

汉语声调感知的研究进展及应用

牟志伟, 陈卓铭, 杨伟, 王春花

【关键词】 汉语; 声调; 感知

【中图分类号】 R49 【DOI】 10.3870/zgkf.2016.06.005

汉语语言是一种形声调的语言,不仅象形的文字含有语义而且韵调也包含语义。20世纪初至今使用实验的方法对汉语声调的感知进行了许多研究,本文通过对汉语声调感知的理论研究与新进展进行总结,为听力言语康复提供新的思路与方法,对设计与改进目前使用的一些仪器与设备提供理论支持。汉语声调的识别具有偏侧性脑功能定位特征,具有范畴性特征,汉语声调在失语症患者语音感知及康复中起到重要作用。此外,由于目前设计的人工耳蜗对于汉语声调感知存在缺陷,过少的频域信息导致植入后的儿童四声感知发育普遍存在问题,因此对提升人工耳蜗的声调感知能力的改进重点应该是提供必要的频域信息。

1 汉语声调语言特征与探索

汉语语言是一种形声调的语言,不仅象形的文字含有语义而且韵调也包含语义。如一个脑卒中患者把“我不要 da 针”中“da”的三声读成四声,护士就可能把“打针”误解成“大针”。因此,对言语障碍患者进行语言康复时不仅要注意词汇语言的康复,还要注意声调的康复。

早在公元5世纪,南朝宋明帝时期的著名文人周顒按照汉语声、韵、调的特征,提出了汉语“古代四声”的平、上、去、入四种声调的论述^[1]。明清出现现代汉语四声,即阴平(第一声)、阳平(第二声)、上升(第三声)、去声(第四声)。1922年赵元任^[2]的《中国言语学调底实验研究法》中,首先提出使用实验的方法研究声调。1963年梁之安^[3]引入了声学仪器,《汉语普通话中声调的听觉辨认依据》的文章中基于物理学的手段测量和分析汉语声调,将言语信号通过低通或高通滤波,提取基频及其谐波信息。近年来,随着影像技

术的迅猛发展,出现使用功能性磁共振技术研究声调感知的心理学方法。

随着20世纪声学实验的兴起,汉语声调的理论发展也经历了四个时期。20世纪20~30年代的音高论,是关于声调的自然属性研究,其核心理论认为声调是一种连续渐变的相对音高,在音节内具有分辨语义的作用。50~60年代的音位论,是关于声调的语言属性研究,认为声调是音系中的和元音音位、辅音音位一样的独立音位。70年代后是自主音段论的发展阶段,主要论述声调自身构成以及声调与音段关系,即非线性音系学的研究。1976年麻省理工学院的 Goldsmith^[4]提出超音段成分是一个自主平面,与另一平面的音段音位平行,它们按一定规则非线性关联。声调特征和调型构成了汉语声调。第四个时期是90年代中期,开始用优选论研究声调,提出声调组合过程中一些规则。1933年美国人 Prince^[5]和 Paul 在《优选论——生成语法中制约条件的交互作用》中首次正式提出:①在语音生成过程中的存在一系列制约条件,各种制约条件间存在优先层级顺序;②高等级的制约条件比低等级的制约条件起优先作用。优选论可以通过有限的规则解释各种汉语声调特征^[6]。

2 汉语声调的声学特征与感知

汉语声调与语音均是以声波形式通过空气传递到外耳,经过人的听觉器官、听语神经系统加工处理后使听话人感知。声波可以提取出基频(F0)、音强(dB)、最长声时(MPT)、基频微扰、振幅微扰、音域、S/Z 比等各种声学特征参数。基频与其谐波携带最主要的声调信息^[3]。

汉语声调的感知过程分为两个步骤:一是,听觉系统对语音信号的察知与传递,二是大脑皮层的处理。首先是对语音信号声学特征的听觉感知与识别,汉语的声调的声学信号主要携带于频域信息之中,频域信息包含频域精细结构和包络。频域精细结构,包括基频及其谐波,包络是指频谱信息中共振峰所形成的外围包络结构。目前研究认为,与声调听觉感知相关的

基金项目:国家自然科学基金“熟练普通话-粤语双言失语者康复语种选择模式研究”(81372113)

