

# 德宏傣语单元音和复元音感知范畴研究\*

王璐 孔江平

[提要] 本文采用语音合成和感知实验的研究方法,探讨了德宏傣语两组单元音—复元音/a/-/aai/、/u/-/iu/的感知范畴模式。研究表明:(1)两组元音均显示出范畴感知特征,不同于以往对单元音感知为连续感知的认识;(2)两组元音范畴化程度有所不同,/a/-/aai/的范畴化程度高于/u/-/iu/,这与元音内部共振峰变化量有关,变化量越大,感知范畴化程度越高;(3)两组元音感知边界位置存在一定差异,/a/-/aai/的感知边界偏向单元音/a/,/u/-/iu/的边界位于二者中间,感知边界的偏移是受到元音内部音强分布的影响。

[关键词] 德宏傣语 单元音 复元音 元音感知 范畴感知

## 一 引言

范畴感知是人类重要的认知方式,语音感知研究人如何将物理上连续的语音刺激分成离散的范畴。Liberman(1957)首次提出范畴感知研究范式,并确认塞辅音/b/-/d/-/g/属于范畴感知模式。Fry(1962)最早对元音进行感知实验,发现单元音/t/-/ε/-/æ/不符合范畴感知模式的特征,提出元音为连续感知模式。很长一段时间内人们普遍认为辅音是范畴感知、元音是连续感知,实际上元音的感知模式不能一概而论。

正如Stevens(1969)强调的,只能说稳态的孤立单元音被证明是连续感知模式。后续研究探讨了具有外部动态线索的元音感知,可分为两类。一类是听辨不同语音环境下的近似元音对,研究发现处在CVC中的近似元音听辨正确率要高于孤立条件,说明辅音环境确实对元音感知产生了影响(Strange et al. 1976, 1983; Gottfried & Strange 1980; Verbrugge 1979)。另一类研究将元音范畴感知实验置于不同语音环境下进行,结果表明语音环境影响感知边界的位置(Lindblom & Studdert-Kennedy 1967),同时有语境条件下的范畴化程度要高于孤立条件(Repp et al. 1979)。以上研究均反映出辅音环境下协同发音造成的动态特征对元音感知产生了影响。Nearey & Assmann(1986)、Nearey(1989)也指出,声学动态特征包括辅音环境的外部因素和复合元音的内部因素,都可能影响元音感知。前者已有不少讨论,而元音内部因素,即复元音的研究还相当不足。先前对复元音的感知研究多为切音听辨实验,讨论复元音各部分对感知的贡献(Gay 1970; Bond 1982; Bladon 1985)、听辨为复元音所需时间(贺宁基 1985)、听辨的临界点位置(王萍 2008)等问题,鲜有涉及感知模式的研究,无从回答

\* 本文得到了教育部基地重大项目“基于语言多模态的语言本体研究(17JJD740001)”“环南海国家语言生态研究及语言资源库建设(16ZDA211)”的支持。实验过程得到了陈焕茂、杨晓华、沙喊银、杨相洼等人的热心帮助,《民族语文》匿名评审专家给出了宝贵的修改意见。在此并致谢忱!

内部因素对元音感知模式的具体影响。陈飞等(2019)通过比较普通话中具有内部变化的二合元音和单元音/aɪ/-a/、/ia/-a/之间与单元音/a/-i/之间的感知范畴化程度,发现前者感知范畴化程度更高,初步证明了元音内部变化对感知的影响。不过,陈飞等(2019)考察的样本比较有限,仅仅围绕普通话元音a和i之间的感知进行了实验,其结论是否适用于其他语言,内部共振峰变化量不同的元音之间的感知情况是否一致,以及是否还有其他声学参数影响感知等问题还需要进一步的实验探讨。

由于普通话韵母数量有限,本研究选择韵母系统相对发达的德宏傣语为研究对象。傣语系属汉藏语系壮侗语族,德宏傣语是其主要方言,芒市话为德宏傣文的标准音点(喻翠容、罗美珍 1980)。我们选取德宏傣语两组单元音和复元音进行感知实验研究,并通过声学分析寻求感知差异的解释,以期回答:(1)德宏傣语中具有内部变化的元音感知是否体现出范畴感知模式特征;(2)不同单元音和复元音间的感知范畴化程度有无差异及原因;(3)不同单元音和复元音间的感知边界有无差异及原因。

