

HANYU YUYIN QUBIE
TEZHENG FENXI

汉语语音 区别特征分析

孟子厚 刘亚丽 章斯宇 冯晓亮 王孟杰 著



国防工业出版社

National Defense Industry Press

本书研究内容得到国家自然科学基金资助项目 NSFC11174257 支持

汉语语音区别特征分析

孟子厚 刘亚丽 章斯宇 冯晓亮 王孟杰 著

国防工业出版社

· 北京 ·

内 容 简 介

本书以相关基金课题的研究工作为基础,针对语音评测中存在的区别对立现象,从语音的区别特征思想出发,研究了适合汉语语音发音质量评测的汉语普通话声韵调系统的参数化区别特征体系,提出了区别特征体系的框架,介绍了具体的特征参数的计算,验证了运用区别特征进行声母识别和检测的准确性。书中的具体内容包括汉语普通话水平测试与区别特征体系的论述和分析、声母参数化区别特征的研究结果、韵母参数化区别特征的研究结果、汉语普通话听觉感知实验研究结果与区别特征体系的分析、自然音节状态下汉语普通话声母的区别特征检测等。书中所提出的汉语普通话区别特征体系和区别特征参数的检测方法,对从解析分析的角度研究普通话语音评测技术提供了可行性和理论基础。

本书适合于从事语音学、言语工程、语音评测技术等领域的科技人员、相关专业和研究领域的大专院校师生参考。

图书在版编目(CIP)数据

汉语语音区别特征分析/孟子厚等著. —北京:
国防工业出版社,2016. 6
ISBN 978 - 7 - 118 - 10782 - 1

I. ①汉... II. ①孟... III. ①汉语 - 语音 - 研究
IV. ①H11

中国版本图书馆 CIP 数据核字(2016)第 122836 号

※

国防工业出版社出版发行
(北京市海淀区紫竹院南路 23 号 邮政编码 100048)
三河市众誉天成印务有限公司印刷
新华书店经售

*

开本 710 × 1000 1/16 印张 16 $\frac{3}{4}$ 字数 318 千字
2016 年 6 月第 1 版第 1 次印刷 印数 1—2000 册 定价 68.00 元

(本书如有印装错误,我社负责调换)

国防书店:(010)88540777
发行传真:(010)88540755

发行邮购:(010)88540776
发行业务:(010)88540717

前 言

目前的主流语音识别技术是以统计模式识别为技术基础的,计算机辅助的汉语语音发音质量评测技术是建立在语音识别的技术平台之上的,并不完全适用于语音发音水平,特别是发声缺陷的评测。统计模式识别需要大量的数据支持,而且并不涉及语音的本质,因此难以与语音学的解释发声直接的关联,而这是在语音质量评测中所期望的。为了突破传统的模式识别的思路,着眼于语音的更本质的特点和机理,探寻更加有效的语音评测方法,自2003年开始,中国传媒大学传播声学研究所开始了对普通话区别特征的研究,此项工作得到了国家自然科学基金项目“汉语普通话声韵调参数化区别特征体系研究”的资助。

在此项研究中,拟建立的区别特征体系以发音缺陷判定或发音质量检测为目的。在构建过程中,一方面借鉴传统区别特征的思想,另一方面要考虑到汉语“声韵调”音位系统的特点。针对汉语语音评测,结合语音的生理发音、语言声学和知觉心理特点,从声韵调三个方面建立适合评测的参数化的区别特征体系。本书的内容即为中国传媒大学传播声学研究所汉语普通话区别特征研究领域工作的总结。

全书共分为六章,第一章为阅读本书所需要的汉语语音学的基础;第二章论述了普通话水平测试与区别特征的关系,以及区别特征研究的概况,这是本书研究工作的背景和出发点;第三章和第四章分别介绍了汉语普通话声母和韵母参数化区别特征体系的构建方法和具体的特征参数分析;第五章介绍了在听觉感知实验的基础上,基于言语知觉特性的区别特征的构建过程,建立了一个相对成形的汉语普通话区别特征体系;第六章初步讨论了在自然音节状态下进行区别特征检测的问题,这是将区别特征方法运用于语音检测的一个初步尝试。

