

汉语元音对声调感知边界的影响

郑秋晨

(香港中文大学语言学与现代语言系; 香港理工大学中文及双语学系, 香港)

摘要 本研究基于现代汉语普通话声调感知具有范畴性这一结论(Wang, 1976), 以普通话中 6 个单元音 (/a//ɔ//ɤ//i//u//y/, 即拼音 a、o、e、i、u、ü)阴平(T1)到阳平(T2)调的连续统(continuum)为实验语料, 用专业语音学软件 Praat 合成连续统语音样本, 对 20 名平均年龄为 22.90 岁的普通话母语者进行声调感知实验, 即让被试对所听刺激进行 T1 或 T2 的二选一鉴别测试(identification test), 从而探究单元音的内在音高(intrinsic pitch)与声调感知边界(boundary)的关系。本文借鉴 Wang (1976)和 Peng 等(2010)对元音/i/声调连续统的合成方法, 使声调感知研究拓展到普通话中除/i/之外的其他 5 个单元音, 对以上两项研究起到一定的对照作用。研究发现, 元音会对 T1 到 T2 的声调范畴感知边界位置产生影响, 低元音/a/的内在音高较低, 其边界位置亦始终低于其他元音, 且具有显著性差异, 实验结果证明, 内在音高会对感知边界产生影响, 从而支持了 Ohala 和 Eukel (1976)的“舌位拉伸”假说。

关键词 声调; 范畴感知; 内在音高; 连续统; 元音
分类号 B842

1 前言

范畴感知(Categorical perception, CP)源于人类对连续物理属性的非连续感知, 关于这种声学上与听觉上的不匹配现象, Reetz 和 Jongman (2009)提出过如下定义: “物理上连续的量经过等分之后得到的刺激, 被感知后会归入一个个离散的范畴, 而不再是连续的物理量函数值。”早在上世纪 50 到 80 年代, 就有一些实验表明人类对某些连续物理变化的感知是离散的, 不连续的, 因为“人能把无限多的语音归纳为有限的音位范畴。归纳范畴, 就需要划分范畴的界限。”(王士元, 彭刚, 2007)物理上连续的量经过等分之后得到的便是连续统, 这种等分的方法让人类感知边界的值变得可以测量。

在经典的范畴感知实验中, 被试需参加两项测试, 分别为鉴别测试和区分测试(discrimination test), 随着声学数值的连续变化, 鉴别曲线与区分曲线若在某处剧烈变化, 此处便被界定为范畴感知的边界。其中的鉴别测试采用的是“被迫选择法”, 即要求被试对所听刺激进行有限选项内的单选操作。

由于这种实验刺激往往包含了经过均分后的有限多个声学变量, 被试会面临难以判断的混淆选项, 所以该法往往要求被试对刺激进行多次听辨, 而且随机播放刺激内容, 最后通过辨认百分率反映边界位置。由于此法较不容易受到心理物理学的影响(孔江平, 1995), 而且我们的实验基于母语者对普通话声调的感知具有范畴性这一结论而设计, 所以只采用了鉴别测试。而且, Liberman (1957)也曾提到, 在区分测试中, 被试很容易根据与实验目的无关的其他声学参数进行判断, 而并非仅仅依靠声调, 所以区分测试的实验结果并不是特别可靠。

范畴感知模型最早由 Liberman (1957)提出, 他通过实验证明人类对辅音的感知是具范畴性的, 他之后的学者则证明元音的感知是连续的(Fry, Abramson, Eimas, & Liberman, 1962), 因为元音的鉴别测试曲线平滑, 它与声学数值变化一样, 是等距变化, 所以属于非范畴感知。

早期的范畴感知研究集中于音段层面, 最早将范畴感知模型运用到超音段层面, 对声调领域进行研究的文献是 Abramson (1961)对泰语的研究以及

Wang (1976)对普通话的研究。

Abramson (1961) 以泰语(暹罗话) naa 为负载音节, 合成了从中调(mid)到升调(high)的 5 个连续统, 通过泰语母语者对以上连续统的感知实验, 作者得出了泰语声调感知有一定的范畴性这一结论。11 年之后, Abramson (1977)将自己原先的实验进行了优化, 连续统数目从 5 个增加到 16 个, 并改用泰语音节 khaa 附着声调作为刺激, 结果却显示泰语母语者的感知不存在范畴性的特质。但该次实验未设计对照组加以控制, 影响了结果的可靠性。

同是声调语言, Francis, Ciocca 和 Ng (2003)对广东话声调感知的研究却并未发现确切证据可以证明粤语母语者的声调感知具有范畴性, 因为虽然鉴别测试有较为显著的感知边界, 区分测试却无明显的区分高峰; 汉藏语系中, 范畴感知研究成果并不丰硕。孔江平早在 1995 年便对藏语声调的音高、调型和长短三方面进行过研究, 实验结果均可证明藏语(拉萨语)母语者对藏语声调的感知是具有范畴性的。

现代汉语普通话有 T1~T4 四个声调, 基频(F0)是声调最重要的声学特征, 其中 T1 是平调, T2 则为升调, 分别可以用五度标记法的 55 与 35 表示。声调与声母、韵母一样, 在语音系统中缺一不可。普通话声调具有区别意义的特性, 如“妈” (ma1)与

“麻” (ma2), 其元音、辅音与音节的构成均完全相同, 但因为声调分属 T1 与 T2 而使其代表了截然不同的意义(Wang, 1976; Peng et al., 2010)。