## 二 实验材料及合成

德宏傣语一大特点是元音a后接韵尾时区分长短,长元音记作aa、短元音记作a。我们遵循常用词、实义词的原则,从《德宏傣语同音词典》(周耀文等 2005)中选取了两组“单元音—复元音”最小对立的材料:(1)/a<sup>33</sup>/“姑妈”,/aai<sup>33</sup>/“气体”<sup>①</sup>;(2)/u<sup>35</sup>/“祖母”,/iu<sup>35</sup>/“叫喊”。发音人为土生土长的芒市人,录制在德宏广播电视台播音室内进行,设备包括笔记本电脑(Thinkpad X1)、调音台(XENYX 302 USB)、麦克风(Sony ECM-44B)、外置声卡(Creative Labs Model No. SB109S),录制软件为Adobe Audition,双声道录制,采样率44kHz。录音时用负载句“我说X这个词”进行录制,合成材料仅截取单词部分。

为合成单元音和复元音之间的共振峰过渡样本,我们首先参照统计音节时长的结果(均值为398.4ms),将样本时长归一化为400ms,接着用Winn(2016)基于LPC算法编写的Praat软件共振峰过渡脚本,过渡元音的前三个共振峰F1、F2、F3进行批量合成。样本的共振峰参数过渡值如图1所示<sup>②</sup>,纵轴为共振峰(Hz),横轴为点数,图中深色曲线为复元音的共振峰曲线,浅色曲线为单元音的共振峰曲线,每个样本取30个共振峰点。如图1(1)所示,我们以/a<sup>33</sup>/为母本,合成了从/a<sup>33</sup>/共振峰值过渡到/aai<sup>33</sup>/共振峰值的10个连续体,含母本共11个刺激;我们以/aai<sup>33</sup>/为母本,合成了从/aai<sup>33</sup>/共振峰值过渡到/a<sup>33</sup>/共振峰值的10个连续体,含母本共11个刺激。同样地,如图1(2)所示,我们合成了/u/过渡到/iu/、/iu/过渡到/u/各10个连续体,总共得到了11×4组=44个刺激。

<sup>①</sup> 德宏傣语中/aai/与/ai/中的a音质上有较大差异,其中/aai/中的a更接近单元音/a/。

<sup>②</sup> 有审稿专家指出实验材料选用/a/-/aai/、/u/-/iu/组配而非/i/-/iu/组配的合理性问题,本文主要出于对控制元音时间结构的考虑。如图1所示,/aai/过渡段中点约位于音节总长60%处,/iu/过渡段中点约位于40%处,因此a和u在各自音节总长中占比相当(均为60%)。

