特殊性、能力先天性和脑机制的特异性. 本文围绕这3个标准, 对语言和音乐的进化起源进行论述. 在语言方面, 阐述了人类语言的领域特异性, 个体语言发展理论以及语言进化的脑基础, 并指出未来研究应针对语言进化的争论焦点, 采取多种技术, 多学科交叉合作进行研究; 在音乐方面, 发现已有音乐心理学研究结果既不能完全支持音乐适应论, 也不能支持非适应论, 未来研究应从音乐能力的不同维度考察音乐进化起源, 并以原始母语为切入点深化音乐与语言的对比研究.

关键词 语言, 音乐, 进化起源, 人类特殊性, 先天性, 领域特异性

语言和音乐是人类社会的普遍现象, 存在于不同的民族和文化群体之中. 然而, 从进化视角来看, 语言和音乐进化的起源是什么? 这个问题至今尚未定论. 达尔文最早论述了语言和音乐的进化起源, 在他看来, 语言的演变与不同物种的演变类似, 经历了渐进变化的过程[1]; 人类的音乐能力是一种被赋予的神秘能力之一, 音乐源于性选择[2]. 自达尔文始, 学界就开始对语言和音乐进化起源的问题展开讨论. 如果语言和音乐都是人类社会独有的, 对二者进化起源的探究无疑将为理解人类的进化提供一个独特的视角.

下文将围绕种系进化、遗传与个体发展等方面分别对语言和音乐的进化起源进行阐述. 如果语言和音乐是自然选择的结果, 那么, 语言或音乐能力必然是人类独有的, 且先天具备的. 相应地, 大脑必然存在专门的"语言脑区"或"音乐脑区"[3].

1 语言与进化

语言的进化研究主要关注两个问题: 语言是何时在何地如何起源的? 语言产生之后是如何进化和发展的? 相应地, 关于语言的起源存在两种观点: 一种观点认为语言是人类长期进化与自然选择的结果, 遵循渐进的连续的模式; 另一种观点则认为语言是随着脑容量的增大在较晚时期迅速出现的, 遵循突发型模式. 下文将围绕人类语言的领域特异性, 个体语言发展理论以及语言进化的脑基础阐述语言进化的起源.

1.1 人类语言的领域特殊性

在人类的语言出现之前, 原始人类已经出现了

引用格式: 蒋存梅, 张清芳. 语言与音乐进化的起源. 科学通报, 2016, 61: 682–690
Jiang C M, Zhang Q F. The evolutionary origins of language and music (in Chinese). Chin Sci Bull, 2016, 61: 682–690, doi: 10.1360/N972015-01144

类似语言的交流系统, 称之为前语言起源期[4,5]. 研究发现, 与人类相似, 非人灵长类也能使用符号表征自然界物体, 尽管这种能力较为有限[6], 例如, 南非猿猴 (*Australopithecus*) 的警报呼叫并不特指蛇(*Dendroaspis polylepis*)、鹰(*Accipiter trivirgatus*)或者豹子(*Panthera pardus*)等动物, 但它能引发不同的逃跑反应[7]; 黑猩猩(*Pan troglodytes*)能够利用手势交流[8]. 尽管语言起源于口语还是手势尚存在争论, 但是, 口语和手势语交流系统的发展为人类语言的产生奠定了基础[9]. 在前语言起源期, 非人灵长类所具备的注意[4,5]、动作模仿以及觉察他人意图的能力[4]是语言出现的先决条件.

进化心理学认为, 语言的进化并不存在一般的适应性问题, 人类的语言具有领域特异性, 语言交流系统与其他动物物种的交流存在3个本质的区别. 首先, 人类语言交流具有符号象征性. 语言符号是人类交流共享的符号表征系统, 个体间的交流依赖于符号系统, 通过与他人分享注意和意图, 从而达到交流的目的[10,11]. 其他物种不会运用语言符号进行交流, 这主要缘于它们不能分享意图和目的. 其次, 人类语言的交流具有语法性. 这种语法性实质上体现了语言结构, 该结构本身隐含着意义. 人类语法结构的复杂性是其他物种的交流系统不可比拟的. 例如, Chomsky(乔姆斯基)[12]认为语法结构包括了表层结构和深层结构. 从心理语言学的观点来看, 语言结构包括了单词、单词的形态标识、词序以及语调/重音[13], 儿童通过有意识阅读和文化学习[14]、图式化和类比[15]、防御和竞争[16], 以及基于功能的分布式分析[17]等机制习得语言. 这种独特的语法化过程贯穿在人类语言进化的过程之中. 人与其他物种的第3点差异体现在人类的语言交流系统具有多样性, 各类语言系统在语法结构和构成规则方面完全不同. 如果要获得一门语言, 人类要比其他动物物种学习大量的语言规则.

1.2 人类语言能力的先天性: 个体语言的获得与发展

根据生物进化论的观点, 个体的发展会重演整个动物系统从低级到高级的发展过程, 因此研究儿童获得语言及其发展的过程可能会发现人类创造语言的过程. 在语言习得方面, 存在着两类理论趋向: 一种是以乔姆斯基学派的生成语法理论为出发点, 从语言形式的角度研究语言及语言习得; 另一种以认知功能语言学的研究角度为出发点, 注重发展变化空间的儿童中心方法, 倾向于从功能和应用的角度来研究语言习得.

