

# 婴儿对语音和非语音的感知:认知和神经机制\*

于文汶<sup>1,2</sup>, 陈淑美<sup>2</sup>, 沈钧石<sup>2</sup>, 张丹丹<sup>1</sup>

(1. 四川师范大学脑与心理科学研究院, 成都 610066; 2. 深圳大学心理学院, 深圳 518060)

**摘要:** 婴儿听觉感知能力的发展对于他们未来的语言学习和社会化都具有重要意义。过去大量的研究主要关注语音感知方面, 只有较少的研究将非语音感知纳入考虑之中, 但了解非语音感知的特征和机制将有助于增加研究者对听觉加工以及儿童发育的认识。该文分别介绍了婴儿语音感知中的三种偏好——对语音、“婴儿语”和母语的偏好, 并尝试着将非语音分为音乐、人类的非言语发声、环境声音三类进行阐述。通过对比这两大类声音的感知得到婴儿可能存在语音感知的左脑偏侧化和音乐感知的右脑偏侧化现象, 但这也尚存争议, 目前有特定领域模型、特定线索模型和脑网络模型三种理论对偏侧化现象的认知机制进行解释。

**关键词:** 语音感知; 非语音感知; 婴儿; 听觉加工

**中图分类号:** B842.5

**文献标识码:** A

**文章编号:** 1003–5184(2022)03–0201–09

## 1 引言

早在人类出生之前听觉系统就已开始工作, 在出生后视觉尚未发展成熟的婴儿时期, 听觉对于婴儿感知周遭环境起着重要的作用。当婴儿同时进行多种认知任务, 尤其是视觉和听觉通道同时参与时, 通常表现为听觉优势效应 (Robinson & Sloutsky, 2019)。因此, 研究婴儿的听觉加工是非常有必要的。然而过去此领域大量的研究主要关注语音感知方面, 只有较少的研究将非语音感知纳入考虑之中, 而了解婴儿非语音感知的认知特征和神经机制有助于增加我们对听觉加工和儿童发育的认识。例如, 婴儿对语音和非语音的偏好可以预测他们未来的语言能力 (Sorcinelli et al., 2019), 还可以用于区分学习障碍孩子的类型 (Kuuluvainen et al., 2016) 以及作为孤独症早期诊断的辅助指标 (Galilee et al., 2017)。因此, 同时将语音感知和非语音感知进行探讨和评估具有重要意义。在婴儿的生活环境中存在着各种各样的声音, 这些不同的声音对婴儿听觉系统的发展均具有重要作用 (Hamalainen et al., 2011)。听觉加工分为声音识别和定位, 它们有着不同的加工通路 (Alain et al., 2001), 在这里主要探讨的是婴儿对不同声音的识别, 而考虑到语音的独特性和重要性, 因此分为语音和非语音感知两大类分别进行论述, 婴儿对这两类声音的感知有哪些特点? 其背后的认知和神经机制是否有所差异呢? 该文将对这些问题进行探讨。

## 2 语音感知

语言是人类有别于动物最本质的能力之一, 而婴儿期语音感知的发展对未来语言学习起着至关重要的作用, 因此, 语音感知一直是研究者关注的重点。人类在出生时乃至出生前就处在复杂的听觉环境中, 在万千声音中对语音的偏好奠定了言语交流的基础, 成长初期独特的“婴儿语”交流方式得以让人们快速地把握这种交流的特征, 对母语的特异性感知适应了言语交流的社会性发展 (周玉, 张丹丹, 2017), 因此, 在婴儿期语音感知的发展中, 这三种偏好具有重要意义。

### 2.1 对语音的偏好

从出生开始, 人脑对语音的感知就具有特异性。行为学研究发现, 新生儿听语音时吮吸频率会明显增多 (Vouloumanos & Werker, 2007), 即便是与声学特征相似的非语音相比, 两岁婴儿仍然对语音表现出偏好 (Vouloumanos & Werker, 2004)。在脑成像研究中也得到了相同的结果, May 等 (2018) 使用近红外光学技术 (near-infrared spectroscopy, NIRS) 发现, 与用哨声形式朗读故事相比, 新生儿的颞区前部对语音朗读故事的反应更大。不过有研究也发现, 与恒河猴叫声相比, 新生儿却并没有表现出对语音的偏好, 这可能是因为新生儿还不能区分同物种和其他物种发出的声音; 但可在 3 月龄的婴儿中观察到对语音的选择性注意 (Vouloumanos et al., 2010)。新生儿和小婴儿表现出的这种对语音的偏好对他们

\* 基金项目: 国家自然科学基金 (32271102、31970980、31920103009), 国家社会科学基金 (20&ZD153) 以及深港脑科学创新研究院基金 (2022SHIBS0003) 支持。

通讯作者: 张丹丹, E-mail: zhangdd05@gmail.com。

未来学习语言具有重要意义 (Shultz & Vouloumanos, 2010)。

为何婴儿甚至新生儿具有语音偏好? 有研究者认为是因为在母亲怀孕期间, 胎儿更长时间地暴露在语音环境中, 所以一出生时就对语音有着偏好 (DeCasper & Spence, 1986)。但也有研究者并不认可这个观点, 他们认为婴儿一出生时对于语音的敏感性和较成熟的语音处理能力, 主要依赖于特异化的听觉神经回路, 由于宫内的听觉环境不足以产生这个复杂的大脑听觉系统, 因此这些学者认为语音加工系统的遗传性在新生儿和婴儿的语音偏好中起着决定性的作用 (Dehaene - Lambert et al., 2006)。