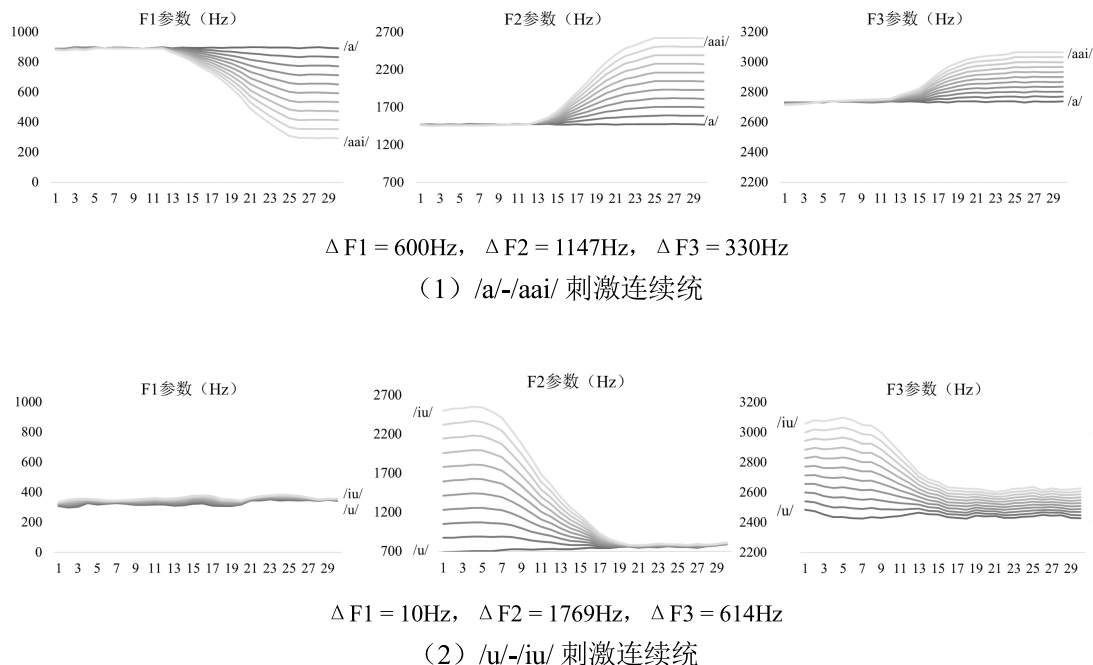


图 1 两组合成刺激连续体共振峰参数值

### 三 感知实验方法

#### (一) 实验被试

实验选择 30 名德宏傣语母语者作为被试，均为土生土长的芒市人，能用傣语、汉语流利地进行日常交流，能识傣文，有 11 名男性，19 名女性，年龄为 19-24 岁，平均 21.6 岁，所有被试均完整参与了感知实验的全部任务，并在实验完成后获得一定报酬。

#### (二) 实验程序

感知实验使用笔记本电脑 (Thinkpad Yoga)、E-prime 1.1 软件、Sony 耳机进行，所有实验程序均为强迫选择 (forced-choice paradigm)，即被试必须要在规定时间内，在提供的选项中做出单选。感知实验包括确认实验和区分实验两个部分，总时长约 30 分钟。

确认实验要求被试判定刺激为某个特定音。实验程序随机播放刺激材料，被试要在屏幕上出现的两个词 (单元音词 / 复元音词) 中二选一。如前所述，共有 11×4 组刺激，每个刺激出现两次，被试需在刺激播放完毕 5 秒内做出按键反应，共计 2 次×11×4 组 = 88 次。

区分实验是为了检验被试能否觉察声学参数上等距的刺激连续体间的差异，为控制时间成本，本实验采用最简单便利的 AX 法，即要求被试判断听到的刺激对是否相同。刺激对由每组 11 个刺激间隔一个单位组成，包括相异刺激和相同刺激，相异刺激有 1-3, 2-4, …, 8-10, 9-11, 共 9 对，相同刺激有 2-2, 3-3, …, 9-9, 10-10, 共 9 对，相异刺激正序、倒序播放各两遍，相同刺激播放两遍，刺激对播放顺序随机，一组刺激共计 9×2×2+9×2 = 54 对，被试需在刺激播放完毕 5 秒内做出按键反应，共计 54 次×4 组 = 216 次按键反应。

