



中国语言学会语音学分会
Phonetic Association of China



第5辑

Chinese Journal
of Phonetics

中国语音学报

主 编：李爱军

中国社会科学出版社

第5辑

Chinese Journal of Phonetics

中国语音学报

主 编：李爱军



中国社会科学出版社

图书在版编目(CIP)数据

中国语音学报. 第5辑/李爱军主编. —北京: 中国社会科学出版社,
2015. 12

ISBN 978-7-5161-7346-6

I. ①中… II. ①李… III. ①汉语—语音学—文集 IV. ①H11-53

中国版本图书馆 CIP 数据核字 (2015) 第 313133 号

出 版 人 赵剑英
选题策划 张 林
责任编辑 陈雅慧
责任校对 高建春
责任印制 戴 宽

出 版 中国社会科学出版社
社 址 北京鼓楼西大街甲 158 号
邮 编 100720
网 址 <http://www.csspw.cn>
发 行 部 010-84083685
门 市 部 010-84029450
经 销 新华书店及其他书店

印 刷 北京君升印刷有限公司
装 订 廊坊市广阳区广增装订厂
版 次 2015 年 12 月第 1 版
印 次 2015 年 12 月第 1 次印刷

开 本 787×1092 1/16
印 张 7.75
插 页 2
字 数 205 千字
定 价 36.00 元

凡购买中国社会科学出版社图书, 如有质量问题请与本社营销中心联系调换
电话: 010-84083683

版权所有 侵权必究

《中国语音学报》编辑委员会

主编：李爱军

编委：孔江平、陶建华、李蕙心、胡 方、熊子瑜

中国语言学会语音学分会顾问委员会委员

鲍怀翘

曹剑芬

方棣棠

黄泰翼

林茂灿

吕士楠

王理嘉

王仁华

徐云杨

中国社会科学院民族学与人类学研究所

中国社会科学院语言研究所

清华大学

中国科学院自动化研究所

中国社会科学院语言研究所

中国科学院声学研究所

北京大学

中国科学技术大学

香港城市大学

中国语言学会语音学分会专业委员会委员

艾斯卡尔·艾木都拉

白音门德

蔡莲红

曹 文

陈 桦

陈轶亚

陈忠敏

党建武

丁红卫

新疆大学

内蒙古大学

清华大学

北京语言大学

南京大学

荷兰莱顿大学

复旦大学

天津大学

上海交通大学

顾文涛
呼 和
胡 方
贾 媛
孔江平
李爱军
李蕙心
李敬阳
李永宏
凌震华
刘俐李
刘新中
路继伦
马秋武
冉启斌
石 锋
陶建华
王洪君
王英利
王韞佳
魏建国
吴民华
谢 磊
熊子瑜
许 毅
颜永红
杨玉芳
翟红华
张 华
张劲松
郑 方
郑秋豫
朱维彬
朱晓农
祖漪清

南京师范大学
中国社会科学院民族学与人类学研究所
中国社会科学院语言研究所
中国社会科学院语言研究所
北京大学
中国社会科学院语言研究所
香港城市大学
公安部物证鉴定中心情报信息技术处
西北民族大学
中国科学技术大学
南京师范大学
暨南大学
天津师范大学
同济大学
南开大学
南开大学, 北京语言大学
中国科学院自动化研究所
北京大学
广东省公安厅刑事技术中心
北京大学
天津大学
香港大学
西北工业大学
中国社会科学院语言研究所
伦敦大学学院
中国科学院声学研究所
中国科学院心理研究所
山东农业大学
同仁医院耳鼻喉科
北京语言大学
清华大学
台湾中研院语言学研究所
北京交通大学
香港科技大学
安徽科大讯飞信息科技股份有限公司

目 录

鲍怀翘 再谈语音量子理论	(1)
郑秋豫 从口语韵律的功能与分析看语音对立	(11)
阎锦婷 王 萍 石 锋 普通话是非问句语调的起伏度	(21)
冉启斌 顾 倩 余慧敏 进贤七里赣语有多少个声调? ——汉语方言极限声调清单声学实验分析之二	(28)
秦 鹏 葛佳登 词内语境和刺激类型对普通话阴平—阳平感知的影响	(41)
谢郴伟 温宝莹 梁嘉乐 广州话双字组平调的听感实验初探	(53)
刘掌才 石 锋 普通话元音/a/和/ə/的听感边界初探	(62)
桑 塔 姚 云 兰正群 安多藏语塞音的 VAT 研究	(75)
梁嘉莹 熊子瑜 刘新中 粤方言肇庆端州话的声调系统实验研究	(82)
时秀娟 成都话响音的鼻化度 ——兼论其/n、l/不分的实质及类型	(92)
王毓钧 贾 媛 李爱军 语篇中生理呼吸与韵律切分的关系研究 ——基于汉语母语者和汉语学习者的分析	(101)
高玉娟 夏全胜 中国学生法语元音发音中“反向”现象的实验研究	(110)

再谈语音量子理论

鲍怀翘

摘要 1989年Stevens在Journal of Phonetics发表了关于语音量子理论的长文(1989),该杂志同期刊登了15位当时世界顶尖语音学家的11篇评论文章,这在语音学研究史上也是不多见的。该期语音学报最后发表了Stevens的“Response to commentaries”(1989),在这篇文章中Stevens回答了对他的评论,进一步阐述了他的量子理论思想。2010年他重申了他的量子理论,并作了重要补充(Steven & Keyser 2010)。

Stevens的语音量子理论已在上期《中国语音学报》(李智强 2014)和最近出版的《实验语音学概要增订版》(鲍怀翘、林茂灿 2014)作了简要介绍。为了深入了解Stevens的语音量子理论以及语音学界对此理论的讨论,本文将对此作出较为翔实的介绍和评论。

关键词 语言量子理论,言语产生理论,零点效应,非单调性,平台

Revisit the Quantal Nature of Speech

BAO Huaqiao

Abstract In 1989 K. N. Stevens' *Quantal Nature of Speech* (1989) was published, along with 11 commentaries written by 15 world leading phoneticians in the same issue, which was unusual in the history of phonetics study. The journal later published K. N. Stevens' "Response to commentaries" (1989), in which he made replies and elaborated his Quantal theory. In 2010, he reiterated that the QNS and made an important supplements. (Steven & Keyser 2010)

K. N. Stevens' Quantal Theory has been introduced briefly in latest issue of *Chinese Journal of Phonetics* (李智强 2014) and in the *Outline of Experimental Phonetics (Revised and Enlarged Edition)* (鲍怀翘、林茂灿 2014). In order to further explore the theory and its commentaries, this paper will make a detailed introduction and commentary.

