

黔东苗语新合话三升调的声学与感知研究*

刘文

[摘要] 声调范畴化感知是语音学和认知神经科学等领域关注的焦点。前人研究指出，调型相异的声调间的感知模式通常是范畴感知，调型相同的声调间的感知模式则是连续感知，然而有关声调范畴感知和连续感知的成因一直没有得到解决。本文以黔东苗语新合话的三个升调为例探讨这一问题。研究发现，无论调型异同，当声调间的基频斜率存在显著性差异时，感知结果就会表现出范畴感知的特性；当声调间的基频斜率无显著性差异时，感知结果就会表现出连续感知的特性。这一发现解决了声调感知理论中的一个根本问题，即基频斜率是导致声调范畴感知和连续感知区分的一项决定性因素。在此基础上，本文提出声调范畴化感知是个连续统的观点。

[关键词] 声调感知 范畴感知 黔东苗语新合话 三升调 基频斜率

一 引言

语音范畴化感知指的是听者能够将连续性的语音刺激离散为数量有限的范畴的能力，听者在不同的音位范畴中可以感知声学连续统上的细微变化，但这种细微差异在同一个音位范畴中就不能被感知出来 (Liberman 1957; Liberman et al. 1957)。经典范畴感知实验包含辨认（判断听到的一个声音属于哪个音位范畴）和区分（判断听到的两个声音的异同）两项任务。通常来说，如果听者对某个语音单位的感知是范畴性的，那么辨认和区分任务应该满足以下条件：一是辨认函数在连续统上有一个比较陡峭的边界；二是区分函数在范畴边界处有区分峰值，但在范畴内区分正确率则处于随机水平；三是区分函数可以通过辨认函数预测出来 (Studdert-Kennedy et al. 1970; Repp 1984)。据此，语音范畴化感知有范畴感知和连续感知两种常见模式 (Liberman et al. 1967)。

在语言学研究中，音段层面的塞音的感知属于范畴感知，如 /b/-/d/-/g/ (Liberman et al. 1957) 和 /d/-/t/ (Liberman et al. 1961)，而单元音之间的感知则是连续感知，如 /i/-/ɛ/-/æ/ (Fry et al. 1962)，单元音与复合元音之间通常会呈现出范畴感知特征 (陈飞等 2019; 王璐、孔江平 2021)。在超音段层面，声调的感知研究受到学界广泛关注，尤其是声调是否属于范畴感知成为言语感知研究领域的一个核心问题。前人研究表明，当声调间的调型不同时，感知结果通常呈现为范畴感知。例如，汉语普通话的高平调和高升调 (Chan et al. 1975; Wang 1976; Hallé et al. 2004; Xu et al. 2006; 杨若晓 2009; 席洁等 2009; Peng et al. 2010; 王韫佳、李

* 本文得到国家社会科学基金重大项目“中华民族语言文字接触交融研究(22&ZD213)”和一般项目“民族语的声调类型研究(20BYY178)”资助。感谢《民族语文》匿名审稿专家惠赐宝贵意见。

美京 2010; 张林军 2010a, 2010b; 覃夕航 2012; 刘思维 2015; 于谦 2017; Si et al. 2017)、高平调和高降调(杨若晓 2009; 覃夕航 2012; 刘思维 2015; Shen 2015; 于谦 2017; Si et al. 2017)、高升调和高降调(杨若晓 2009; Xi et al. 2010; 于谦 2017); 藏语拉萨话的声调(孔江平 1995); 粤语的高平调和高升调、低降调和高升调(Francis et al. 2003); 上海话中不同调型的声调(高云峰 2004); 台湾闽南话的高平调和高降调(Sun & Huang 2012); 合肥话的相异调型(蔡雯清 2016); 黔东苗语新寨话的高平调和高升调(刘文 2020, 2021:227)。然而, 当声调间的调型相同时, 感知结果通常呈现出连续感知的特点。例如, 泰语的三平调(Abramson 1979); 紫云苗语的五平调(孔江平 1992); 粤语的低平调和高平调(Francis et al. 2003); 台湾闽南话的高平调和中平调(Sun & Huang 2012); 黔东苗语新寨话的五平调(Liu 2020; 刘文 2021:213); 粤语的低升调和高升调(Francis et al. 2003); 上海话的两个升调(高云峰 2004)。也有研究就相同调型的声调得到了范畴感知的结果。例如, 广州话和东海话的平调(金健 2010)、谷饶方言的降调(金健、施其生 2010)和杭州方言的升调(金健 2015)。除范畴感知和连续感知, Hallé et al. (2004) 提出准范畴感知的概念, 认为范畴感知的程度存在一个梯度, 梯度等级则取决于听者对声调的熟悉程度。在汉语普通话中, 王韫佳、李美京(2010) 和王韫佳、覃夕航(2015) 发现阳平和上声的感知模式介于范畴感知和非范畴感知之间。

同为声调, 为什么调型异同会呈现截然不同的感知结果? 这种表现背后的成因是什么? 前人研究给出了造成声调感知范畴化程度不同的假设。例如, 调型假设(Francis et al. 2003; Xi et al. 2010)和声调音高曲拱相似性假设(王韫佳、覃夕航 2015), 详见吴倩、王韫佳(2018)和王韫佳等(2021)的评述。我们需要进一步追问的是: 即使是调型相同或相似的声调, 它们之间是否全部表现为连续感知? 以往研究多聚焦于不同调型的声调, 较少关注同型声调。需要说明的是, 异型声调指的是音高运动轨迹不同的声调(如平调和升调、平调和降调、升调和降调等), 同型声调指的是音高运动轨迹相同的声调(如平调和平调、降调和降调、升调和升调)。需注意的是, 同型声调的斜率并非完全相同。考虑到调型异同可以导致完全不同的感知结果, 这就使得深入挖掘同型声调感知的模式及其背后的成因变得十分必要。

本文以苗语新合话为例, 探究声调范畴化感知的模式及其成因。新合位于贵州省黔东南苗族侗族自治州剑河县观么镇, 该村居民全部为苗族, 通行苗语。新合话属于苗语黔东方言(王辅世 1983, 1985:103)。根据田野调查^①, 黔东苗语新合话有 32 个声母、18 个韵母和 8 个声调。表 1 中的调类 T1—T8 是基于原始苗语的构拟(王辅世 1994:3), 调值采用“五度标调法”(Chao 1930) 描写。

表 1 黔东苗语新合话的声调系统

调类	T1	T2	T3	T4	T5	T6	T7	T8
	中平	高降	高升	中低平	中升	低平	低升	中降
调值	33	53	35	22	24	11	13	31

本文选择黔东苗语新合话的原因在于它有三个升调, 这在声调类型中是比较罕见的。下

^① 2017 年 9 月, 笔者第一次调查黔东苗语新合话。2018 年 5 月进行补充调查和核对工作, 并完成了语音数据采录和感知实验。

文首先通过声学分析探讨新合话声调声学分布空间以及三升调对立的声学基础（第2部分），进而设计行为学感知实验考察三个升调间的感知模式（第3部分），最后结合发音和感知的研究来解释声调范畴化感知模式的形成原因（第4部分）。