正式实验开始前,有一段练习让被试熟悉实验流程,被试可以反复进行练习来熟悉实验程序,在实验进行过程中,被试可以通过空格键暂停或继续实验。

### (三) 数据分析

感知实验结果的数据处理使用 SPSS 20.0 和 Excel 2019 软件进行,要考察的参数包括确认实验的边界位置和边界宽度、区分实验的正确率峰值等。

根据确认实验结果,可以计算出被试将某个刺激判断为单元音或复元音的确认率  $P_1$ ,将其代入二分变量逻辑回归方程 (Xu et al. 2006):  $\log_e(P_1 \div (1-P_1)) = b_0 + b_1 \times x_0$  公式中  $b_0$ 、 $b_1$  分别为回归直线的截距和斜率,  $x$  表示在第几个样本位置,要计算确认边界的位置  $x_0$ ,就是要计算确认率  $P_1$  为 50% 时的样本位置。令在第  $x_0$  个样本处,  $P_1 = 50%$ , 代入方程可得:  $x_0 = -b_0 \div b_1$ 。要计算确认边界的宽度,即确认曲线上确认率值在 25% 和 75% 时样本位置的距离。令  $P_1$  值为 25% 时样本位置为  $x_1$ ,  $P_1$  值为 75% 时样本位置为  $x_2$ , 代入方程得:  $w = x_2 - x_1 = \log_e 9 \div b_1$ 。根据区分实验结果可以统计出每位被试区分每个刺激对的正确率  $P$ , 根据 Xu et al. (2006) 的公式  $P = P("S"|S) \times P(S) + P("D"|D) \times P(D)$ , 其中 S 代表相同 (same), D 代表不同 (different),  $P$  由两部分组成,  $P("S"|S)$  是将相同刺激对正确判断为相同的概率,  $P("D"|D)$  是将不同刺激对正确判断为不同的概率,  $P(S)$  是相同刺激对在所有刺激对中的比重,本研究为 1/3,  $P(D)$  是不同刺激对的比重,本研究为 2/3。接着,为找到区分实验正确率的显著峰值,使用 SPSS 软件的单因素方差分析 (One-way ANOVA) 和 Tukey's HSD (Honestly Significant Difference) 事后检验法对区分正确率进行两两比较,得出组内无显著差异、组间有显著差异的数据分组,从而判断区分正确率是否存在具有显著差异的峰值。如果区分峰值存在,可进一步比较区分峰值与确认实验的边界位置对应情况。

## 四 感知实验结果与分析

### (一) 实验结果

首先要考察声源对感知实验的影响,我们对 /a/-/aai/、/aai/-/a/ 两组实验结果作主体间效应检验,得到  $F(1, 58) = 1.788 < F_{crit}$ ,  $p = 0.186 > 0.05$ , 无显著差异,排除了声源中特殊发生类型等因素的影响; /u/-/iu/、/iu/-/u/ 两组实验结果作主体间效应检验,得到  $F(1, 58) = 0.236 < F_{crit}$ ,  $p = 0.629 > 0.05$ , 无显著差异,声源同样没有造成影响,因此对不同声源的感知数据进行合并计算。合并后感知实验结果如图 2、图 3 所示。

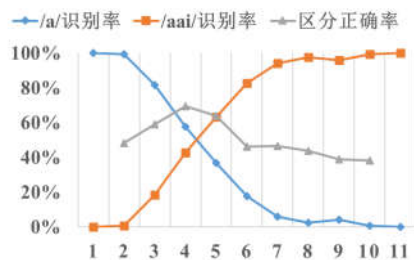


图 2 /a/-/aai/组感知实验结果

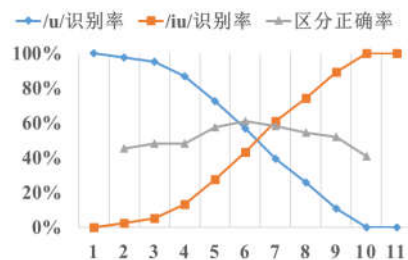


图 3 /u/-/iu/组感知实验结果

计算 /a/-/aai/、/u/-/iu/ 两组确认曲线的回归方程,得到斜率  $b_1$  分别为 -1.078、-0.900, 截

距  $b_0$  分别为 4.827、5.541，计算出确认边界位置为 4.478、6.156，边界宽度为 2.038、2.441。区分实验正确率经 Tukey-HSD 检验结果见表 1、表 2，输出子集内部无显著差异，不同子集间有显著差异。

表 1 /a-/aai/ 区分正确率 Tukey-HSD 结果

	序号	Alpha=0.05 的子集			
		1	2	3	4
Tukey HSD	10	0.3833			
	9	0.3889	0.3889		
	8	0.4361	0.4361		
	6	0.4611	0.4611		
	7	0.4639	0.4639		
	2		0.4833		
	3			0.5889	
	5			0.6378	0.6378
	4				0.6944
显著性		0.202	0.067	0.827	0.678

表 2 /u-/iu/ 区分正确率 Tukey-HSD 结果

	序号	Alpha=0.05 的子集			
		1	2	3	4
Tukey HSD	10	0.4056			
	2	0.4528	0.4528		
	3	0.4806	0.4806	0.4806	
	4	0.4806	0.4806	0.4806	
	9		0.5194	0.5194	0.5194
	8		0.5444	0.5444	0.5444
	5			0.575	0.575
	7			0.5806	0.5806
	6				0.6083
	显著性		0.336	0.11	0.055