Key words Quantal Nature of Speech, Theory of speech production, Zero effect, Non-monotonic, Plateau

1. 发音器官参数—声学参数计算模型问题

众所周知,发音器官参数与声学参数之间的关系是语音学研究中最基本的“言语产生理论”问题。涉及什么样的发音器官形状会产生什么样的声学结果,进而利用声学结果(参数)进行听辨感知的研究,同时还涉及由声学参数反推发音器官形状的问题。假若“走通”了这个循环,那么可以说“言语产生理论”的研究真正实现了。

1.1 Stevens的言语产生模型:

Stevens的言语产生模型是他的语音量子理论的基础和依据,这个模型可以简要地归结为“规则声管耦合”模型(见图1)(Stevens 1989),这一模型成为语音量子理论的基础。

l_1 为后腔管长, A_1 为后腔截面积; l_2 为前腔管长, A_2 为前腔截面积。 A_c 、 l_c 分别为收紧段的面积和长度。当 $A_c \ll A_1$ 和 A_2 (因此 A_c 可视为零), 这种情况下前后腔之间没有“耦合”, 此时后腔的自然频率为:

$$0, \frac{c}{2l_1}, \frac{c}{l_1}, \frac{3c}{2l_1} \dots$$

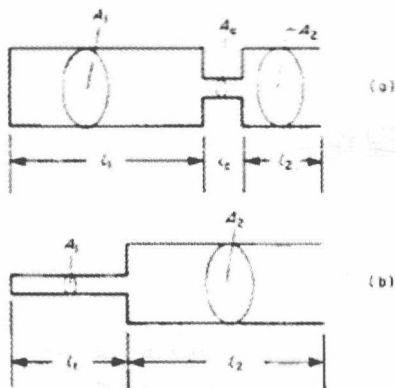


图 1 双管耦合声道模型

其中 c 为声速。前腔的自然频率为：

$$\frac{c}{4l_1}, \frac{3c}{4l_1}, \dots$$

假定共鸣器壁是硬性的，而且不考虑辐射的作用，对应的自然频率如图 2 中虚线所示（图中标记为 0 处为 $A_c=0$ ），此时，最低频率为 0。

但是，当 A_c 为非零，但比起 A_1 和 A_2 仍然算是小的（这里假设 $A=A_1=A_2$ ），也就是这个系统为“耦合”时，前后腔产生互相影响，则 F_1 为非“0”：

$$F_1 = \frac{c}{2\pi \sqrt{A l_1 \left(\frac{l_c}{A_c} + \frac{l_2}{A} \right)}} \quad (1)$$

假定共鸣器壁是硬性的，那么其他自然频率会向高变化。当 $l_1+l_c+l_2=16\text{cm}$ ， $l_c=2\text{cm}$ ，截面积 A_1 、 A_2 为 3cm^2 时，得到 4 个自然频率如图 2。图中实线和短虚线分别表示 $A_c=0.2\text{cm}^2$ 和 0.5cm^2 ，后腔长度 l_1 从 0cm 改变 14cm 时四个自然频率走向。

对前高元音 /i/ 和 /e/，Stevens 给出了具体计算。图 3 是这两个元音的发音器官矢状面，其模拟的声管模型见图 4。

改变图 4 中 l_1 和 l_c 可以获得不同的前高元音声管构型，它们的自然频率随后腔和前腔长度的改变而变化，见图 5。 $l_1 > 6\text{cm}$ 的构型对应着非低前元音。

后高元音 /u/、/o/ 的发音矢状面见图 6。舌根抬高后缩，因此后腔明显缩小而前腔有较大的容积，双唇合拢形成圆唇，它的声管模型见图 7。图 7 与图 1 (a) 类

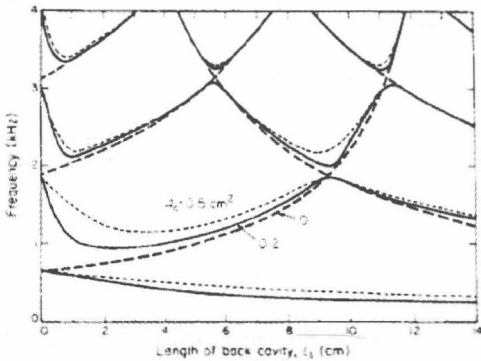


图 2 图 1 (a) 构型对应的自然频率 (Stevens 1989)



图 3 两个前高元音的矢状面 (Stevens 1989)

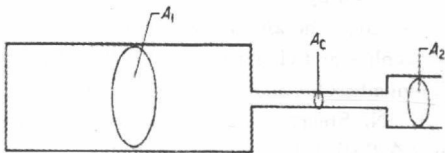


图 4 非低前元音声管逼近 (Stevens 1989)

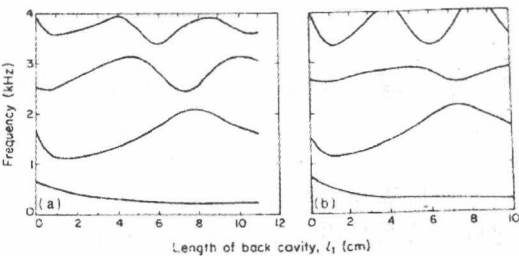


图 5 图 4 声管构型的自然频率。(a) 相当于 $l_c=5\text{cm}$ ，而 (b) $l_c=6\text{cm}$ 。面积 $A_c=0.3\text{cm}^2$ ， $A_1=3\text{cm}^2$ ， $A_2=1.0\text{cm}^2$ 。声管总长为 16cm。(Stevens 1989)

似，但开口端声管变窄，对应“圆唇”，其面积为 A_m 。这种构型的声管参数是：

$A_1=A_2=3\text{cm}^2$, $A_m=0.3\text{cm}^2$, 收紧点 A_c 长度 2cm, A_m 长度 1cm, 声管总长度为 16cm, 此时可得到该声管随后腔长度变化而改变的自然频率 (见图 8)。

