

## 载瓦语母语者普通话声调感知研究\*

陆 尧

**提 要：**普通话是我国的国家通用语言，然而，不同民族母语者在习得普通话声调的过程中，由于各自母语声调系统或音位系统的影响，会遇到一定的困难。本文通过语音感知实验，研究了载瓦语母语者习得普通话声调系统的过程，并与普通话母语者感知结果进行了对比，探讨了载瓦语母语者建立普通话声调系统的困难和内在规律。研究结果显示载瓦语母语声调系统密切影响对汉语普通话声调的感知：1）载瓦语母语被试对普通话大部分声调对立组的感知范畴化程度小于普通话母语被试；2）载瓦语母语被试在行为学意义上对 T2-T3（T2 母本）、T3-T4（T4 母本）声调对立组的感知模式是连续感知，对其他各声调对立组的感知模式是范畴感知，且在合成母本为带有紧嗓音性质的 T3 时完全无法区分 T2 和 T3 这两个声调音位范畴；3）声调的基频模式和发声类型是载瓦语母语被试感知普通话声调系统的两个重要的声学特征，如果发声类型缺失，就会影响对普通话声调的感知。以上研究表明，除了基频模式，声调中的发声类型对二语声调感知与习得也具有重要的语言学意义。

**关键词：**普通话声调 载瓦语声调 感知范畴 声调习得

### ○ 引言

汉语普通话是我国的国家通用语言，在我国的政治、经济和日常生活中起着重要的作用，然而，我国有众多的民族语言和方言，由于各自母语声调系统的影响，这些民族语和方言母语者在学习普通话的过程中会遇到一定的困难，其中普通话声调系统的习得问题最为突出，解决这一问题需要对声调进行感知研究。

20 世纪 70 年代，王士元（Wang 1976）最早对声调感知进行了实验，认为母语者对普通话声调阴平和阳平的感知是范畴感知。随后，学界围绕声调感知问题进行了更多探讨。如 Abramson（1979）认为，母语者对泰语三平调的感知是连续的，Francis et al.（2003）认为，母语者对粤语平调的感知是连续感知，对粤语平调与曲拱调的感知结果是范畴感知。孔江平（1992）发现母语者对紫云苗语五个平调的感知范畴性比较弱。孔江平（1995）认为母语者对藏语拉萨话的感知是范畴感知。对于普通话的声调感知研究，研究者们普遍认为，母语者对普通话阴平和阳平、阴平和去声、阳平和去声、上声和去声之间的声调感知是范畴感知（Xu et al. 2006；Peng et al. 2010；高云峰 2004；覃夕航 2012；Chen et al. 2017），对普通话阳平和上声之间的感知模式存在争议（杨若晓 2009；于谦 2017；王韞佳、李美京 2010；Chen 2012）。在声调感知的判定标准上，学界普遍认为主要包括如下两点：1）识别曲线上存在陡峭的识别边界；2）区分曲线出现峰值且能够对应识别曲线的边界。两个标准同时满足属于范畴感知；反之，则属于连续感知。（Liberman et al. 1957；

收稿日期：2023-08-29；定稿日期：2023-11-15。

\* 本文得到国家社科基金重大项目“中华民族语言文字接触交融研究”（22ZD&213）和北京大学引导专项“国家通用语言感知研究”（7101303305）的支持。部分内容曾于 2019 年在墨尔本召开的第十九届国际语音科学大会（19th International Congress of the Phonetic Sciences, ICPHs）上报告。文中数据采集得到了德宏州检察院杨晓华，瑞丽市户育乡宣传委员王珏，瑞丽市户育乡户育村支部书记翁草鲁，村长腊端，村民夺石业、关立、董木鲁、杨果等人的大力支持，在此一并表示感谢。

Liberman et al. 1961; Fry et al. 1962; Studdert-Kennedy et al. 1970) 但是, 也有学者提出“准范畴感知”“范畴感知梯度”等概念, 认为存在介于范畴感知与连续感知中间的感知模式。(Schouten and van Hessen 1992; Hallé et al. 2004; 杨若晓 2009)

对二语习得者感知声调语言的情况, 不少学者指出, 语言经验密切影响着语音范畴感知的形成。部分研究认为, 非声调语言母语者在学习声调语言的初级阶段很难建立起对普通话声调的范畴感知, 如 Xu et al. (2006), Peng et al. (2010), 张林军 (2010a、b), 王韞佳、李美京 (2011) 等的研究。部分研究认为, 母语的声调语言经验有助于学习者更好地感知二语声调, 如 Wayland and Guion (2004) 发现母语为普通话的学习者在感知泰语声调时比英语母语者具有优势。Qin and Mok (2011) 采用 AX 识别任务, 考察普通话、英语和法语母语者感知粤语声调的差异, 结果显示普通话被试的成绩要显著优于英语和法语被试。但也有一些研究表明, 母语的声调语言经验反而会对感知二语声调产生消极影响, 如 Wang (2006、2013) 选取在美国生活但母语背景为“苗语—声调语言、日语—音高重音语言、英语—语调重音语言”的汉语普通话初级学习者作为被试, 考察他们对普通话声调的感知情况, 结果发现, 日语被试和英语被试的表现几乎一致, 且显著优于苗语被试, 尤其是对普通话阴平和阳平感知的准确性, 这说明具有声调语言经验的苗语被试在学习的初级阶段受到了声调经验的困扰。还有一些研究指出, 母语为声调语言和非声调语言的学习者感知二语声调时, 在总体正确率上并没有显著差别, 而是表现出不同的偏误模式, 如 Francis et al. (2008) 的研究。总的来说, 目标语言和母语之间的语言特征差异、音位地位差异在感知非母语声调时发挥着重要的作用 (康宇等 2017)。但由于研究方法、研究对象和研究范式等的差异, 上述学者的研究中, 声调语言经验对于二语学习者的声调感知是否存在影响以及存在何种影响, 目前尚无一致结论。