收稿日期:2016-05-27

作者单位:暨南大学附属第一医院康复科,广州 510630

作者简介:牟志伟(1978-),男,主治医师,主要从事失语症与构音障碍方面的研究。

通讯作者:陈卓铭, zml20tchzm@qq.com

频域信息有基频、谐波、共振峰^[7-9]。梁之安等^[3]通过语音听觉辨认的实验发现,仅保留基频、第一共振峰、第二共振峰中的1~2项时,仍可很好识别汉语声调。如果有基频及其谐波信息,四声的声调感知可高达100%;即使在没有基频信息仅有谐波信息存在的情况下,声调仍可得到较好的感知。林茂灿^[10]对声调研究认为基频频率及其变化在汉语四声调识别过程中起主要作用,而声调的时长和振幅变化起到的作用较小。单独依靠共振峰识别声调,识别率较低。王硕等^[11]将声音经过LPC信号处理后,过滤掉频域基频信息,仅保留共振峰信息。结果表明声调识别率只有36%。但是共振峰仍能对音调感知提供一些有用信息,尤其是对三声的识别^[12]。另外一些学者对包络信息进行的研究结果表明,仅靠频域包络信息识别汉语声调存在较大差异,准确性大约在40%~70%之间^[13-15]。Liu等^[13]在评价正常人的声调识别功能时采用LPC信号处理技术提取单音节中的频域包络信息。结果显示,一声识别率低于机会值,二声与四声的识别率在30%~45%之间,三声的正确识别率约为60%。Kong等^[14]使用LPC技术分离频域包络与精细结构信息的实验中,包络信息可以提供平均约65%的声调识别率。Wang等^[15]的研究与Liu等^[13]的研究结果基本一致,但与Kong等^[14]的研究结果相差较大,可能是每个研究使用的LPC参数设置不同造成的。

我们认为在声调感知的声学特征中,基频与其谐波携带最主要的声调信息。近年来通过人工智能的方法识别声调成为研究热点。首先用软件采集声调信息后提取其基频曲线,然后对提取的基频曲线进行平滑、重采样、归一化处理后可以得到便于分析的基频曲线。为了便于临床应用、指导声调的康复治疗,在分析汉语声调特点的基础上建立声调识别模型,从而可以对汉语声调进行人工智能识别^[16]。常用的声调识别模型包括有基于规则的方法、模糊识别方法、神经网络识别方法、基于隐马尔科夫模型方法和高斯混合模型等。

3 汉语声调识别的脑功能定位特征具有偏侧性

与语言的偏侧性类似,声调的感知也具有偏侧性。随着影像学的发展,许多学者使用功能性磁共振研究各种语言机制。陈卓铭等^[17]对语言障碍患者的声调感知偏侧化问题研究的结果,认为左半球损伤对声调知觉的影响很小,但是大脑对声调的控制仍然具有双侧性。从心理学角度看,声调的加工分为注意阶段和注意前阶段,这种两阶段论也是声调感知偏侧性的体现。2006年罗昊^[18]的研究表明汉语母语受试者在注意前阶段汉语声调的加工主要在右侧大脑半球,而辅

音的加工却在左侧大脑半球。Tallal等^[19]发现对快速变化的声音信息的加工主要在左侧半球,对慢速变化的声音信息加工则无半球偏侧化效应。由于汉语声调和辅音的不同可表达不同语义,虽然它们有类似的功能性特性,但是两者的声学特征却有很大的差别。声调变化相对缓慢,主要以基频特征变化为特征,而辅音以发音起始时间快速变化为特征。可见声调和辅音不同的大脑半球偏侧化效应从侧面说明了在语言的早期听觉加工过程中,语言样本的声学属性决定了大脑半球优势效应,但与功能属性无关。因此我们在康复左侧大脑损伤的语言障碍患者时应该可以遵循优先康复声调,再是元音和辅音的顺序。

从自主音段论看,声调本质是一种超音段音位。Hsieh^[20]发现,汉语母语被试进行汉语声调和音段音位的感知时均会激活左侧前运动皮层、额下回岛盖部和三角部;不过在左侧额叶的激活模式有差异,且声调加工比音段音位加工有更多的右侧脑区激活^[21]。说明汉语声调和音段音位加工既有相似之处,又有不同。

从以上研究我们可以发现对于声调的大脑偏侧性研究结论存在差异,但是对于声调感知的注意前阶段的右脑偏侧优势是比较肯定的,因此对于左脑损失的语言障碍患者可以考虑利用较好的声调感知能力设计康复计划。

