

·综述·

听觉语言加工的双流模型*

汪洁¹ 吴东宇^{2,3} 袁英¹ 张大华¹ 张甜甜¹ 王卓¹ 程亦男¹

随着神经影像学技术和皮质直接电刺激研究的发展,人类对语言加工的神经机制有了更多的认识。特别是近十年,对听觉语言加工的双流模型(dual-stream model of language processing)涉及的脑区功能和结构联系有了更深入的了解,为语言理解的连贯性功能解剖模型提供了良好的基础。2007年,Hickok和Poehpel^[1]提出了听觉语言加工的双流模型。其基本内容包括:听觉语言加工的执行由两个解剖分离和功能特异的平行加工流,腹侧流和背侧流完成。腹侧流由颞叶上中、前部和额叶前部的结构组成,参与加工言语理解信号(言语识别)。背侧流由额叶后部和颞叶后背侧部,以及顶岛盖的结构组成,其功能涉及了在颞叶后部把听觉言语信号转换为发音表征。背侧流具有听觉运动整合功能,它是言语发展和正常言语产生的基础。

1 听觉语言加工的腹侧流

腹侧流将听觉言语输入映射到概念和语义表征。颞叶初级听皮质被分为核心、带、带旁区。前腹侧流由带区头部和中部开始,经头部带旁、颞上沟和前颞叶,在多个额区终止,包括额极(10区)、46区和腹侧前额叶(12、45区)。位于核心的神经元对简单听觉刺激,如纯音反应最好,带区和带旁区的神经元对更复杂的声音反应最好(如发声)。大脑皮质的分层组织将系列成分及平行加工结合起来,腹侧流加工语音的频谱时间属性,涉及听觉模式识别^[2]。邻域激活(neighbourhood activation)模型提出从声音映射到意义按若干阶段序列发生,典型的操作是声学的、音位的、词汇水平表征^[3]。具有简单接收功能的“低级”区,如感觉核心区,投射到具有增加复杂反应性质的“高级”区,如带、带旁和前额叶腹侧皮质区。

与语言加工主要在左半球的观点不同,双流模型提出腹侧流为双侧组织(虽然两半球之间有重要的功能差异)。因此,腹侧流本身构成了平行加工流。这可以解释为什么单侧颞叶损害没有发现真正的言语识别缺陷。而且,双侧颞上回

损害伴有严重的言语识别缺陷即词聋,这与言语识别系统是双侧组织的观点是一致的^[4]。在罕见病例,词聋也可以发生在局限的单侧损伤^[5];但是这种事件发生的频率非常小,说明这种情况是普遍规律的例外。

1.1 音位分析

音位是言语信号的基本成分,皮质电刺激显示音位分析涉及了初级听皮质的前部和后部^[6]。功能影像学研究显示听觉音位分析在颞上回中部,于Heschl回外侧^[7-8]即初级听皮质,以及颞上回后部激活^[9]。颞上回背侧(后部)编码复杂的频谱时间特征,将更简单的声学特征进行整合;颞上回腹侧(前部)整合来自颞上回背侧的信息,建立与音位模式有关的更复杂的表征。频谱变化慢信息在颞上回前部和后部加工,而频谱变化慢与快信息在颞上回中部加工^[8]。这就可以解释为什么功能影像学 and 皮质电刺激研究音位加工涉及了颞上回前、后部以及中部。

1.2 词形识别

对听觉词形识别的加工脑区存在两种观点。一种观点认为听觉腹侧流的颞上回前部支持听觉词形识别^[10],短语加工开始于颞上沟前部。左颞上回前部被看做是腹侧流的“听觉词形区”^[11]。颞上回后部和顶下小叶损伤的患者表现为听觉词识别保留、听觉句子理解障碍,这种缺陷归因于句法加工缺陷^[12],即颞上回后部和顶下小叶参与句法加工。这个结果提示颞上回前部完整,保留了听觉词汇识别。

另一种观点认为词形加工位于左颞上回后部,如Roux等^[13]在术中电刺激发现,听觉词汇辨别涉及了左颞后部,与双音节辨别区一致。皮质电图记录显示左颞上回后部涉及了从声学信号到更高层次语音水平整合^[14-15]。