中国传媒大学多位研究生先后参与了区别特征的研究工作,包括刘亚丽博士、章斯宇博士、冯晓亮博士、王孟杰博士、贾玉博、徐慧、徐益华等同学,他们对此项研究做出了很多贡献,本书的许多具体内容来自他们的学位论文工作。这项研究工作是一项基础性很强的工作,短期内看不到直接的应用效益,对做语音信号处理的研究者来说,能静下心来进行如此基础的探索和研究工作是很难得的,在此非常感谢这项工作的参与者。

鉴于作者水平有限,特别是在语言学和语音学理论方面的学识浅薄,工作中肯定有不足之处甚至是错误,但是本着求实创新和积极探索的精神,我们认为还是有必要把这些粗浅的工作总结出来供汉语语音学和语音技术的研究者们参考,请大家批评指正,这也有助于我们精益求精、深入研究。

作者

二〇一六年二月十六日

目 录

第0章 绪论	1
第1章 汉语普通话语音学基础	6
1.1 语音的生成与感知	6
1.1.1 语音的物理属性	6
1.1.2 语音的发声机制	7
1.1.3 语音的听觉感知	8
1.2 汉语普通话的语音结构	9
1.2.1 元音	10
1.2.2 辅音	15
1.2.3 声调	18
1.2.4 音节结构	21
1.2.5 语流音变	22
1.2.6 轻重音	26
1.3 汉语普通话语音的区别特征	28
1.3.1 音位与区别特征	28
1.3.2 音节结构与区别特征	32
1.3.3 区别特征的应用与可操作性	38
参考文献	39
第2章 汉语普通话水平测试与区别特征体系	41
2.1 汉语普通话的音位分析	41
2.1.1 音位	41
2.1.2 归纳音位的原则	45
2.1.3 音位的聚合和组合	46
2.1.4 语素音位	47
2.2 基于音位分析的区别特征	48
2.2.1 从音位到区别特征	48
2.2.2 结构音系学的区别特征	49
2.2.3 生成音系学的区别特征	50

2.3 汉语普通话水平测试的音素分析	52
2.3.1 汉语普通话水平测试的规则要求	52
2.3.2 汉语普通话水平测试中的常见发声缺陷	52
2.3.3 汉语普通话水平客观测试的可操作性	58
2.4 语音评测的区别特征体系	64
2.4.1 汉语普通话的声韵调区别特征体系的建立	65
2.4.2 多元分析	68
2.4.3 构建决策树	69
2.4.4 预测评分	69
参考文献	69
第3章 汉语普通话声母的参数化区别特征	71
3.1 辅音声母区别特征体系设计	71
3.1.1 辅音的区别特征	71
3.1.2 辅音的声学特征	72
3.1.3 区别特征参数的选择	76
3.1.4 区别特征体系参数的完备性和正交性	79
3.2 基于发音方式的辅音区别特征参数	81
3.2.1 清擦音、塞擦音与其他辅音的区别特征参数	82
3.2.2 非送气塞音、鼻音、边音、浊擦音与送气辅音和舌根音的区别特征参数	84
3.2.3 非送气塞音与鼻音、边音、浊擦音的区别特征参数	86
3.2.4 浊擦音与鼻音、边音的区别特征参数	87
3.2.5 鼻音与边音的区别特征参数	88
3.2.6 送气塞擦音与非送气塞擦音、清擦音的区别特征参数	90
3.2.7 送气塞擦音与清擦音的区别特征参数	95
3.3 基于发音部位的辅音区别特征参数	97
3.3.1 唇齿音 f 与舌尖音、舌面音、卷舌音的区别特征参数	97
3.3.2 舌尖音与舌面音、卷舌音的区别特征参数	98
3.3.3 卷舌音与舌尖音的区别特征参数	99
3.3.4 送气塞音与舌根音的区别特征参数	101
3.3.5 双唇塞音与舌尖塞音的区别特征参数	102
3.3.6 双唇鼻音 m 与舌尖鼻音 n 的区别特征参数	103
3.4 基于参数区别特征体系的辅音检测	104
3.4.1 参数化区别特征体系的构建与辅音检测结果	104
3.4.2 与 HMM 统计识别方法的比较	106