## 2.2 对“婴儿语”的偏好

“婴儿语”(IDS)是指成人同婴儿交流时不自主地产生的一种特殊语言形式 (Saint - Georges et al., 2013), 其具有语速较慢、音调较高、以及语音波形包络变化幅度大的特点 (Narayan & McDermott, 2016)。婴儿对 IDS 存在着偏好 (Outters et al., 2020; Sulpizio et al., 2018), 早在 1990 年时, 就有行为学研究发现相较于“成人语”(ADS), 新生儿在听 IDS 时, 对视觉刺激的注视时间更长 (Cooper & Aslin, 1990)。随着认知神经科学研究技术的发展, Haden 等 (2020) 使用 EEG 技术发现以 ADS 呈现的单词会在新生儿大脑顶部中心位置引起一个 200 ~ 600 毫秒时间窗内缓慢的正波, 而以 IDS 呈现的单词则会在其额叶引起一个 700 ~ 900 毫秒时间窗内微小的负波, 这说明在刚出生时婴儿就能够有效地区分 ADS 和 IDS。Saito 等 (2007) 和 Sulpizio 等 (2018) 使用具有较高空间分辨率的 NIRS 技术发现, 相较于 ADS, 新生儿和四月龄婴儿的额叶对 IDS 的激活更强。IDS 对于婴儿未来语言学习具有重要意义, 通过促进对语音的有效编码 (Kalashnikova et al., 2018), 从而使婴儿能够更好地学习词语 (Ma et al., 2020)。同时, “婴儿语”中丰富的声调变化也更能吸引婴儿的注意力, 增加他们与说话者的互动, 而互动有利于他们的社会性发展 (周玉, 张丹丹, 2017)。过去有研究发现患有抑郁症的母亲较少用 IDS 与她的孩子交流 (Kaplan et al., 2001), 而使用 IDS 的人会比使用 ADS 的人表现出更大兴趣去与婴儿互动, 从而给予更多的关注和照顾, 因此, 婴儿还可以通过 IDS 来判断照顾者是否有兴趣或能力去照顾他们, 从而通过投以更长的注视时间去选择更合适的照顾者, 这种选择性奠定了他们社会认知的基础 (Schachner & Hannon, 2011)。

## 2.3 对母语的偏好

婴儿对母语有着偏好 (Moon et al., 1993; Soley

et al., 2015; Sato et al., 2011)。语音感知开始是一个普遍的过程, 婴儿在出生时对母语和非母语的音位信息普遍敏感, 但随着年龄的增长, 通过社会互动和持续接触语音, 他们逐渐丧失了对非母语音位的区分能力, 同时对母语音位的区分能力增强, 然后变得专门感知母语的声音 (Tsao et al., 2006; Werker & Hensch, 2015), 而这一过程主要发生在生命第一年的后半年中 (Ramirez et al., 2017; Kuhl et al., 2006; Werker & Tees, 1984)。行为学研究发现 5、6 月龄的婴儿就已经会选择去看那个说母语而不是外语的人, 而 10 月龄的婴儿更偏向于接受母语者的玩具 (Kinzler et al., 2007)。许多的 ERP 研究也可以支持这一观点。在成年人的听觉研究中常将失匹配负波 (mismatch negativity, MMN) 作为大脑是否成功识别新旧两种声音之间差异的指标 (Marklund et al., 2018)。而在婴儿的听觉研究中也发现了一个类似于 MMN 的 ERP 成分, 称为失匹配反应 (mismatch response, MMR), 这个成分通常被用来反映普遍音位感知到母语音位特异性感知的转变 (Näätänen et al., 1997), 它分成两个部分, 早期成分通常发生在刺激后约 100 至 260 毫秒之间, 标志着不太成熟的声学水平的神经编码和识别。而晚期成分是刺激后约 260 至 460 毫秒的典型潜伏期, 标志着更成熟的音位水平的分析 (Shafer et al., 2010), 在婴儿的普遍音位感知到母语音位特异性感知的转变过程中, 呈现出对母语和非母语声音反应的早期 MMR 成分减少, 同时对母语声音晚期 MMR 成分选择性增加的现象 (Kuhl et al., 2008; Rivera - Gaxiola et al., 2005)。婴儿对母语的敏感性一方面有利于他们获取到更多的语言线索, 奠定语言学习的基础, 另一方面也有助于他们在母语的环境中获得更多的知识和资源, 同时帮助他们形成社会认同。

## 3 非语音感知

该文所指的非语音并不是简单的纯音, 而是现实环境中复杂的有意义的声音, 过去只有少量的研究者尝试对非语音进行分类, 目前这一分类体系可能还不完善。例如, Lewis 等 (2004) 简单地将那些除了语音外的声音都称为环境声音。而 Fröhholz 等 (2017) 将其分为人类的非言语发声 (human non - speech vocalizations)、环境声音和人造物体的声音 (human - produced sounds), Salvari 等 (2019) 又将非语音分为自然声音 (natural sound), 音乐声 (musical sound) 和人造声音 (artificial sound)。考虑到分类的完整性和准确性, 该文最终参照 (Boatman, 2006) 中对复杂声音的分类, 将非语音分为音乐声、人类的非言语发声和环境声音三类。

### 3.1 音乐

该文中的音乐是指不含歌词的乐曲。音乐对婴儿具有重要意义,它不仅与婴儿语言的发展有关(Politimou et al., 2019),还可以促进大脑的发育(Lordier et al., 2019)。婴儿对音乐的感知是与生俱来的,给胎儿播放高强度音乐会让他们心率加速以及运动反应增强,而低强度音乐则显示出相反的效果(Deliège, 1996),给新生儿呈现一段音乐,他们自发的吮吸、哭闹和肢体运动会与给定的节奏同步(Provasi et al., 2014)。这一行为表现同时也出现在其他年龄段的婴儿身上(Zentner et al., 2010),还有研究发现他们天生就对音乐和歌声有着偏好(Sarkamo et al., 2013),婴儿在母亲唱歌时比母亲说话时表现出更加专注和投入(Nakata & Trehub, 2004)。除了行为研究的结果,ERP研究也为婴儿的音乐感知提供了证据,给新生儿呈现一系列有节奏的音乐声,在一个违反声音序列节奏的拍子出现时,会在新生儿大脑中诱发MMR成分,表明他们能够分辨出该偏差刺激(Winkler et al., 2009)。这一现象在6月龄婴儿对一段旋律中的偏差刺激检测中也能看到(Tew et al., 2009)。由于音乐所涉及到的脑区远不止听觉皮层,还包括广泛的双侧网络,与注意力、运动功能、记忆相关的额叶、颞叶、顶叶和皮层下区域以及情绪处理相关的边缘系统(Koelsch, 2011; Patel, 2003)。因此,还可以将音乐作为婴儿多感官刺激的有效工具(Standley, 2012; Chorna et al., 2019),现在临床上已有大量有关于音乐治疗的研究(Haslbeck & Bassler, 2020),比如有fMRI研究发现对早产儿进行音乐训练,会增强他们听觉皮层与丘脑和背侧纹状体之间的连接性,这不仅会促进了他们对音乐节奏的加工,而且还会使他们产生愉悦感(Lordier et al., 2019)。这足以看出,音乐在婴儿发展过程中具有重要意义。