由于功能磁共振是在语言任务操作时观察脑区的血氧代谢水平,在语言任务下激活的脑区在它损伤时不一定造成该任务的缺陷。由此可见,这些技术只能间接提供大脑解剖与语言加工的关系。而在神经外科手术前或术中,对清醒患者的不同脑区实施直接电刺激,产生短暂的虚拟损伤。因

DOI:10.3969/j.issn.1001-1242.2017.07.025

*基金项目:国家自然科学基金资助课题(81171011;81272173;81572220);首都临床特色应用研究(Z121107001012144、Z171100001017028、Z171100001017111);国家科技支撑计划(2013BAH14F03)

1 首都医科大学宣武医院康复医学科,北京,100053; 2 中国中医科学院望京医院; 3 通讯作者

作者简介:汪洁,女,副主任治疗师;收稿日期:2016-06-03

此,直接电刺激为确定语言功能的脑结构提供了一个独特的、更准确的(约5mm)、可重复实施的方法。因此,直接电刺激的发现可能提供更可靠的证据。

1.3 词汇语义的加工

一旦语音词形得到确认,它的语义信息必须在随后的阶段得到提取。词汇语义加工阶段的脑区没有明确定位。在功能性磁共振成像(functional magnetic resonance imaging, fMRI)研究,词汇语义加工主要发生在颞中回,但是它们不局限在这一区域,它们也包括左右半球的联合皮质^[16]。它们可能是沿着左颞上回中部或更分散的位于缘上回周围和颞中回后部,由短纤维在它们之间传递神经冲动^[17]。虽然,早期 Geschwind 的研究把词汇语义加工定位于后颞顶区,但近些年的一些研究指出颞上回前部在听理解加工的作用。对灵长类的电生理研究显示,颞上回平面具有层级组织的听觉加工流,伴有逐步增加的刺激选择性^[18]。在人类,颞上回前部涉及了词和句子水平的语义整合^[19],它从音位(加工于左颞上回中部)到整合为词汇(左颞上回前部)、到句子的层级性解码梯度^[20]。脑激活研究报告了相似的该区听觉词形解码功能^[9]。皮质电刺激研究显示刺激颞上回前部、缘上回和颞中回后部损害了词汇语义系统^[21]。

Hickok 等^[11]认为颞中回后部以语音到意义界面网络的形式支持词汇和语义通达。颞中回后部区域涉及了颞上沟的语音表征与广泛分布的语义表征之间的映射。词汇和语义通达在某种程度上是双侧加工。目前,广泛认为构成词义的概念表征由外侧裂区以外的更广泛的脑区表征。当代研究显示,涉及词汇概念的脑区包括颞下回(前部)、颞中回和角回^[20-21]。

一些功能影像学研究显示句子水平加工涉及了前颞区,提示句法或联合加工可能更多地激活前颞叶。颞叶前部参与了词汇语义和句子水平加工的句法和语义整合加工。语义性痴呆患者涉及了双侧颞叶前部萎缩,词汇任务的缺陷包括命名、语义联系、词理解,其特征是有关事物概念知识的丧失,尤其是不常见事物(如,患者把斑马误认为马)^[22]。神经电生理的研究提示,颞极皮质具有在线概念加工的作用^[23]。功能影像研究发现当个体听或读句子,而不是无结构的词或语音,颞叶前部更为活跃。但是,大量的证据表明左侧颞叶后部受损导致听理解障碍。Hickok 等认为可能是腹侧投射通路向后部和前部扩展。

言语理解是一个复杂的多功能任务,它不仅要求听觉信息的加工,而且需要注意、工作记忆和整合加工,以便理解词汇、语句和叙述。因此,腹侧流是言语理解的主要加工网络,它支持词汇、语义和句法加工^[24]。沿着这条通路,腹侧流变得越来越左侧化,形成从后颞叶向前颞叶的“侧化梯度”^[25-26]。随着加工阶段,腹侧流形成了层级组织,即输入到一个水平,

依赖于前一个输入的输出。这种侧化梯度反映了操作加工的语言复杂性的增加,听知觉加工在后颞上回,语义和句法加工主要在前颞叶。这种层级性加工不是线性的和单向的,而是受到半球间交互加工,如注意和期待。