普通话母语者对 T1 与 T2 的感知是范畴性的(Wang, 1976), Wang 针对普通话 T1 与 T2 的感知边界进行过著名的实验研究, 实验报告在 1975 年发表, 并于次年在文章《Language Change》中进行了详细阐述, 如图 1 所示, Wang (1976)的实验设计是人工合成相邻起点相差 3 Hz、终点 F0 相同的 11 个刺激, 即第一个音节 yi2(姨)从 105 Hz 为起点上升到 135 Hz, 第二个音节的起点比第一个上升 3 Hz, 终点相同, 第三个音节的起点又比第二个上升 3 Hz, 以此类推, 第 11 个音节则是起终点频率均为 135 Hz 的平调 yi1(衣)。鉴别测试与区分测试的实验结果均显示, 普通话母语者的感知边界虽然会因个体差异而略有不同, 但其共同点在于母语者对它们的感知具有鲜明的范畴性。Peng 等 (2010), Wang, Wang 和 Peng (2012)以及王韞佳和李美京(2010)的研究虽各有侧重, 但都通过类似实验佐证了以上观点, 其中王韞佳实验的边界位置高于其他文献, 该学者推测, 这样的实验结果或可用调阶进行解释。Hallé, Chang 和 Best (2004)则选取了音节/pi//pa/和/kwo/与各种声调组合作为实验的刺激材料, 考察了台湾国语为母语者与法语为母语者对国语声调的感知, 从而提出了准范畴性

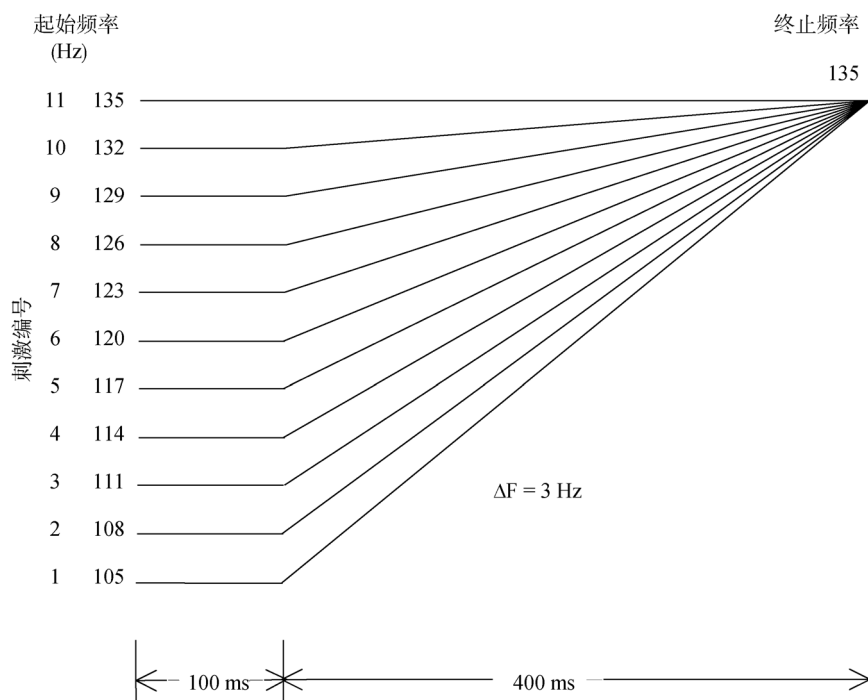


图 1 改变起点音高的升调-平调连续统基频曲线 (Wang, 1976)

(quasi-categorical)的概念。

在不同的语音系统中, 音位范畴的界限与数目均不相同(王士元, 彭刚, 2007), 但人类语言共有的这种音位概念并非生来就有。实验证明 1 岁前的婴儿对所有辅音的感知都具有一定的范畴性(Mehler & Dupoux, 1994), 而 1 岁后, 由于受到后天语言环境的影响, 一些辅音归于同一音位, 婴儿对它们之间的差异便不再敏感, 这才有了类似日语为母语者 r 与 l 不分以及汉语一些方言区 n 与 l 不分的现象。同样, 声调感知的音位界限也不可能先天就有(王士元, 彭刚, 2007), 所以后天语言环境是影响声调范畴感知的重要因素。

众多研究结果均表明, 声调的感知会受到被试语言背景的影响。Zue (1976), Hallé 等(2004)和 Peng 等(2010)都对此进行过深入的探究, 此类研究大多倾向于按语言背景将被试分为声调语言组与非声调语言组, 从而进行相应的对比研究。较多的实验选择把声调连续统附着在以爆破音为声母的完全音节(辅音+元音)上, 其中, Zue (1976)以母语为普通话者以及母语为英语者作为研究对象, 通过人工合成普通话 T2 到 T3 的连续统, 并把 9 个刺激声调分别附着在完全音节 bao 上, 对被试进行感知实验, 结果发现两组被试的感知都呈现出范畴性的特点。Hallé 等(2004)则以台湾国语母语组与法语母语组为研究对象, 他们的实验结果证明了台湾组对声调的感知是准范畴性的, 而法语组的感知则显示出更多的心理物理学特性。Peng 等(2010)不但对声调与非声调为母语者进行了考察, 还加入了对粤语母语组的考量, 由于粤语的声调体系比普通话复杂, 所以由此为切入点对他们各自的范畴感知进行探究是很有意义的。覃夕航(2012)不单证实了语言背景会对声调感知产生影响, 还通过对比普通话母语者与粤语母语者的声调感知特点, 提出了可能导致此类影响的三个因素。