生成语法理论假设儿童天生就具有一套普遍抽象的语法, 由此得以建构世界上任一种语言. 语言习得包括两个过程: 先学习语言单词、习惯用语和特殊结构, 然后将所学到的语言与普遍语法相联系. 普遍语法具有先天性, 终生保持不变, 即连续性假设[18]. 与之相对, 认知功能语言学认为语言结构来自于语言的使用[5,10,19], 语言的本质是符号, 在语言发展的历史进程中派生出各种语法规则和结构. 这些语法规则和结构是通过学习而习得. 儿童语言发展的早期阶段尚未建立起类似成人语法的抽象概念和图式, 他们通过注意的参与理解他人的意图和相应话语的意义; 通过归类、图式建构、统计学习和类比等方式, 建立起语言的抽象语法结构. 这种学习方式体现出渐进性. 可见, 生成语法理论和认知功能语言学分别代表先天与后天习得观念, 但这两种理论都未涉及儿童如何将话语或语句与交流情境相联系. 认知发展领域对儿童如何理解别人的意图做了大量研究, 提出了儿童发展的心理理论, 儿童理解他人错误信念的能力是心理理论发展的标志.

遗传学的发展支持人类的语言获得具有先天性, 与特定的基因有关. 行为遗传学研究发现7号染色体上的叉头框P2基因(*forkhead box P2, FOXP2*)基因与语言遗传缺陷有关. KE家族大约有1/2的家庭成员表现出语言上的障碍, 主要表现为不能获得形态句法学(morphosyntatic)规则[20], 乔姆斯基学派的理论中认为*FOXP2*可能是"语法基因"[21]; 遗传缺陷与发音和非言语运用也有关[22]. 最近发现KE家族语言障碍的核心是音韵单元的序列发音[23].

1.3 语言进化的脑基础

对人类语言加工脑机制的研究必须考虑哪些核心计算(core computation)成分是普遍适用的, 哪些成分存在个体差异[24]. 在语言加工的神经网络中, 包括了背侧通路和腹侧通路. 背侧通路有两条: 一条通路负责核心句法的计算, 包括了Broadman 44区(BA 44)和颞上皮层(superior temporal cortex, STC)后部[25]; 另一条通路负责语言与感觉运动皮层之间的交互, 包括了前运动皮层和STC[26]. 腹侧通路所包括

的脑区与语义加工和理解概念信息密切相关, 具体涉及额下皮层的BA 45区和颞叶皮层[27].

背侧通路中的第二条通路连结了STC的听觉感觉区域与中央前回中的前运动皮层, 与其他背侧通路相比, 这条通路在婴儿出生时就存在, 而且终生保持不变[28]. 成人的这条背侧通路则与言语的口语重复密切相关[26], 婴儿大脑中运动皮层之间的联结与出生后几个月内的语音学习密切相关[29,30], 婴儿依赖这条通路觉察音韵编码规则. BA 44和STC的背侧联结在儿童7岁时完全发展成熟, 从而为加工复杂的语言句法结构奠定了脑基础. 成人所具备的背侧通路和腹侧通路分别在语言加工中起着至关重要的作用, 至于这两个系统之间如何交互作用是下一步研究的重点[24].

研究者对于布洛卡区域在语言进化中的作用与*FOXP2*基因之间的关系存在争论[9,31]. 在非人灵长类中语言进化的前起源阶段, 镜像神经元系统为动物之间互相理解其运动目的以及模仿行为提供了基础. 布洛卡区域在口语和手势语中都有激活, 这与认为语言是从手部姿势发展而来的观点是一致的[32~35]. 研究发现*FOXP2*和非人灵长类布洛卡区域的镜像神经元系统相关[36], 表明布洛卡区域在语言进化中可能起了一定的作用. 研究者由此认为, 人类进化过程中选择的基因和增大的布洛卡区域都是为了加工复杂的语言句法结构. KE家族的发音障碍与布洛卡区域相关, 因为后者具有发音的功能. 研究者试图教会大猩猩(Chimpanzees)学会人类语音的尝试未能获得成功[37], 但猩猩却能够利用手势[38]或符号[39]表达一些简单的句法, 句法可能比声音更早地在进化历程中产生, 而人类进化中所选择的*FOXP2*基因可能具有负责言语发声的功能[38].

2 音乐与进化

尽管在达尔文的时代, 其音乐适应论就已经引发争议[3], 但是, 直到20世纪音乐进化的问题才重新引起学界的关注, 尤其在近10年间, 随着认知神经科学的发展, 音乐适应论和非适应论再度成为研究的热点. 音乐适应论者支持达尔文音乐行为的生物学起源观点[40~42]. 其理论主要建立在如下假说上: 性选择、母婴交流、群体凝聚力以及节奏相关的奖赏和情绪系统. 非适应论者认为, 音乐是语言或听觉的附属品[43,44]. 其代表人物Pinker[43,44]认为, 从生物学角度说, 音乐是没有用处的, 它仅是令人愉悦的听觉附属品. 与Pinker观点不同, Patel[3,45]认为, 尽管大脑没有存在特定的“音乐脑区”, 但是音乐具有生物学的意义, 这是因为, 音乐行为能影响非音乐的脑功能, 音乐是一种具有变革意义的发明.