3.4.3 参数化的辅音区别特征体系分析	114
参考文献	117
第4章 汉语普通话韵母的参数化区别特征	119
4.1 韵母的区别特征检测及分类	119
4.1.1 韵母的区别特征	119
4.1.2 特征参数	123
4.1.3 区别特征参数的选择	126
4.1.4 区别特征的检测	129
4.1.5 普通话韵母的分类模型	132
4.2 基于参数距离的韵母决策树	134
4.2.1 特征计算与决策树生成	135
4.2.2 与感知实验结果的比较	140
4.2.3 节点计算	142
4.2.4 韵母的分类结果	145
4.3 基于参数的普通话韵母区别特征	146
4.3.1 基于参数的易混淆韵母	146
4.3.2 聚类逻辑	147
4.3.3 参数化的韵母区别特征	147
4.3.4 韵母分类	149
4.4 韵母对补关系的分析	150
4.4.1 有无介音韵头的对立	151
4.4.2 韵母发音中的前后对立	153
4.4.3 韵母发音中的宽窄对立	155
参考文献	157
第5章 汉语普通话的听觉感知特性	159
5.1 混响环境下的普通话听觉感知实验	159
5.1.1 实验设计与方法	159
5.1.2 语音清晰度分析	161
5.1.3 语音知觉特征分析	169
5.2 噪声环境下的普通话听觉感知实验	177
5.2.1 实验设计与方法	177
5.2.2 语音清晰度分析	179
5.2.3 语音知觉特征分析	187
5.3 言语知觉特性与区别特征	194
5.3.1 声母的聚类逻辑与区别特征构建	195

5.3.2	韵母的聚类逻辑与区别特征构建	201
5.3.3	声调的聚类逻辑与区别特征构建	212
5.4	区别特征系统的有效性验证和参数化分析	214
5.4.1	验证实验的设计	214
5.4.2	语音知觉特征的有效性验证	215
5.4.3	区别特征系统的参数化分析	222
	参考文献	225
第6章	自然音节状态下汉语普通话声母的区别特征检测	226
6.1	汉语普通话声母的区别特征参数分析	226
6.1.1	声母的区别特征参数体系	226
6.1.2	基于言语知觉特性的声母的区别特征	228
6.1.3	声母区别特征初步参数化的失效性分析	230
6.1.4	声母区别特征参数化结果的修正	230
6.2	汉语普通话声母的区别特征检测	243
6.2.1	音节中检测声母的区别特征的探针设计	243
6.2.2	音节中声母的区别特征检测结果	251
6.3	检测探针的有效性分析	252
6.3.1	两种状态下检测效果总体对比	252
6.3.2	两种状态下检测效果在各区别特征上的对比	253
	参考文献	259

第0章 绪 论

汉语语音水平评测主要包括普通话水平测试(Putonghua Shuiping Ceshi, PSC)、对外汉语水平测试(Hanyu Shuiping Kaoshi, HSK)、艺术发音评测等几个方面。对于汉语语音评测,应试人按照特定的测试文本发音,其重点是判断发音样本与标准样本的接近程度,并按照测试大纲的要求做出发音错误或缺陷的判定。汉语普通话水平测试当前仍主要采用人工打分的评测方法。语言测试中主观性测试的优点是容易测试出应试者的实际语言能力,特别是语音规范即语音标准程度。但是由于主观性测试评分是由主试人主观判断,尽管有评分标准,可由于评分者主观看法、观点、印象以及自身语言水平等诸多因素的影响,会出现评分差异,影响测试成绩的稳定性,同时使得测试成本很高,测试员劳动强度大。随着实验语音学、语音心理学的发展,以及语音识别等相关技术的进步,计算机辅助发音质量评测成为了可能。基于生理发音、语言声学 and 知觉心理特点,借鉴言语工程中的相关技术,利用计算机为应试者提供汉语语音发音质量的参考评分和评分的可信度级别,是具有重要的学术价值和实用意义的研究。