不同于辅音与声调(范畴感知), 人类对元音本身的感知则是连续的(Fry et al., 1962)。元音的音色主要由舌位与唇形决定: 舌位在声学参数上的体现是共振峰, 第一、二共振峰分别与舌位的高低、前后紧密联系; 音色还会受到唇形圆展的影响(易克初, 2000)。Meyer-Eppler (1957)认为, 在耳语时, 元音的共振峰不同, 会起到区别声调的作用。现代汉语普通话中有 6 个单元音(/a//ɔ//ɤ//i//u//y/), 它们均可独立构成音节, 与不同的声调组合可表达不同的意

义, 所以以它们作为负载音节, 得到的声调属于言语性声调(speech tone), 与非言语性的声调(non-speech tone)相对立。

元音的基频与舌位高低正相关这一现象已在多种语言的研究中得到了证实, 并由此生发出元音的内在音高(IPV)理论。反之, 从感知角度来看, 内在音高则体现了听觉的非线性特点, 即元音的基频虽然不同, 但人耳无法识别, 反而认为其音高相同。Black (1949)与 Peterson 和 Barney (1952)均对美语中的[a][i][u]进行过相应的测量, 结果发现母语者在发音时, 基频值均呈现[a] < [i] < [u]的趋势。其中, Black 选取了 16 个被试, 得到的平均基频数值如下: [a] 132.7 Hz, [i] 145.7 Hz, [u] 153.0 Hz。Peterson 和 Barney 的实验结果与 Black 类似, [a][i][u]的内在音高数值分别为 124Hz, 136Hz 和 141Hz, 同样与舌位的高低关系一致。

对现代汉语普通话内在音高探究的文献相对较少, 早期文献中做得比较到位的有 Shi 和 Zhang (1987)对普通话[a][i][u]的测量, 他们将[a][i][u]的四声分男声与女声均进行了细致的 F0 值测量, 其实验结果显示, [a][i][u]的内在音高趋势与上文提及的 Black (1949)与 Peterson 和 Barney (1952)对美语的测量结果一致: 譬如同是后元音, 较高舌位的元音[u], 其基频就比低舌位的[a]高。

到目前为止, 对内在音高进行跨语言系统性探究的文献需追溯到 Whalen 和 Levitt (1995)的研究, 他们的统计结果建立在前人实验数据的基础上, 收集了含普通话在内的 31 种语言中[a][i][u]的 F0 值, 探讨了性别、舌位与内在音高之间的关系, 通过统计学上的相关测量后得出结论: 舌位高低才是影响元音内在音高的主要因素, [i]与[u]的内在音高并无显著性差异。

以往关于声调感知的研究多数选用“辅音+元音”作为负载音节, 用单元音作为负载的研究至今为止尚局限于高元音/i/, 未有研究者对其他单元音进行过系统探究, 所以单元音与感知边界之间的关系至今尚不清晰。本项实验是针对以上问题设计的, 欲对普通话不同元音的感知边界进行定量研究, 基于范畴感知理论, 利用语音分析软件 Praat 人工合成具有上升趋势的声调连续统, 运用经典的鉴别测试, 欲探究单元音与感知边界之间的关系。实验假设如下:

元音会对感知边界位置产生影响, 此影响最有可能由元音的内在音高所决定。

2 研究方法

2.1 参与者*

香港中文大学学生 20 人, 平均年龄为 22.90 岁 ($M = 22.90$, $SD = 1.86$), 其中男生 9 人, 女生 11 人, 被试均为普通话母语者, 非语言学、心理学专业学生, 无听力障碍与口、咽部疾病。实验结束后, 所有被试都得到了 50 元的被试费。被试的详细背景如表 1 所示。

2.2 实验刺激

本实验语料取自北京大学出版社“汉语拼音速成”光盘中普通话单元音的标准发音, 该组语料的声调均为阴平, 由男声朗读。本项研究刺激的制作研究模仿 Wang (1976) 的实验设计, 图 1 即是实验

刺激的基频曲线示意图, 它是呈上升趋势的 11 个连续统, 分别编为 $\alpha 1$ 号到 $\alpha 11$ 号(α 为变量, 如刺激为元音[ɤ], 则编号为 $\alpha 1$ 、 $\alpha 2$ 、 $\alpha 3$ …… $\alpha 11$ 11 个刺激。因为有 6 组单元音, 所以本实验一共将呈现 $6 \times 11 = 66$ 项不同刺激)。在以上实验设计的条件下, $i1$ 是个阳平调, 在普通话中的代表字是“姨”; $i11$ 则为阴平调, 在普通话里可以被理解为“衣服”的“衣”。图 2 呈现的则是 6 种负载元音刺激, 即音节 /a//ɔ//ɤ//i//u//y/ 在阴平状态下的宽带频谱。所有的声音刺激都是利用 PSOLA (Pitch-synchronous Overlap and ADD) 合成器合成的, 借助语音学软件 Praat (version 5.3.18 by Boersma & Weenink, 2009) 实现。

合成语音刺激的主要步骤如下:(1)截出 500 ms 的目标音节, 将声调曲线调节成平拱, 音高值设置

表 1 实验参与者语言背景简表

被试编号	性别	出生地	成长地	母语	家庭使用语言	年龄(岁)
01	男	广东梅州	安徽合肥	普通话	普通话	22
02	男	河南周口	河南周口	普通话	周口话	29
03	女	河南项城	河南项城	普通话	普通话	18
04	男	北京	北京	北京话	北京话	23
05	女	黑龙江哈尔滨	黑龙江哈尔滨	普通话	东北话	23
06	女	山西太原	山西太原	普通话	普通话	23
07	女	北京	北京	北京话	北京话	23
08	女	辽宁旅顺	辽宁大连	普通话	普通话	23
09	男	北京	北京	北京话	北京话	21
10	男	河南安阳	河南安阳	普通话	普通话&安阳话	23
11	男	山西晋城	山西晋城	普通话	晋城话&普通话	23
12	女	河北唐山	河北唐山	普通话	普通话	23
13	男	陕西西安	陕西西安	普通话	普通话	23
14	女	江苏宝应	北京	普通话	普通话	23
15	男	陕西西安	陕西西安	普通话	西安话	23
16	女	安徽淮南	安徽淮南	普通话	普通话	23
17	男	河南郑州	河南洛阳	普通话	普通话	23
18	女	山西太原	山西太原	普通话	普通话	23
19	女	福建福州	福建福州	普通话	普通话	23
20	女	黑龙江齐齐哈尔	辽宁大连	普通话	东北话	23