音乐适应论与非适应论之争尚未结束, 其核心差异可能缘于音乐的复杂性. 音乐加工包含低层级和高层级两个方面. 前者指要素(如音高、节奏等)的加工, 后者则指音乐调性、句法、情绪以及意义等方面的加工[46,47]. 相应地, 人类音乐能力必然分为这两个层面的加工能力. 下文将结合已有研究成果, 围绕音乐适应论与非适应论之争, 从音乐高低层级加工视角对此进行论证.

2.1 人类特殊性

在音高方面, 绝对音高(absolute pitch)与相对音高(relative pitch)被认为是音高知觉的两种模式. 绝对音高者拥有超常的音高命名能力, 即在没有参照音的情况下, 他们可以对孤立的音高进行准确命名[48,49]; 而相对音高者则需要参照音才能对音高进行命名. 相对音高在人类中是普遍存在的, 即便婴儿都能以这种策略加工旋律[50,51]. 尽管一些鸟类具有较强的绝对音高感[52], 然而, 动物缺乏相对音高感[53]. 在节奏节拍方面, 尽管猴子(*Zoology*)[54]和黑猩猩[55]呈现出与音乐同步运动的迹象, 但是它们的同步运动并不源自结构性的时间预期或不具有速度灵活性[56]. 此外, 动物声音学习激活的神经网络(前脑运动区-基底神经节-听觉网络)[57]与人类(顶上回-听觉连接网络)[58]也存在差异.

在高层级加工方面, 研究发现, 猴子能分辨出旋律转调, 但无法分辨出单音及其八度音[59]; 没有表现出对协和音程的偏好[60]. 然而, 在情绪交流方面, 灵长类动物在幼年就能通过声音频率[61]、时长[62]、强度[63]的变化向母亲或同伴表达情绪, 并诱发相应的生理变化[64]. 人类与类人猿在声音情绪加工的相似性可能缘于二者在情绪表达的共性[65].

歌唱经常也被认为是论证人类与动物音乐能力相似性的论据. 的确, 动物(如鸣禽或某些鲸类动物)通过歌唱吸引异性, 使物种得以繁衍. 然而, 与人类相比, 动物歌唱具有4个特点: (ⅰ) 限定在雄性动物中; (ⅱ) 具有季节性, 受到机体生物调控; (ⅲ) 动物学习歌曲能力受到更多的限制; (ⅳ) 动物歌曲的结

构不够丰富[3]. 此外, 神经生物学的研究结果也验证了在歌唱机制上动物与人类的差异[66].

可见, 在音高与节奏方面, 与人类不同, 动物不具备相对音高感或内隐的节拍感, 无法调节节奏的变化. 动物歌唱与人类歌唱也存在本质的差异. 但就声音情绪而言, 类似人类的声音情绪加工能力至少在类人猿身上得到体现.

2.2 音乐能力的先天性

绝对音高是一种特殊的音高知觉能力. 拥有该能力被视作可能获得卓越音乐成就的征兆. 遗传学研究表明, 绝对音高在家族成员中汇聚[67]; 家庭中有一人拥有绝对音高感, 其同胞拥有绝对音高感的比例比普通人高7.8~15.1倍[68]; 8q24.21染色体能够预测绝对音高能力[69]. 与绝对音高感相反, 先天失歌症(congenital amusia)是一种对音乐音高加工的障碍. 研究发现, 先天失歌症者对音高知觉困难可能缘于遗传[70,71]. Tan等人[72]的研究也表明, 音乐知觉、歌唱、音乐记忆和聆听可能都有相关的基因表达.

遗传为个体发展提供了前提. 胎儿在6个月就开始知觉声音[73], 出生时已具备音乐经验和记忆[74]; 新生儿能够知觉时长[75], 并对乐音进行分组[76]; 2个月婴儿能分辨音高的半音变化[77], 识别出熟悉旋律[78], 并表现出对协和音程的偏好[60]; 4个月婴儿能够区分乐器的音色[79], 分辨简单的节奏变化[80]; 6个月婴儿已经具备相对音高能力[50], 并能识别出调外音[81].

以上研究在一定程度上论证了音乐能力的先天性. 然而, 如上所述, 先天失歌症存在对音乐音高加工的障碍, 一些先天失歌症者还存在对节奏加工的困难[82]. 研究发现, 先天失歌症者的音乐障碍体现在分辨音高的细微差异[83~85]、判断旋律轮廓[83,86]和不协和和弦[87]、外显知觉调性和句法结构[46,88]. 然而, 他们可以内隐地加工和弦片段[89]. 在一项未发表的研究中, 本课题组发现, 在伴随节奏、强度等声学线索情况下, 先天失歌症者可以对音乐意义进行加工. 音乐情绪识别的研究[90]也发现类似的结果. 这些研究暗示着先天失歌症者对音乐情绪意义加工可能不存在困难.

可见, 人类对音乐低层级的要素加工(如音高、节奏等)可能生来就存在差异, 但在音乐情绪和意义等高层级加工方面, 音乐能力可能是普遍存在的. 后者也可以从跨文化研究中获得支持[91,92], 即音响的声学线索可能对音乐高层级加工造成影响[93].