由表知，/a-/aai/组区分正确率峰值位于第 4 个样本位置，第 4 个样本与其他子集交集不大（1 个），第 4 和第 5 个样本位置可看作为区分峰值范围；确认实验边界落在区分峰值位置区间内，因此，/a-/aai/组的确认实验结果与区分实验基本对应。/u-/iu/组区分正确率峰值在第 6 个样本位置上，它属于第 5 到第 9 个样本组成的最高值子集，其区分正确率峰值范围较大，且与次高值子集交集较多（4 个），没有出现典型的尖峰型峰值；确认实验边界落在区分峰值位置的区间内，可以认为 /u-/iu/ 组确认实验结果与区分实验基本对应。表 3 从确认实验、区分实验以及二者对应情况几个维度整理了两组感知实验结果。

表 3 /a-/aai/、/u-/iu/ 两组感知实验结果汇总

组别	确认实验		区分实验		确认、区分对应情况
	边界位置	边界宽度	峰值位置	与其他子集交集数	
/a-/aai/	4.478	2.038	4, 5	1 个	基本对应
/u-/iu/	6.156	2.441	5, 6, 7, 8, 9	4 个	基本对应

两组感知实验结果共同点在于，确认实验和区分实验结果基本对应。不同点包括：（1）确认边界位置和区分峰值位置，可统称为感知边界位置；（2）确认实验边界宽度、区分实验正确率峰值所在子集的个数及其与其他子集的交集数，同确认、区分实验的对应情况一样，都是范畴感知模式的重要特征，可用以衡量范畴感知模式的显著程度，即范畴化程度。下面从感知边界位置和范畴化程度两个方面展开讨论。

### （二）感知边界位置差异分析

由表 3 可知，/a-/aai/组的感知边界明显偏向单元音/a/一侧，/u-/iu/组的感知边界则大约位于二者中间。为考察两组感知边界位置差异的元音，我们从内部成分的时长、音强两方面

进行分析。时长上,由图1,两组复元音/aai/和/iu/中,强势成分为a和u,弱势成分为i,强势成分和弱势成分时长比值均为6:4,/aai/过渡到/a/和/iu/过渡到/u/,变化的是时长占比相同的弱势成分i,可见两组时长因素相同,因此时长不是构成两组感知边界位置不同的影响因素。音强上,我们提取了两组复元音/aai/、/iu/的音强(intensity)数据,画成图4的两条曲线,纵轴为音强,单位是分贝(dB),横轴为30个音强点。

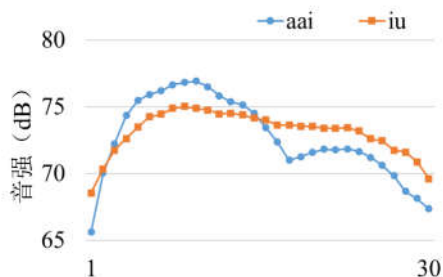


图4 /aai/和/iu/的音强曲线图

如图4所示,一方面,从成分的音强峰值来看,/aai/中a和i的音强峰值相差较大,约6个分贝,而/iu/中i和u的音强差值是非常小的,不足1个分贝;另一方面,从音强模式看,/aai/的音强曲线在i处有明显的下降,两个元音成分音强对比体现出了前响二合元音的特征,/iu/的音强曲线则基本呈马鞍型,相对而言,没有在某个成分上明显上升或下降,两个元音成分强度差异很小,可见二者在音强模式上确实表现出了差异,那么音强会对感知边界造成怎样的影响呢?