*匿名审稿人认为, 是次实验在被试的遴选时未严格锁定为“北京出生并长大, 母语为北京话的被试”, 在被试成长地诸多方言中, 声调差别很大, 可能是影响本次实验结果的无关变量。对此, 笔者做出回应并对被试成长地的方言类型做出简要介绍:

在被试的遴选时, 我们就严格控制了“只招募普通话母语者”, 而且, 实验开始前实验员有与被试交谈的环节, 剔除了说话有明显口音的被试后, 才签订“实验同意书”, 之后方进入正式的人机实验环节。而且被试均为大学生, 在故乡居住以及与持方言者的交谈的时间不多, 平时学习与人际交往所用的多为普通话与英文。

更重要的是, 收到审稿意见后, 笔者按被试家中使用的语言将他们分为普通话(含东北话)与非普通话组, (其中普通话组有 15 人, 非普通话组有 5 人)并进行了独立样本 t 检验, 发现除了元音 u 的感知边界稍有差异($p = 0.026 < \alpha = 0.05$), 其他元音均无显著性差异。可见被试家中使用语言对感知边界的影响不大。

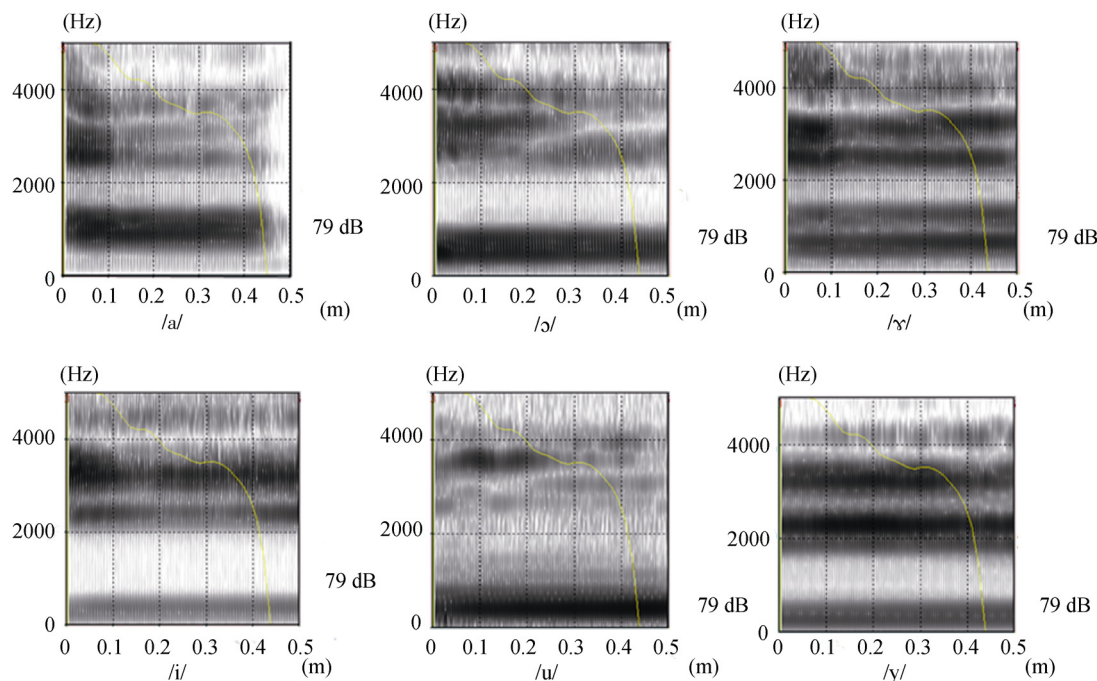


图2 六个元音 11 号刺激宽带频谱图

为 135 Hz。(2)将频率节点减少到只剩下起点处、100 ms 处以及终点处 3 点。(3)通过调整前两个点的频率值,合成出所需实验刺激,即同一元音相邻序号刺激的频率起始点均相差 3 Hz。

为了使不同元音刺激间的响度有可比性,参考相关文献的响度值参数后,我们将本实验所有刺激的响度调节到约 79 dB。

2.3 实验步骤

实验程序用软件 DMDX (version 4.1.2.2 by Forster, 2000)编写,以实现实验中刺激材料的呈现以及实验后数据的收集。为降低误按对实验的影响,每号刺激会无规律地重复播放 9 遍,每 22 个刺激间会有 1 个间隔,被试可以自行选择继续实验或是休息片刻,休息后按空格键方可继续实验。实验全程在安静的录音室进行,被试被要求坐在计算机前,眼睛距屏幕约 50 cm。实验开始后,屏幕中央先出现两个粗号的注视点“+”,持续时间为 500 ms,刺激项目与注视点同时出现。被试在听到刺激后则可以开始进行 T1、T2 二选一调类鉴别的强迫选择操作:若其认为该声音是 T1 按 A 号键盘,是 T2 按 B 号。每一个项目的呈现时间均为 1500 ms, 500 ms 的声音刺激消失后将进入 1000 ms 的间隔期,然后自动进入下一号刺激的听辨。正式实验前有一个实验前测环节:被试被要求仔细阅读实验说明,并进行含有 20 个刺激与 1 次休息的模拟测试,以便使被