4 汉语普通话声调感知特征具有范畴性

汉语普通话属于汉藏语系是声调语言,其声调感知具有范畴性;而印欧语系的语言是非声调语言,其语音声调只表示语气,无区别词汇意义,其语音声调感知是非范畴性的。汉语母语的成人对声调和嗓音启动时间的知觉都是范畴性的,一般6岁时已经具备和成人相似的声调加工模式^[22]。另外,声调的感知会受到被试者语言背景的影响^[23-24]。Wang等^[25]发现普通话的阳平和阴平之间存在范畴感知,但是感知界线具有个体差异。Halle等^[26]以台湾大学生和巴黎大学生作为被试,对汉语中的阴平-阳平、阳平-去声、上声-去声三组声调分别进行了听觉感知试验。台湾学生被试对这些声调的听辨呈现出准范畴性感知模式;而巴黎学生被试的声调听辨以心理物理学现象为基础。2010年王韞佳^[27]研究表明:汉语普通话母语者对阳平和上声的分辨呈现出一定的范畴化倾向,但不如阴平和阳平的范畴化显著。Zheng等^[28]用ERP对汉语普通话母语者和粤语母语者进行听觉实验,证明了不同声调系统对声调范畴感知存在差异。与汉语普通话有相关性的少数民族语言和其它国家语言(语种)的声调范畴性研究。孔江平^[29]证实了藏语拉萨话的声调在音高

和时长这两个声学特征上都具有范畴性。Abramson^[30]使用16个合成的平调音节连续体进行的选择适应实验发现辨认部分泰语暹罗话具有明显范畴性,而区分实验中高中低三个声调的感知是非范畴的连续感知。

由于范畴化的实质是将连续的物理量感知成突然变化的心理量,因此汉语声调的范畴性承载着语音知觉模式,我们在设计听力评估量表、仪器、康复设备时要将范畴性的评估作为一个考量因素。

5 脑损伤患者的声调感知

脑损伤后对语言的影响已经有大量的研究与结论,近年来对脑损伤后声调感知障碍研究也越来越多。汉语普通话为母语的脑损伤患者左右脑损伤后造成声调感知损伤的机制不同,因此对其声调康复的机制也不同。关于汉语失语的研究表明左脑损伤患者伴有声调损伤时,患者的声调识别模式很接近母语为非声调语的听辨人,并伴有较长的识别时间,也就是说脑损伤后声调感知模式会发生改变。汉语是声调语,其功能属性,也就是说对声调的语义处理可能在左脑。基于对失语症和声调感知两方面的研究,梁洁^[31]提出汉语失语患者的脑损伤对声调影响机制的假设:一是损伤导致失去声调底层表达,此时丧失音高辨义的功能,患者的声调感知模式类似于非声调语言的听辨人,患者声调系统需要重新构建;二是非底层损伤,仅破坏了具体语言的声调系统,但音高辨义的功能依然保留,此时患者的声调感知过程好像是在感知另一种声调语言。此实验结果支持综合型语言感知模式,即音高的处理分别由两侧大脑完成,左侧大脑主要处理与语言密切相关的音高信息,而声学信息则由右侧作为优势区来完成。结论同采用音位识别和区分实验分析失语患者的言语损伤的假设基本一致。

单纯失语症患者的声调错误相对于音位错误而言较少。晏培等^[32]对传导性失语复述障碍患者进行语音分析,发现虽然有很多音位性错误,但仅发现少数的声调错误。田鸿等^[33]对国内16例汉语普通话失语症患者进行四声检查实验,发现失语症患者除了朗读时的四声错误外,四声听理解能力比四声朗读更差。汪洁^[34]在对1例左顶叶、左顶深部梗死失语症患者进行语言评估时,发现以声调感知与表达障碍为主要特征的语言障碍。认为汉语声调感知和表达障碍可以独立于音素而单独出现,左顶叶及左顶深部参与声调感知与加工。高素荣^[35]提出声调错语的概念,报道了左颞顶区损伤的汉语普通话传导性失语患者出现的声调错误,认为左侧大脑损伤会导致声调障碍。以上脑损伤

患者语音感知的研究成果提示左侧大脑损伤会影响声调感知的深层加工,并能够指导汉语失语症声调障碍患者康复训练的计划制定。

6 人工耳蜗植入儿童声调感知

目前国内外已经广泛使用多通道人工耳蜗植入技术治疗重度感音神经性耳聋患者。由于其言语编码设计基于西方非声调语言特点,因此对汉语声调的感知能力差。西方多通道人工耳蜗可以很好地体现信号时域信息,但是没有考虑声调语言特点,只有很低的频域分辨率,所以只能反映相对较少的频域信息。正常听力儿童可以很容易地感知声调、元音和辅音,但是由于人工耳蜗的缺陷,重度感音神经性耳聋患儿没有良好的对声调感知能力^[36-37]。人工耳蜗植入后的儿童四声感知发育普遍存在问题,对声调及声调组合的识别能力普遍低于正常儿童,还会受到背景噪音的干扰。