### 3.2 人类的非言语发声

人类的非言语发声指:除了语音,人们所发出的一切用于或不用于交流的声音,包括哭、笑或叹气声等。刚出生的新生儿就能够区分自己的哭声和其他同龄婴儿的哭声(Dondi et al., 1999)。人类的非言语发声在3月龄婴儿的双侧颞叶皮层有着相似的广泛激活模式(McDonald et al., 2019)。而到了4到7月龄,婴儿的双侧颞叶皮层进一步分化并特异性地开始加工和处理人类的声音(Grossmann & Friederici, 2012)。婴儿可以区分不同的非言语发声,尤其对于哭声最为敏感。Crespo - Llado 等(2018)记录了8月龄婴儿在听其他婴儿哭闹、大笑和咳嗽时的额叶ERP数据。结果表明,与正性和中性声音相

比,同伴哭声诱发了幅度更大的N100和晚期正成分(LPP)。同样地,Missana 等(2017)也发现婴儿在听到另一个婴儿的哭声而不是笑声或成人哼唱声时会产生N200的增强。一项fMRI研究发现,与中性发声相比,3至7月龄的婴儿听哭叫声会导致他们的脑岛和眶额皮层更活跃,而对笑的反应与其对中性发声的反应并无差异(Blasi et al., 2011)。这个现象可能与情绪性语音加工中的负性偏向现象一致(Cheng et al., 2012),婴儿对负性情绪声音更为敏感,从进化的角度来看,这有利于他们避开危险,选择更安全的环境,适应人类的生存和发展。

### 3.3 环境声音

环境声音分为两类,包括自然声音和人造声音。自然声音指的是自然界的各种声音,如动物叫声、风声等,而人造声音指的是在进化和发展过程中人们所制造出来的一些声音,如汽车鸣笛声、手机铃声等。已有的成年人研究表明,无论人们是否能够将环境声音正确进行分类,大脑可以非常快速地辨别环境声音(Lucia et al., 2012)。一项对成年人的fMRI研究发现,同时让被试关注各种环境声音,与不可识别的、逆序播放的声音不同,可识别的环境声音激活了左侧与语义加工相关的神经网络,以及双侧颞中回的后部(Lewis et al., 2004)。另一项成年人的fMRI研究发现,分类动物发声主要激活双侧颞上回的中间部分,而分类工具声音则主要激活左半球的“镜像网络”(Lewis et al., 2005),后者表明当人们注意到工具声音时会有效地将声音与其可能产生的动作联系起来(Lewis et al., 2004)。Salvari 等(2019)使用脑磁图考察了不同种类声音加工过程中的脑神经反应。结果显示,与自然声音相比,人造声音在颞上皮层、后扣带回、顶叶下叶和双侧前额叶皮层引起了更强的神经活动,同时这些脑区间的功能连接也得到增强。

目前有关婴儿对环境声音加工的研究非常少,这可能是由于大部分环境声音需要学习才能够进行加工和区分,而婴儿的后天学习经验较少,难以对环境声音进行识别和分类。不过,有研究者提出,声音的进化意义会影响人们对其的加工(Sharda & Singh, 2012)。例如,给9月龄的婴儿听进化恐惧相关的声音(蛇嘶嘶的声音),现代恐惧相关的声音(爆炸声)和愉快的声音,与其他两类声音相比,婴儿在听与进化恐惧相关的声音时,心率减慢,眨眼幅度更大,视觉定向能力更强(Erlich et al., 2013),这说明婴儿会更加关注具有进化恐惧意义的环境声音,因此,在未来的研究中还可以将声音的进化意义纳入讨论之中。

#### 4 婴儿语音和非语音感知的对比

婴儿在加工语音和非语音时,双侧颞叶显著激活(Sato et al., 2011; Minagawa - Kawai et al., 2011; McDonald et al., 2019; Shultz et al., 2014),这可能是由于它们作为一个听觉刺激,都共同激活了初级听觉皮层。