试熟悉键盘,并使实验员及时得到反馈结果。为避免被试猜到实验意图,模拟测试设计了干扰项,加入了对双元音与带辅音韵尾的韵母听辨;为了使实验员更容易获取反馈结果,模拟测试中语音刺激的负载声调 70%是标准的 T1 与 T2,如果被试不能对标准读音进行正确的声调判断,实验员将进行辅助解释,并重做模拟测试。

2.4 数据处理

为了考察不同元音对鉴别测试结果的影响,在此我们将主要探究范畴感知的边界位置与边界宽度:

首先记录下每个被试对于每组元音中每号刺激按 A 键与按 B 键的次数,并计算出反应百分率。进而采用统计分析软件 SPSS (version 20.0 by IBM, 2012)中的概率分析(Probit Analysis)功能对每位被试每个元音的辨认实验反应曲线进行处理从而得到范畴感知边界与宽度。根据 Finney (1971)的感知边界计算方法,边界位置即升调与降调的听辨曲线交叉点在 x 轴上的所对应的数值;范畴感知宽度则被定义听辨反应百分率为 25%与 75%间的直线距离,这个数值由概率分析中的平均偏差与标准差所决定。

3 实验结果

3.1 鉴别曲线

6 个元音的鉴别实验曲线如图 3 所示,通过概

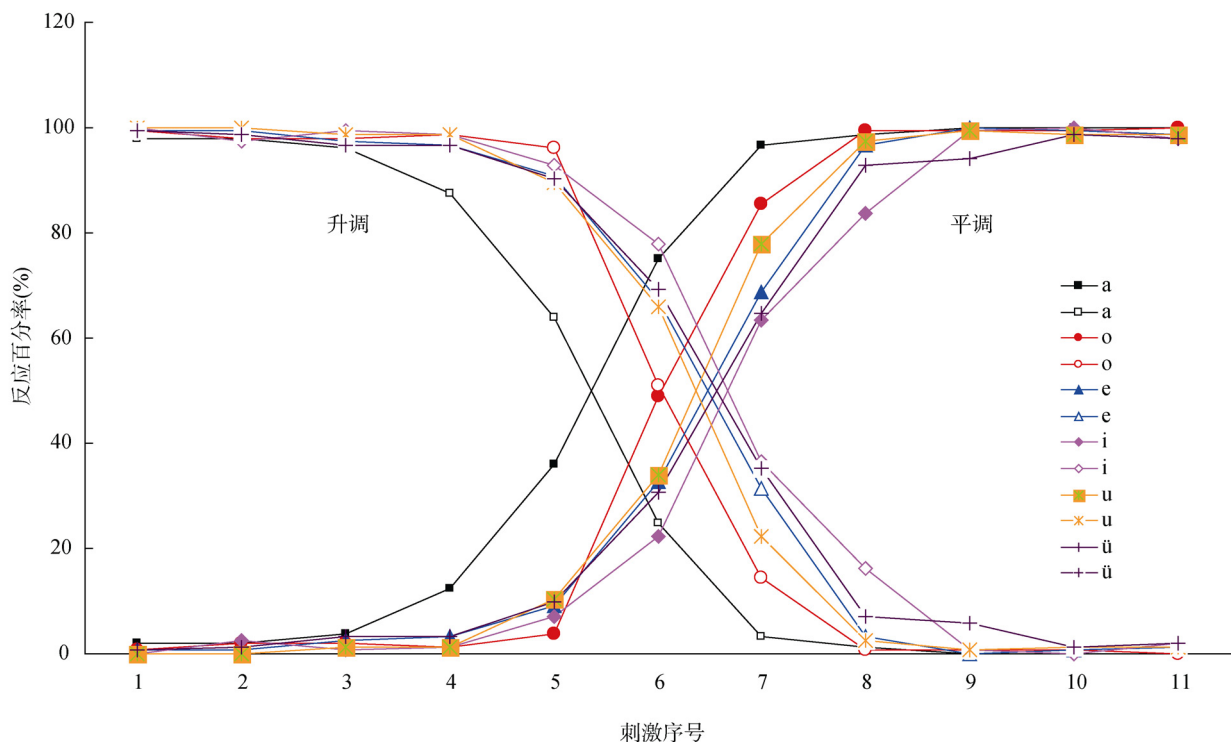


图3 汉语六个单元音的鉴别曲线

表2 六个单元音范畴感知边界位置与宽度

元音类型	边界位置	边界宽度
/a/	5.13	1.67
/ɔ/	6.02	1.55
/ɤ/	6.31	1.91
/i/	6.59	1.53
/u/	6.27	1.52
/y/	6.54	1.85

率分析而计算出的平均边界位置与边界宽度见表2。其中/i/的相对边界位置在6与7之间, 低于Wang (1976)的实验结果。

3.2 范畴边界的位置与宽度

为考察不同元音对感知边界的影响, 我们用SPSS 软件对实验反应百分率数据进行重复测量(repeated measures)方差分析, 选用格林豪斯-盖斯尔检验矫正结果, 当 $p < 0.05$ 时, 实验结果方被定义为有显著性影响。我们对20个被试进行的是6(元音类型: /a//ɔ//ɤ//i//u//y/) $\times$ 2(边界类型: 边界位置、边界宽度)两因素的重复测量方差分析。并未对反应时间进行统计分析。