崔丽丽^[38]发现人工耳蜗植入后儿童的声调感知与普通儿童有区别。①对声调的识别能力显著低于普通儿童;②背景噪音会影响人工耳蜗植入后的声调识别;③儿童人工耳蜗植入的年龄对声调发育起到决定性的作用,越早植入声调发育越接近正常水平。进一步发现植入儿童与普通儿童对声调组合识别的难度顺序相似。人工耳蜗植入后儿童识别一声最容易,其次是二声和三声,四声最难,普通儿童也是一声最容易,接着是四声、三声、二声。徐立^[39]做了进一步研究,研究表明人工耳蜗植入后进行汉语普通话声调识别时,时域和频域的信息可以互为补充。那么时域和频域信息哪个更重要?实验显示植入者会优先使用频域信息。Xu等^[40]在此基础上对用于汉语母语人群的人工耳蜗设计提出改进方案,报道正在相关公司进行实验。此外,在人工耳蜗植入的成人患者中,声调感知能力与对音乐的欣赏能力正相关。植入者对声调感知能力越强,对歌曲音乐的细节就把握得越多^[41]。这可能是由于声调识别和音乐识别在听觉感知中有着类似的识别机制。如何利用人工耳蜗使用者的声调和音乐感知相关性进行听力康复将是未来科研和临床运用的一个新探索。

综上所述,汉语声调对汉语普通话的理解具有重要作用,汉语声调具有独立研究价值,研究其语音信息特征、偏侧性脑功能定位特征、范畴性特征,对听力言语障碍患者的康复治疗方法具有指导与借鉴作用,对听力言语仪器开发与改进提供理论支持。今后,我们需要将理论构建、实验研究与先进的科学手段相结合,使声调感知研究更具系统化,促进声调感知研究与临床结合进一步深入。