2.3 音乐脑机制的特异性

如果存在专门的音乐脑区, 则音乐适应论就得到支持. 尽管针对脑损伤病人和正常人的研究都发现与音乐加工相关的脑区, 例如, 颞上沟负责旋律加工[94]; 颞上回负责节奏加工[95]; 左侧额下回负责音高记忆[96]、音乐句法[97]和情绪[98], 但这些相关均不能说明音乐脑机制的特异性. 因此, 有研究从绝对音高视角探究专门的“音乐脑区”, 如颞平面[99], 双侧颞上回、左侧额下回和右侧缘上回[100]以及颞叶结构的特异性连接[101], 然而, 近年研究发现, 绝对音高的形成与音乐训练相关[102], 且这种能力可能与音乐加工无关[103]. 因此, 绝对音高研究仍不能验证音乐脑机制的特异性.

如果失歌症的音高障碍是先天具有的[71], 则可能存在特异性的音乐脑区. 研究发现, 失歌症者大脑右侧额下回与听觉皮层连接异常[104,105]、左侧颞上沟和额下回[106]也出现异常, 但研究还发现, 无论在音高层面[86,107], 还是在韵律理解层面上[108,109], 失歌症者在音乐和语言加工上都存在障碍, 由此难以确证, 失歌症者的异常脑区属于音乐加工独有的, 还是与语言加工共享的.

3 总结与展望

语言进化的研究得到了语言学、生物学、遗传学、人类学、认知科学、神经科学、心理学等多个领域研究者的关注[4,5,110], 采取多种技术进行多学科交叉研究将是探索语言进化起源的有效途径. 同时, 研究方法的多样化也是未来研究的趋势, 例如, 计算建模方法[5,111], 脑成像方法[24]等. 此外, 由于心理学多个分支的研究都涉及语言的进化起源, 其研究成果必然对语言进化的探究具有积极的推动作用.

另一方面, 从音乐进化来看, 以上论据既不能完全支持音乐适应论, 也不能完全论证音乐的非适应论. 本研究认为, 由于音乐结构的复杂性, 且各要素加工之间并不完全存在相关, 人类的音乐能力必然体现出多样化的特点, 因此, 音乐进化研究应从音乐能力的不同维度出发进行探究, 例如, 哪些音乐能力可能体现出自然选择的结果? 哪些音乐能力则体现出非适应论的观点?

音乐是否与语言共享特定的脑机制? 这是语言与音乐对比研究的核心问题, 也是判断音乐进化起源的重要标准[3]. 目前存在两种观点: 模块化观点和资源共享框架理论. 模块化观点认为, 大脑对音乐和语言的加工是分离的[112,113]. 该理论建立在脑损伤病人音乐和语言加工研究的基础上. 从知觉层面说, 脑损伤导致病人失去语言识别能力, 但音乐知觉能力仍然完好[114]; 或病人失去对音乐的知觉能力, 但仍保留正常的语言识别能力[115]. 从产生角度来说, 失语症病人可以歌唱熟悉的歌曲或学习新的曲调, 但无法以歌唱或说话形式表达歌词[116], 或病人失去歌唱能力, 但是言语产生能力是正常的[117]. 相反, 资源共享框架理论[3,118]则依赖于正常人对音乐和语言句法的加工研究[119~121], 认为音乐与语言共享特定的神经机制.

模块化观点与资源共享框架理论的问题还在讨论之中. Patel[45]认为, 模块化观点所倡导的分离是缘于储存在长时记忆中的音乐或语言的知识具有特异性, 但对这些知识或表征的加工却是共享的. 的确, 不论是基因研究[122], 还是认知机制的研究[123], 都发现音乐与语言加工的交集, 由此很难判断究竟音乐是语言的附属品, 还是语言是音乐的附属品. 音乐与言语加工的共享也许正印证了达尔文关于音乐与言语起源的原始母语假说[2]. 未来研究有必要对原始母语假说进行验证.