婴儿加工语音的脑区主要为颞叶和额叶,并且呈现出左侧激活更强的现象(Minagawa - Kawai et al., 2011; Cabrera & Gervain, 2020; Perani et al., 2011; Pena et al., 2003)。Dehaene - Lambertz 等(2002)使用 fMRI 技术研究三月龄婴儿在听语音时的脑活动,发现其激活范围为颞上回(包括赫氏回)到颞上沟和颞极周围区域。Minagawa - Kawai 等(2011)使用 NIRS 技术发现四月龄的婴儿在听语音时颞上回和额下回显著激活。最新的一篇元分析发现,新生儿就已存在相对完善的语音感知神经系统,加工语音的脑区主要为额叶(尤其是左侧额下回)、颞叶(尤其是颞上回),其中左侧额下回在检测语音结构中发挥关键作用,而颞上回在检测偏差刺激中扮演重要角色(陈钰等, 2020)。婴儿的语音加工存在左半球处理优势。利用 NIRS(Pena et al., 2003)和 fMRI(Dehaene - Lambertz et al., 2002)的脑成像技术均得到了婴儿的左侧颞叶在加工正序播放语音时的激活水平明显高于逆序播放语音和安静条件,这与在成人中所发现的左侧大脑激活更强的模式相似(Pena et al., 2003)。不过这些实验中所使用的逆序播放的语音材料是无实际意义的,Shultz 等(2014)进一步将语音与有实际意义的非语音(人类的非言语发声)作对比,使用 fMRI 对 1 至 4 月龄的婴儿进行记录,发现与那些咳嗽、打喷嚏声相比,颞叶皮层对语音有左偏侧化反应,并且在出生后的第一个月就对语音具有高度的选择性,同时在接下来的 3 个月里,随着听觉神经网络对其他非语音的反应越来越弱,左侧颞叶皮层对语音的选择性也越来越强。然而,也有一些研究并没有得到这种偏侧化的结果(Price et al., 2005; Kotilahti et al., 2010)。这可能与婴儿的发育年龄有关,大脑半球间的相互作用是大脑发育的一个重要方面,这使得两个半球产生的活动得以协调和整合,从而形成大脑半球偏侧化的成人模式(Perani et al., 2011),但早期两个半球的功能连接还没有得到很好的发展。因此,年龄更大的婴儿比小婴儿的左偏侧化现象更明显,不过由于大脑具有可塑性,随着在语音环境的暴露越多,从儿童期到青春期,语音加工向左半球偏侧化的现象也有所增加(Holland et al., 2001; Szaflarski et al., 2006; Brauer et al., 2008)。

目前有关婴儿加工环境声音、人类非言语发声的研究较少,其背后的神经机制也尚不清楚,因此这里主要关注婴儿加工音乐的神经机制。一项 fMRI 研究测量了 1 到 3 日龄的新生儿在听到西方音乐节选时的大脑活动,研究发现新生儿在听音乐时,右侧颞上回激活显著,其中初级听觉皮层(颞横回)激活最强,并延伸至次级听觉皮层、颞极平面(planum polare)、颞平面(planum temporale)、颞顶联合区和顶下小叶。此外,在右侧的脑岛和右侧杏仁核-海马复合体(amygdala - hippocampal complex)中还观察到了激活。该研究还发现了新生儿在加工音乐时右半球的初级和次级听觉皮层的激活强于左半球(Perani et al., 2010)。这种对音乐加工的大脑右侧优势与在失音症患者(Sarkamo et al., 2013)和健康成人(Zatorre et al., 2002; Tervaniemi & Hugdahl, 2003)中的现象一致,研究者认为右初级听觉皮层在音高分析和整合中起着关键作用(Zatorre & Belin, 2001),而这是加工音乐的先决条件。

对于这种偏侧化现象,过去存在着特定领域模型和特定线索模型尝试着去进行解释(Zatorre & Gandour, 2008)。特定领域模型认为语音是一种独特的声音输入,其加工依赖于一种特定的神经回路,而不适用于其他类型声音的处理(Lieberman & Whalen, 2000),音乐可能也同样如此(Peretz, 2001)。而特定线索模型认为两个半球的听觉皮层的功能是相对专门化的,左听觉皮层对时域信息的分辨率更好,而右听觉皮层对频域信息的分辨率更好,低水平的声学特征在不同声音的感知中起着关键作用。根据这个模型可得左半球主要处理快速变化的听觉信息,而右半球主要加工频域信息丰富的听觉信息(如声音的韵律线索)(Talkington et al., 2013)。语音和音乐有着不同的时域和频域特征,比如语音和音乐的时间调制分别在 5 和 2Hz 左右显示宽而分离良好的峰值(Ding et al., 2017)。此外,它们依赖于声学线索的程度也不同,语音高度依赖于快速变化的宽带声音(Zatorre et al., 2002),而在区分语音中的语调和音乐中的音高方面,频域信息显得更为关键(Reybrouck & Podlipniak, 2019)。因此特定线索模型认为这种半球差异可能是由于声音特征线索的差异而导致的(Zatorre et al., 2002; Telkemeyer et al., 2011)。该模型是基于声音的声学特性提出的,而婴儿已具有处理这些声学信息的能力(Häden et al., 2015; Carral et al., 2005; Sambeth et al., 2008),因此这个模型同样也适用于去解释婴儿中发现的偏侧化现象。不过,也有研究者质疑这种简单的分工观点,他们认为声音识别涉及到复杂

的神经计算,大脑皮层内在的神经元振荡在解码不同声音中起着重要作用(Poeppel,2014)。

除此之外,也有研究者提出了脑网络模型,即这些声音的加工不依赖于具体的、离散的皮层区域,而是由于不同区域的功能连接水平不同,但每个区域都参与语音和非语音任务的处理过程(Price et al., 2005)。例如,在加工语音、环境声音和音乐中频率和振幅变化的神经基础可能是相同的,但在语音过程中可能对其连接性的水平要求更高。不只聚焦于某一区域的作用,而从更宏观的脑网络层面探讨不同区域的功能连接、协同作用,正是当前研究的大势所趋,其方法学上的发展,还能够进一步地帮助研究者去探讨复杂脑网络功能和结构之间的关系(Bullmore & Sporns,2009)。

## 5 总结与展望

综上所述,首先,从出生开始,婴儿就对语音存在加工偏好,目前有环境和遗传影响两种解释。其次,相较于ADS,婴儿会更喜欢IDS,IDS不仅能够促进婴儿对词汇的学习,为他们未来语言学习打下基础,还可以增加他们的社会互动,有利于社会性的发展。除此之外,IDS还可以作为婴儿选择合适照顾者的有效信号,并奠定他们社会认知的基础。再者,婴儿还对母语有着偏好,语音感知开始是一个普遍的过程,但在生命第一年尤其是后半年中婴儿就逐渐丧失了对非母语音位的区分能力,同时对母语音位的区分能力增强,实现了知觉重构。就非语音感知而言,婴儿对音乐有特殊偏好,他们对于音高、音色和节奏的检测能力似乎是与生俱来的,由于音乐在婴儿的成长过程中的重要作用,所以音乐训练和治疗在临床中被广泛应用;婴儿对人类的非言语发声也非常敏感,并且似乎对哭声的反应最为强烈,这可能与情绪性语音加工中的负性偏向现象一致。目前有关婴儿对环境声音加工的研究还非常少,这可能与婴儿在此方面知识经验匮乏而导致他们无法分辨和精细加工该类声音有关。