本文讨论载瓦语母语者对汉语普通话声调的感知。以载瓦语母语者作为研究对象有特殊意义。首先, 载瓦语是中国景颇族中自称“载瓦”的人使用的语言, 属于汉藏语系藏缅语族缅语支, 当地人称“小山话”, 其声调系统与普通话的声调系统有许多相似之处。其次, 陆尧、孔江平 (2019) 的研究发现, 载瓦语的发声类型对其感知母语声调具有重要作用。此外, 由于普通话在中国西南地区的推广, 大部分载瓦语母语者都能够日常生活中使用普通话进行交流。因此, 本文在范畴感知框架下研究载瓦语母语者对普通话声调的感知, 能够进一步揭示母语声调系统如何影响对二语声调的感知, 研究成果能够为言语感知理论和语言演化问题的研究提供新的证据, 对少数民族地区的普通话声调教学有较强指导意义。

## 一 载瓦语声调系统分析

根据《景颇族语言简志(载瓦语)》(徐悉艰、徐桂珍 1984) 的声调系统, 戴庆厦 (1989)、何勒腊 (2016) 对载瓦语声调的研究, 以及陆尧、孔江平 (2019) 的声学分析, 提取男女发音人的基频数据作平均值, 再将其转换为五度值,  $\text{五度值} = [(\lg f_x - \lg f_2) / (\lg f_1 - \lg f_2) * 4] + 1$  (李洪彦等 2006), 其中  $f_1$  为调域内的基频最高值,  $f_2$  为调域内的基频最低值,  $f_x$  为每一点的实测基频值。最后得出载瓦语声调基频及五度值情况, 如图 1 所示。载瓦语有 6 个声调, 其中 2 个平调、1 个升调、3 个降调, 调值记为: 55, 44, 35, 51, 31, 21, 具体基频值和五度值如图 1 所示。从时长上来看, 55 调与 31 调为促声调, 其他几个声调为舒声调。

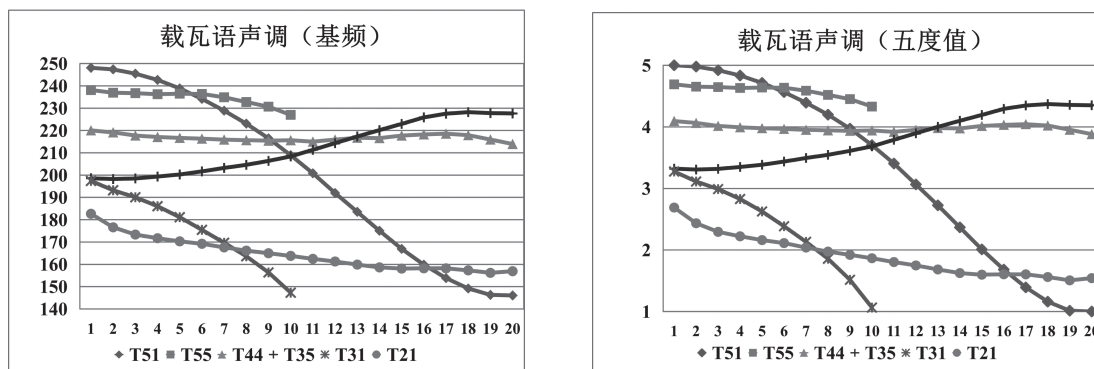


图1 载瓦语声调基频和五度值图

此外，载瓦语的声调与元音发声类型是相互配合的。根据孔江平（2001）给出的嗓音区别性特征矩阵，载瓦语的发声类型可以称为正常嗓音和紧嗓音。其中升调 35 的出现需要一定的条件，即声母是不送气音（清擦音除外），韵母是正常嗓音。而紧嗓音韵母只出现在高平调 55 调或 44 调中，高平调与高升调声韵母具体配合情况如表 1 所示。其他声调音节中的元音可以是紧嗓音，也可以是正常嗓音。