参考文献

1 Darwin C. On the Origin of Species. London: John Murray, 1859
2 Darwin C. The Descent of Man, and Selection in Relation to Sex. London: John Murray, 1871
3 Patel A D. Music, Language, and the Brain. New York: Oxford University Press, 2008
4 Cangelosi A, Parisi D. Simulating the Evolution of Language. Berlin: Springer, 2002
5 Christiansen M H, Kirby S. Language Evolution. Oxford: Oxford University Press, 2003
6 Conway C M, Christiansen M H. Sequential learning in non-human primates. Trends Cogn Sci, 2001, 5: 539–546
7 Owren M J, Rendall D. Sound on the rebound: Bringing form and function back to the forefront in understanding nonhuman primate vocal signaling. Evol Anthropol, 2001, 10: 58–71
8 Arcadi A C. Is gestural communication more sophisticated than vocal communication in wild chimpanzees? Behav Brain Sci, 2003, 26: 210–211
9 Corballis M C. Mirror neurons and the evolution of language. Brain Lang, 2010, 112: 25–35
10 Tomasello M. Reference: Intending that others jointly attend. Pragmat Cogn, 1998, 6: 229–243
11 Tomasello M. The Cultural Origins of Human Cognition. Cambridge: Harvard University Press, 1999
12 Chomsky N. Syntactic Structures. The Hague: Mouton, 1957
13 Bates E, McWhinney B. Funetionalist approaches to grammar. In: Wanner E, Gleitman L R, eds. Language Acquisition: The State of the Art. Cambridge: Cambridge University Press, 1982. 173–218
14 Tomasello M. Perceiving intentions and learning words in the second year of life. In: Bowerman M, Levinson S C, eds. Language Acquisition and Conceptual Development. Cambridge: Cambridge University Press, 2001. 132–158
15 Childers J B, Tomasello M. The role of pronouns in young children's acquisition of the English transitive construction. Dev Psychol, 2001, 37: 739–748
16 Brooks P J, Tomasello M. How children constrain their argument structure constructions. Language, 1999, 75: 720–738
17 Lieven E, Behrens H, Speares J, et al. Early syntactic creativity: A usage-based approach. J Child Lang, 2003, 30: 333–370
18 Pinker S. Words and Rules: The Ingredients of Language. New York: Basic Books, 1999
19 Croft W. Radical Construction Grammar. Oxford: Oxford University Press, 2001
20 Gopnik M. Feature-blind grammar and dysphasia. Nature, 1990, 344: 715–715
21 Pinker S. The Language Instinct: How the Mind Creates Language. New York: William Morrow & Company, 1994
22 Vargha-Khadem F, Watkins K, Alcock K, et al. Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. Proc Natl Acad Sci USA, 1995, 92: 930–933
23 Watkins K E, Dronkers N F, Vargha-Khadem F. Behavioural analysis of an inherited speech and language disorder: Comparison with acquired aphasia. Brain, 2002, 125: 452–464
24 Berwick R C, Friederici A D, Chomsky N, et al. Evolution, brain, and the nature of language. Trends Cogn Sci, 2013, 17: 89–98
25 Friederici A D, Bahlmann J, Heim S, et al. The brain differentiates human and non-human grammars: Functional localization and structural connectivity. Proc Natl Acad Sci USA, 2006, 103: 2458–2463
26 Saur D, Kreher B W, Schnell S, et al. Ventral and dorsal pathways for language. Proc Natl Acad Sci USA, 2008, 105: 18035–18040
27 Friederici A D. The cortical language circuit: From auditory perception to sentence comprehension. Trends Cogn Sci, 2012, 16: 262–268

28 Perani D, Saccuman M C, Scifo P, et al. Neural language networks at birth. Proc Natl Acad Sci USA, 2011, 108: 16056–16061
29 Friederici A D, Mueller J L, Oberecker R. Precursors to natural grammar learning: Preliminary evidence from 4-month-old infants. PLoS One, 2011, 6: e17920
30 Mueller J L, Friederici A D, Männel C. Auditory perception at the root of language learning. Proc Natl Acad Sci USA, 2012, 109: 15953–15958
31 Corballis M C. *FOXP2* and the mirror system. Trends Cogn Sci, 2004, 8: 95–96
32 Rizzolatti G, Arbib M A. Language within our grasp. Trends Neurosci, 1998, 21: 188–194
33 Arbib M A. From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. Behav Brain Sci, 2005, 28: 105–124
34 Corballis M C. The origins of modernity: Was autonomous speech the critical factor? Psychol Rev, 2004, 111: 543–552
35 Fogassi L, Ferrari P F. Mirror neurons and the evolution of embodied language. Curr Dir Psychol Sci, 2007, 16: 136–141
36 Liégeois F, Baldeweg T, Connelly A, et al. Language fMRI abnormalities associated with *FOXP2* gene mutation. Nat Neurosci, 2003, 6: 1230–1237
37 Ploog D. Is the neural basis of vocalization different in nonhuman primates and Homo sapiens? In: Crow T J, ed. The Speciation of Modern Homo Sapiens. Oxford: Oxford University Press, 2002. 121–135
38 Gardner R A, Gardner B T. Teaching sign language to a chimpanzee. Science, 1969, 165: 664–672
39 Savage-Rumbaugh S, Shanker S G, Taylor T J. Apes, Language, and the Human Mind. New York: Oxford University Press, 1998
40 Balter M. Seeking the key to music. Science, 2004, 306: 1120–1122
41 Wallin N L, Merker B, Brown S. The Origins of Music. Cambridge: MIT Press, 2001
42 Wang T. A hypothesis on the biological origins and social evolution of music and dance. Front Neurosci, 2015, 9: 30
43 Pinker S. How the Mind Works. New York: W. W. Norton & Company, 1997
44 Pinker S. Toward a consilent study of literature. Philos Literature, 2007, 31: 162–178
45 Patel A D. Music, biological evolution, and the brain. In: Levander C, Henry C, eds. Emerging Disciplines. Houston: Rice University Press, 2010. 91–144
46 Jiang C, Liu F, Thompson W F. Impaired explicit processing of musical syntax and tonality in a group of Mandarin-speaking congenital amusics. Music Percept, in press
47 Koelsch S. Brain and Music. Oxford: John Wiley & Sons, 2013
48 Miyazaki K I. Absolute pitch identification: Effects of timbre and pitch region. Music Perception, 1989, 7: 1–14
49 Ross D A, Gore J C, Marks L E. Absolute pitch: Music and beyond. Epilepsy Behav, 2005, 7: 578–601
50 Plantinga J, Trainor L J. Memory for melody: Infants use a relative pitch code. Cognition, 2005, 98: 1–11
51 Trehub S E, Bull D, Thorpe L A. Infants' perception of melodies: The role of melodic contour. Child Dev, 1984, 55: 821–830
52 Weisman R G, Njegovan M G, Williams M T, et al. A behavior analysis of absolute pitch: Sex, experience, and species. Behav Processes, 2004, 66: 289–307
53 McDermott J, Hauser M. The origins of music: Innateness, development, and evolution. Music Percept, 2005, 23: 29–59
54 Zarco W, Merchant H, Prado L, et al. Subsecond timing in primates: Comparison of interval production between human subjects and rhesus monkeys. J Neurophysiol, 2009, 102: 3191–3202
55 Hattori Y, Tomonaga M, Matsuzawa T. Spontaneous synchronized tapping to an auditory rhythm in a chimpanzee. Sci Rep, 2013, 3: 1566
56 Patel A D. The evolutionary biology of musical rhythm: Was Darwin wrong. PLoS Biol, 2014, 12: e1001821
57 Petkov C I, Jarvis E. Birds, primates, and spoken language origins: Behavioral phenotypes and neurobiological substrates. Front Evol Neurosci, 2012, 4: 12
58 De Schotten M T, Dell'Acqua F, Valabregue R, et al. Monkey to human comparative anatomy of the frontal lobe association tracts. Cortex, 2012, 48: 82–96
59 Wright A A, Rivera J J, Hulse S H, et al. Music perception and octave generalization in rhesus monkeys. J Exp Psychol Gen, 2000, 129: 291–307
60 McDermott J, Hauser M. Are consonant intervals music to their ears? Spontaneous acoustic preferences in a nonhuman primate. Cognition, 2004, 94: B11–B21
61 Schrader L, Todt D. Contact call parameters covary with social context in common marmosets, *Callithrix jacchus*. Anim Behav, 1993, 46: 1026–1028
62 Notman H, Rendall D. Contextual variation in chimpanzee pant hoots and its implications for referential communication. Anim Behav, 2005, 70: 177–190
63 Cheney D L, Seyfarth R M, Palombit R. The function and mechanisms underlying baboon "contact" barks. Anim Behav, 1996, 52: 507–518
64 Levine S, Wiener S G, Coe C L. Temporal and social factors influencing behavioral and hormonal responses to separation in mother and infant squirrel monkeys. Psychoneuroendocrinology, 1993, 18: 297–306
65 Leinonen L, Laakso M L, Carlson S, et al. Shared means and meanings in vocal expression of man and macaque. Logoped Phoniatr Vocol, 2003, 28: 53–61