通过对比婴儿的语音和非语音感知可知,他们似乎天生就具有处理时域和频域信息的能力。但是由于婴儿的头围小,头皮上无法密集安放脑电或近红外探头,且婴儿进行磁共振扫描一般需要服用或注射镇静剂(有副作用),因此目前对婴儿语音和非语音加工的神经机制了解得还比较粗略,已掌握的大脑功能分化的知识还基本处于“脑半球”的水平,即婴儿可能存在语音感知的左脑偏侧化和音乐感知的右脑偏侧化现象,但这也尚存争议,还需要未来研究进一步阐明其背后的神经机制,不过目前有特定领域模型、特定线索模型和脑网络模型三种理论对

偏侧化现象的认知机制进行解释。

未来该领域的研究可以从以下方面展开。第一,同时纳入语音和非语音进行对比研究。婴儿早期的发育障碍会影响人们语言和交流的发展,而语音和非语音的感知在检测这些疾病中起着重要作用。例如,有研究表明,比起语音,孤独症儿童比正常儿童更偏好计算机合成的声音(Kuhl et al., 2005),除此之外,孤独症患者往往还表现出对情绪性声音的加工异常(Frühholz & Staib,2017),他们在区分语音和非语音时无法激活右半球相关脑区,这可能与语音检测的社会或情绪功能障碍有关(Galilee et al.,2017)。另外,提高我们对处理非言语发声的神经机制早期发展的理解,还有望为沟通障碍儿童提供更好的诊断和早期干预策略(Talkington et al.,2013)。因此,同时探讨人类发展早期的语音和非语音加工是很有必要的。

第二,准确划分非语音的种类,并在研究中多选用现实生活中的复杂声音。由于不同研究中使用了截然不同的声音材料,这使得难以对各项研究的结果进行对比和归纳总结。因此,明确和细化非语音的种类,选用标准的实验材料非常有必要。此外,尽管不少研究已将非语音声音纳入考察,目前许多实验仅聚焦于声音的声学特性,例如使用纯音作为实验材料,但实际生活中的声音都是时域和频域信息非常丰富的复杂声音,如果将这些声音刺激简单拆分成单一的声学特征进行考察,很可能会丢失情绪(Frühholz et al.,2016)、进化(Sharda & Singh,2012)意义等重要信息。因此我们建议直接选取语音和非语音中的一些具有代表性的声音作为实验材料,正如Shultz等(2014)在研究中直接使用语音、人类非言语发声以及恒河猴的叫声进行对比。

## 参考文献

- 陈钰,莫李澄,毕蓉,张丹丹.(2020). 新生儿语音感知的神经基础:元分析. *心理科学进展*,28(8),1273-1281.
- 周玉,张丹丹.(2017). 婴儿情绪与社会认知相关的听觉加工. *心理科学进展*,25(1),67-75.
- Alain, C., Arnott, S. R., Hevenor, S., Graham, S., & Grady, C. L. (2001). 'what' and 'where' in the human auditory system. *Proceedings of the National Academy of Sciences of the United States of America*,98(21),12301-12306.
- Blasi, A., Mercure, E., Lloyd-Fox, S., Thomson, A., Brammer, M., Sauter, D., ... Murphy, D. G. (2011). Early specialization for voice and emotion processing in the infant brain. *Current Biology*,21(14),1220-1224.
- Boatman, D. F. (2006). Cortical auditory systems: Speech and other complex sounds. *Epilepsy & Behavior*,8(3),494-503.