从对反应率进行的两因素重复测量方差分析中我们发现:

(1)元音类型与边界类型的交互作用显著,

$F(5,95) = 8.54, p < 0.001$; 元音类型对边界位置与边界宽度二者有着非平行性的影响。其中边界位置起了主效应, $F(5,95) = 7.73, p < 0.001$; 元音对边界宽度则无显著性的影响, $F(5,95) = 1.24, p = 0.303$ 。

(2)元音类型主效应显著, /a/的边界位置最低(见表2), 与其他单元音相比, /a/的边界位置差异显著, $F(5,95) = 18.41, p < 0.001$, /ɔ//ɤ//i//u//y/的边界位置则比较接近, 差异不显著。

4 讨论

本实验考察了6组普通话单元音对范畴感知边界的影响, 实验结果基本符合实验前假设, 证实了不同元音不会对边界宽度造成影响, 却会影响边界位置。但其中只有低元音/a/的边界位置与其他元音呈现显著性差异, 其余5个单元音的边界位置虽略有差异, 却未达到统计上的显著性。笔者对此进行以下讨论:

4.1 元音内在音高与边界位置的关系

实验结果之所以显示元音对感知边界产生了影响, 其主效应来自于低元音/a/, 而除/a/外的其他元音间的边界位置均无显著性差异。/a/的边界位置始终低于其他元音且具有显著性差异, 笔者认为这一现象可用元音的内在音高进行解释。因为根据Ohala 和 Eukel (1976)在解释内在音高现象提出的

“舌位拉伸”假说, 元音的内在音高与舌位的高低呈正相关趋势, 而且 Shi 和 Zhang (1987) 的实验结果也显示, 除了在女性 T3 与 T4 的发音中, /a/ 的内在音高始终低于 /i/ 与 /u/。

内在音高依据人类的自然发音测量而得, 它与元音的发音生理状态不无关系, 而且一些一线声乐教师也发现, 学员在歌唱时, 发元音 /i/ 时, 声带亦趋向于紧绷, 从而容易影响音准。而发 /a/ 时, 声带则较松弛。从发音的角度上来看, 人类发 /a/ 音时 F0 普遍较低, 长期的语言输入经验极有可能影响普通话母语者对声调的感知, 从而使范畴感知边界也随之整体下移。

除了 /a/, 其他元音的感知边界位置均未达到显著性差异, 原因或可有如下解释: 跨语言的研究结果 (Whalen & Levitt, 1995) 显示, /i/ 与 /u/ 的内在音高并无显著性差异, 所以二位作者认为, 舌位前后对于内在音高而言, 非关键性因素, 舌位的上下, 方能影响元音的内在音高, 所以 /a/ /i/ /u/ 内在音高差异的主效应来源于 /a/ 这一实验结果便不难解释: 在本实验测试的 6 个元音中, /a/ 的舌位最低, 范畴感知边界也随之较低且与其他各元音均存在显著性差异, 而 /i/ /u/ /y/ 均为高元音, 故在感知边界方面亦无显著性差异, 所以实验现象用内在音高理论解释合情合理。

然而, 从图 3 可知, 元音 /ɔ/ 的感知边界略低于 /a/, 并与其他元音有一定距离, 但并未显示出统计上的显著性差异。笔者在此从语言输入频率角度试做以下推测: 虽同为单元音, 且均有词汇意义, /ɔ/ 在普通话中与其他 5 个单元音的出现频率却有比较大的差异。因为拼音 o 只在零声母环境下读作 /ɔ/, 在唇音声母后则读成一个双元音。从音系学角度看, /ɔ/ 是否应该归为独立的音位在学术界尚未达成共识, 普通话母语者对它的感知频率远远不及其他单元音, 而且本项实验的被试均为大学生, 有二语学习的背景, 此处不可排除以下可能性: 被试在实验中并非凭借普通话的经验去感知元音 /ɔ/, 而凭借的是对英语元音的感知经验。

此外, 本项研究要讨论不同元音对边界的影响, 便无法对共振峰因素进行有效控制, 所以从本项实验结果所显示出的边界位置差异并无法排除与不同元音的共振峰模式有关。譬如舌位相距最远的元音 /i/ 与元音 /a/, /a/ 的第一第二共振峰间的距离比 /i/ 近, 其中, /a/ 的第一共振峰数值高于 /i/, 第二共振峰却与之相反。

4.2 与前人研究结果的比较

本项实验结果表明, 元音 /a/ 的边界位置与其他元音有显著性差异, 而边界宽度则无, 这与 Hallé 等 (2004) 统计结果的显著性分布不尽相同。Hallé 等 (2004) 以 /pi/ /pa/ 为负载音节, 将合成好的连续统嵌入句子中让被试进行听辨。其实验结果显示 /pi/ 与 /pa/ 的感知边界位置无显著性差异, 而边界宽度却有, 具体体现为: /pa/ 的边界宽度陡于 /pi/。

导致显著性分布不同的原因笔者认为可以从两方面进行讨论:

① Hallé 等 (2004) 探究的是完全音节 V+C 的边界位置与宽度, 而本次实验仅仅着眼于单元音, 所以该差异很有可能是由 /i/ 与 /a/ 之前的辅音作用。

② Hallé 等 (2004) 的听辨实验中有句子作为负载, 被试很有可能受前后音节的影响。而且, 李幸河和石锋 (2010)、薛鑫和石锋 (2010) 和荣蓉和石锋 (2013) 的实验中均证明, 双音节中的前后字会对范畴感知听感界限造成影响, 所以笔者推测目标音节位置也会对阴平与阳平的感知边界造成相似的影响。