2 听觉语言加工的背侧流

背侧流是强左侧优势,它解释了为什么背侧颞叶和额叶损害出现显著的言语产生缺陷。听觉背侧流具有感觉运动整合功能。早在十九世纪,Wernicke 的语言神经环路的经典模型将言语感觉和运动表征之间构建了直接联系,并明确地提出感觉系统参与言语产生。言语感知的运动理论也认为,在感觉输入和运动性言语系统之间存在着连接。然而,言语的听觉与运动相互作用的证据来自言语发展。学习说话主要是运动学习任务,对它的初级输入是语音知觉,尤其是言语。因此,必定存在一个编码和保持言语声音的神经机制,能够应用这些感觉印迹指导言语音姿的调整,使得语音准确再现。

言语发展是形成背侧听觉运动整合环路的初级和关键时刻,在成年时它也继续起到作用。后者的证据包括改变听觉反馈对言语产生的破坏作用,当晚期出现聋可以造成发音能力以及获得新词汇能力的减退。这种听觉运动环路为语音短时记忆提供了基本的神经机制。

在听觉运动交互作用存在两个水平,一个涉及了言语音段,另一个涉及了音段序列。音段水平加工涉及了基本发音音位技能的获得和保持。在音段序列水平的听觉运动加工涉及了新词汇的获得和言语序列的实时指导。新词汇获得时,听觉运动相互作用涉及了产生新词的感觉表征,它编码音段或音节序列。这种感觉表征可用来指导运动发音序列,它可能涉及了真正的前馈机制(言语序列感觉码被转换成运动言语序列)、反馈监测机制或两者。当词变得熟悉,感觉运动交互作用的性质改变了。新的、低频或更复杂的词需要增量的运动编码,因此需要更多的感觉指导;高频或更简单的词作为运动组块变得更自动化,它需要极少的感觉指导。

2.1 来自脑损伤的证据

左半球听觉相关区域的损伤常导致言语产生障碍,显示感觉系统参与了运动性言语。特别是损伤了左侧颞上回后部或颞顶联合区伴有传导性失语症,特点是理解好,但言语产生时出现大量的音位性错语。传导性失语症传统上被认为是失联综合征,涉及了弓状束损伤。然而,现在的证据表明皮质功能损害也可以造成复述障碍。言语产生缺陷是负荷敏感,即错误更多的是长词、低频词和无语义的言语串的复述。功能上,传导性失语症的特征是言语输出时编码语音信息的能力受损。

Hickok 等提出传导性失语症代表了听觉运动界面的破

坏。其言语理解保留,是因为损伤没有破坏腹侧流通路,或因为右半球言语系统可以代偿左半球言语知觉系统的损害。语音错误的发生是因为言语的感觉表征妨碍了提供言语语音序列的实时指导。这一作用对长词、低频词、新词更为突出,因为这些词比短的、高频词依赖于更大的感觉介入。功能影像学研究显示造成传导性失语症的脑区活动常常受到隐性命名任务词长的影响。

2.2 背侧流的脑区及其皮质下白质

Hickok 和 Poeppel^[1]提出的双流模型中强调背侧流是强左侧优势。Cogan 等^[27]在患者操作言语(词汇和非词汇)复述的感觉运动任务时应用直接神经记录,观察到感觉运动转换发生在双侧。在词复述任务时,双侧额下回、顶叶下部、颞上回、前运动皮质和本体感觉皮质电极显示感觉运动神经皮质。听觉活动主要位于颞上回和颞中回(74%的电极),言语产生活动发生在前运动皮质和运动皮质、本体感觉皮质和顶下小叶(79%的电极)。在所有的患者观察到感觉运动活动发生在双侧缘上回、颞中回、颞上回、体感皮质、运动皮质、前运动皮质和额下回。他们提出双侧亚词汇感觉运动转换可以支持基于感觉和运动表征联合形成的感觉运动界面,这在双侧言语感觉运动功能和一侧化高层次语言加工之间是有区别的。Berthier 等^[28]的研究提出弓状束和其他白质束是言语复述的主要解剖结构,个体之间两半球白质束一侧化的差异可以解释为什么有些左半球弓状束损伤的患者存在复述之间的差异。因此,除了考虑白质束和皮质损伤的大小外,还要考虑自身两半球白质束之间的差异可能对复述功能后的影响。