- Brauer, J. , Neumann, J. , & Friederici, A. D. (2008). Temporal dynamics of perisylvian activation during language processing in children and adults. *Neuroimage*, 41(4), 1484 – 1492.
- Bullmore, E. , & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3), 186 – 198.
- Cabrera, L. , & Gervain, J. (2020). Speech perception at birth: The brain encodes fast and slow temporal information. *Science Advances*, 6(30), eaba7830.
- Carral, V. , Huotilainen, M. , Ruusuvirta, T. , Fellman, V. , Naatanen, R. , & Escera, C. (2005). A kind of auditory ‘primitive intelligence’ already present at birth. *European Journal of Neuroscience*, 21(11), 3201 – 3204.
- Cheng, Y. W. , Lee, S. Y. , Chen, H. Y. , Wang, P. Y. , & Decety, J. (2012). Voice and emotion processing in the human neonatal brain. *Journal of Cognitive Neuroscience*, 24(6), 1411 – 1419.
- Chorna, O. , Filippa, M. , De Almeida, J. S. , Lordier, L. , Monaci, M. G. , Huppi, P. , . . . Guzzetta, A. (2019). Neuroprocessing mechanisms of music during fetal and neonatal development: A role in neuroplasticity and neurodevelopment. *Neural Plasticity*, 2019, 3972918.
- Cooper, R. P. , & Aslin, R. N. (1990). Preference for infant – directed speech in the first month after birth. *Child Development*, 61(5), 1584 – 1595.
- Crespo – Llado, M. M. , Vanderwert, R. E. , & Geangu, E. (2018). Individual differences in infants’ neural responses to their peers’ cry and laughter. *Biological Psychology*, 135, 117 – 127.
- DeCasper, A. J. , & Spence, M. J. (1986). Prenatal maternal speech influences newborn’s perception of speech sounds. *Infant Behavior and Development*, 9(2), 133 – 150.
- Dehaene – Lambertz, G. , Dehaene, S. , & Hertz – Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science*, 298(5600), 2013 – 2015.
- Dehaene – Lambertz, G. , Hertz – Pannier, L. , & Dubois, J. (2006). Nature and nurture in language acquisition: Anatomical and functional brain – imaging studies in infants. *Trends in Neurosciences*, 29(7), 367 – 373.
- Deliège, I. , & Sloboda, J. A. (1996). *Musical beginnings: Origins and development of musical competence*. Oxford University Press.
- Ding, N. , Patel, A. D. , Chen, L. , Butler, H. , Luo, C. , & Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience and Biobehavioral Reviews*, 81 ( Pt B ), 181 – 187.
- Dondi, M. , Simion, F. , & Caltran, G. (1999). Can newborns discriminate between their own cry and the cry of another newborn infant? *Developmental Science*, 35(2), 418 – 426.
- Erlich, N. , Lipp, O. V. , & Slaughter, V. (2013). Of hissing snakes and angry voices: Human infants are differentially responsive to evolutionary fear – relevant sounds. *Developmental Science*, 16(6), 894 – 904.
- Frühholz, S. , & Staib, M. (2017). Neurocircuitry of impaired affective sound processing: A clinical disorders perspective. *Neuroscience & Biobehavioral Reviews*, 83, 516 – 524.
- Frühholz, S. , Trost, W. , & Kotz, S. A. (2016). The sound of emotions—towards a unifying neural network perspective of affective sound processing. *Neuroscience & Biobehavioral Reviews*, 68, 96 – 110.
- Galilee, A. , Stefanidou, C. , & McCleery, J. P. (2017). Atypical speech versus non – speech detection and discrimination in 4 – to 6 – yr old children with autism spectrum disorder: An ERP study. *Plos One*, 12(7), e0181354.
- Grossmann, T. , & Friederici, A. D. (2012). When during development do our brains get tuned to the human voice? *Social Neuroscience*, 7(4), 369 – 372.
- Háden, G. P. , Honing, H. , Török, M. , & Winkler, I. (2015). Detecting the temporal structure of sound sequences in newborn infants. *International Journal of Psychophysiology*, 96(1), 23 – 28.
- Haden, G. P. , Mady, K. , Torok, M. , & Winkler, I. (2020). Newborn infants differently process adult directed and infant directed speech. *International Journal of Psychophysiology*, 147, 107 – 112.
- Hamalainen, J. A. , Ortiz – Mantilla, S. , & Benasich, A. A. (2011). Source localization of event – related potentials to pitch change mapped onto age – appropriate MRIs at 6 months of age. *Neuroimage*, 54(3), 1910 – 1918.
- Haslbeck, F. B. , & Bassler, D. (2020). Clinical practice protocol of creative music therapy for preterm infants and their parents in the neonatal intensive care unit. *Journal of Visualized Experiments*, (155), e60412.
- Holland, S. K. , Plante, E. , Byars, A. W. , Strawsburg, R. H. , Schmithorst, V. J. , & Ball Jr, W. S. (2001). Normal fMRI brain activation patterns in children performing a verb generation task. *Neuroimage*, 14(4), 837 – 843.
- Kalashnikova, M. , Peter, V. , Di Liberto, G. M. , Lalor, E. C. , & Burnham, D. (2018). Infant – directed speech facilitates seven – month – old infants’ cortical tracking of speech. *Scientific Reports*, 8(1), 13745.
- Kaplan, P. S. , Bachorowski, J. A. , Smoski, M. J. , & Zinser, M. (2001). Role of clinical diagnosis and medication use in effects of maternal depression on infant – directed speech. *Infancy*, 2(4), 537 – 548.
- Kinzler, K. D. , Dupoux, E. , & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences of the United States of America*, 104(30), 12577 – 12580.
- Koelsch, S. (2011). Toward a neural basis of music perception – a review and updated model. *Frontiers in Psychology*, 2, 1 – 20.