20 世纪 70 年代,IBM 的 Fred Jelinek 和卡内基·梅隆大学的 Jim 与 Janet Baker 分别独立地提出隐马尔可夫模型(Hidden Markov Model, HMM),该统计模型在模式识别领域的成功应用推动了包括语音识别在内的相关技术的发展,也对后来的语音评测技术发展产生了较大的影响。自 20 世纪 90 年代以来,国内外许多学者在语音评测方面进行了广泛的研究:美国斯坦福大学的 SRI、日本京都大学、英国的剑桥大学和爱丁堡大学等对带口音的英语、法语、西班牙语和日语的语音评测进行了研究;清华大学、北京大学、中国科学院、中国科技大学、中国传媒大学、哈尔滨工业大学和台湾清华大学等对汉语和英语的发音质量评测进行了研究。这些工作基本上都采用基于统计模型的方法。尽管统计方法在语音技术中起着重要作用并且得到不断改善,但其本身仍存在固有的缺陷和难以避免的局限性。统计的方法过于依赖数据驱动,忽视了对理论模型的研究,缺少对语音的声学模型和语音的知觉模型内部细节问题的研究,只是通过加强训练和扩大数据库来提升系统性能。语言具有相对的稳定性和封闭性,任何一种语言的基本构造单位都是有限的。而统计的方法为解决数据稀疏问题和协同发音问题需要不断扩大语料的范围,使语料库的规模呈指数性的增长。同时,语言交流是个体行为,个人的感情、心理状态等都会对其产生影响,语言本质与统计模

型之间必然会存在一定的差别,不是用统计模型就可以完全描述的了。

在统计模型方法发展的同时,也有学者在积极探索另一种基于语音特征的研究思路,试图寻找到构成语音信号的最小基本单元,以便对各种语言进行统一描述。在音系学中,从听感上分析,认为音素是语音的最小单位,是人能听出的最短语音片段。如果考虑语音的辨义作用,那么对语音知觉而言,语音最重要的基本单位是音位。然而,从生理发音或物理声学角度考虑,音位并不是不可分割的最小单位,可以看作是更小的发音或声学要素(特征)的组合。对于整个语音系统来说,每个音位都可以用若干语音特征的集合来表达,各音位均有与其他音位对立的起区分作用的语音特征。根据特征的相似性可以推测不同音位间的听觉混淆情况,反之,通过听感实验可以推导出知觉过程中起区别性作用的语音特征。

20 世纪 50 年代, Jakobson, Fant 和 Halle 在继承音系学中重视语音特征的传统上,结合信息论中常使用二元选择序列分析通信过程的特点,通过二分法原理描述语音中普遍存在的区别对立现象,提出了区别特征(Distinctive Feature)的概念,区别特征是构成语音信号的基本元素,是区分音位的最小差别。Jakobson 认为言语是一个二元对立的系统,人在辨识语音的过程中总是面临一系列的二元选择,利用区别特征所传递的区别作用进行二元选择,从而完成语音的识别。

语音的区别特征理论,提供了一种不同于统计方法的新思路,不论是对语言学或语音学的理论研究,还是对包括语音合成、语音识别技术在内的言语工程应用都有着重要的意义。由于语音识别任务的复杂性,区别特征理论仍很少在成熟的语音识别系统中得到实际应用。然而对语音评测来说,由于其特定任务中所针对的发音错误或发音缺陷大多发生在最小对立体之间,在客观评测算法的研究中利用区别特征的理论,不但可以加快搜索速度,提高语音判别的准确率,也能够提高系统的鲁棒性。有学者认为区别特征与语音信号中的声学线索相互对应,但目前都还没有找到一套可实际操作的描述汉语发音质量的客观参数。尽管已经提出了一些区别特征体系,但都还有不足之处。真正有声学参数支持、保持二元对立关系、能够应用于言语工程的区别特征体系还少有研究,这也是区别特征研究的瓶颈。

区别特征的原始概念起源于 20 世纪 30 年代,作为生成音系学源头之一的布拉格学派对特征尤为重视。其代表人物捷克的 Trubetzkoy 首先突破了语音分析到音素为止的观念,创立了音位对立二分(区别特征)的概念。Jakobson 进一步把区别特征具体化,认为语音中任一最小有区别的信息,都给予听者二选一的选择,在此基础上建立了偶值特征(Binary Features)理论。区别特征研究的是音位间的对立区别,即只在某项特征上区别对立,其余特征都相同的情况。Jakobson、Fant 和 Halle 在对语音进行物理声学分析的基础上,结合其生理发音特点,提出所有人类语言的语音都可以用 12 对“最小对立体”进行描述,把区别特征