二 声调的声学研究

（一）方法

基于田野调查，本研究选取了86个常用单音节语素作为录音样本（见附录）。为避免协同发音，所选语素声母要求为零声母或不送气塞音声母及其腭化形式（/p、t、t^j、k、k^j/），韵母为单元音韵母（/i、e、ɛ、æ、a、o、u/）和复合元音韵母（/ei、eɪ、ai、ao、əu、ou、ui/）。

声学实验录制了六位母语发音人（三男三女，平均年龄19岁，标准差为2.68）。录音在剑河县一个安静的房间内进行，录音软件为Audition 2.0，采样频率为44.1 kHz，精度是16 bit。录音要求被试以一种自然和舒适的方式朗读词表，每个词重复两遍，间隔为两秒。为避免语流中的声调协同发音，所有词均要求单念。录音过程中，邀请调查时的主要发音合作人监听被试发音的准确性和自然流畅度。

在参数选取和数据分析方面，本文选取了与研究目的密切相关的声学和嗓音频谱参数，即基频起点与终点、基频斜率（[基频终点-基频起点]/时长）、时长、H1*-H2*（第一谐波与第二谐波的振幅差值）、H1*-A1*（第一谐波与第一共振峰位置处谐波的振幅差值）、H1*-A2*（第一谐波与第二共振峰位置处谐波的振幅差值）、H1*-A3*（第一谐波与第三共振峰位置处谐波的振幅差值）、CPP（cepstral peak prominence，倒频谱峰值凸显）。“*”表示采用元音共振峰校正的形式（Iseli et al. 2007）。在跨语言研究中，H1-H2被认为是测量发声类型最有效的参数（Gordon & Ladefoged 2001），它与声门周期内的开相时长相关，开相时间越长，H1-H2就越大（Holmberg et al. 1995）。H1-A1、H1-A2、H1-A3反映频谱中高频能量的强度，它们被认为与声门关闭的速率有关（Stevens 1977）。这些谐波振幅参数的数值越大，说明整个频谱的能量衰减得越快，语音信号越倾向于气嗓性发声；相反，数值越小，说明整个频谱（尤其是高频）的能量越强，语音信号越倾向于挤喉音或紧嗓音。CPP用来测量信号的周期性和信噪比，它被认为是判定气嗓音的一个可靠指标（Hillenbrand et al. 1994）。CPP越小，说明信号的噪音成分越多；CPP越大，说明信号的信噪比越高，即信号的周期性就越好。以上参数通过VoiceSauce（Shue et al. 2011）从语音信号中自动提取，每个元音被分为等长的11份。

（二）结果

尽管男、女声带生理上的差异导致他们的基频绝对值存在差异，但是新合话不同性别发音人的声调音高曲线与时长分布上的表现模式一致（刘文 2019:74）。囿于篇幅，下文将三名男性和三名女性的数据合并作图（图1）。图中对基频的平均处理只是为了展示的需要，后续分析所使用的是经过归一化处理后的原始基频数据。黔东苗语新合话的八个声调根据调型可以分为三类：一是平调（中平调T1、中低平调T4、低平调T6）；二是降调（高降调T2、中降调T8）；三是升调（高升调T3、中升调T5、低升调T7）。时长方面，黔东苗语新合话声调的时长由长到短依次为T7、T6、T4、T5、T1、T8、T2、T3。以声调作为因子的方差分析显示声调在时长上存在显著性主效应（($F(7, 1014) = 26.848, p < 0.001$)），采用多重假设检验校正方法Bonferroni校正的事后成对比较发现T7显著长于其他声调（ $p < 0.05$ ），T1、T4、T5

和 T6 两两之间无显著性差异($p > 0.05$)，T2、T3 和 T8 两两之间也无显著性差异($p > 0.05$)。基于时长统计分析的结果可以将这八个声调分为三组：一是 T7；二是 T1、T4、T5、T6；三是 T2、T3、T8。从发音的视角看，黔东苗语新合话的低升调占时最长，其次是平调和中升调，降调和高升调占时最短。

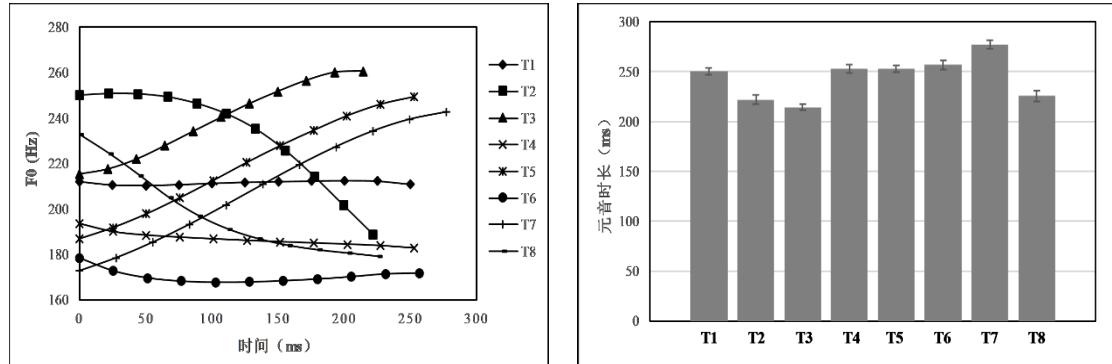


图 1 黔东苗语新合话声调曲线及时长均值(含标准误差)

同一种语言含有三个最小对立升调的现象在声调类型中是比较罕见的，在上文分析基础之上，我们接下来着重考察这三个升调的声学对立基础。鉴于单独考察男、女发音人参数的结果与将男、女发音人参数汇总的结果表现一致，在此我们只报道男、女发音人参数平均后的数据（表 2）。

表 2 三升调的参数均值与标准误差

声学参数	T3 (35)	T5 (24)	T7 (13)
基频起点 (Hz)	215.61 (5.44)	184.86 (3.84)	173.03 (4.47)
基频终点 (Hz)	260.67 (6.42)	247.36 (5.17)	238.53 (5.91)
基频斜率	0.22 (0.0069)	0.25 (0.0073)	0.24 (0.0082)
时长 (ms)	214.31 (3.18)	252.76 (3.12)	277.10 (4.08)
H1*-H2* (dB)	5.35 (0.40)	5.40 (0.23)	5.89 (0.32)
H1*-A1* (dB)	17.15 (0.49)	17.50 (0.32)	20.03 (0.64)
H1*-A2* (dB)	16.33 (0.61)	17.58 (0.49)	19.24 (0.75)
H1*-A3* (dB)	9.25 (0.84)	9.91 (0.64)	9.44 (0.98)
CPP (dB)	23.12 (0.15)	22.40 (0.11)	21.00 (0.17)