我们认为,以/aai/为母本过渡到目标音/a/,是母本/aai/中的音强弱势成分i的音值过渡到a的音值,由于母本/aai/中的a成分本身强势,两成分有音强区别,这就要求弱势成分i发音要更加到位,因此从/aai/过渡到/a/的步数就相对更多,从而造成了感知边界向单元音a方向偏移;反之,以/a/为母本过渡到/aai/,由于目标音/aai/中的i成分本身弱势,发音不需那么到位,也能被判别为目标音/aai/,这同样导致了感知边界向单元音a方向的偏移。而/iu/的两个元音成分i和u音强上基本相当,感知边界没有受到成分音强差异的影响而发生偏移。因此,我们可以从元音内部的音强模式的角度来解释/a/-/aai/、/u/-/iu/两组元音感知边界位置的差异。

### (三) 范畴化程度差异分析

我们从三个维度考察两组元音感知的范畴化程度:(1)确认实验边界宽度,(2)区分实验正确率峰值所在子集的个数及其与其他子集的交集数,(3)两种感知边界,即确认边界与区分峰值位置的对应情况。第一个维度,/a/-/aai/组的确认边界宽度小于/u/-/iu/组,宽度越小,意味着边界附近确认率下降越陡峭,即范畴化程度更高;第二个维度,/a/-/aai/的区分正确率峰值子集及与其他子集交集都比较小,说明/a/-/aai/存在显著的区分峰值,/u/-/iu/组的区分正确率峰值所在子集有5个,为数据量一半之多,且与其他子集交集较多,即不存在显著的区分峰值,因此/a/-/aai/在这一维度显示出显著的范畴感知模式特征,而/u/-/iu/组没有显示出相应特征;第三个维度,两组确认边界都落在区分峰值范围,感知边界均基本对应,符合范畴感知模式特征,不过/u/-/iu/组虽符合边界对应,但因其区分峰值范围过大,这一维度的考量值得商榷。

以上两组感知实验中,被试都能够区分音位范畴,亦显示出了一定的范畴感知模式特征,

/a/-/aai/组的感知范畴化程度更高，属于典型的范畴感知模式；/u/-/iu/组范畴化程度较低，不属于典型的范畴感知模式。其中，/a/-/aai/的感知结果与陈飞等（2019）的研究结论一致，验证了元音的内部变化对感知的影响，我们需要进一步讨论的是为什么/u/-/iu/的感知范畴化程度要明显低于/a/-/aai/？

先前的二合元音听辨实验研究（贺宁基 1985）提示我们，共振峰变化率与其听辨所需时长阈值存在补偿关系，即共振峰变化率越大，所需听辨时长就越短。我们可以据此试着作出这样的假设，就复元音感知而言，共振峰变化越大的复元音，越容易被感知到差异，其范畴化程度也就越高。由图 1，本实验两组复元音的起始点、收尾点前两个共振峰 F1、F2 的共振峰变化量  $\Delta F1$ 、 $\Delta F2$  为：/a/-/aai/组， $\Delta F1 = 600\text{Hz}$ ， $\Delta F2 = 1147\text{Hz}$ ；/u/-/iu/组， $\Delta F1 = 10\text{Hz}$ ， $\Delta F2 = 1769\text{Hz}$ 。由于人耳对低频的变化比高频敏感，因此 F1 对语音感知的贡献要高于 F2（Flanagan 1965）。不过 F1、F2 对感知作用的数值关系尚不明晰，这里暂且遵循贺宁基（1985）的平方根法对共振峰变化量  $\Delta F$  进行计算，可得：/a/-/aai/组  $\Delta F \approx 634\text{Hz}$ ；/iu/组  $\Delta F \approx 52\text{Hz}$ 。可见，范畴化程度更高的/a/-/aai/组，其共振峰变化量要远大于/u/-/iu/组，我们有理由推断，同样是具有内部动态变化的复元音，共振峰变化量大的复元音，其感知的范畴化程度更高。

此外，共振峰的概念除了已作分析的共振峰频率值，还有共振峰振幅，即共振峰能量。下面通过 LPC 包络分析两组元音的共振峰能量分布情况，如图 5 所示，是/aai/中 a 和 i 的稳定段以及/iu/中 i 和 u 的稳定段的 LPC 包络，横轴为频率，纵轴为能量。