66 Konishi M. An outline of recent advances in birdsong neurobiology. Brain Behav Evol, 1994, 44: 279–285
67 Baharloo S, Johnston P A, Service S K, et al. Absolute pitch: An approach for identification of genetic and nongenetic components. Am J Hum Genet, 1998, 62: 224–231
68 Baharloo S, Service S K, Risch N, et al. Familial aggregation of absolute pitch. Am J Hum Genet, 2000, 67: 755–758
69 Theusch E, Basu A, Gitschier J. Genome-wide study of families with absolute pitch reveals linkage to 8q24. 21 and locus heterogeneity. Am J Hum Genet, 2009, 85: 112–119
70 Drayna D, Manichaikul A, de Lange M, et al. Genetic correlates of musical pitch recognition in humans. Science, 2001, 291: 1969–1972
71 Peretz I, Cummings S, Dubé M P. The genetics of congenital amusia (tone deafness): A family-aggregation study. Am J Hum Genet, 2007, 81: 582–588
72 Tan Y T, McPherson G E, Peretz I, et al. The genetic basis of music ability. Front Psychol, 2014, 5: 658
73 Lecanuet J. Prenatal auditory experience. In: Deliege I, Sloboda J, eds. Musical Beginnings: Origins and Development of Musical Competence. Oxford: Oxford University Press, 1996. 3–34
74 Partanen E, Kujala T, Tervaniemi M, et al. Prenatal music exposure induces long-term neural effects. PLoS One, 2013, 8: e78946
75 Winkler I, Háden G P, Ladinig O, et al. Newborn infants detect the beat in music. Proc Natl Acad Sci USA, 2009, 106: 2468–2471
76 Winkler I, Kushnerenko E, Horváth J, et al. Newborn infants can organize the auditory world. Proc Natl Acad Sci USA, 2003, 100: 11812–11815
77 He C, Hotson L, Trainor L J. Maturation of cortical mismatch responses to occasional pitch change in early infancy: Effects of presentation rate and magnitude of change. Neuropsychologia, 2009, 47: 218–229
78 Plantinga J, Trainor L J. Melody recognition by two-month-old infants. J Acoust Soc Am, 2009, 125: EL58–EL62
79 Trainor L J, Lee K, Bosnyak D J. Cortical plasticity in 4-month-old infants: Specific effects of experience with musical timbres. Brain Topogr, 2011, 24: 192–203
80 Lewkowicz D J. Learning and discrimination of audiovisual events in human infants: The hierarchical relation between intersensory temporal synchrony and rhythmic pattern cues. Dev Psychol, 2003, 39: 795–804
81 Lynch M P, Eilers R E, Oller D K, et al. Innateness, experience, and music perception. Psychol Sci, 1990, 1: 272–276
82 Foxton J M, Nandy R K, Griffiths T D. Rhythm deficits in “tone deafness”. Brain Cogn, 2006, 62: 24–29
83 Foxton J M, Dean J L, Gee R, et al. Characterization of deficits in pitch perception underlying “tone deafness”. Brain, 2004, 127: 801–810
84 Jiang C, Hamm J P, Lim V K, et al. Fine-grained pitch discrimination in congenital amusics with Mandarin Chinese. Music Percept, 2011, 28: 519–526
85 Jiang C, Lim V K, Wang H, et al. Difficulties with pitch discrimination influences pitch memory performance: Evidence from congenital amusia. PLoS One, 2013, 8: e79216
86 Jiang C, Hamm J P, Lim V K, et al. Processing melodic contour and speech intonation in congenital amusics with Mandarin Chinese. Neuropsychologia, 2010, 48: 2630–2639
87 Ayotte J, Peretz I, Hyde K. Congenital amusia: A group study of adults afflicted with a music-specific disorder. Brain, 2002, 125: 238–251
88 Omigie D, Pearce M T, Stewart L. Tracking of pitch probabilities in congenital amusia. Neuropsychologia, 2012, 50: 1483–1493
89 Tillmann B, Gosselin N, Bigand E, et al. Priming paradigm reveals harmonic structure processing in congenital amusia. Cortex, 2012, 48: 1073–1078
90 Gosselin N, Paquette S, Peretz I. Sensitivity to musical emotions in congenital amusia. Cortex, 2015, 71: 171–182
91 Balkwill L L, Thompson W F, Matsunaga R. Recognition of emotion in Japanese, Western, and Hindustani music by Japanese listeners. Jpn Psychol Res, 2004, 46: 337–349
92 Fritz T, Jentschke S, Gosselin N, et al. Universal recognition of three basic emotions in music. Curr Biol, 2009, 19: 573–576
93 Balkwill L L, Thompson W F. A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. Music Percept, 1999, 17: 43–64
94 Lee Y S, Janata P, Frost C, et al. Investigation of melodic contour processing in the brain using multivariate pattern-based fMRI. Neuroimage, 2011, 57: 293–300
95 Grahn J A, Brett M. Rhythm and beat perception in motor areas of the brain. J Cogn Neurosci, 2007, 19: 893–906
96 Watanabe T, Yagishita S, Kikyo H. Memory of music: Roles of right hippocampus and left inferior frontal gyrus. Neuroimage, 2008, 39: 483–491
97 Sammler D, Koelsch S, Friederici A D. Are left fronto-temporal brain areas a prerequisite for normal music-syntactic processing? Cortex, 2011, 47: 659–673
98 Koelsch S, Fritz T, V Cramon D Y, et al. Investigating emotion with music: An fMRI study. Hum Brain Mapp, 2006, 27: 239–250
99 Zatorre R J, Perry D W, Beckett C A, et al. Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. Proc Natl Acad Sci USA, 1998, 95: 3172–3177
100 Dohn A, Garza-Villarreal E A, Chakravarty M M, et al. Gray-and white-matter anatomy of absolute pitch possessors. Cereb Cortex, 2015, 25: 1379–1388
101 Jäncke L, Langer N, Hänggi J. Diminished whole-brain but enhanced peri-sylvian connectivity in absolute pitch musicians. J Cogn Neurosci, 2012, 24: 1447–1461
102 Vanzella P, Schellenberg E G. Absolute pitch: Effects of timbre on note-naming ability. PLoS One, 2010, 5: e15449