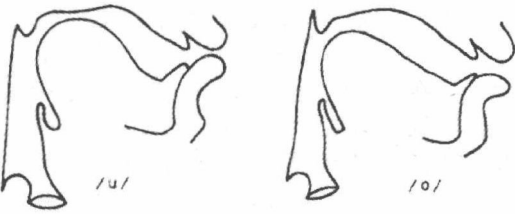


图 6 后高元音 /u/、/o/ 的构型 (Stevens 1989)

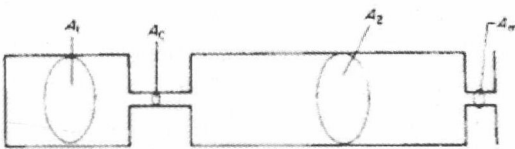


图 7 后高元音 /u/、/o/ 声管模型

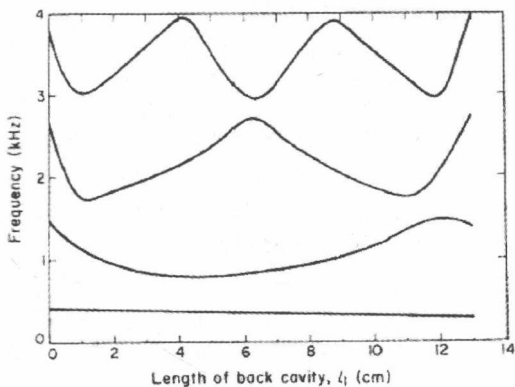


图 8 后高元音 /u/、/o/ 声管模型的自然频率是后腔长度 L_i 的函数 (Stevens 1989)

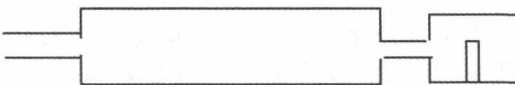


图 9 擦音声管模拟 (Stevens 1989)

当 $6\text{cm} > L_i > 2\text{cm}$ 时, 对应着非低后圆唇元音。此时 F_1 和 F_2 的变化空间为较低的 400—500Hz, F_1 不会获得最大值, 变化范围也只有 80Hz。

其他元音都可通过改变声管参数来

近似。

产生擦音的声管较特别 (见图 9), 前腔有一阻碍, 此处截面积 A 大大缩小; 而后腔的后边耦合一段管子 (气管), 声门是打开的, 呼出气流直接在前腔阻碍处产生摩擦噪声 (擦音声源), 其振幅 P_s 可由公式 (2) 逼近

$$P_s = KU^3 A^{-5/2} \quad (2)$$

其中 K 为常量, U^3 为体积流速度, A 为截面积。

1.2 对此模型评论及回应

图 10 是 Stevens 给出的语音量子理论 QNS 的理想模型 (Stevens 1989), 此模型认为发音—声学参数之间为非线性的关系, I 与 III 为两个相对稳定区域, 称之为“平台” (plateau), II 区的声学参数快速变化而发音参数变化相对较小。具体解释见 (李智强 2014; 鲍怀翹、林茂灿 2014)。

针对 QNS 中元音的发音—声学关系, 有几个评论 (Fant 1989; Ladeforged & Lndau 1989) 反复以不同的形式指出, Stevens 所选择的发音器官空间是经过精心挑选的, 而选择的特定声学参数是为了说明平台 (图 10 中 I 和 III) 和不连续 (discontinuous) 的存在 (图 10 中 II)。基本的问题是 (1) 模型中的发音器官参数是否与发音人真实的参数相当? (2) 与这些发音器官改变相应的声学参数是否被用于听音人对语音的感知?

Stevens 认为, QNS 中给出的许多例子都是实际的物理维度 (A_i , L_i), 如 1.1 中引用的那些例子。图 10 中, 平台和不连续性是耦合共鸣器的直接结果, 也是机械振荡器不稳定性、管子收紧点和阻碍处气流的不稳定性的直接结果。QNS 前提之一是这些原理被利用来选择发音器官模型, 而这些模型的声学结果可以用于语言的语音。

某些评论指出 QNS 缺乏声学参数与发音参数之间非单调性关系的系统性例证, 需要区别哪些是有实证意义的非单调关系, 哪些又不是。

Stevens 认为, 系统讲述和建立选择

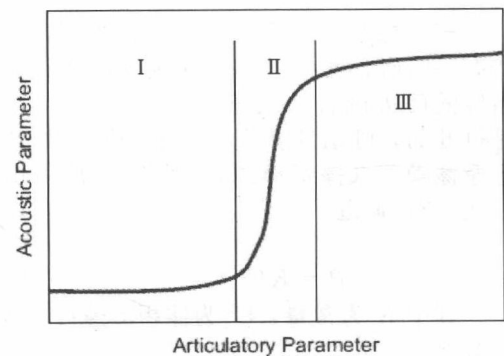


图 10 QNS 给出的发音—声学关系假设, 横轴为发音参数, 纵轴为声学参数。I 与 III 为 2 个相对稳定区域, II 区的声学参数快速变化而发音参数变化相对较小。(Stevens 1989)

发音器官参数的标准, 以及选择描述声学 与发音关系的标准不是一件容易的事, 受 发音人控制的一组发音器官参数是变化万 千的, 相关声学参数也是变化多端的。

声学 与发音参数之间最直接、最清晰的 证据来自于与发音方法和声带振动相关 的发音器官和喉的调整。例如, 考察非 鼻元音—辅音—元音音段 (VCV) 的产 出, 发辅音时声门上收紧点截面积会成系 统地减小, 而声源声学特性会出现突然改 变; 当辅音收紧面积小于平均声门面积 时, 收紧点后面的气压会积聚起来, 于是 收紧点附近产生了扰动噪声, 同时声带振 动足够地弱化, 甚至完全停止振动。其结 果是在语流中出现明显的不连续性, 而不 是发音器官和喉部的相对连续运动。在 这些声学不连续之间, 声源特性仍然是稳 定的、持续的 (指元音部分)。