2.3 自身言语产生的监测和运动程序的纠正

一些研究提出颞上回后部,即听觉背侧流的一部分具有监测自身言语产生和纠正言语运动程序的功能^[29-30]。为了协调言语产生,运动控制理论提出自身产生言语的语音表征映射到言语发音器官的参考框架^[31]。多模态发音器官编码言语表征,激活期待的运动表征序列。在期待的和反馈(言语错误)之间的差异传递到额叶皮质,用于更新运动输出。颞顶部与听觉反馈有强联系,言语产生与颞平面后部、颞上回后部和缘上回这些听觉背侧流有关。因此,错语的产生可能因背侧流损伤破坏了涉及修正非意向言语的输出环路,以及不能监测自身言语,从而错语先于显性言语产生。

总之,皮质听觉语言加工的最早阶段涉及了频谱时间分析,它在双侧颞平面听皮质加工。语音水平加工和表征(语音网络)涉及了双侧颞上沟的中至后部,这一加工水平有左半球微弱的偏移。随后分为两支信息流,背侧通路将感觉或语音表征映射到发音运动表征,腹侧通路将感觉或语音表征映射到词汇概念表征。腹侧流更后部区域,颞叶后中、下部对词汇接口有反应,它连接着语音和语义信息。前颞叶对语

义信息进行整合,并将信息传递到 Broca 区前部。背侧流是强左侧优势。背侧流的后部区域对应于颞顶叶外侧裂区,它是感觉运动接口,而额叶更前部涉及了 Broca 区,前额叶背侧部对应于发音网络。背侧通路支持语音到运动的映射,以及自身产生言语的监测和言语运动程序的纠正。腹侧流和背侧流对语言不同加工成分更准确的、确定性的功能定位有待进一步的深入研究。

目前,失语症的语言康复治疗除了进行常规的行为治疗外,采用非侵入性脑刺激技术,如经颅直流电刺激、经颅磁刺激也是对行为治疗的补充。例如,当语义系统受损时,可以对腹侧流的神经位点进行相应的兴奋性脑刺激,改善语义障碍;当发音网络受损时,对相应的网络结点的脑区进行刺激,有助于改善言语运动功能,如言语失用症。充分理解听觉语言加工机制,有助于我们针对语言障碍产生的原因进行干预,更好地应用非侵入性脑刺激技术促进失语症患者的语言功能的恢复。

参考文献

- [1] Hickok G, Poeppel D. The cortical organization of speech processing[J]. *Nat Rev Neurosci*, 2007, 8(5):393—402.
- [2] Arnott SR, Alain C. The auditory dorsal pathway: orienting vision[J]. *Neurosci Biobehav Rev*, 2011, 35(10):2162—2173.
- [3] Luce PA, Pisoni DB. Recognizing spoken words: the neighborhood activation model[J]. *Ear Hear*, 1998, 19(1):1—36.
- [4] Stefanatos GA, Gershkoff A, Madigan S. On pure word deafness, temporal processing, and the left hemisphere[J]. *J Int Neuropsychol Soc*, 2005, 11(4):456—470.
- [5] Buchman AS, Garron DC, Trost-Cardamone JE, et al. Word deafness: one hundred years later[J]. *J Neurol Neurosurg Psychiatry*, 1986, 49(5):489—499.
- [6] Molholm S, Mercier MR, Liebenthal E, et al. Mapping phonemic processing zones along human perisylvian cortex: an electro-corticographic investigation[J]. *Brain Struct Funct*, 2014, 219(4):1369—1383.
- [7] Miglioretti DL, Boatman D. Modeling variability in cortical representations of human complex sound perception[J]. *Exp Brain Res*, 2003, 153(3):382—387.
- [8] Humphries C, Sabri M, Lewis K, et al. Hierarchical organization of speech perception in human auditory cortex[J]. *Front Neurosci*, 2014, (8):406—418.
- [9] Dehaene-Lambertz G, Pallier C, Serniclaes W, et al. Neural correlates of switching from auditory to speech perception[J]. *Neuroimage*, 2005, 24(1):21—33.
- [10] DeWitt I, Rauschecker JP. Phoneme and word recognition in the auditory ventral stream[J]. *Proc Natl Acad Sci USA*, 2012, 109(8):E505—514.