注：表中为不分性别的汇总数据。除斜率的标准误差外，其他参数均保留小数点后两位。

为了确定黔东苗语新合话三个升调的声学差异，本部分使用 R 语言（R Core Team 2019）的 lmerTest 包（Kuznetsova et al. 2017）对表 2 中的各项参数进行了一系列线性混合响应分析。数据预处理过程先对各个参数的原始数据进行归一化处理，使得到的用于后续统计的新数据的范围在 [-1, 0] 或 [0, 1] 之间。在所构建的线性混合响应模型中，自变量分别为基频起点、基频终点、基频斜率、时长、H1*-H2*、H1*-A1*、H1*-A2*、H1*-A3* 和 CPP，调类和性别作为固定因子，说话者为随机截距。统计分析发现新合话三个升调在基频起点 ($F(2, 396) =$

894.409, $p < 0.001$)、终点 ($F(2, 396) = 89.923, p < 0.001$)、斜率 ($F(2, 396) = 24.144, p < 0.001$)、时长 ($F(2, 396) = 171.354, p < 0.001$)、H1*-A1* ($F(2, 396) = 18.714, p < 0.001$)、H1*-A2* ($F(2, 396) = 11.068, p < 0.001$) 和 CPP ($F(2, 396) = 112.317, p < 0.001$) 上都存在显著性的主效应，而在 H1*-H2* ($F(2, 396) = 1.144, p = 0.32$) 和 H1*-A3* ($F(2, 396) = 0.675, p = 0.51$) 上则没有显著性的主效应。采用多重假设检验校正方法 Bonferroni 校正的事后成对比较来确定哪些声调之间存在显著性的不同（见表 3）。

表 3 三升调两两间的统计结果

声学参数	T3 (35) vs. T5 (24)	T3 (35) vs. T7 (13)	T5 (24) vs. T7 (13)
基频起点 (Hz)	$p = 0.184$	$p < 0.001$	$p < 0.001$
基频终点 (Hz)	$p = 0.824$	$p = 0.034$	$p = 0.295$
基频斜率	$p = 0.698$	$p = 0.067$	$p < 0.001$
时长 (ms)	$p < 0.001$	$p < 0.001$	$p < 0.001$
H1*-A1* (dB)	$p < 0.001$	$p < 0.001$	$p = 1.000$
H1*-A2* (dB)	$p = 0.151$	$p = 0.005$	$p = 0.409$
CPP (dB)	$p < 0.001$	$p < 0.001$	$p < 0.001$

注：表中数据保留小数点后三位，显著性水平 $p = 0.05$ 。

具体来说，在时长和 CPP 上，三个升调两两之间都存在显著性差异 ($p < 0.001$)。在基频斜率上，T5 (24) 和 T7 (13) 之间存在显著性差异 ($p < 0.001$)，T3 (35) 和 T7 (13) 之间存在边缘性显著差异 ($p = 0.067$)，而 T3 (35) 和 T5 (24) 之间则不存在显著性差异 ($p = 0.698$)。需要注意的是，T3 (35) 与 T7 (13) 之间的斜率差异虽然未达到 0.05 的显著水平，但是简单效应分析可以获得一个显著的效应，在这种情形下我们可以说这二者之间存在边缘性显著差异，否则就是不显著，文献中有关边缘性显著差异可接受的 p 值范围通常为 0.05 ~ 0.10 (Pritschet et al. 2016)。此外，T3 (35) 和 T7 (13) 在 H1*-A1* ($p < 0.001$) 和 H1*-A2* ($p = 0.005$) 上也存在显著性差异。上述统计结果似乎表明在三升调的区分中，除基频外，嗓音质量也可能起到一定的贡献。

为了更加直观地展现三个升调的声学对立特征，本部分采用多维尺度分析方法，分别计算了由声学参数和声学、噪音参数所共同构建的发音空间（图 2）。在由声学参数构建的发音空间中，每个声调由基频起点、终点、斜率和时长四个参数表达；在由声学和噪音参数共同构建的发音空间中，每个声调由基频起点、终点、斜率、时长、H1*-H2*、H1*-A1*、H1*-A2*、H1*-A3* 和 CPP 九个参数表达。多维尺度分析方法的功能是在一个高维度空间内计算不同声调间的距离，并将这个距离投射到一个低维度、可解释的空间中。由于图 2 是基于母语者的发音数据计算而来的，所以这个空间实质上是一个发音空间，而非感知空间。

对图 2 的直观解释是空间中声调间的距离越远，声调间的对立性就越充分。统计模型显示，仅基于声学参数的发音空间可以解释 99.95% 的变异，基于声学和噪音参数的发音空间可以解释 99.99% 的变异。对比这两个发音空间，不难发现三个升调在这两个空间中的分布表现一致。由此可见，相对于仅有声学参数所构建的发音空间而言，噪音参数的加入并没有改变这三个升调在声学空间中的分布，且仅解释了 0.04% 的变异，这就说明噪音发声类型对这三

个升调的区分不起作用。另外，多维尺度分析显示，在这些声调的声学分布中，基频均值和斜率是最主要的区分因素。除此之外，时长也起到一定的作用。图2中的二维空间可以根据音高和时长特征分为三个部分：低音高、时长长（13），中音高、时长中长（24），高音高、时长短（35）；维度1将高升调（35）和非高升调（13、24）区分开，维度2则将中升调（24）和处于极端位置的两个升调（13、35）区分开。

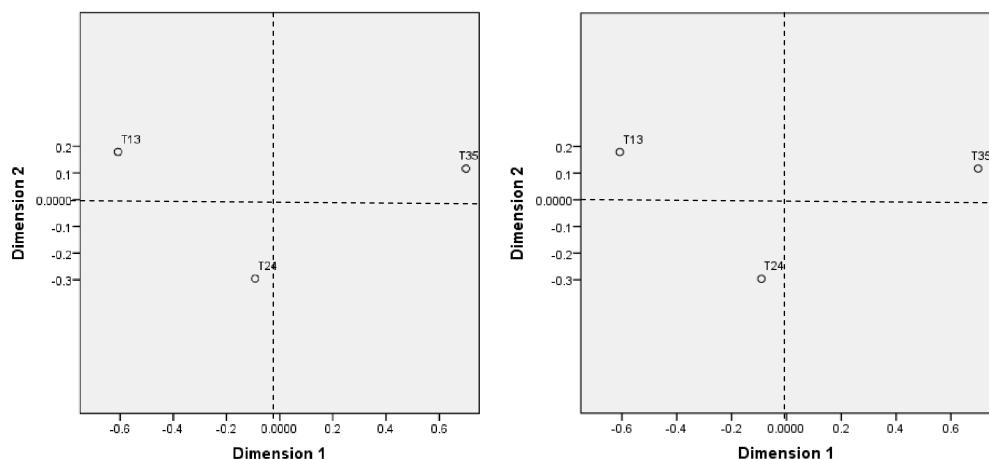


图2 基于多维尺度分析的三升调的发音空间（左：声学参数；右：声学和噪音参数）

综上，从言语产生的视角看，声学分析和多维尺度分析的结果显示黔东苗语新合话三个升调的区别要素有基频、斜率和时长。具体来说，T5（24）和T7（13）、T3（35）和T7（13）这两对声调在斜率上存在显著性差异 ($p < 0.001$) 或边缘显著性差异 ($p = 0.067$)，而T3（35）和T5（24）在斜率上则没有显著性差异 ($p = 0.698$)。在时长方面，三个声调对之间均存在显著性差异 ($p < 0.001$)。