- Kotilahti, K., Nissila, I., Nasi, T., Lipiainen, L., Noponen, T., Merilainen, P., ... Fellman, V. (2010). Hemodynamic responses to speech and music in newborn infants. *Human Brain Mapping*, 31(4), 595 – 603.
- Kuhl, P. K., Coffey – Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: Behavioral and electrophysiological measures. *Developmental Science*, 8(1), 9 – 20.
- Kuhl, P. K., Conboy, B. T., Coffey – Corina, S., Padden, D., Rivera – Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM – e). *Philosophical Transactions of The Royal Society B Biological Sciences*, 363(1493), 979 – 1000.
- Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9(2), F13 – F21.
- Kuuluvainen, S., Alku, P., Makkonen, T., Lipsanen, J., & Kuja-la, T. (2016). Cortical speech and non – speech discrimination in relation to cognitive measures in preschool children. *European Journal of Neuroscience*, 43(6), 738 – 750.
- Lewis, J. W., Brefczynski, J. A., Phinney, R. E., Janik, J. J., & DeYoe, E. A. (2005). Distinct cortical pathways for processing tool versus animal sounds. *The Journal of Neuroscience*, 25(5), 5148 – 5158.
- Lewis, J. W., Wightman, F. L., Brefczynski, J. A., Phinney, R. E., Binder, J. R., & DeYoe, E. A. (2004). Human brain regions involved in recognizing environmental sounds. *Cerebral Cortex*, 14(9), 1008 – 1021.
- Lieberman, A. M., & Whalen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Sciences*, 4(5), 187 – 196.
- Lordier, L., Loukas, S., Grouiller, F., Vollenweider, A., Vassung, L., Meskaldij, D. E., ... Huppi, P. S. (2019). Music processing in preterm and full – term newborns: A psychophysiological interaction (PPI) approach in neonatal fMRI. *Neuroimage*, 185, 857 – 864.
- Lucia, M. D., Tzovara, A., Bernasconi, F., Spierer, L., & Murray, M. M. (2012). Auditory perceptual decision – making based on semantic categorization of environmental sounds. *Neuroimage*, 60(3), 1704 – 1715.
- Ma, W., Fiveash, A., Margulis, E. H., Behrend, D., & Thompson, W. F. (2020). Song and infant – directed speech facilitate word learning. *Quarterly Journal of Experimental Psychology*, 73(7), 1036 – 1054.
- Marklund, E., Lacerda, F., & Schwarz, I. C. (2018). Using rotated speech to approximate the acoustic mismatch negativity response to speech. *Brain and Language*, 176, 26 – 35.
- May, L., Gervain, J., Carreiras, M., & Werker, J. F. (2018). The specificity of the neural response to speech at birth. *Developmental Science*, 21(3), e12564.
- McDonald, N. M., Perdue, K. L., Eilbott, J., Loyal, J., Shie, F., & Pelphrey, K. A. (2019). Infant brain responses to social sounds: A longitudinal functional near – infrared spectroscopy study. *Developmental Cognitive Neuroscience*, 36, 100638.
- Minagawa – Kawai, Y., van der Lely, H., Ramus, F., Sato, Y., Mazuka, R., & Dupoux, E. (2011). Optical brain imaging reveals general auditory and language – specific processing in early infant development. *Cerebral Cortex*, 21(2), 254 – 261.
- Missana, M., Altvater – Mackensen, N., & Grossmann, T. (2017). Neural correlates of infants' sensitivity to vocalexpressions of peers. *Developmental Cognitive Neuroscience*, 26, 39 – 44.
- Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two – day – olds prefer their native language. *Infant Behavior and Development*, 16(4), 495 – 500.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., et al. (1997). Language – specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, 385(6615), 432 – 434.
- Nakata, T., & Trehub, S. E. (2004). Infants' responsiveness to maternal speech and singing. *Infant Behavior And Development*, 27(4), 455 – 464.
- Narayan, C. R., & McDermott, L. C. (2016). Speech rate and pitch characteristics of infant – directed speech: Longitudinal and cross – linguistic observations. *Journal of the Acoustical Society of America*, 139(3), 1272 – 1281.
- Outters, V., Schreiner, M. S., Behne, T., & Mani, N. (2020). Maternal input and infants' response to infant – directed speech. *Infancy*, 25(4), 478 – 499.
- Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, 6, 674 – 681.
- Pena, M., Maki, A., Kovacic, D., Dehaene – Lamberts, G., Koizumi, H., & Bouquet, F. (2003). Sounds and silence: An optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences*, 100, 11702 – 11705.
- Perani, D., Saccuman, M. C., Scifo, P., Anwander, A., Spada, D., Baldoli, C., ... Friederici, A. D. (2011). Neural language networks at birth. *Proceedings of the National Academy of Sciences*, 108(38), 16056 – 16061.
- Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., ... Koelsch, S. (2010). Functional specializations for music processing in the human newborn brain. *Proceedings of the National Academy of Sciences*, 107(10), 4758 – 4763.
- Peretz, I. (2001). Brain specialization for music. *Annals of the New York Academy of Sciences*, 930(1), 153 – 165.
- Poeppel, D. (2014). The neuroanatomic and neurophysiological infrastructure for speech and language. *Current Opinion Neurobiology*, 28, 142 – 149.