当收紧点缩到完全关闭, 声音的特性 将更进一步改变。关闭突然被释放, 就会 形成声谱上振幅的不连续性。不连续性也 会发生在腭咽部的突然打开或关闭。

观察辅音收紧点大小变化在一定意义 上说明了声学 与发音参数之间的原始 (典 型) 形态。当辅音收紧点尺寸大于声门平 均面积时声源特性相对不变化, 在声门上 收紧处可观察到小的气流扰动, 或根本不 会出现扰动噪声。当某部位收紧得更狭窄 时, 收紧处就会出现扰动噪声, 此时噪声

特性对收紧点面积大小保持相对不灵敏。 当辅音内破裂 (implosion) 时, 收紧点上 声源特性和扰动噪声振幅会突然改变。辅 音关闭的内破裂和释放产生了另一类不同 的不连续。

在不同发音方法的情况下 (被认为是 continuant 和 sonorant 特征), QNS 量子 关系处在图 1 的 II 区, 发音—声学关系的 这种不连续, 对应到语音信号中具有明确 定义的界标 (landmarks) 或事件 (events) 上。Lindblom & Engstrand (1989) 举例 ([^xpi | ki: p]) 说明发音的连续性 (指发 音器官处在连续运动状态) 但在声学上却 表现为离散型。

下面再来讨论阻碍辅音 (obstruent) 不同发音方法的情况。我们限定仅涉及收 紧点附近扰动噪声的声学属性, 无论辅音 是塞音还是擦音, 辅音收紧部位 (收紧 点) 从唇—齿龈—腭—咽方向移动, 前腔 长度依次递增, 由受阻部位激发的噪声的 共鸣频率也依次降低。受阻部位在唇部, 如唇擦音, 可以说没有前腔, 因此不存在 强的声谱凸显, 而齿音或齿龈擦音, 共鸣 会很高, 会出现最低共鸣与邻接元音 F_4 接近; 对硬腭、软腭或咽擦音, 最低共振 峰往往与邻接元音的 F_2 、 F_3 靠近。在 这种情况下, 相关的声学参数不是共振峰 的绝对频率或绝对声谱形状, 而是选择与 邻接元音声谱凸显 (即元音共振峰) 相联系 的噪声谱。Fant (1989) 的评论使我们想 起阻碍部位的破裂声谱和摩擦声谱由声源 与收紧点耦合产生的零点效应。

还有几个评论指出收紧点由唇移到齿 后位置 (postdental) 不代表连续, 因为 这类移动包括了主动发音器官结构由唇改 为舌尖。针对这点, Stevens 指出: 发音 器官本身是量子, 没有必要去调用 QNS 图中的声学—发音关系。无论是塞音、擦 音还是鼻音部位由双唇改变到齿和齿龈, 与其声学结果是一致的, 如图 11 所示, 图中曲线标记 0cm 为双唇辅音, 1.5 或 2.2cm 为齿背辅音。在鼻音情况, 在辅音 除阻进入后接元音处频谱表现明显不同 (注: 指元音过渡起始点), 因为双唇和齿 后辅音的零点是不同的, 因此发音部位被

看成“量子”上的不同。

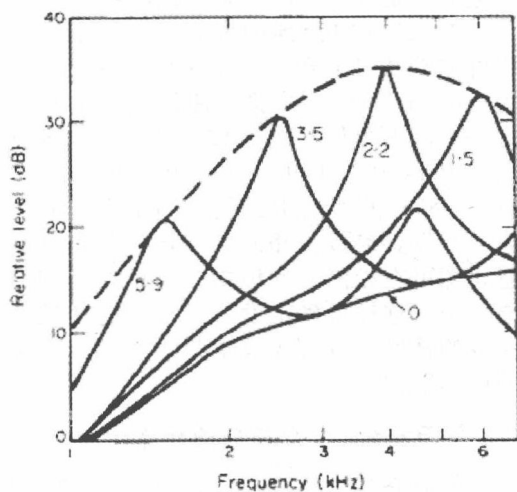


图 11 擦音的辐射声压谱。不同的曲线对应不同的前腔长度 (cm)。假设声源在距离收紧点 1cm 处成阻, 前腔面积 1cm^2 , 收紧部分 0.2cm^2 , 声门下压力为 $8\text{cm H}_2\text{O}$ 。图中虚线给出了振幅谱峰的包络。(Stevens 1989)

下面再来讨论声门动作的声学结果。由于缺少合适的声带振动模型和与声带不同构型相关的量化数据, 分析产生困难, 举步维艰。通过一些零碎的知识, QNS 提出了一个可能的分析办法。假设当声带处在振动周期中, 声门尽量内收但仍不能完全闭合, 使其基频分量相对于高频得到增强, 而高频部分被弱化。这一点不难理解, 因为气流的直流成分被叠加在交流成分上, 其结果第一谐波分量得到增强, 而高频成分被抑制。

发元音时调整腭咽开度会引向另一类声学—发音关系。当腭咽口截面积增加使元音低频分量发生改变, 这种改变超过一定界限时被听音人感知为“鼻化”, 此时在频谱中低频部分出现一个声谱峰。在这种情况下, 发音参数是腭咽口部尺寸大小, 听音人对鼻化的灵敏度依赖于元音高度 (House & Stevens 1956)。

最后我们回过头来看元音声学—发音参数之间的各种关系。我们知道由于舌头控制声道中的收紧, 可能还有唇部圆展的配合, 收紧由前向后改变, 共振峰频率也

会成系统地变化。收紧位置在前部的话, F_2 是高的并接近 F_3 , 在这个区域两个共振峰接近对收紧点位置是相对灵敏的, 假定收紧截面和收紧长度相对固定的话。当舌体后缩, 即收紧点向后移动, F_2 降低接近 F_1 , 此时共振峰频率对收紧点位置的灵敏度就降低。当收紧点位置处在两端点之间, F_2 也会处于 F_1 与 F_3 之间。当收紧点越过中间位置, F_2 不会突然改变。在收紧点位置不变的条件下, F_2 的大小还与唇的圆展度相关; 在相似区域, 声道前部张开 (如开元音) 而后部相对狭小, 但收紧点狭窄部位的长度受到操控的话, 可观察到 F_2 的最大值或最小值。