三 三升调的感知研究

（一）材料

本部分基于黔东苗语新合话三升调最小对立的单音节语素（“翅膀”[ta¹³]；“杀”[ta²⁴]；“手指”[ta³⁵]）构建了三组声调连续统（T13-T24、T13-T35、T24-T35）。三个语素的原始样本采用的是调查时主要发音人的语音，选择这三个单音节语素的原因在于它们都是日常交流中的高频词，并且参与实验的母语者都能轻松自如地发出这三个语素。

表4 三升调原始样本的11个基频(Hz)和时长(ms)

调值	1	2	3	4	5	6	7	8	9	10	11	时长
T13	116.7	121.1	125.6	130.1	134.6	139.0	143.5	148.0	152.4	156.9	161.4	358
T24	145.4	152.4	159.4	166.4	173.4	180.5	187.5	194.5	201.5	208.5	215.5	348
T35	185.3	190.5	195.6	200.7	205.9	211.0	216.1	221.2	226.4	231.5	236.6	323

注：限于篇幅，基频值保留小数点后1位。

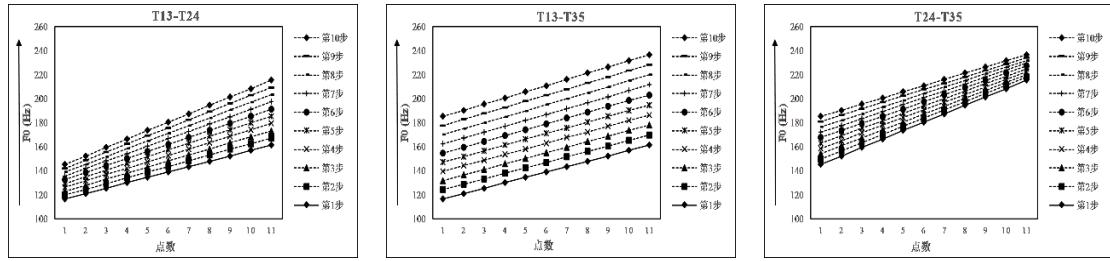


图 3 声调连续统分十步的基频调节
(实线代表原始样本, 虚线代表合成样本, 箭头指示合成方向)

三组声调连续统是在 Praat (Boersma & Weenink 2016) 中合成的, 算法是基音同步叠加法, 该方法在合成过程中只改变信号的基频, 并不会对影响嗓音质量的频谱特性作出改变 (Moulines & Charpentier 1990; Moulines & Laroche 1995)。合成刺激样本的步骤如下: (1) 提取出原始样本归一化的 11 个基频和时长 (表 4)。(2) 根据合成方向选择所用原始样本的时长。具体来说, 基于 T13 合成 T13-T24、T13-T35 的连续统时采用 T13 的时长, 基于 T24 合成 T24-T35 的连续统时采用 T24 的时长。这一方案与前人采用固定时长的处理不同, 理由有二: 一是原始样本的时长相差较小 (T13: 358ms; T24: 348ms; T35: 323ms), 并且在一定范围内改变时长的听辨测试不会影响母语者对这三个升调的判断, 也就是说, 三个升调虽然在时长上存在差异, 但是时长在它们的感知区分中不起作用^①; 二是在合成过程中保持母语者发音样本的时长可以更好地保留原始样本的细节信息, 避免因不必要的调节而损伤原始样本的自然度。(3) 通过调节原始样本的相应基频来合成新的刺激样本。具体来说, 通过内插的方法分十步来调节特定声调的 11 个基频, 从一个声调到另一个声调合成含 10 个刺激样本的声调连续统 (图 3)。上述三组声调连续统共合成 30 个刺激样本。需要说明的是, 本文在合成刺激时选用的单位是赫兹 (Hz), 其目的是保证每一个步长的声学距离都相等, 便于后续研究中与前人的研究 (Wang 1976; Abramson 1979) 进行比较。(4) 将所有刺激的均方根振幅值 (root mean square amplitude, RMS) 设定为 75 dB (相对于听觉阈值)。

(二) 被试

感知实验共招募了 25 名母语者被试 (13 男 12 女, 平均年龄 18.3 岁, 标准差为 1.7), 被试都能熟练运用黔东苗语新合话进行日常交流, 且苗语为日常用语。实验时, 没有被试报告有言语、语言或听力障碍。所有被试均需签署知情同意书, 且在实验结束后会得到一定额度的报酬。

(三) 程序

实验是在剑河县一个安静的房间内进行的, 刺激通过 E-Prime 2.0 呈现。对所有被试而言, 刺激播放的音量都是固定的、舒适的。考虑到三升调的原始语音样本来自主要调查发音合作人, 所以实验之前先让被试熟悉刺激, 目的是为了消除刺激样本的性别差异对后续实验结果的影响。在正式实验之前, 我们还要求被试参加练习实验, 目的是为了让被试熟悉实验流程

^① 需要说明的是, 文章区分了声学和感知两个层面上的时长表现。第 2 节声学分析显示三个升调的时长存在显著性差异, 但是这种差异没有被母语者利用来区别意义。换句话说, 在三个升调的区分中, 时长的差异仅是发音时的一种伴随性特征, 而非区别性特征。

和任务内容，以保证后续正式实验中的稳定表现。此外，练习实验还有助于筛选出那些不符合实验要求的被试。在本项实验中，如果被试在辨认实验中对声调连续统两端的刺激样本的判断正确率小于 80%，那么该被试的数据将会在接下来的分析中剔除。

实验分为两个部分：一是辨认任务，二是区分任务。实验中刺激以随机的方式呈现，每名被试都要完成上述两项任务，程序会自动记录下被试的反应按键和反应时间。辨认任务要求被试通过在电脑上按键来判断每个声调连续统中的刺激样本，本文采用二选一的强迫性选择 (two-alternative forced choice, 2AFC)，要求被试在给出的两个声调中必须选择一个，每个刺激重复七次。区分任务采用 AX 范式 (步长为两步)，一对样本中的两个刺激间隔时间为 500 ms，因为这个时长可以最大化范畴间和范畴内的区分差异 (Pisoni 1973)。对于每一个声调连续统 (含十个刺激样本，编号为 1 到 10)，区分任务共有 26 个试次，其中 16 个试次是包括两个不同的刺激 (正序：1-3, 2-4, 3-5, 4-6, 5-7, 6-8, 7-9, 8-10；逆序：3-1, 4-2, 5-3, 6-4, 7-5, 8-6, 9-7, 10-8)，10 个试次是相同的刺激样本对 (1-1, 2-2, 3-3, 4-4, 5-5, 6-6, 7-7, 8-8, 9-9, 10-10)，任务要求被试判断所播放的两个声音是相同的还是不同的，每对刺激重复三次。