103 Miyazaki K. Recognition of transposed melodies by absolute-pitch possessors. Jpn Psychol Res, 2004, 46: 270–282
104 Hyde K L, Lerch J P, Zatorre R J, et al. Cortical thickness in congenital amusia: When less is better than more. J Neurosci, 2007, 27: 13028–13032
105 Hyde K L, Zatorre R J, Peretz I. Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. Cereb Cortex, 2011, 21: 292–299
106 Mandell J, Schulze K, Schlaug G. Congenital amusia: An auditory-motor feedback disorder? Restor Neurol Neurosci, 2007, 25: 323–334
107 Nan Y, Sun Y, Peretz I. Congenital amusia in speakers of a tone language: Association with lexical tone agnosia. Brain, 2010, 133: 2635–2642
108 Thompson W F, Marin M M, Stewart L. Reduced sensitivity to emotional prosody in congenital amusia rekindles the musical protolanguage hypothesis. Proc Natl Acad Sci USA, 2012, 109: 19027–19032
109 Jiang C, Hamm J P, Lim V K, et al. Amusia results in abnormal brain activity following inappropriate intonation during speech comprehension. PLoS One, 2012, 7: e41411
110 Hauser M D, Chomsky N, Fitch W T. The faculty of language: What is it, who has it, and how did it evolve? Science, 2002, 298: 1569–1579
111 Nowak M A, Komarova N L. Towards an evolutionary theory of language. Trends Cogn Sci, 2001, 5: 288–295
112 Peretz I. The nature of music from a biological perspective. Cognition, 2006, 100: 1–32
113 Peretz I, Coltheart M. Modularity of music processing. Nat Neurosci, 2003, 6: 688–691
114 Mendez M F. Generalized auditory agnosia with spared music recognition in a left-hander: Analysis of a case with a right temporal stroke. Cortex, 2001, 37: 139–150
115 Ayotte J, Peretz I, Rousseau I, et al. Patterns of music agnosia associated with middle cerebral artery infarcts. Brain, 2000, 123: 1926–1938
116 Hébert S, Racette A, Gagnon L, et al. Revisiting the dissociation between singing and speaking in expressive aphasia. Brain, 2003, 126: 1838–1850
117 Peretz I, Kolinsky R G, Tramo M, et al. Functional dissociations following bilateral lesions of auditory cortex. Brain, 1994, 117: 1283–1301
118 Patel A D. Language, music, syntax and the brain. Nat Neurosci, 2003, 6: 674–681
119 Maess B, Koelsch S, Gunter T C, et al. Musical syntax is processed in Broca's area: An MEG study. Nat Neurosci, 2001, 4: 540–545
120 Patel A D, Gibson E, Ratner J, et al. Processing syntactic relations in language and music: An event-related potential study. J Cogn Neurosci, 1998, 10: 717–733
121 Koelsch S, Gunter T C, v. Cramon D Y, et al. Bach speaks: A cortical "language-network" serves the processing of music. Neuroimage, 2002, 17: 956–966
122 Gingras B, Honing H, Peretz I, et al. Defining the biological bases of individual differences in musicality. Philos Trans R Soc Lond B Biol Sci, 2015, 370: 20140092
123 Koelsch S, Gunter T C, Wittfoth M, et al. Interaction between syntax processing in language and in music: An ERP study. J Cogn Neurosci, 2005, 17: 1565–1577