QNS 指出, 许多情况下, 精心控制收紧点前后位置 (不变) 或者控制狭管管子的截面, 声学参数 (如 F_1) 和发音参数之间不存在单调性关系。Diehl (1989) 在他的评论文章中观察到了元音的收紧程度对共振峰频率最灵敏, 收紧点前后位置对 F_2 不太灵敏。Lindblom & Engstrand (1989) 指出, F_1 对收紧度的灵敏性在高元音中要差一些, 因为受到口腔壁阻尼的影响。Perkell & Cohen (1989) 文章指出, 在前高元音和后低元音中, 生物机制约缩 (特别是肌肉的活性与位置之间的关系) 使其减小了 F_1 对肌肉收紧度的灵敏性, 而肌肉控制了舌体的位置与形状。在高元音中生物机制的影响效能也在 Kakita & Fujimura (1977) 和 Abry, Boe & Schwartz (1989) 文章中谈到了。

Fant (1989) 认为元音共振峰与收紧点位置之间关系的研究和结论有必要做进一步推敲。Fant (1989), Lindblom & Engstrand (1989), 还有 Ladeforged & Lindau (1989) 等提出了另一种模型和构建发音—声学关系。例如 Fant 指出, 瑞典语中与非圆前元音声学参数关系最大的不是 F_2 , 也不是 F_1 与 F_3 之间的间距, 而是由 F_2 、 F_3 和 F_4 组成的谱重心 (凸显), 调整向前收紧, 可获得最大 (频率最高) 的重心。此时 F_3 非常接近 F_4 , 而不是 F_2 最高。当然, 最前位置收紧 F_2 必然达到最大值。

Ladeforged & Lindau (1989) 使用

UCLA 声道模型进行的研究, 认为前高元音 [i] 附近的声学—发音关系是 F_2 随着发音参数的调整不存在平台 (plateau) 换句话说, 随着发音参数的调整, F_2 会引起较大的改变。对此评论, Stevens 指出: 在没有评估他们的面积函数之前是难以解释这一结论的。收紧点截面积出现大的改变 (近似 Stevens 模型中三个因素中的一个, 假定他们 (UCLA) 的模型中已包括了壁阻抗), 这种改变可能是 F_2 缺少平台的一个原因。问题之一是选择发音参数如此大的改变是否有道理。当控制收紧位置前后移动时, 有理由去假定舌边表面静止地对着硬腭形成了一个相对稳定的面积。这种稳定性使 F_4 成为声学参数之一, 由此可以获得 Fant 所说的高频重心。

以上对声学—发音参数的回顾绝非全部, 但已看出发音参数的调整有很宽的变化, 却与其产生的声学结果之间的关系却为非单调性。对辅音来说, 发音参数包括声道的截面积控制在 0.1—0.3cm; 对类元音声道面积为 0.2—1cm²。辅音舌叶或舌尖的变化宽度 1—2cm, 或舌体抬高 3—4cm。对前后元音声道收紧位置约 4cm 或更多一些。上面提到的评论一个共同的目标就是独立操控发音参数, 然而到目前为止, 我们缺乏合适的标准去独立调整发音参数。对声门外展、内收参数独立操控标准预期将得到实现, 然而其他参数, 如辅音的关闭程度, 仍然是不清楚的。

图 10 的 QNS 中三个区发音参数变化依赖于发音器官的结构。对这种变化的精度在 Studdert-Kennedy (1989) 的评论中给出了意见。舌体运动的精度大大小于声带外展—内收的精度。

对以上评论, Stevens 认为在声学—发音关系中声学参数变化是相当大的, 包括在不同频率范围内频谱振幅和谐凸显频率。假设这些简单的声学参数是跟随发音器官参数定位在类平台区间 (QNS 图 1 中由 I 区到 III 区), 从而使听音人得到分辨, 也就是必须区别发音—声学关系坐标上的参数与听觉系统得到不同响应的那些特征。正是这些特征形成了语音特征 (features) 的声学相关。毋庸讳言, 我们的听

觉处理知识还没有达到定量的方法去说明响应的区别模式 (distinctive pattern)。

在 QNS 中提到的发音—声学证据其中许多涉及对语音产生系统模型各成分行为 (behavior) 的检验。使用模型的理由之一是能够很好控制参数。另一种逼近方法在 Perkell & Cohen (1989) 和 Fujimura (1989) 的评论中提到, 测量发音时的声学 and 发音参数在相同的语声进行重复测量, 即可在相同的环境中, 也可在不同的语境中, 这样可确定发音人控制这些发音参数的精度, 也有助收集与发音参数相关的声学数据。

1.3 关于发音—声学转换的计算

上面介绍了 Stevens 的“声管耦合”模型以及一些语言学家的评论, 本文作者在 1983 年初在一台内存只有 16K 的微小型计算机上完成了普通话元音发音参数与声学参数 (共振峰) 之间关系的系列实验 (鲍怀翘 1984), 1989 年将实验结论集成在《实验语音学概要》中 (鲍怀翘、林茂灿 2014)。发音器官参数是真实的, 完全依据中国社会科学院语言研究所 60 年代初完成的五个发音人的 X 光片。在实验中我们也进行了类似 Stevens 均匀声管模型计算中改变前后腔长度的研究, 即收紧点沿声道长度前后移动。但从得到的各共振峰频率走向来看, 与 Stevens 的结果是有较大差距的。我们同时还计算了改变前、后腔面积、改变唇形面积和改变声管总长度等发音参数得到共振峰频率的变化数据。整体看来, 各共振峰变化的斜率是不同的, 也没有发现明显的平稳段。当然这两种实验虽然都是由面积函数求共振峰 (自然频率) 参数, 但由于研究目的不同, 因此具体执行方法模型和算法是不同的。我们的目的在于探究声腔各参数变化与共振峰变化之关系, 也就是收紧点 (舌位) 前后移动, 前、后腔体积改变 (近似舌位高低) 和唇开度大小对共振峰的影响, 找出它们之间总的倾向, 见图 12。

通过实验我们发现, 发音器官参数变化对 F_1 、 F_2 是灵敏的, F_3 次之, F_4 、 F_5 的变化不明显。具体来说, 舌位前后移动



图 12 舌位与头两个共振峰频率关系

对 F_2 比较敏感, 而收紧点面积的改变会使 F_1 有较大的变化。这与我们在实际语音声学分析中获得的的结果是相同的, 也得到合成实验的应证。

2. 感知对立和发音器官稳定性及其精度

QNS 中的非单调性发音—声学关系对揭示语音区别特征提供了一个重要基础。非单调性关系中存在平台的假设影响到言语产生中对发音参数处置的正确性。对这个问题, 方面一些学者有如下评论:

(1) 在定义名词如 contrast, stability 和 precision 方面缺乏严谨性 (Studdert-Kennedy 1989); 定义非单调性的发音和声学参数定位在静态构型和特性上, 很少关注言语产生的运动和姿态方面 (Fujimura 1989, Goldstein 1989, Studdert-Kennedy 1989)。

(2) 为了得到声学上最终的产出只需要有限发音器官精度, 而这假设与在语流中通常出现的弱化、同化现象不一致 (Goldstein 1989, Lindblom & Engstrand 1989, Fant 1989)。

(3) 在 QNS 中倾向于限定讨论发音—声学关系包括单一的发音维度, 而其他维度保持恒定。但是发出一个语音音段要求众多发音维度协同动作 (Abry 1989, 等)。

(4) 除了 QNS 文章中所考虑的非单调性或量子关系外的其他因素可能对语言中实际出现的语音清单 (inventory) 具有不可忽视的影响 (Ladefoged & Lindau 1989, Lindblom & Engstrand 1989, ten Bosch & Pols 1989, Diehl 1989)。

Stevens 对以上问题作了如下解答。

2.1 对立 (contrast), 稳定性 (stability) 和精确度 (precision) 的定义

“对立”的概念是一个面向听音人的概念。在 QNS 中 I 区和 III 区定义的声学参数可能是两组值, 它们代表十分不同的声学参数, 使听觉系统触发了不同的响应模式。例如, 元音 F_2 与 F_3 相关, 而非与 F_1 相关; 声音是否存在低频周期性 (注: 关系到元音是否鼻化); F_1 凸显形状微扰程度是否超过一定阈值; 辅音受噪声激励的共振峰是否处在邻接元音 F_2 、 F_3 范围内, 或者它在这个范围之上能量得到凸显等。

在某些声学—发音关系中声学改变从 I 区到 III 区是非常突然的, 如定义 sonorant 到 non-sonorant 区域的改变, 或者声道由完全关闭到突然打开。结果 II 区被定义为 contrast (对立), 因为发音系统在 II 区突然改变。但在某些关系中穿越 II 区不是突然的, 如低后元音与低前元音之间的区别, 是由较低的 F_2 逐渐向较高的 F_2 转化。因此, 对立的显著性或程度在量化测量上是不易建立的。

“稳定性”这个概念与精确度是相关的。“稳定”的基本思想是在一个发音参数范围内 (相当于 I 区和 III 区), 其参数的设定是相当宽松的, 只要声学参数在这期间没有发生有意义的改变。例如, 在语境中取得某个声学特性, 允许发音参数的某些变化或补偿, 而不要求发音十分精确。在声学—发音关系中, 如果产出了一个对应平台的持续发音, 那么, 其声学参数在整个音长中将保持稳定而发音参数却有变化。如发擦音允许收紧点面积有些变化而摩擦噪声的振幅却没有发生本质意义的改变。

2.2 言语产出中的静态 (static) 与运动 (kinematic)

在 QNS 中提出了相当于平台的发音参数 (发音结构的目标构型), 但这并不暗示语音特征的发音可能出现静态的发音构型和稳定的声学特性这种单调性关系。

在特定的话语中当发音参数穿越两个平台之间的区域时, 声学参数有一个相当突然的变化。一些与语音特征相关的声学特性是通过Ⅱ区前后的不同时间点的声学采样之间关系确定的。例如, 一个非连续辅音(如塞音)的目标构型对应于声道某个点的完全关闭, 区别特征[-continuant]的声学表现为无声段(gap)后的除阻或者一个低振幅(voice bar)后的除阻。突然释放或者前加内破裂(implosive)被认为是发音参数穿越Ⅱ区的结果。另一类短时声谱的变化出现在塞音除阻处(附近), 同时连续谱之间的关系可以用来认定发音部位(指过渡音)。于是, 非单调性或非连续性的声学—发音关系即可定义与平台相关联的声学参数和发音参数, 也可定义时变的或由两个相邻平台之间发音器官移动造成的相关声学特性。

探究声学—发音关系的目的在于建立具有区别意义(对立)的声学特性清单。有的评论指出, 用特征矩阵方式对言语感知来说可能不是一个有用的办法, 同时也不能捕捉到发音过程的动态方面, 倒不如用姿态叠加来表示可能是一条更有成效的途径, 因为这样的表示以更自然的方式考虑到语境的变化。Stevens认为, 对这个问题的讨论其核心涉及词汇的模型, 这已超出了QNS的范围。在QNS中只是试图去给出各种声学—发音关系, 在这些关系中包括了平台和过渡区域, 这些关系在形成语音范畴上起到了决定性作用。这些关系也提供了语音特性清单, 包括非连续性和能用来指示作为样品的特性, 它们对语音感知是有效的。也就是说, QNS假定言语感知的初始阶段包括了对这些事件响应的处置, 并赋值声学特性给这些事件和这些事件定义的区域。并认为这些处置过程是推论话语更抽象表示的基本步骤, 它可能是姿态系列, 离散音段, 区别特征, 或者其他形式。不过QNS认识到由声音到更抽象转换可能会出现不一致。

2.3 量子关系会降低发音的精度吗?

Goldstein (1989), Lindblom & Engstrand (1989), Fant (1989) 和 Fujimura

(1989)指出, 在连续语流中常常不可能得到声学特性的稳定段。提到的例子是辅音的弱化使塞音变成连续音, 持续的清擦音变成响音, 特别在重读音节之后的位置上。此外, 还有很多同化现象, 由此声学特性也改变了原来的范围。面对相当宽松的发音精度, 为什么得不到理想的、预期的声学结果呢? 对以上提出的问题 Stevens认为是语音信号特性常常存在冗余度, 不是全部声学特性都用来定义一个词的常规形式, 某些变化(如一个擦音去替代塞音)不会干扰这个词的认定, 比如“spokeman”被读成“sposman”。因为英语中软腭塞音与擦音是不对立的, 当一个词放在一定的语境中, 那么冗余度就会增加, 很多因素都会参与发音补偿, 甚至离声学—发音关系中平台定义的精度很远。

2.4 单一发音维度是否限制了QNS的通用性?