(四) 数据分析

为了研究音高对辨认和区分表现的影响，本文基于范畴感知范式的基本特性获得了每一名被试的辨认得分、边界位置、边界宽度、区分得分、区分范畴内得分和区分范畴间得分。

根据实验设计，辨认得分的定义是给定刺激被判断为声调 A 或 B 的百分比。辨认曲线的边界位置和边界宽度可以通过对辨认曲线进行概率分析获得，边界位置指的是辨认曲线的交叉点处 (50%)，边界宽度被定义为辨认曲线上通过平均值和标准差得到的 25% 到 75% 之间的线性距离 (Best & Strange 1992)。

区分得分的计算采用 Xu et al. (2006) 的公式。每一组包含四种类型的对比对 (AB, BA, AA, BB)，其中 AB 和 BA 属于不同样本对，AA 和 BB 属于相同样本对，并且相邻比较组包含重叠的 AA 和 BB。每一个对比组的得分 P 被定义为： $P = P(\text{"S"}|\text{S}) P(\text{S}) + P(\text{"D"}|\text{D}) P(\text{D})$ ，其中 $P(\text{"S"}|\text{S})$ 代表相同样本对被判断为“相同”的比率， $P(\text{"D"}|\text{D})$ 代表不同样本对被判断为“不同”的比率， $P(\text{S})$ 和 $P(\text{D})$ 分别代表每组样本中相同对和不同对的比率。在获得每一个被试的区分数据后，我们进而可以计算区分曲线范畴内得分和范畴间得分。区分范畴间得分测量的是与辨认曲线的范畴边界相对应的区分刺激对的得分，区分范畴内得分测量的是除与辨认曲线范畴边界相对应的区分刺激对得分之外的其他刺激对的平均得分。

(五) 结果

根据被试辨认实验的表现，本文在 T13-T24 这个连续统中剔除了两名女性被试的数据，在 T13-T35 这个连续统中剔除了一名女性被试的数据。图 4 显示的是有效被试的平均辨认曲线和区分曲线。

对每一名被试的辨认曲线进行统计便可以得到每个声调连续统的辨认得分、辨认边界位置和边界宽度 (表 5)。表 5 中的辨认得分指的是被试对声调连续统中前一个声调的判断正确率，即 T13-T24 和 T13-T35 中被判断为 T13 的正确率、T24-T35 中被判断为 T24 的正确率。从声调连续统的平均辨认得分来看，T13-T24 最小，T13-T35 处于中间，T24-T35 最大。再观察边界宽度，通常来说，边界宽度越大，辨认曲线在边界位置处就越平缓；反之，边界宽度越小，辨认曲线在边界位置处就越陡峭。图 4 和表 5 显示，T13-T24 的边界宽度最大，T24-T35

次之, T13-T35 最小。

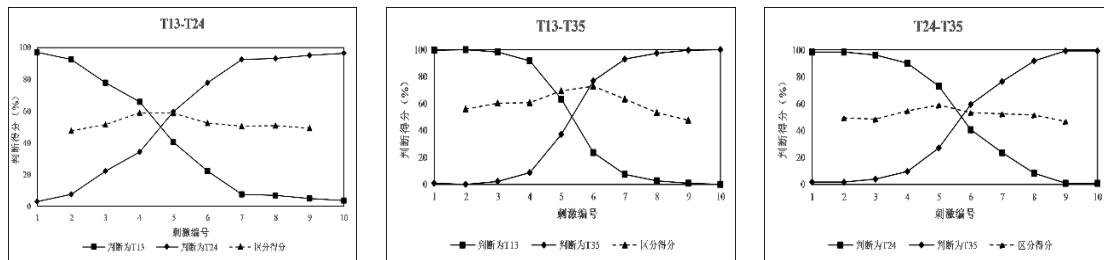


图 4 三升调的辨认曲线(实线)和区分曲线(虚线)

表 5 声调连续统之间的各项感知参数(单位: %)

声调连续统	辨认得分	边界位置	边界宽度	区分正确率	范畴间区分率	范畴内区分率
T13-T24	41.86 (1.91)	4.64 (0.20)	2.56 (0.25)	52.45 (1.44)	58.62 (2.80)	50.39 (1.39)
T13-T35	48.57 (1.10)	5.36 (0.11)	1.19 (0.13)	60.38 (1.82)	71.08 (2.85)	56.81 (2.09)
T24-T35	52.91 (1.63)	5.78 (0.16)	1.74 (0.14)	51.96 (1.71)	56.12 (2.46)	50.57 (1.81)

就区分曲线而言, 图 4 显示每一个声调连续统的区分正确率在范畴内多处于随机水平(50%),但是区分正确率在范畴边界处呈现出强弱不同的区分峰值。为了观察这个区分峰值是否具有统计学上的意义,本文分别计算了范畴间和范畴内的区分正确率(表 5)。需要说明的是,由于感知实验的数据量较小,在进行统计分析之前,我们首先考察了数据的分布趋势, Shapiro-Wilk 检验的 p 值均大于 0.05,说明数据接近正态分布。后续统计结果显示,在 T24-T35 这个声调连续统中,尽管范畴间的区分正确率高于范畴内(5.55%),但是配对样本 T 检验的结果显示区分正确率在范畴间和范畴内并不存在统计上的显著性差异($t(48) = 1.818, p = 0.075$)。同理,在 T13-T24 和 T13-T35 这两个声调连续统中,范畴间的区分正确率也高于范畴内(分别为 8.23% 和 14.27%),配对样本 T 检验的结果表明这两个声调连续统的区分正确率在范畴间和范畴内存在统计上的显著性差异,分别为($t(44) = 2.628, p = 0.012$)和($t(46) = 4.045, p < 0.001$)。

结合辨认曲线和区分曲线的表现,在黔东苗语新合话三升调的感知中,本文发现 T13-T24、T13-T35 这两个声调连续统的感知结果符合范畴感知的定义,即辨认曲线在边界位置处存在一个陡峭的变化(体现为较小的边界宽度),区分曲线存在区分正确率峰值(表现为区分正确率在范畴间和范畴内存在显著性差异),且辨认边界位置对应于区分曲线上的区分峰值。然而 T24-T35 则不具备上述特点,换句话说, T24-T35 的感知结果属于连续感知。

四 讨 论

声调范畴化感知是言语感知研究中的一个核心问题。尽管前人研究揭示出了声调感知的常见模式,即异型声调间的感知通常是范畴感知,同型声调间的感知则属于连续感知,并给出了造成这种差异的不同假设。然而,以往研究并未解决造成这种感知模式差异的原因(见

引言部分的综述)。在语言学研究中,探究声调范畴化感知的成因不仅对确定声调的音位范畴具有重要参考价值,还为解释声调库藏的类型与数量、洞悉声调的起源与演变等提供一些直接或间接的证据。