- Politimou, N. , Dalla Bella, S. , Farrugia, N. , & Franco, F. (2019). Born to speak and sing: Musical predictors of language development in pre - schoolers. *Frontiers in Psychology*, 10, 948.
- Price, C. , Thierry, G. , & Griffiths, T. (2005). Speech - specific auditory processing: Where is it? *Trends in Cognitive Science*, 9(6), 271 - 276.
- Provasi, J. , Anderson, D. I. , & Barbu - Roth, M. (2014). Rhythm perception, production, and synchronization during the perinatal period. *Frontiers in Psychology*, 5, 1048.
- Ramirez, F. N. , Ramirez, R. R. , Clarke, M. , Taulu, S. , & Kuhl, P. K. (2017). Speech discrimination in 11 - month - old bilingual and monolingual infants: A magnetoencephalography study. *Developmental Science*, 20(1), 1 - 16.
- Reybrouck, M. , & Podlipniak, P. (2019). Preconceptual spectral and temporal cues as a source of meaning in speech and music. *Brain Sciences*, 9(3), 53.
- Rivera - Gaxiola, M. , Silva - Pereyra, J. , & Kuhl, P. K. (2005). Brain potentials to native and non - native speech contrasts in 7 and 11 - month - old American infants. *Developmental Science*, 8(2), 162 - 172.
- Robinson, C. W. , & Sloutsky, V. M. (2019). Two mechanisms underlying auditory dominance: Overshadowing and response competition. *Journal of Experimental Child Psychology*, 178, 317 - 340.
- Saint - Georges, C. , Chetouani, M. , Cassel, R. , Apicella, F. , Mahdhaoui, A. , Muratori, F. , ... Cohen, D. (2013). Motherese in interaction: At the cross - road of emotion and cognition? (a systematic review). *Plos One*, 8(10), e78103.
- Saito, Y. , Aoyama, S. , Kondo, T. , Fukumoto, R. , Konishi, N. , Nakamura, K. , ... Toshima, T. (2007). Frontal cerebral blood flow change associated with infant - directed speech. *Archives of Disease in Childhood Fetal & Neonatal Edition*, 92(2), F113 - 116.
- Salvari, V. , Paraskevopoulos, E. , Chalas, N. , Muller, K. , Wollbrink, A. , Dobel, C. , ... Pantev, C. (2019). Auditory categorization of man - made sounds versus natural sounds by means of MEG functional brain connectivity. *Frontiers in Neuroscience*, 13, 1052.
- Sambeth, A. , Ruohio, K. , Alku, P. , Fellman, V. , & Huotilainen, M. (2008). Sleeping newborns extract prosody from continuous speech. *Clinical Neurophysiology*, 119(2), 332 - 341.
- Sarkamo, T. , Tervaniemi, M. , & Huotilainen, M. (2013). Music perception and cognition: Development, neural basis, and rehabilitative use of music. *Wiley Interdisciplinary Reviews Cognitive Science*, 4(4), 441 - 451.
- Sato, H. , Hirabayashi, Y. , Tsubokura, H. , Kanai, M. , Ashida, T. , Konishi, I. , ... Maki, A. (2011). Cerebral hemodynamics in newborn infants exposed to speech sounds: A whole - head optical topography study. *Human Brain Mapping*, 33(9), 2092 - 2103.
- Schachner, A. , & Hannon, E. E. (2011). Infant - directed speech drives social preferences in 5 - month - old infants. *Developmental Psychology*, 47(1), 19 - 25.
- Shafer, V. L. , Yu, Y. H. , & Datta, H. (2010). Maturation of speech discrimination in 4 - to 7 - yr - old children as indexed by event - related potential mismatch responses. *Ear and Hearing*, 31(6), 735 - 745.
- Sharda, M. , & Singh, N. C. (2012). Auditory perception of natural sound categories - an fMRI study. *Neuroscience*, 214, 49 - 58.
- Shultz, S. , & Vouloumanos, A. (2010). Three - month - olds prefer speech to other naturally occurring signals. *Language Learning and Development*, 6(4), 241 - 257.
- Shultz, S. , Vouloumanos, A. , Bennett, R. H. , & Pelphrey, K. (2014). Neural specialization for speech in the first months of life. *Developmental Science*, 17(5), 766 - 774.
- Soley, G. , & Sebastian - Galles, N. (2015). Infants prefer tunes previously introduced by speakers of their native language. *Child Development*, 86(6), 1685 - 1692.
- Sorcinelli, A. , Ference, J. , Curtin, S. , & Vouloumanos, A. (2019). Preference for speech in infancy differentially predicts language skills and autism - like behaviors. *Journal of Experimental Child Psychology*, 178, 295 - 316.
- Standley, J. (2012). Music therapy research in the NICU: An updated meta - analysis. *Neonatal Network*, 31(5), 311 - 316.
- Sulpizio, S. , Doi, H. , Bornstein, M. H. , Cui, J. , Esposito, G. , & Shinohara, K. (2018). fNIRS reveals enhanced brain activation to female (versus male) infant directed speech (relative to adult directed speech) in Young Human Infants. *Infant Behavior and Development*, 52, 89 - 96.
- Szaflarski, J. P. , Holland, S. K. , Schmithorst, V. J. , & Byars, A. W. (2006). fMRI study of language lateralization in children and adults. *Human Brain Mapping*, 27, 202 - 212.
- Talkington, W. J. , Taglialatela, J. P. , & Lewis, J. W. (2013). Using naturalistic utterances to investigate vocal communication processing and development in human and non - human primates. *Hearing Research*, 305, 74 - 85.
- Telkemeyer, S. , Rossi, S. , Nierhaus, T. , Steinbrink, J. , Obrig, H. , & Wartenburger, I. (2011). Acoustic processing of temporally modulated sounds in infants: Evidence from a combined near - infrared spectroscopy and EEG study. *Frontiers in Psychology*, 62(1), 1 - 14.
- Tervaniemi, M. , & Hugdahl, K. (2003). Lateralization of auditory - cortex functions. *Brain Research Reviews*, 43(3), 231 - 246.
- Tew, S. , Fujioka, T. , He, C. , & Trainor, L. (2009). Neural representation of transposed melody in infants at 6 months of age. *Annals of the New York Academy of Sciences*, 1169, 287 - 290.
- Tsao, F. M. , Liu, H. M. , & Kuhl, P. K. (2006). Perception of n-



- ative and non – native affricate – fricative contrasts; Cross – language tests on adults and infants. *Journal of the Acoustical Society of America*, 120(4), 2285 – 2294.
- Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth; Evidence for a bias for speech in neonates. *Developmental Science*, 10(2), 159 – 164.
- Vouloumanos, A., Hauser, M. D., Werker, J. F., & Martin, A. (2010). The tuning of human neonates' preference for speech. *Child Development*, 81(2), 517 – 527.
- Vouloumanos, A., & Werker, J. F. (2004). Tuned to the signal: The privileged status of speech for young infants. *Developmental Science*, 7(3), 270 – 276.
- Werker, J. F., & Tees, R. C. (1984). Cross – language speech perception; Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49 – 63.
- Werker, J. F., & Hensch, T. K. (2015). Critical periods in speech perception; New directions. *Annual Review of Psychology*, 66, 173 – 196.
- Winkler, I., Haden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences*, 106(7), 2468 – 2471.
- Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, 11, 946 – 953.
- Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, 16(1), 37 – 46.
- Zatorre, R. J., & Gandour, J. T. (2008). Neural specializations for speech and pitch; Moving beyond the dichotomies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363, 1087 – 1104.
- Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences*, 107(13), 5768 – 5773.

## Infant Perception of Speech and Non – speech: Cognitive and Neural Mechanisms

Yu Wenwen<sup>1,2</sup>, Chen Shumei<sup>2</sup>, Shen Junshi<sup>2</sup>, Zhang Dandan<sup>1</sup>

(1. Institute of Brain and Psychological Sciences, Sichuan Normal University, Chengdu 610066;

2. School of Psychology, Shenzhen University, Shenzhen 518060)

**Abstract:** The development of auditory perception in infancy is of great significance to their future language learning and socialization. In the past, a large number of studies have focused on speech sound perception, and only a few have taken non – speech sound perception into consideration. However, learning the characteristics and mechanisms of non – speech sound perception can deepen our understanding of auditory processing and children's development. According to previous studies, Infants have a preference for speech sound, mother tongue and infant directed speech. By comparing the two kinds of sound perception, there may be left brain lateralization of speech perception and right brain lateralization of music perception in infants. However, this is still controversial. At present, there are three theories, namely, domain – specific model, cue – specific model and brain network model to explain the cognitive mechanism of the phenomenon.

**Key words:** speech perception; non – speech perception; infant; auditory processing