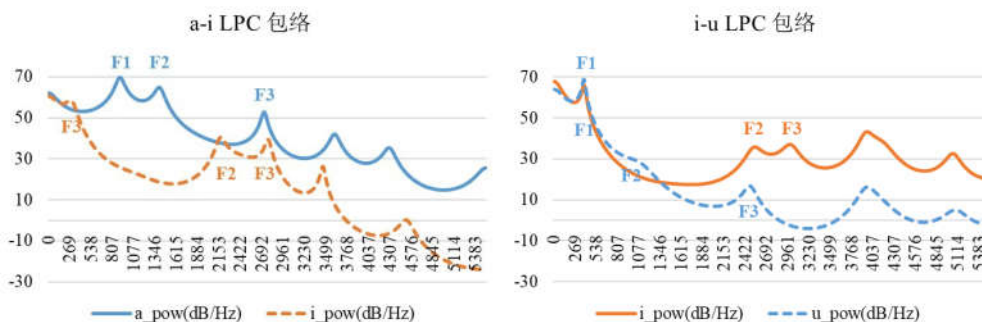


图 5 /aai/中 a 和 i、/iu/中 i 和 u 的 LPC 包络对比

图 5 中标出了元音的前三个共振峰，可以看出，第一张图中 a 和 i 的 LPC 包络差异更大，F1、F2 频率值和能量都相差很远；第二张图中 i 和 u 的 LPC 包络中 F1 频率值和能量都很接近，F2 的频率值虽然相差较大，但由于 u 的 F2 能量很低，导致 i 和 u 的 LPC 包络整体相差不大。另外，因为 u 的 F2 能量很低，人们在合成元音 u 的时候，可以用一个中心频率与其 F1、F2 的合峰相当的单个共振峰代替，因此元音 u 又被称为单共振峰元音。从 LPC 包络差异的对比可以确认，/aai/中的两个元音成分 a 和 i 的共振峰差异要大于/iu/中的两个元音成分 i 和 u，即/a/-/aai/的共振峰变化量大于/u/-/iu/。这提示我们，探讨元音的声学特征、声学距离时，仅测量元音第一、第二共振峰 F1、F2 的频率值是不够的，直接比较数值差异也是不准确的，我们还需考察共振峰能量的分布情况，作进一步比较，才是考察元音更为完备的方法。

以上，我们从共振峰频率、能量分布两个方面确认了/a/-/aai/的共振峰变化量要大于/u/-/iu/，这与我们的假设相符，内部共振峰变化量越大的单元音和复元音间体现出的范畴感

知模式特征更显著,即感知的范畴化程度更高。

## 五 结 论

本研究得出3个结论:(1)两组德宏傣语单元音—复元音的感知实验中,被试均能区分音位范畴,均显示出了一定的范畴感知模式特征,体现了元音内部动态变化对感知模式的影响。(2)两组元音的范畴化感知程度存在差异,其中/a/-/aai/组的范畴化程度高于/u/-/iu/组,结合感知差异和声学分析可以推知,元音内部共振峰变化量是其决定性因素,变化量越大,其感知范畴化程度越高。值得注意的是,要考察共振峰变化情况,不能简单计算声学元音图上F1、F2频率值的距离,还要考虑F1、F2能量上的差异。(3)两组元音的感知边界位置有所不同,/a/-/aai/组的感知边界偏向单元音/a/,/u/-/iu/组的感知边界位于二者中间,声学分析表明,感知边界的偏移受到元音内部音强分布的影响。

由于材料所限,我们根据常用词、实义词的标准,筛选单元音—复元音最小对立对,没有找到相同调值的组合,感知实验所用两组材料虽然符合组内最小对立,但调值不同可能对研究结果造成影响。另外,本文对两组单元音与二合元音的感知情况作了初步探讨,至于其他二合元音、乃至三合元音之间的感知模式如何,值得进一步的研究。