在黔东苗语新合话中,同为升调,低升调和中升调(T13-T24)、低升调和高升调(T13-T35)之间表现出范畴感知的特性,而中升调和高升调(T24-T35)之间则呈现出连续感知的特点。这一结果表明除调型的异同会导致不同的感知模式外,即使是调型相同的声调,它们之间的感知模式也并非全部表现为连续感知。造成这种现象的原因是什么呢?如果仅依据感知实验的结果,我们无法找到这种表现背后的成因。当检视这三个升调的声学基础时,通常来说可以考虑的因素有音色、音强、时长、发声类型和音高斜率^①。

首先,用于合成刺激的原始样本具有相同的声韵母结构[ta],并且它们的共振峰结构也是相同的,所以可以排除音色的影响。其次,音强也可以被排除,因为用于听辨的刺激样本的振幅都被设定为75 dB。再者,从发音的视角看,多维尺度分析的结果显示这三个升调的主要区别特征是平均基频及斜率。虽然时长也起一定的作用(见第2部分),但本文感知实验中用于合成刺激的三个升调的原始语音样本在时长上的差异较小,尤为关键的是,在一定范围内改变时长的听辨测试也不会影响母语者对它们的判断,即时长在这三个升调的感知区分中不起作用。另外,无论时长是区别性特征还是伴随性特征,它都不会改变声调感知的模式(刘文等2021)。即,时长的差异不会将原本为连续感知的结果改变成范畴感知,反之亦然。基于上述分析,我们也可以排除时长对三升调之间感知结果差异的影响。第四,造成三升调之间感知结果不一致的因素有可能是噪音发声类型的差异?从发音的视角看,第2部分声学研究显示这三个升调之间的噪音发声类型并无明显差异,因为有无噪音参数并不会影响这三个升调在声调空间中的分布(图2)。此外,从感知的视角看,跨语言研究已表明发声类型的不同只会影响到辨认曲线上的辨认得分和边界位置,并不会影响声调的感知模式(杨若晓2009;张锐锋、孔江平2014;Yang 2015; Liu et al. 2024)。在排除上述几种可能后,结合前人研究中得到的同型声调通常为连续感知、异型声调为范畴感知的基本结论,本文考察了黔东苗语新合话三升调的基频斜率。声学分析显示,T13和T24($p < 0.001$)、T13和T35($p = 0.067$)这两对声调在基频斜率上存在显著性差异或边缘显著性差异,而T24和T35这两个升调尽管在基频斜率上也存在差异,但是这种差异没有达到显著水平($p = 0.698$)(表3)。由此可见,尽管黔东苗语新合话三个升调的调型相同,但是只有当两个升调之间的斜率存在显著性或边缘显著性差异时,听者在感知二者之间的声学连续统时才会呈现出范畴感知的特性,

^① 特别需要强调的是,声调的声学表现和感知表现是两个层面的问题。声调是一个语言学概念,是感知出来的范畴。因此,在研究声调时,除了关注它的生理和声学基础,更为重要的是关注它的感知基础,因为在说话人所表现出的生理和声学信号的差异中,只有那些能够被母语者感知到的差异才能作为语言对立要素的候选对象。换句话说,语音单位(例如声调、辅音、元音)的生理和声学变异是客观存在的,感知就是一个归一化过程,它可以消除说话人、语音环境和语速等影响因素,进而将这些生理、物理存在变异的信号归派到同一个音位范畴内。举例来说,笔者曾以汉语普通话为例做过相关实验,实验要求一名女性母语者在其自然舒适的发音状态下尽可能保持音量、音高和口腔形状不变的条件下产出100个元音[a]。结果显示,尽管发音人和听话人都认为这100个[a]属于同一个音位/a/,但是它们在基频和共振峰分布上均存在变异。无论基频还是共振峰数据,没有任何两个[a]的数据是完全重叠的。这一实验结果表明,语言的音位是范畴性的,其分布在一个区域或空间内,而非单一一个点或几个点。就黔东苗语新合话的三个升调而言,尽管它们的时长和斜率在声学上存在差异,但是这种声学的差异与感知上的表现并非一一对应。鉴于声学与感知存在非对等的关系,本文才从言语产生和感知两个层面设计实验来解决三升调的问题。

即跨范畴的刺激对的差异被放大，而同范畴的刺激对就没有这种放大效应。相反，如果调型相同或相似的升调在斜率上没有显著性差异，那么听者对它们的区分就与平调的感知类似，呈现出连续感知的特点。

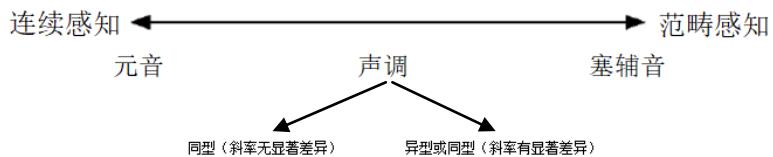


图 5 声调感知范畴化程度连续统

由此可见，就声调而言，其感知结果既有与塞辅音类似的范畴感知，亦有与元音相近的连续感知（见第 1 节引言部分的综述）。不过以往研究大多采用二分的方法来区分感知模式：范畴感知和连续感知。在这种二分的框架之外，Hallé et al. (2004) 指出范畴感知的程度具有梯度性。此外，来自眼动实验的证据也同样支持汉语普通话的声调感知是一个渐进的过程 (Shen et al. 2013)。结合前人的研究成果和本文的研究发现，我们提出声调感知的范畴化是处于一个连续统之上的观点，并且范畴化程度受制于所研究的对象（图 5）。具体来说，当声调间的调型相异或声调间的调型相同但斜率之间存在显著性差异时，感知结果与塞辅音相似，呈现出范畴感知的特点，即辨认曲线上有一个比较陡峭的边界，区分曲线上有区分峰值；相反，当声调间的调型相同且斜率之间没有显著性差异时，感知结果就与元音相似，呈现出连续感知的特点，即辨认曲线在边界位置处比较平缓，区分曲线上没有区分峰值。本文提出的这一观点可以合理地解释前人的研究结果。例如，引言部分所提到的平调和升调、平调和降调、升调和降调之间正是由于斜率上存在显著性差异才会导致范畴感知模式的出现；相反，斜率没有显著性差异的平调之间就属于连续感知模式。由此，本文认为声调范畴感知在本质上是由基频斜率这一声学参数的突变导致的。当然，这一结论只是基于黔东苗语新合话升调的个案研究所得，后续研究还需要更多的同型声调的语言事实来验证。

五 结 语

声调是人类语言演化出的提高语言信息量的独特方式，它的特征不仅由其物理属性决定，同时还受制于听者的生理和心理特性。声调范畴化感知一直是言语感知研究的焦点，已有研究指出声调感知存在范畴感知和连续感知两种常见模式，然而有关声调范畴化感知的成因一直没有得到解决。黔东苗语新合话三升调的声学分析和感知实验结果表明声调间的基频斜率是声调范畴化感知的成因。在此基础上，本文提出声调感知的范畴化是处于一个连续统之上的观点。本文的研究成果不仅阐释了声调范畴化感知的物理基础，同时也有助于我们从言语感知的视角更好地认识和了解声调。