蒋存梅

上海师范大学音乐学院教授. 博士毕业于中央音乐学院, 曾在中国科学院心理学研究所从事博士后研究. 研究领域集中在音乐与语言音高对比、音乐情绪与意义加工等, 研究对象涉及先天失歌症(五音不全)、自闭症和绝对音高等人群. 已在国内外学术期刊发表 40 余篇论文.

张清芳

博士, 中国人民大学心理学系教授, 中国英汉语比较研究会中国心理语言学专业委员会常务理事, 中国心理学会普通心理和实验心理专业委员会委员. 在汉语语言产生及其与西方语言对比方面有较多研究, 主持多项国家自然科学基金和社会科学基金项目, 发表国内外期刊论文 40 余篇.

The evolutionary origins of language and music

JIANG CunMei[1] & ZHANG QingFang[2,3]

[1] *Music College, Shanghai Normal University, Shanghai 200234, China;*
[2] *Department of Psychology, Renmin University of China, Beijing 100872, China;*
[3] *Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China*

Language, as well as music, is ubiquity across cultures. The evolutionary origins of language and music have been studied since Charles Darwin's day. According to Darwin, human bodies and brains have been shaped by natural selection for language and music, and musical abilities must be ranked among the most mysterious with which endowed. However, whether language and music are targets of natural selection still remains uncertain. If language and music are unique to human society, exploring the evolutionary origins of language and music would be helpful to shed light on human evolution.

From an evolutionary perspective, Patel argues that adaptationist theory of music should meet criteria: human-specificity, innateness, and domain-specificity. Analogously, the theory of "natural selection for language" should satisfy these three criteria. Based on these, we reviewed the existing psychological literature on language and music to clarify the evolutionary origins of language and music.

The uniqueness of speech to humans is indisputable, but the question of how it came to be in humans and no other animal remains a source of contention. The development of simple speech and sign language provides basis for the origins of human speech. There are three distinct properties between human and other animal's communication system. Human speech is a system of symbols, and the communication among people is completed by joint attention and sharing intentions. Human Language bears grammatical structures and there are diverse language systems with distinct grammatical rules all over the world. From the perspectives of language acquisition and development, researchers proposed two theoretical frameworks to interpret the acquisition of human speech: Theory of Generative Grammar and Theory of Cognitive Functional Linguistics. Behavioral genetics study found that a gene Forkhead box P2 is related to a specific language deficits, reflecting that the innateness of human language. Study in cognitive neuroscience of language indicates that different pathways connecting frontal and temporal cortex. The dorsal pathway from the posterior portion of Broca's area to the superior seems to be of particular importance for high-order language functions, and this pathway probably is crucial for the evolution of human language.

In the past decade, advances in musical cognitive and neuroscientific research have leaded to renew interest in the debate of adaptationist and nonadaptationist theories of music. Proponents of the adaptationist theory of music support the biological origins of music proposed by Darwin. In contrast, Proponents of the nonadaptationist theory of music consider music as a technology or "transformative invention" that has important consequences for human culture and biology, or "auditory cheesecake"—a mere pleasure-producing substance. To address the evolutionary origins of music, we reviewed the existing literature from a perspective of both a high and low-level music processing. The current findings revealed that although animals are unable to perceive pitch and rhythm, unlike human, they have comparable processing of vocal emotion with human. Genetic and developmental studies also suggest that there may be individual difference on low-level music processing, but the capacities for music emotion and meaning may be ubiquity in human. Moreover, it is hard to verify there is brain areas or networks specialized for music since whether or not music and language share neural resources is still uncertain.

Taken together, further investigation on evolutionary language needs interdisciplinary collaborations and new methodologies (e.g., computational modeling). On the other hand, the current findings on music psychology would not provide sufficient evidence to support adaptationist or nonadaptationist view of music. A multidimensional perspective on musical processing is needed to further clarify the evolutionary origin of music. Moreover, protolanguage would provide a new way to further explore the processing of language and music.

language, music, evolutionary origin, human-specificity, innateness, domain-specificity

doi: 10.1360/N972015-01144