本项实验结果始终指向低元音 /a/ 具有较低的感知边界, 并与其他单元音的感知边界有显著性差异, 然而在高云峰 (2004) 的博士论文中, 曾对 /i/ /u/ /a/ 三个元音阳平到阴平的连续统进行过探究, 实验设计与本实验大同小异, 但是高的实验每号刺激之间相差 0.4 半音 (semi-tone), 其结果表明, /i/ 的边界位置最低, 在第 7 和第 8 号刺激之间, 而 /u/ /a/ 的边界位置略高, 均在第 8 和第 9 号刺激之间, 与内在音高并无正比关系。但是, 作者主要考量的是斜率值, 未对感知边界的显著性进行测量, 而且值得一提的是, 作者只选取了两名被试 (男女各一人), 在统计上是否有效, 这一点值得怀疑。

5 结论

本项研究着眼于探究不同单元音与声调范畴感知边界的关系, 通过经典的声调感知鉴别实验, 让 20 名被试对普通话 6 个单元音 T1 到 T2 的连续统进行听辨, 对实验结果进行统计分析后我们清楚地知道: 不同元音对感知边界会带来显著性的影响, 该影响主要来源于低元音 /a/, 实验结果或可用元音内在音高进行解释, 并受到共振峰模式的影响。

本文未能用统计方法有效分析反应时间的差异, 且由于实验地点的客观条件限制, 对被试的语言背景未控制到“北京出生并长大, 并以北京话为

母语”的理想状态, 对被试间差异(如性别差异)的考察也较为薄弱, 另外, 元音/ɔ/为何在本实验中未见显著性差异, 其原因尚不明晰, 今后的研究或可增加被试数量, 对本次实验进行有效修正。以上是本次研究留下的遗憾, 但同时也为后续的研究埋下了伏笔。今后的研究可以对不同语言背景或音乐能力的被试进行对比研究, 本实验未对非言语性声调的感知边界进行探究, 可以考虑以此为切入点, 探究声调与音乐之间的关系。

致谢: 本文是作者在香港中文大学攻读硕士学位的毕业论文, 蒙硕导彭刚教授一次次地耐心指正, 谨致谢忱。同时感谢匿名审稿人的宝贵意见, 还有参加实验全体同学的鼎力支持。文中错误一概由作者负责。

参 考 文 献

- Abramson, A. S. (1961). Identification and discrimination of phonemic tones. *The Journal of the Acoustical Society of America*, 33, 842.
- Abramson, A. S. (1977). The noncategorical perception of tone categories in Thai. *The Journal of the Acoustical Society of America*, 61, 66.
- Black, J. W. (1949). Nature frequency, duration, and intensity of vowels in readings. *Journal of Speech and Hearing Disorders*, 14, 216–221.
- Finney, D. J. (1971). *Probit analysis* (3rd ed). Cambridge, UK: Cambridge University Press.
- Francis, A. L., Ciocca, V., & Ng, B. K. C. (2003). On the (non)categorical perception of lexical tones. *Perception & Psychophysics*, 65, 1029–1044.
- Fry, D. B., Abramson, A. S., Eimas, P. D., & Liberman, A. M. (1962). The identification and discrimination of synthetic vowels. *Language and Speech*, 5, 179–189.
- Gao, Y. F. (2004). *The research of tonal perception* (Unpublished doctoral dissertation). Shanghai Normal University.
- [高峰. (2004). 声调感知研究 (博士学位论文). 上海师范大学.]
- Hallé, P. A., Chang, Y. C., & Best, T. C. (2004). Identification and discrimination of mandarin Chinese tones by mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32, 395–421.
- Kong, J. P. (1995). The research of Tibetan tonal perception. *Minority Languages of China*, (3), 56–64.
- [孔江平. (1995). 藏语(拉萨话)声调感知研究. 民族语文, (3), 56–64.]
- Li, X. H., & Shi, F. (2010). *The perception boundary of tone 1 and tone 2 in mandarin*. Paper presented at the Symposium of 9th Phonetics Conference of China. Tianjin, China.
- [李幸河, 石锋. (2010). 普通话阴平与阳平之间的听感分界. 发表于第九届中国语音学学术会议论文集. 天津, 中国.]
- Liberman, A. M. (1957). Some results of research on speech perception. *Journal of the Acoustical Society of America*, 29, 117–123.
- Mehler, J., & Dupoux, E. (1994). *What infants know: The new cognitive science of early development*. Oxford, UK: Blackwell.
- Meyer-Eppler, W. (1957). Realization of prosodic features in whispered speech. *The Journal of the Acoustical Society of America*, 29, 760–760.
- Ohala, J. J., & Eukel, B. W. (1976). Explaining the intrinsic pitch of vowels. *Journal of the Acoustical Society of America*, 60, 207–215.
- Peng, G., Zheng, H. Y., Gong, T., Yang, R. X., Kong, J. P., & Wang, S. Y. (2010). The influence of language experience on categorical perception of pitch contours. *Journal of Phonetics*, 38, 616–624.
- Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, 24, 175–184.
- Qin, X. H. (2012). *Role of mother language experience in categorical perception of mandarin* (Unpublished master's thesis). Beijing University.
- [覃夕航. (2012). 母语经验对汉语普通话声调范畴化感知的影响 (硕士学位论文). 北京大学.]
- Reetz, H., & Jongman, A. (2009). *Phonetics: transcription, production, acoustics, and perception*. Chichester, UK: Wiley-Blackwell.
- Rong, R., & Shi, F. (2013). Perception of the boundary between mandarin tone 1 and tone 3: Effects of pitch and duration. *Linguistic Sciences*, 12, 17–26.
- [荣蓉, 石锋. (2013). 音高和时长对普通话阴平和上声的听感影响. 语言科学, 12, 17–26.]
- Shi, B., & Zhang, J. (1987). *Vowel intrinsic pitch in standard Chinese*. Paper presented at the 11th International Congress of Phonetic Science. Tallinn, Estonia.
- Wang, D. Z., Wang, X. X., & Peng, G. (2012). *Effects of carriers on mandarin tone categorical perception*. Paper presented at the 8th International Symposium on Chinese Spoken Language Processing. Kowloon, Hong Kong.
- Wang, W. S. (1976). Language change. *Annals of the New York Academy Sciences*, 208, 61–72.
- Wang, S. Y., & Peng, G. (2007). *Language, phonetics, and speech technology*. Hong Kong, China: City University of Hong Kong Press.
- [王士元, 彭刚. (2007). 语言, 语音与技术. 香港: 香港城市大学出版社.]
- Wang, Y. J., & Li, M. J. (2010). The effects of tone pattern and register in perceptions of tone 2 and tone 3 in mandarin. *Acta Psychologica Sinica*, 42, 899–908.
- [王韞佳, 李美京. (2010). 调型和调阶对阳平和上声知觉的作用. 心理学报, 42, 899–908.]
- Whalen, D. H., & Levitt, A. G. (1995). The universality of intrinsic F0 of vowels. *Journal of Phonetics*, 23, 349–366.
- Xue, X., & Shi, F. (2010). *The perception boundary of tone 1 and tone 4 in mandarin*. Paper presented at the Symposium of 9th Phonetics Conference of China, Tianjin, China.
- [薛鑫, 石锋. (2010). 普通话阴平和去声之间的听感分界. 发表于第九届中国语音学学术会议论文集. 天津, 中国.]
- Yi, K. C. (Ed). (2000). *Speech signal processing*. Beijing: National Defence Industry Press.
- [易克初(主编). (2000). 语音信号处理. 北京: 国防工业出版社.]
- Zue, V. W. (1976). *Acoustic characteristics of stop consonants: A controlled study* (Unpublished doctoral dissertation). Massachusetts Institute of Technology.