作为分析语音的标尺。20 世纪 60 年代后,区别特征理论逐渐得到补充和发展。比较有代表性的是 N. Chomsky 和 Halle 的区别特征体系,其特点是除声学特征之外,还根据生理发音部位增加了若干对立体,对二元特征进行修订,允许多元特征的存在。Ladeford 从语言学的角度看待语音特征,强调音系的建立应考虑音位在特定语言中的作用。在对多种语言调查分析的基础上,把区别特征扩充为 20 项。20 世纪 90 年代,Halle 对区别特征理论进一步扩展,把区别特征分成发音器官不定(articulator-free)特征和发音器官约束(articulator-bound)特征两类。

国内关于汉语普通话区别特征的研究开始于 20 世纪 80 年代。原有的区别特征理论一般不包括超音段特征(如轻重音、音调等),但对汉语而言,某些超音段特征也有区别意义的作用。王士元把汉语声调归纳为 13 种调型,用 7 个特征进行分类。吴宗济建立了普通话区别特征矩阵,并探讨了它们之间的相互关系模型,认为汉语声调是区别特征的要素之一,与元音、辅音同属“本质特征”。这与西方学者认为声调属于语音的“韵律特征”或“次要特征”范畴而不是区别意义的要素的看法不同。吴宗济在区别特征矩阵中给出了 7 对元音的区别特征,但是在矩阵中存在较大的冗余,造成冗余的原因可能是把“生理现象”和“声学现象”对立起来了。此外,还有些音位在矩阵中没有被区别开来。其后,叶蜚声、徐通锵列出了“北京话音位的区别特征表”,将元音和辅音统一由 9 对区别特征来描述,比吴宗济的工作前进了一步。但是,区别特征理论指出区别特征的拟定需要有语音学的基础,前人在定义区别特征时,这方面表现出一定的随意性。上述研究都是从传统音位学角度出发的,但是前者从音的响度和发音两方面来观察,后者则完全从发音角度来定义特征。二者存在共同的问题:虽然都明确指出是用“二分的方法”来建立区别特征矩阵的,但实际上并没有始终贯彻偶分法则。陆致极进一步采用 12 对区别特征来建立普通话音位的区别特征矩阵,初步建立了能描写音位层次的区别特征体系。

张家骥根据汉语语音听感实验的群集分析结果确定了汉语的声韵调音位系统。继而按照区别特征划分的二分法原理,提出了汉语的声韵调的区别特征体系。与前人的研究有所区别的是,这里首次把声、韵、调系统,而不是元音、辅音系统作为汉语的音位系统。根据区别特征的概念,要求在建立的区别特征体系中,声母间和韵母间是可以两两完全区分开的,即任意两个声母或韵母至少在某一特征上是区别对立的。但是通过分析张家骥的声母和韵母区别特征矩阵可以发现,由于对于声母中的 x 和 f、x 和 h,以及 r 和 l,韵母中的 er 和 ao 以及 er 和 ou 的区分出现问题,该区别特征体系还有待完善。另外,虽然这套区别特征体系在提出时是设想以声学参数为基础的,但是这个体系提出时没有给出实际可操作的语音特征参数能与之对应,同时该区别特征体系也无法覆盖普通话水平测试中的错误类型和缺陷类型。对于错误类型和缺陷类型的判断需要进一步探

索构建一个全面完备、易操作的理论框架。

中国传媒大学传播声学研究所自 2003 年开始,在针对汉语语音发音质量评测研究的基础上,为了探寻更加有效的评测方法,开始了对普通话区别特征的研究。研究目的是通过借鉴传统区别特征理论,从语音的生理发音特点,声学参数分析和语音知觉心理实验多方面考虑以构建面向汉语语音评测的区别特征体系,为汉语语音评测提供新的思路和方法。拟建立的区别特征体系是以发音缺陷判定或发音质量检测为目的的。在构建过程中,一方面借鉴传统区别特征的思想,另一方面要考虑到汉语“声韵调”音位系统的特点。针对汉语语音评测,结合语音的生理发音、语言声学和知觉心理特点,从声韵调三个方面建立适合评测的参数化的区别特征体系。

语音评测有其特殊性,不同于通常的语音识别问题。在实际评测时,应试人是按照特定的试卷文本发音的,即使出现了语音错误,这种错误也与正确的发音样本有一定的内在关联(最小对立体)。在传统的语音评测的研究中,采用和语音识别类似的思路,利用统计模型的方法实现。这样的方法没有充分结合语音评测的特点,忽视了对语音学先验知识的应用,在评测结果上也很难满足对识别发音错误类型和发音缺陷的要求。本书的研究工作试图从语音本质特征的角度出发,通过构建语音区别特征体系,以实现汉语语音的解析化评测。其中区别特征体系的构建要综合考虑语音的生理发音、语言声学和知觉心理特点等多方面因素,最终建立起汉语语音的声韵调区别特征体系,这是研究工作的核心。主要的研究内容包括:

- ① 结合区别特征理论和汉语语音评测的特点,提出基于声韵调的对补关系;
- ② 语音信号分析,寻找与对补关系相对应或能起区分性作用的声学特征参数;
- ③ 通过多声学传递条件下的听觉感知实验,分析语音知觉混淆特点,提出区分性的知觉特征;
- ④ 对区别特征模型进行多元统计分析,使区别特征有完备性和操作性;
- ⑤ 构建语音评测的区别特征决策树;
- ⑥ 将区别特征体系用于语音评测,对其合理性和实用性进行检验。

由于区别特征与语音的生理发音、语言声学和听觉感知特性存在紧密的联系,在构建面向语音评测的区别特征体系时,从总结语音评测中的声韵调的对补关系、声学参数分析和听觉感知实验三个角度进行。

对补关系是研究的出发点,给出了研究需要解决的问题,设定了研究的范围。普通话水平测试时测试员在语音判定时常碰到的语音错误主要有:误读产生的读音错误,受汉语方言语音系统影响,说普通话时字音遗留某类声母、某类韵母或某类声调的读音。常见的语音缺陷也主要指的是遗留汉语方言系统的某

类声母、某类韵母或某类声调的读法。由具体的常见声母语音错误类别可归纳出声母的对补关系。结合张家骥提出的区别特征体系,同时考虑到对声母语音错误和语音缺陷的评测需求,在我们的工作中尝试了从声母的对补关系出发寻找声母的区别特征体系。由于韵母的构成较为复杂,韵母对补关系的梳理比较麻烦,不像声母那么明显。大部分语音错误和语音缺陷都直接和韵母的具体构成有关系,难以系统归类处理。本书从辅音的、过渡的等 17 对特征的对补关系出发尝试建立韵母的区别特征。

声学参数分析是从信号处理的角度研究区别特征,为区别特征在语音评测中的应用提供可操作性。基于上述对补关系,寻求有声学参数支持的区别特征的技术路线可以从自上而下和自下而上两个方面尝试。除归纳发音上的对补关系和信号分析声学特征参数的方法外,从语音的知觉心理特点入手,通过感知实验分析听感上起区分性作用的语音知觉特征,为确立可靠的区别特征提供依据,也是构建区别特征体系的重要途径。语音知觉是一项复杂的加工过程,涉及到外部刺激的物理属性,也涉及到听者内部的生理、心理过程。通过对语音知觉的研究,发现声音中某些对准确、稳定识别语音起关键作用的特征以及影响语音知觉的因素,对包括语音识别、语音评测在内的相关技术的突破都是至关重要的。

在多种声学传递条件下实施听觉感知实验以寻找稳定的语音知觉特征,实验方法借鉴的是汉语单音节清晰度的主观评价方法。考虑到影响语言清晰度的外部因素主要有三个方面:环境的噪声、环境的混响、系统的失真。其中系统无明显失真是一个基本要求,一般情况下都能达到。因此主要研究噪声造成的掩蔽作用、混响带来的混叠效应,以及噪声掩蔽和混响混叠共同作用时对语言清晰度和语音知觉混淆情况的影响,最终得到多传递条件综合时的语音知觉特征。而关于汉语普通话的知觉混淆特点和语音知觉心理空间的研究以前不多,只有张家骥给出的在噪声条件下的语言清晰度实验的分析结果。不同的噪声特性和混响特性对语音知觉混淆规律有着怎样的影响以及在噪声或混响环境下起区分性作用的汉语语音知觉特征还鲜有人研究,探索声学环境与语音知觉心理空间的联系对建立恰当的区别特征体系有着积极的启发意义。

第1章 汉语普通话语音学基础

1.1 语音的生成与感知

语音和其他声音不同,是由人类发音器官发出的,在社会约定俗成的条件下负载意义的声音。它是语言的物质载体,与语义一起描述语言。语音既具备物理属性,是一种声波,同时又具有心理属性,是声波在听觉产生的主观感觉。