- [11] Cohen L, Jobert A, Le Bihan D, et al. Distinct unimodal and multimodal regions for word processing in the left temporal cortex[J]. *Neuroimage*, 2004, 23(4):1256—1270.
- [12] Thothathiri M, Kimberg DY, Schwartz MF. The neural basis of reversible sentence comprehension: evidence from voxel-based lesion symptom mapping in aphasia[J]. *J Cogn Neurosci*, 2012, 24(1):212—222.
- [13] Roux FE, Miskin K, Durand JB, et al. Electrostimulation mapping of comprehension of auditory and visual words[J]. *Cortex*, 2015, (71):398—408.
- [14] Chang EF, Rieger JW, Johnson K, et al. Categorical speech representation in human superior temporal gyrus[J]. *Nat Neurosci*, 2010, 13(11):1428—1432.
- [15] Edwards E, Soltani M, Kim W, et al. Comparison of time-frequency responses and the event-related potential to auditory speech stimuli in human cortex[J]. *J Neurophysiol*, 2009, 102(1):377—386.
- [16] Binder JR, Desai RH, Graves WW, et al. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies[J]. *Cereb Cortex*, 2009, 19(12):2767—2796.
- [17] Brugge JF, Volkov IO, Garell PC, et al. Functional connections between auditory cortex on Heschl's gyrus and on the lateral superior temporal gyrus in humans[J]. *J Neurophysiol*, 2003, 90(6):3750—3763.
- [18] Kikuchi Y, Horwitz B, Mishkin M. Hierarchical auditory processing directed rostrally along the monkey's supratemporal plane[J]. *J Neurosci*, 2010, 30(39):13021—13030.
- [19] Rogalsky C, Hickok G. Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex[J]. *Cereb Cortex*, 2009, 19(4):786—796.
- [20] Roelofs A. A dorsal-pathway account of aphasic language production: the WEAVER++/ARC model[J]. *Cortex*, 2014, (59):33—48.
- [21] Price CJ. The anatomy of language: a review of 100 fMRI studies published in 2009[J]. *Ann N Y Acad Sci*, 2010, (1191):62—88.
- [22] Ranganath C, Ritchey M. Two cortical systems for memory-guided behaviour[J]. *Nat Rev Neurosci*, 2012, 13(10):713—726.
- [23] Nobre AC, McCarthy G. Language-related field potentials in the anterior-medial temporal lobe: II. Effects of word type and semantic priming[J]. *J Neurosci*, 1995, 15(2):1090—1098.
- [24] Vouloumanos A, Kiehl KA, Werker JF, et al. Detection of sounds in the auditory stream: event-related fMRI evidence for differential activation to speech and nonspeech[J]. *J Cogn Neurosci*, 2001, 13(7):994—1005.
- [25] Shannon RV, Zeng FG, Kamath V, et al. Speech recognition with primarily temporal cues[J]. *Science*, 1995, 270(5234):303—304.
- [26] Remez RE, Rubin PE, Pisoni DB, et al. Speech perception without traditional speech cues[J]. *Science*, 1981, 212(4497):947—949.
- [27] Cogan GB, Thesen T, Carlson C, et al. Sensory-motor transformations for speech occur bilaterally[J]. *Nature*, 2014, 507(7490):94—98.
- [28] Berthier ML, Lambon Ralph MA, Pujol J, et al. Arcuate fasciculus variability and repetition: the left sometimes can be right[J]. *Cortex*, 2012, 48(2):133—143.
- [29] DeWitt I, Rauschecker JP. Wernicke's area revisited: parallel streams and word processing[J]. *Brain Lang*, 2013, 127(2):181—191.
- [30] Rauschecker JP, Scott SK. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing[J]. *Nat Neurosci*, 2009, 12(6):718—724.
- [31] Dhanjal NS, Handunnetthi L, Patel MC, et al. Perceptual systems controlling speech production[J]. *J Neurosci*, 2008, 28(40):9969—9975.