## 参考文献

- [1] 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.
- [2] 贺宁基. 1985.《北京话二合元音感知中的时间因素》,载林焘、王理嘉编《北京语音实验录》第196-224页,北京:北京大学出版社.
- [3] 王 萍. 2008.《北京话二合元音/ai/、/au/、/ia/、/ua/韵腹/a/的听感实验研究》,《语言文字应用》第3期.
- [4] 周耀文、方峰和、郗卫宁. 2005.《德宏傣语同音词典》,北京:民族出版社.
- [5] 喻翠容,罗美珍. 1980.《傣语简志》,北京:民族出版社.
- [6] Bladon, A. 1985. Diphthongs: A case study of dynamic auditory processing. *Speech Communication*, 4 (1-3):145-154.
- [7] Bond, Z. S. 1982. Experiments with synthetic diphthongs. *Journal of Phonetics*, 10: 259-264.
- [8] Flanagan, J. L. 1965. *Speech Analysis Synthesis and Perception*. Berlin: Springer-Verlag.
- [9] Fry, D. B., A. S. Abramson, P. D. Eimas, & A. M. Liberman. 1962. The identification and discrimination of synthetic vowels. *Language and Speech*, 5(4): 171-189.
- [10] Gottfried, T. L., & W. Strange. 1980. Identification of coarticulated vowels. *The Journal of the Acoustical Society of America*, 68(6): 1626-1635.
- [11] Gay, T. 1970. A perceptual study of American English diphthongs. *Language and Speech*, 13(2): 65-88.
- [12] Liberman, A. M., K. S. Harris, H. S. Hoffman, & B. C. Griffith. 1957. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54(5): 358-368.
- [13] Lindblom, B., & M. Studdert-Kennedy. 1967. On the role of formant transitions in vowel recognition, *The Journal of the Acoustical Society of America*, 42: 830-843.
- [14] Nearey, T. M., & P. F. Assmann. 1986. Modeling the role of inherent spectral change in vowel identification. *The Journal of the Acoustical Society of America*, 80(5): 1297.



- [15] Nearey, T. M. 1989. Static, dynamic, and relational properties in vowel perception. *The Journal of the Acoustical Society of America*, 85(5): 2088-2113.
- [16] Repp, B. H., A. F. Healy, & R. G. Crowder. 1979. Categories and context in the perception of isolated steady-state vowels. *Journal of Experimental Psychology: Human Perception and Performance*, 5(1): 129-145.
- [17] Stevens, K. N., A. M. Liberman, M. G. Studdert-Kennedy, & S. E. G. Öhman. 1969. Cross-language study of vowel perception. *Language and Speech*, 12(1): 1-23.
- [18] Strange, W., R. Verbrugge, D. P. Shankweiler, & T. R. Edman. 1976. Consonantal environment specifies vowel identity. *The Journal of the Acoustical Society of America*, 60: 213-224.
- [19] Strange, W., J. J. Jenkins, & T. L. Johnson. 1983. Dynamic specification of coarticulated vowels. Vowel Inherent Spectral Change. *The Journal of the Acoustical Society of America*, 74: 695-705.
- [20] Verbrugge, R. R., & D. Shankweiler 1979. Context-conditioned specification of vowel identity. *The Journal of the Acoustical Society of America*, 65: S1, S6-S6.
- [21] Winn, M. B. 2016. [http://www.mattwinn.com/praat/Make\\_Formant\\_Continuum\\_v37.txt](http://www.mattwinn.com/praat/Make_Formant_Continuum_v37.txt).
- [22] Xu, Y., J. T. Gandour, & A. L. Francis, 2006. Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *The Journal of the Acoustical Society of America*. 120(2): 1063-1074.

## **A Study on the Perceptual Category of Monophthong-Diphthong Sets in the Dehong Dai Language**

**WANG Lu and KONG Jiangping**

**[Abstract]** This paper uses the speech synthesis approach and the categorical perception paradigm to explore the perceptual patterns of two monophthong-diphthong sets, viz. /a/-/aai/ and /u/-/iu/, in the Dehong Dai language. The results turn out as follows: (1) both sets show categorical perception characteristics, different from the previous view that the perception of monophthongs is continuous perception; (2) the categorization degree of /a/-/aai/ is higher than that of /u/-/iu/; (3) the perceptual boundary of /a/-/aai/ is biased towards /a/, while the perceptual boundary of /u/-/iu/ lies in the middle. Based on acoustic analysis, it can be inferred that the shift of perceptual boundary is related to the intensity distribution of the diphthongs, and the degree of categorization is related to the variation of formant, namely, the greater the variation, the higher the categorization degree.

**[Keywords]** the Dehong Dai language monophthong diphthong vowel perception categorical perception

(通信地址: 王璐 100871 北京 北京大学对外汉语教育学院  
孔江平 100871 北京 北京大学中国语言文学系)

【本文责编 木再帕尔】