1.1.1 语音的物理属性

语音和其他各种声音一样,具有声音的物理属性。传统语音学认为语音的4个要素分别为:音高、音强、音长和音色。

(1) 音高的物理属性对应为声带振动的基频,也就是语音信号中浊音的基音频率,单位为 Hz。音高也是一种主观心理量,其心理属性就是音调的高低。音高与声音的频率有关,频率高的声音听起来感觉音调“高”,而频率低的声音,听起来感觉音调“低”。“音高”的含义随不同的应用场合而定:在声学上它等同于“基频”;在听感方面有不同的量纲美(mel)或巴克(bark);在音系学上指的是声调或语调,是一个相对音高的概念。在声调语言或其他各种语言中,基音频率随时间的变化模式构成了声调或语调。以汉语普通话为例,普通话有4种声调类型:①阴平:基音频率最高而且平直;②阳平:基音频率走势由中等上升到最高;③上声:基音频率走势先向下下降到最低而后上升到次高;④去声:基音频率走势从最高一直下降到很低。汉语普通话声调在语音结构中最为敏感,它具有很强的抗干扰能力,在多种失真条件下(例如,噪声掩蔽、非线性失真、频率失真等)保有很高的清晰度,同时有助于提高言语可懂度。

(2) 音强是指声音的强弱,它是由声波振动幅度决定的。人耳对声音强弱的主观感受和客观实际并不一致。在声学上,语音的强弱用声压级(dB)来表示;响度是其主观感受量(sones),其与声压和频率均有关系。

(3) 音长又叫时长,是指声音信号的长短,单位为 s。在声学上,经常采用“时长”;“音长”更多见于听感,但两者之间并没有严格的区分。在音系学上有长音和短音的区别。在汉语普通话里,一个音节中的各个音素,因为作用的不同,其音的长短不一样。一个多音节的词,各个音节的轻重不同,其长短就不同。此外,不同音长还可以表达出不同的语气和情态。

(4) 音色又称为音质,是一种声音区别于其他声音的基本特征。音色更多用于语言学和听感。音色的不同取决于三个因素:声道形状和尺寸不同,发出的音的音色不同;声带振动发出的音与声带不振动而由其他器官发出的音的音色不同;相同的发音器官而采用送气的与采用不送气的方法发出的音的音色不同。语音的声源、发音器官的变化,导致了多种多样、丰富多彩的语音音质。

1.1.2 语音的发声机制

语音是由人调节呼吸器官所产生的气流通过发音器官发出来的声音。任何一个语音都是发音器官活动的结果,气流通过的部位、方式不同,形成的声音也就不同^[1]。因此语音具有生理特征。了解发音器官的构造以及各部分器官在构成语音过程中的作用,可以更好地帮助我们正确发出和辨别各种不同的声音。语音学对语音的定性描写和分类通常是以发音器官的部位及其活动方式作为基础的。

人的发音器官及其工作原理如图 1-1 所示,主要是由肺、气管、喉(含声带)、咽、鼻和口组成。这些器官形成一条由肺至唇的形状复杂的通道,其中,喉以上的部分称之为声道。发声器官的功能结构或参与发声的部件如图 1-1(b)所示。

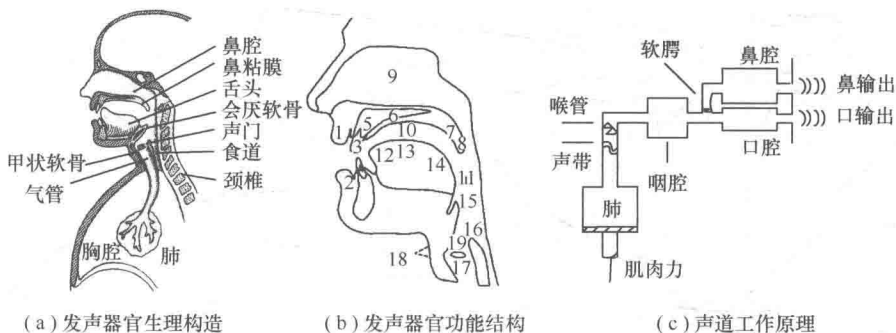


图 1-1 人的发音器官及其工作原理

- 1—上唇; 2—下唇; 3—上齿; 4—下齿; 5—齿龈; 6—硬腭; 7—软腭; 8—小舌; 9—鼻腔;
10—口腔; 11—咽头; 12—舌尖; 13—前舌尖; 14—后舌面;
15—会厌; 16—食道; 17—气管; 18—喉头; 19—声带。