【参考文献】

- [1] Han XJ. The first man to find the Chinese tones[J]. *Education for the aged*, 2007, 10(2): 28.
- [2] Chao YR. Experimental methodology on Chinese tone values[J]. *Science*, 1922, 7(9):871-882.
- [3] 梁之安. 汉语普通话中声调的听觉辨认依据[J]. *生理学报*, 1963, 26(2):85-91.
- [4] Goldsmith J. An overview of autosegmental phonology[J]. *Linguistic Analysis*. 1976, 2, 23-68.
- [5] Prince A, Smolensky P. Optimality Theory: Constraint Interaction in Generative Grammar[J]. USA: Blackwell Publishing Ltd, 2008, 20-25.
- [6] 刘俐李. 二十世纪汉语声调理论的研究综述[J]. *当代语言学*, 2004, 6(1): 45-56.
- [7] Peng SC, Tomblin JB, Cheung H, et al. Perception and production of Mandarin tones in prelingually deaf children with cochlear implants[J]. *Ear Hear*, 2004, 25(3): 251-264.
- [8] Feng YM, Xu L, Zhou N, et al. Sine-wave speech recognition in a tonal language[J]. *The Journal of the Acoustical Society of America*, 2012, 131(2): 133-138.
- [9] Wang J, Shu H, Zhang L, et al. The roles of fundamental frequency contours and sentence context in Mandarin Chinese speech intelligibility[J]. *The Journal of the Acoustical Society of America*, 2013, 134(1): 91-97.
- [10] 林茂灿. 普通话声调的声学特性和知觉征兆[J]. *中国语文*, 1988, (3):182-193.
- [11] 王硕, Robert M, Philip Ne. 共振峰信息在汉语声调感知中的作用[J]. *中国耳鼻咽喉头颈外科*, 2012, 19(1): 8-11.
- [12] Hu YM. A Review on the Brain Mechanism of Chinese Tonal Processing[J]. *Psychological Science*, 2011, 34(1): 196-200.
- [13] Liu S, Samuel AG. Perception of Mandarin lexical tones when F0 information is neutralized[J]. *Lang Speech*, 2004, 47(2): 109-138.
- [14] Kong YY, Zeng FG. Temporal and spectral cues in Mandarin tone recognition[J]. *The Journal of the Acoustical Society of America*, 2004, 120(5): 2830-2840.
- [15] Wang S, Mannell R, Newall P, et al. Contribution of spectral cues to mandarin lexical tone recognition in normal-hearing and hearing-impaired Mandarin Chinese speakers[J]. *Ear & Hear*, 2011, 32(1): 97-103.
- [16] Chen ZM, Ling WX, Zhao JH, et al. Consonant recognition of dysarthria based on wavelet transform and fuzzy support vector machines[J]. *Journal of Software*, 2011, 6(5): 887-893.
- [17] 陈卓铭. 利用汉语语言特点设计语言康复[J]. *新医学*, 2001, 32(9): 526-527.
- [18] 罗昊. 听觉系统自动加工汉语声调和辅音时的大脑半球优势[D]. 北京: 中国科学技术大学, 2007, 9-16.
- [19] Tallal, Paula, Steve Miller, and Roslyn Holly Fitch. Neurobiological Basis of Speech: A Case for the Preeminence of Temporal Processing[J]. *Annals of the New York Academy of Sciences*, 1993, 682(1): 27-47.
- [20] Hsieh L, Gandour J, Wong D, et al. Functional heterogeneity inferior frontal gyrus linguistic experience[J]. *Brain Language*, 2001, 76(3): 227-252.
- [21] 张林军, 周峰英, 王晓怡. 言语中的音高信息声学语音学加工的大脑偏侧化[J]. *应用心理学*, 2008, 14(4): 330-335.
- [22] 席洁, 姜薇, 张林军, 等. 汉语语音范畴性知觉及其发展[J]. *心理学报*, 2009, 41(7): 572-579.
- [23] Yang J. Acoustic properties of vowel production in Mandarin-English bilingual and corresponding monolingual children. Doctorate dissertation [D]. Columbus, The Ohio State University, 2014: 224.
- [24] Chuang HF, Yang CC, Chi LY, et al. Speech intelligibility, speaking rate, and vowel formant characteristics in Mandarin-speaking children with cochlear implant[J]. *International Journal of Speech-Language Pathology*, 2012, 14(2): 119-129.
- [25] Wang SY. Language change[J]. *Annals of the New York Academy of Sciences*, 1976, 280(1): 61-72.
- [26] Halle PA, Chang YC, Best CT. Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners[J]. *Journal of Phonetics*, 2004, 32(3): 395-421.
- [27] 王毓佳, 李美京. 调型和调阶对阳平和上声知觉的作用[J]. *心理学报*, 2010, 42(9): 899-908.
- [28] Zheng HY, James WM, Peng G, et al. The impact of tone systems on the categorical perception of lexical tones: An event-related potentials study[J]. *Language and Cognitive Processes*, 2010, 1(1): 1-31.
- [29] 孔江平. 藏语(拉萨话)声调感知研究[J]. *民族语文*, 1995, (3): 56-64.
- [30] Abramson, Arthur S. The Coarticulation of Tones: An Acoustic Study of Thai[J]. *Speech Research*, 1975, (10-12): 119-125.
- [31] 梁洁. 汉语失语声调损伤感知实验[J]. *当代语言学*, 2012, 14(1): 24-36.
- [32] 晏培, 高素荣. 传导性失语患者的复述和朗读障碍[J]. *中华神经科杂志*, 2000, 33(3): 141-143.
- [33] 田鸿, 李胜利. 失语症患者的四声检查结果分析[J]. *中国康复*, 1996, 11(2): 57-58.
- [34] 汪洁. 失语症声调感知与表达障碍1例分析[J]. *中华物理医学与康复杂志*, 2004, 26(3): 146-147.
- [35] 高素荣. 失语症[M]. 北京: 北京医科大学出版社, 2006, 217-220.
- [36] Moore BCJ. Coding of sounds in the auditory system and its relevance to signal processing and coding in cochlear implants[J]. *Otology & Neurotology*, 2003, 24(2): 243-254.
- [37] Yang J, Brown E, Fox RA, et al. Acoustic properties of vowel production in prelingually deafened Mandarin-speaking children with cochlear implants [J]. *The Journal of the Acoustical Society of America*, 2015, 138(5): 2791-2799.
- [38] 崔丽丽. 人工耳蜗植入儿童的声调识别及发声研究[D]. 上海, 华东师范大学, 2011:17-30.
- [39] 徐立. 言语识别中的时域及频域信息[J]. *中华耳科学杂志*, 2006, 4(4): 335-342.
- [40] Xu L, Tsai Y, Pfingst BE. Features of stimulation affecting tonal-speech perception: Implications for cochlear prostheses [J]. *The Journal of the Acoustical Society of America*, 2002, 112(1): 247-258.
- [41] Wang WQ, Ning Z, Xu L. Musical Pitch and Lexical Tone Perception with Cochlear Implants[J]. *International Journal of Audiology*, 2011, 50(4): 270-278.