附录：黔东南苗语新合话声调例词

调类	例词
T1	[t̪a ³³] 药、[tæ ³³] 厚、[tai ³³] 回答、[tao ³³] 深、[t̪e ³³] 菌子、[qəu ³³] 农活、[k̪ɛ ³³] 虫、[pei ³³] 三、[qe ³³] 鸡、[pi ³³] 送给、[ki ³³] 牛角、[o ³³] 二、[t̪u ³³] 针
T2	[ta ⁵³] 来、[t̪a ⁵³] 茄子、[pæ ⁵³] 花、[tæ ⁵³] 桌子、[qæ ⁵³] 扁担、[t̪e ⁵³] 骑、[tei ⁵³] 直、[pi ⁵³] 落下、[ki ⁵³] 芦笙、[t̪u ⁵³] 门
T3	[pa ³⁵] 雄性、[ta ³⁵] 手指、[qa ³⁵] 犀、[pæ ³⁵] 摆放、[æ ³⁵] 姐姐、[tai ³⁵] 长、[k̪ɛ ³⁵] 饭、[pi ³⁵] 满、[t̪u ³⁵] 酒、[t̪ao ³⁵] 糜粑
T4	[pa ²²] 坏了、[ka ²²] 矮、[pæ ²²] 网、[ta ²²] 丢失、[pao ²²] 刺儿、[tao ²²] 远、[pei ²²] 手、[ki ²²] 干燥、[tou ²²] 火
T5	[ta ²⁴] 杀、[tja ²⁴] 嚼、[qa ²⁴] 公鸡叫、[pæ ²⁴] 猪、[qæ ²⁴] 挑水、[tai ²⁴] 烤火、[pao ²⁴] 响、[təu ²⁴] 树、[qəu ²⁴] 爷爷、[tei ²⁴] 踢、[kei ²⁴] 蛋、[pei ²⁴] 睡、[tju ²⁴] 六、[t̪ao ²⁴] 明白
T6	[ta ¹¹] 死、[t̪a ¹¹] 遇见、[qa ¹¹] 挂、[pæ ¹¹] 逃跑、[qæ ¹¹] 竿、[pao ¹¹] 抱孩子、[tao ¹¹] 喜欢、[təu ¹¹] 跟着、[tou ¹¹] 破了、[t̪ou ¹¹] 筷子
T7	[ta ¹³] 翅膀、[a ¹³] 乌鸦、[æ ¹³] 背孩子、[tao ¹³] 梭子、[tei ¹³] 姑妈、[qe ¹³] 夹菜、[po ¹³] 肉、[o ¹³] 鸭子、[t̪ou ¹³] 刀、[t̪ao ¹³] 笑
T8	[ka ³¹] 刮、[təu ³¹] 豆、[k̪ɛ ³¹] 脸盆、[pei ³¹] 我们的、[qe ³¹] 隔离、[po ³¹] 盐巴、[tou ³¹] 感染、[t̪u ³¹] 十、[tui ³¹] 口袋、[t̪ao ³¹] 跨过去

参考文献

- 蔡雯清. 2016.《合肥话单字调的知觉研究》, 北京大学硕士学位论文.
- 陈飞、张昊、王士元、彭刚. 2019.《内部因素与元音范畴化感知》,《语言科学》第4期.
- 高云峰. 2004.《声调感知研究》, 上海师范大学博士学位论文.
- 金健. 2010.《广州方言和东海方言平调的感知研究》,《方言》第2期.
- 金健. 2015.《杭州方言多个升调的感知研究》,《中山大学学报》第1期.
- 金健、施其生. 2010.《汕头客方言多个降调的声学分析和感知研究》,《中国语文》第6期.
- 孔江平. 1992.《紫云苗语五平调系统的声学及感知研究》, 载中国民族语言学会编《民族语文研究新探》第152-163页, 成都: 四川民族出版社.
- 孔江平. 1995.《藏语(拉萨话)声调感知研究》,《民族语文》第3期.
- 刘思维. 2015.《汉语相异调形和相似调形声调的范畴感知研究》, 北京大学博士学位论文.
- 刘文. 2019.《同型声调感知的多维研究——基于声学、行为学和脑电的证据》, 北京大学博士学位论文.
- 刘文. 2020.《升还是平: 新寨苗语声调个案研究——兼论平调的感知判断》,《民族语文》第1期.
- 刘文. 2021.《语音学田野调查方法与实践——黔东南苗语(新寨)个案研究》, 济南: 山东大学出版社.
- 刘文、吴南开、孙顺. 2021.《闽北赤岩话长短调的声学和感知研究》, 载北京大学中国语言学研究中心编《语言学论丛》第64辑第152-163页, 北京: 商务印书馆.
- 覃夕航. 2012.《母语经验对汉语普通话声调范畴化感知的影响——基于北京话母语者和粤方言母语者的研究》, 北京大学硕士学位论文.
- 王辅世. 1983.《苗语方言划分问题》,《民族语文》第5期.

- 王辅世. 1985.《苗语简志》, 北京: 民族出版社.
- 王辅世. 1994.《苗语古音构拟》, 东京: 东京外国语大学亚非语言文化研究所.
- 王 璐、孔江平. 2021.《德宏傣语单元音和复元音感知范畴研究》,《民族语文》第 1 期.
- 王韫佳、李美京. 2010.《调型和调阶对阳平和上声知觉的作用》,《心理学报》第 9 期.
- 王韫佳、覃夕航. 2015.《普通话单字调阳平和上声的辨认及区分——兼论实验设计对声调范畴感知结果的影响》,《语言科学》第 4 期.
- 王韫佳、吴 倩、刘思维. 2021.《母语和非母语者对北京话相似和相异调拱声调的范畴感知——基于 MMN 的研究》,《当代语言学》第 3 期.
- 吴 倩、王韫佳. 2018.《声调的范畴知觉及其神经机制》,《心理科学进展》第 1 期.
- 席 洁、姜 薇、张林军、舒 华. 2009.《汉语语音范畴性知觉及其发展》,《心理学报》第 7 期.
- 杨若晓. 2009.《基于发声的汉语普通话四声的范畴知觉研究》, 北京大学硕士学位论文.
- 于 谦. 2017.《方言背景与普通话声调范畴感知研究》, 北京大学博士学位论文.
- 张林军. 2010a.《母语经验对留学生汉语声调范畴化知觉的影响》,《华文教学与研究》第 2 期.
- 张林军. 2010b.《日本留学生汉语声调的范畴化知觉》,《语言教学与研究》第 3 期.
- 张锐锋、孔江平. 2014.《河南禹州方言声调的声学及感知研究》,《方言》第 3 期.
- Abramson, Arthur S. 1979. The noncategorical perception of tone categories in Thai. In Bjorn Lindblom and Sven Ohman (eds.), *Frontiers of Speech Communication Research*, pp. 127-134. London: Academic Press.
- Best, Catherine T., and Winifred Strange. 1992. Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20(3): 305-330.
- Boersma, Paul, and David Weenink. 2016. Praat: Doing phonetics by computer. <http://www.fon.hum.uva.nl/praat/>
- Chan, Stephen W., Chiu-Kuang Chuang and William S.-Y. Wang. 1975. Cross-linguistic study of categorical perception for lexical tone. *The Journal of the Acoustical Society of America* 58(S1): S119-S119.
- Chao, Yuen Ren. 1930. A system of “tone-letters”. *Le maître phonétique* 45(30): 24-27.
- Francis, Alexander L., Valter Ciocca, and Brenda Kei Chit Ng. 2003. On the (non)categorical perception of lexical tones. *Perception and Psychophysics* 65(7): 1029-1044.
- Fry, D. B., Arthur S. Abramson, Peter D. Eimas, and Alvin M. Liberman. 1962. The identification and discrimination of synthetic vowels. *Language and speech* 5(4): 171-189.
- Gordon, Matthew, and Peter Ladefoged. 2001. Phonation types: A cross-linguistic overview. *Journal of Phonetics* 29(4): 383-406.
- Hallé, Pierre A., Yueh-Chin Chang, and Catherine T. Best. 2004. Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics* 32(3): 395-421.
- Hillenbrand, James, Ronald A. Cleveland, and Robert L. Erickson. 1994. Acoustic correlates of breathy voice quality. *Journal of Speech and Hearing Research* 37(4): 769-778.
- Holmberg, Eva B., Robert E. Hillman, Joseph S. Perkell, Peter C. Guiod, and Susan L. Goldman. 1995. Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech, Language and Hearing Research* 38(6): 1212-1223.
- Iseli, Markus, Yen-Liang Shue, and Abeer Alwan. 2007. Age, sex, and vowel dependencies of acoustic measures related to the voice source. *Journal of the Acoustical Society of America* 121(4): 2283-2295.