几个评论批评QNS所举声学—发音关系非单调性的例子只包括一个发音参数(见本文1.1)。对此, Stevens在上面已指出, 选择发音维度的目的在于能独立控制, 假若几个维度同时使用会使得发音—声学关系找不到已知的平台和过渡区域。

在QNS中没有去系统地检验每一个发音参数平台和过渡段的位置, 当其他参数固定不变的情况下。对某些关系, 可以预测这些临界区域确实是独立设置其他参数, 而这些其他参数可能会发生相互作用。例如, 声道收紧点截面积由响音(口内建立不起来超压)过渡到非响音是在声道收紧点上独立作用的, 即使它(或多或少)依赖于声门构型(breathy or pressed voicing)和声门下压力。

当声门内收—外展维度发生改变时, 可以不依赖元音舌体构型的改变而使发声类型由气嗓音过渡到正常嗓音, 再到挤压嗓音(pressed voice)。腭咽打开引起鼻化, 其阈值多少与元音构型相关。舌体的位置影响 F_2 的稳定性, 舌体向前向后都会使 F_2 发生变化, 且又与圆唇度有关。在阻碍辅音里(obstruent)舌叶或舌体收紧的位置使其获得稳定的结果, 即声谱峰

与后接元音的 F_2 相连续, 不管辅音是连续音摩擦音还是塞音其结果都是相同的。有意义的是, 无论是阻碍音还是鼻音, 相同的收紧位置可以得到相似的声学输出, 当然进一步的研究是必要的。

这些例子说明平台和过渡区域或多或少独立于其他发音参数的设置, 例外也可能是存在的。在鼻化元音情况下, 发音人降低软腭, 打开鼻咽口的面积是依赖元音高度的 (Henderson, 1984), 显然, QNS 理论的这些方面还需要进一步研究。

也存在这样的情况, 在特定发音一声学关系中, 对某组值 (而其他发音参数固定不变) 可能存在平台, 但对另一组值则不会出现平台。例如, 阻碍音的发音位置在小舌和咽腔时, 发音一声学关系就会呈现出平台。但是这些位置上的鼻辅音不可能产生平台, 舌冠擦音相关的发音位置与对应的声学属性 (建立在声学—发音关系上的平台与过渡区域) 在 QNS 中没有讨论到, 这大概是全部鼻辅音发音部位不存在类似的发音一声学关系。因为对擦音来说建立在声学特性上的关系只能由扰动噪声产生, 与此相类似, 一定的特征组合使用在语言中是绝对无效的。

作为定义语音特征基础的一维非单调性声学—发音关系的有效性问题, 已由 Blumstein (1989) 提到了。其例子是不同部位的塞音、擦音和鼻辅音, 在不同发音方法条件下, 它们的声学参数是不同的。部位相同但发音方法不同其声学特性是否相同? 不同方法的声学 (或感知) 相关都是不同的吗? 当然, 相似的方法会有相似的声学表现, 例如, 齿龈塞音和齿龈鼻音较之双唇塞音和双唇鼻音显示出更高频段的能量 (与邻接元音的入渡 (onset) 关系)。也就是不同方法但部位相同有着相似的语音对立声学相关。这种说法的历史已经很久了, 也是 Jakobson, Fant 和 Halle (1963) 区别特征的中心思想。

Blumstein (1989) 进一步的问题是, 与特定特征 (如定义辅音发音部位的特征) 相联系的声学特性在 CV 环境中是相同的吗? 这方面的工作需要更多的研究来证明。

2.5 语音量子理论局限性

Stevens 声明, QNS 理论发展的目的在于确定语音对立的客观标准。当下的理论是不区别语言中的语音对立, 或仅有少数是由语音对立的, 也不去预测语言中元音和辅音清单的大小和它们是如何被公布的。在量子理论的框架内, 以感知距离的测量去赋值每组的对立的程度或者感知的显著度 (QNS 中的 II 区)。Lindblom 和他的同事已经做了些努力 (Lindblom & Engstrand 1989)。语音清单方面还有很大的领域值得进一步去研究, 但在目前阶段量子理论不注重这个问题。

本文作者认为, 上面提到的“对立”是针对语音感知的。众所周知, 辅音, 如 /b/、/d/、/g/ 是属范畴感知, 它们的发音一声学关系符合 QNS 中平台 (I 区、III 区) 和不连续 (II 区), 但对元音来说就很勉强了, 因为元音属连续感知, 没有明显的范畴。Stevens 也承认, 低后元音与低前元音之间的区别, “对立的显著性或程度在量化测量上是不易建立的”。

从上面对“对立”“稳定性”“精确度”以及“稳定对动态”的讨论中, 我们可以看出 Stevens 提出量子理论, 其目的正如他自己所说的, “探究声学—发音关系的目的在于建立具有区别意义 (对立) 的声学特性清单”, 也就是给区别特征提供发音一声学的解释。这种局限性使他无法对众多评论作出满意的回答。众所周知, 语音底层虽可以用特征矩阵定义一个音, 但当这个音一进入语流中 (音节、词和短语) 就会产生音变甚至脱落, 所赋值该音段 (或音位) 的区别特征及其包含的声学—发音特性也将随之消失其意义。我们认为语言的本质特性, 无论是发音的还是声学的, 存在于语流中, 因而语流中的语音研究应受到特别关注。所以 Stevens 认为语音的动态特性 “这个问题讨论的中心词汇的模型, 这已超出了 QNS 的范围”, 这就是 QNS 问题所在。

为了补偿 QNS 理论的局限性, Stevens & Keyser 在 2010 年的文章中 [7],

针对在语流中音变问题, 提出了“增强”(enhancement)和“叠接”(overlap)的概念来补充 QNS 的不足。我们觉得所举例子及解释无补于 QNS 的局限, 反而觉得用“协同发音”的理论来阐述更能说明问题。我们认为, “协同发音”理论的研究是语音学核心问题之一, 应大力提倡。但无论如何, Stevens 回答中对 VCV 语音串的分析, 包括喉—声腔收紧联动的说明, 无疑值得我们学习; 他提出的语音发声生理和声学分析相结合的研究方法也是语音研究必须遵循的道路。只有深入研究语音生理特征, 才能正确理解语音的声学现象, 这是 QNS 对我们的最大启示。