Effects of Vowels on Mandarin Tone Categorical Perception

ZHENG Qiuchen

(Department of Linguistics and Modern Languages, the Chinese University of Hong Kong;

Department of Chinese and Bilingual Studies, Hong Kong Polytechnic University, Hong Kong, China)

Abstract

Sounds are physically continuous but linguistic phonemes are discrete and limited. Categorical Perception (CP) offers a useful approach to address this mismatching issue between the physical and perceptual features of speech. It is well established that consonants are perceived categorically, while vowels are perceived continuously. Earlier studies showed that the speech CP was restricted to segmental features. However, as Abramson (1961) firstly pointed out that CP mode could be applicable to the suprasegments. His limited investigation revealed that tone perception in Thai is categorical. Likewise, a similar result has been confirmed by Wang (1976) that the perception of Mandarin tones is also categorical, in the sense that native Chinese speakers have a linguistic boundary when perceiving the tones. The studies was based on a synthesized acoustic continuum, which was superimposed on vowel /i/, consisting of 11 tonal variants from Mandarin Tone 2 (yi2 'aunt') to Tone 1 (yi1 'clothing'). There is still no conclusion that what effect vowels have on the boundary of tonal perception. Regarding this, current study is to address this issue.

In this research, tests of tonal CP are extended from /i/ (Wang 1976; Peng et al. 2010) to six Mandarin monophthongs (/a/ɔ/ɤ/i/u/y/) to investigate the effects of vowels' intrinsic pitch on tonal boundary. Firstly, a continuum with a rising tendency was synthesized from Mandarin Tone 2 to Tone 1 using the acoustic software Praat. This continuum was superimposed on six Mandarin monophthongs, which were then used as stimulus in the identification test. Secondly, the DMDX software was used to randomly and with nine repetitions present the stimuli to the subjects in the identification test. Finally, the participants, twenty native Mandarin speakers (nine male and eleven female), were asked to identify the stimulus as either Tone 1 or Tone 2 by pressing different button on the computer.

The mean category boundary position and width in terms of stimulus number were obtained by Probit Analysis of individual identification curves.

To investigate the effects of different monophthongs on categorical boundary position and width, our results were analyzed using repeated measures with SPSS software. The results show that the categorical boundary position differs significantly between Mandarin monophthongs ($p < 0.001$): the boundary position of the tone continuum of vowel /a/ (stimulus step number 5.13) is smaller than that of tone continuum of other vowels (stimulus step numbers of /ɔ/ɤ/i/u/y/ are 6.02; 6.31; 6.59; 6.27 and 6.54 respectively). The finding that boundary position of low vowel /a/, whose intrinsic pitch is low, is consistently lower than the five other monophthongs ($p < 0.001$) supports the validity of the "tongue-pull" hypothesis (Ohala & Eukel 1976: 44). Categorical boundary width across the six vowels was not significantly different ($p = 0.303$).

To sum up, in this study, the influence of different monophthongs on the perception of pitch contours in the framework of CP was examined. The results of this experiment indicate that 1) identification of Tone 1 and Tone 2 is clearly categorical; 2) stimuli of lower vowel tends to require steeper F0 contour slope for subjects to make a "Tone 2" response than the stimuli of higher vowels did; 3) the intrinsic pitch of vowels plays an important role in identification of Tone 1 and Tone 2.

Key words tone; categorical perception; intrinsic pitch; continuum; vowels