- Kuznetsova, Alexandra, Per B. Brockhoff, and Rune H. B. Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13): 1-26.
- Liberman, Alvin M. 1957. Some results of research on speech perception. *The Journal of the Acoustical Society of America* 29(1): 117-123.
- Liberman, Alvin M., Franklin S. Cooper, Donald P. Shankweiler, and Michael Studdert-Kennedy. 1967. Perception of the speech code. *Psychological Review* 74(6): 431-461.
- Liberman, Alvin M., Katherine Safford Harris, Howard S. Hoffman, and Belver C. Griffith. 1957. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 54(5): 358-368.
- Liberman, Alvin M., Katherine Safford Harris, Jo Ann Kinney, and H. Lane. 1961. The discrimination of relative onset-time of the components of certain speech and nonspeech patterns. *Journal of Experimental Psychology* 61(5): 379-388.
- Liu, Wen. 2020. A perceptual study on the five level tones in Hmu (Xinzhai variety). *Proceedings of Interspeech 2020*, pp. 1620-1623. International Speech Communication Association.
- Liu, Wen, Gang Peng, and Jiangping Kong. 2024. The role of breathy voice in Hmu tone perception. *Journal of Chinese Linguistics* 52(1): 138-174.
- Moulines, Eric, and Francis Charpentier. 1990. Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication* 9(5-6):453-467.
- Moulines, Eric, and Jean Laroche. 1995. Non-parametric techniques for pitch-scale and time-scale modification of speech. *Speech Communication* 16(2): 175-205.
- Peng, Gang, Hong-Ying Zheng, Tao Gong, Ruo-Xiao Yang, Jiang-Ping Kong, and William S-Y. Wang. 2010. The influence of language experience on categorical perception of pitch contours. *Journal of Phonetics* 38(4): 616-624.
- Pisoni, David B. 1973. Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics* 13(2): 253-260.
- Pritschet, Laura, Derek Powell, and Zachary Horne. 2016. Marginally significant effects as evidence for hypotheses: Changing attitudes over four decades. *Psychological Science* 27(7): 1036-1042.
- R Core Team. 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Repp, Bruno H. 1984. Categorical perception: Issues, methods, findings. *Speech and Language* 10: 243-335.
- Shen, Guannan N. 2015. Perceptual Learning of Lexical Tone Categories: An ERP Study. PhD dissertation, Columbia University.
- Shen, Jing, Diana Deutsch, and Keith Rayner. 2013. On-line perception of Mandarin tones 2 and 3: Evidence from eye movements. *The Journal of the Acoustical Society of America* 133(5): 3016-3029.
- Shue, Yen-Liang, Patricia Keating, Chad Vicenik, and Kristine Yu. 2011. VoiceSauce: A program for voice analysis. In *Proceedings of the 17th International Phonetic Sciences (ICPhS)*. pp. 1846-1849. City University of Hong Kong.
- Si, Xiaopeng, Wenjing Zhou, and Bo Hong. 2017. Cooperative cortical network for categorical processing of Chinese lexical tone. *Proceedings of the National Academy of Sciences* 114(46): 12303-12308.
- Stevens, Kenneth N. 1977. Physics of laryngeal behaviour and larynx modes. *Phonetica* 34(4): 264-279.

- Studdert-Kennedy, Michael, Alvin M. Liberman, Katherine S. Harris, and Franklin S. Cooper. 1970. Motor theory of speech perception: A reply to Lane's critical review. *Psychological Review* 77(3): 234-249.
- Sun, Kuo-Chan, and Tsan Huang. 2012. A cross-linguistic study of Taiwanese tone perception by Taiwanese and English listeners. *Journal of East Asian Linguistics* 21(3): 305-327.
- Wang, William S.-Y. 1976. Language change. *Annals of the New York Academy of Sciences* 280(1): 61-72.
- Xi, Jie, Linjun Zhang, Hua Shu, Yang Zhang, and Ping Li. 2010. Categorical perception of lexical tones in Chinese revealed by mismatch negativity. *Neuroscience* 170(1): 223-231.
- Xu, Yisheng, Jackon T. Gandour, and Alexander L. Francis. 2006. Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of Acoustical Society of America* 120(2): 1063-1074.
- Yang, Ruoxiao. 2015. The role of phonation cues in Mandarin tonal perception. *Journal of Chinese Linguistics* 43(1b): 453-472.

An Acoustic and Perceptual Study on Three Rising Tones in Hmu (Xinhe Variety)

LIU Wen

[Abstract] The categorization of tone perception has attracted extensive attention in the fields of phonetics and cognitive neuroscience. Previous studies have pointed out that the perception is categorical when the tones have different shapes, while continuous when the tones have the same shape with different pitch levels. However, it has been unresolved concerning the causes of categorical perception and continuous perception. Taking the three rising tones of Hmu (Xinhe Variety) as an example, this study reveals that the perception is categorical when there is significant difference in F0 slope among the tones no matter tone shapes are the same or different, while continuous when there is no significant difference in F0 slope among the tones. This finding solves a fundamental problem in the theory of tone perception, i.e., F0 slope leads to the distinction between categorical and continuous perception. On this basis, the present study proposes that there is a continuum regarding the categorization of tone perception.

[Keywords] tone perception categorical perception Hmu (Xinhe Variety) three rising tones F0 slope

(通信地址: 250100 济南 山东大学文学院/山东大学语言科学实验中心)

【本文责编 胡鸿雁】