参考文献

- Stevens, K. N. (1989) "On the quantal nature of speech", *Journal of Phonetics*, 17: 3—43.
- Stevens, K. N. (1989) "Response to commentaries", *Journal of Phonetics*, 17: 145—157.
- 李智强:《区别特征理论的语音学基础》,《中国语音学报》第4辑,商务印书馆2014年版。
- 鲍怀翘、林茂灿主编:《实验语音学概要》增订版,北京大学出版社2014年版,第107—113页。
- 鲍怀翘:《声道截面函数和元音共振峰计算》,《应用声学学术讨论会文集》,又见《声积函数和共振峰频率的初步报告》,中国社会科学院语言研究所《语音研究报告》光盘版,1984年。
- Jakobson, R., Fant, G. & Halle, M. (1963) *Preliminaries to speech analysis*, Cambridge MA: MIT press.
- Stevens, K. N. & Keyser, S. J. (2010) "Quantal theory, enhancement and overlap", *Journal of Phonetics*, 38: 10—19.
- Henderson, J. B. (1984) *Velopharyngeal function in oral and nasal vowel: across-language study*, Unpublished doctor dissertation, university of Connecticut.
- Lindblom, B. & Engstrand, O. (1989) "In what sense is speech quantal?" *Journal of Phonetics*, 17: 107—121.
- Fant, G. (1989) "Quantal theory and features", *Journal of Phonetics*, 17: 79—86.
- Ladefoged, P. & Lindau, M. (1989) "Modeling articulatory-acoustic relations: A Comment on Stevens' On the quantal nature of speech", *Journal of Phonetics*, 17: 99—106.
- House, A. S. & Stevens, K. N. (1956) "Analog studies of the nasalization of vowels", *Journal of Speech and Hearing Disorders*, 21: 218—232.
- Diehl, R. L. (1989) "Remarks on Stevens' quantal nature of speech", *Journal of Phonetics*, 17: 71—78.
- Perkell, J. S. & Cohen, M. H. (1989) "An indirect test of the quantal nature of speech in production of the vowels /i/, /a/ and /u/", *Journal of Phonetics*, 17: 123—134.
- Kakita, Y. & Fujimura, O. (1977) "Computational model of the tongue: a revised version", *Journal of the Acoustical Society of America*, 62 (51): 515 (A).
- Abry, C., Boe, L. J. & Schwartz, J. L. (1989) "Plateaus, catastrophes and the structuring of vowel systems", *Journal of Phonetics*, 17: 47—54.
- Studdert-Kennedy, M. (1989) "Feature fitting: a comment on K. N. Stevens' On the quantal nature of speech", *Journal of Phonetics*, 17: 135—143.
- Fujimura, O. (1989) "Comments on 'on the quantal nature of speech'", *Journal of Phonetics*, 17: 87—90.
- Goldstein, L. (1989) "On the domain of the quantal theory", *Journal of Phonetics*, 17: 91—97.
- Bosch, L. F. M. & Pols, L. C. W. (1989) "On the necessity of quantal assumptions: Questions to the quantal theory", *Journal of Phonetics*, 17: 87—90.
- Blumstein, S. E. (1989) "Theoretical implications of the quantal nature of speech: a commentary", *Journal of Phonetics*, 17: 55—61.
- 鲍怀翘 中国语言学会语音学分会会员, 中国社会科学院民族学与人类学研究所研究员。研究领域: 语音学。
E-mail: hqbao@163com

从口语韵律的功能与分析看语音对立

郑秋豫

摘要 本文从韵律功能切入语音现象,从对比的角度分析解释口语韵律,以解构汉语字调、小句调及英语词重音的声学感知对比强度,并根据语流中焦点强调的分析结果,以音系之范畴对立思路试论形成对立的条件及意义。本文强调,分析语料需先就其特性,考虑分析单位及方法,厘清口语体中可预测信息及语言体系指派信息的关系,才能透视语言事实并提出正确的语音解释。

关键词 口语,韵律,韵律功能,语段韵律,韵律表达,韵律突出,相对性,对立,对比程度,范畴性

Prosodic Contrast and Categorical Difference

TSENG Chiu-yu

Abstract The paper employs a corpus analysis of speech data to address issues of phonology-and-syntax-specified canonical prosodic differentiations, and provides an acoustic account of why sufficient degree of contrast is necessary to express categorical differentiations and why categorical differentiation is more about realizing systematic discrimination instead of maintaining canonical forms. Our data of Mandarin tones demonstrate how lexically defined tone categories are realized in prosodically discriminative patterns that may deviate from the canonical forms considerably. In the case of syntactically defined sentence, we cite Mandarin evidence of why modulations of intonation contour patterns are predictable instead of random variation. In the case of lexically defined 3-way stress categories in English, our data further demonstrate how the categories are often merged into predictable binary contrasts in realized speech and why. We therefore conclude that more accurate understanding of communicative speech can be achieved from better clarification of how canonical categories are derived in speech output rather than taking canonical forms by face value.

Key words Speech, Prosody, Prosodic function, Segmental prosody, Prosodic expressions, Prominence, Relative contrast, Degree of contrast, Categories and categorical differentiation

1. 前言——语料有类分析有别

近年来由于科技的大幅进步及各式数据库的建置,取得大量的口语语料早已不是难事。口语语料形式繁多,与文本语料最大的差异在于口语中夹带了大量无法以文字表达的信息,比如依照文本朗读、一人即席演讲、二人或多人即兴对话等不同形式的语料,呈现的语音表达及语音信号中的语言信息差异极大,除了语言信息,口语信号中还夹带了许多其他的表达,如大范围的语段关系、来自语义及发话人的意图的焦点信息配置与权重等。更不用说

还有发话人的情绪、口气等。我们认为,区分口语和语言是研究语音必需的知识,无论是收集设计和分析方法都需个别考虑,不可一视同仁。研究语音如能从理解语音信号的视角出发,自然会质疑在方法上以音系单位分析口语信号是否可行,也不易对实际体现的语音信号和抽象语言单位的差异视而不见。若能从多元的口语语料修订分析视角、单位、方法,才能在更贴近口语事实的情况下,提供更全面的解释,口语韵律的研究正是一个极好的例子。

有关韵律的讨论,语音学的文献大多仍未从区别自然口语和诱发语料的角度思考,维持以音系单位描述现象的思路,因