肺的呼吸形成气流作为发音的动力,呼气量的大小和语音的强弱密切相关。气流经过气管到达喉头,引起喉头中间声带的快速张合颤动,把肺气流切成一连串的喷流,产生了像蜂鸣一样的嗡嗡声,也就是声带音。声带的颤动有很强的节奏性,一般人正常说话时声带每秒颤动在 80~400 次之间,这种有节奏的周期波成为语音中的浊音声源。声带的颤动频率与语音的高低密切相关,声带绷越紧,颤动越快,声音就越高;声带越放松,颤动越慢,声音也就越低。这种控制语音高低的能力在语言中起极其重要的作用,尤其是对汉语这种有声调语言,声调的高

低升降完全是由声带的绷紧和放松所决定的。实际上我们无法听到原始的声带音,我们听到的是经过喉腔、咽腔、口腔、鼻腔和唇腔这五个共振声腔调节之后已经发生变化的声波。声音通过声腔时,由于声腔形状的灵活变化,产生不同的共振,形成不同的声音。

1.1.3 语音的听觉感知

人耳的听觉过程是一个复杂的生理过程,它和人耳的机械结构和听觉神经系统有关。人的听觉系统是一个非常精妙的声音信号处理器,对声音信号的处理能力取决于它的生理结构。人处理听觉信息,除了利用听觉系统接受声音外,还需要大脑中枢来处理这些听觉信息,这就形成了听觉通路。听觉通路(auditory pathway)简称听路,是指与听觉产生相关的一系列解剖结构。听觉通路在中枢神经系统之外的部分称为听觉外周,在中枢神经系统内的部分称为听觉中枢或中枢听觉系统。听觉外周,就是我们通常所说的听觉器官,即“耳”。人耳的生理构造及其简化结构如图1-2所示,主要包括外耳、中耳、内耳这三部分。

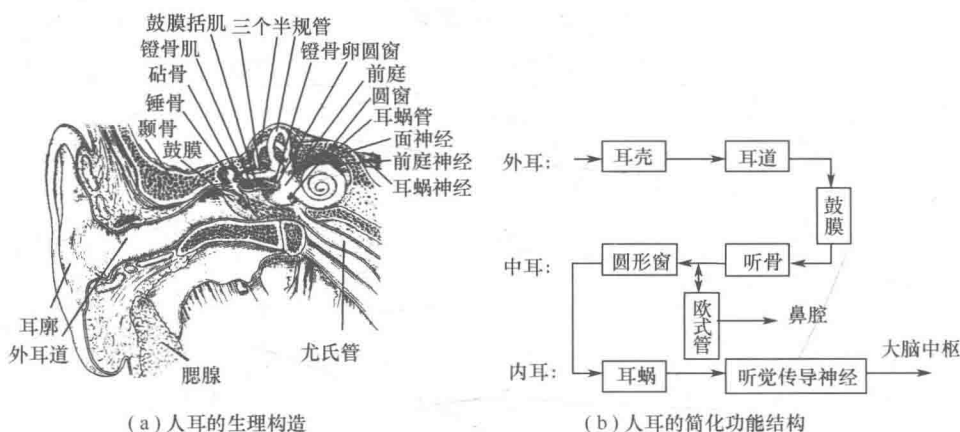


图1-2 人耳的生理构造及其简化结构

外耳包括耳翼、外耳道和鼓膜,负责声音的采集。外耳道是一根一头开,一头闭的管子,长约为2.7cm,直径约为0.7cm。外耳道的自然谐振频率约为3.4kHz,封闭时最低共振频率约为3.06kHz。由于外耳道的共振效应以及人头对声音的衍射作用,会使外耳的传输增益在2~4kHz的频率范围区间较高,这也就是人类听力一般对3kHz~4kHz之间的声音最敏感的一个重要原因。

中耳主要包括锥骨、砧骨和镫骨这三块听小骨构成的听骨链和咽鼓管,负责声音信号的放大和阻抗匹配。同时,在一定声强范围内,听小骨对声音进行线性传递,而在声强特别大时,进行非线性传递,使声强减弱后再传入内耳,以保护内耳。中耳是听觉非线性产生的原因之一。