表1 载瓦语高平调与高升调声韵母配合情况

|    |            |            |          |           |          |
|----|------------|------------|----------|-----------|----------|
| 声母 | 不送气音（除清擦音） | 清擦音 / 送气音  | 所有声母     | 清擦音 / 送气音 | 所有声母     |
| 韵母 | 正常嗓音       | 正常嗓音       | 紧嗓音      | 正常嗓音 + 塞尾 | 紧嗓音 + 塞尾 |
| 声调 | 35         | 44         | 44       | 55        | 55       |
| 例字 | /mau/ 奇怪   | /khou/ 表兄弟 | /mau/ 骗人 | /jut/ 错   | /tsut/ 肺 |

## 二 实验方法

### 2.1 实验被试

本文选取的 28 名被试（14 男，14 女）均来自德宏傣族景颇族自治州瑞丽市户育乡户育村尹山、广帕、芒弄、户育四个村民小组，年龄 18～48 岁。所有被试均以载瓦语为母语，会说普通话。普通话对照组共 25 人（10 男，15 女），均为土生土长的北京人，普通话水平在二级乙等及以上。

### 2.2 实验刺激

实验中所采用的刺激样本基于四个真实的汉语普通话单字调音节合成，即“搭”“达”“打”“大”，这四个音节声韵母相同，声调分别为阴平、阳平、上声、去声（下文称为“T1、T2、T3、T4”）。原始音节由一位女性国家级普通话测试员发声，每个音节平均时长 500ms。录音在安静的室内进行，设备包括：cool edit 录音软件、麦克风（Sony）、笔记本电脑（Thinkpad）、外置声卡和调音台。录音时的采样频率为 22050Hz，采样精度为 16 位。4 个声调对立形式共 6 种：T1-T2、T1-T3、T1-T4、T2-T3、T2-T4、T3-T4。我们通过 PSOLA（基频同步叠加）的方法均匀改变基频，以 T1-T2 组为例，我们会以 T1 为母本（下文统称母本）不断改变基频至 T2 合成一组刺激，也会以 T2 为母本改变基频至 T1 合成一组刺激，最终本文合成了 12 组，每组 11 个，共计 132 个刺激样本。

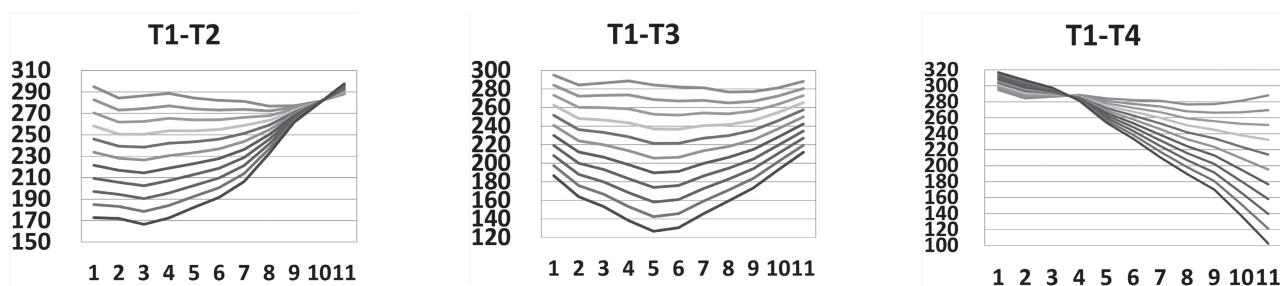


图2 不同声调合成样本设计(1)

图2左图为T1-T2的合成样本,基频最低为166.51Hz,最高为297.58Hz,中图为T1-T3的合成样本,基频最低为126.65Hz,最高为294.91Hz,右图为T1-T4的合成样本,基频最低为102.58Hz,最高为316.92Hz。

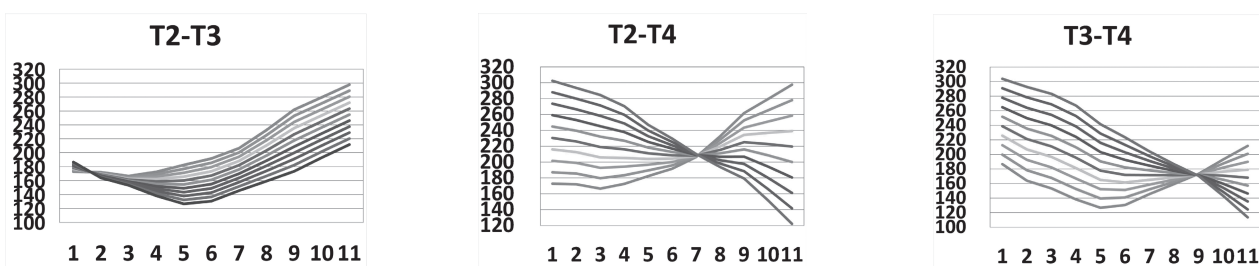


图3 不同声调合成样本设计(2)

图3左图为T2-T3的合成样本,基频最低为126.65Hz,最高为297.58Hz,中图为T2-T4的合成样本,基频最低为122.08Hz,最高为302.5Hz,右图为T3-T4的合成样本,基频最低为113.49Hz,最高为303.89Hz。

## 2.3 实验过程

每个被试都先完成识别实验,再完成区分实验。在识别实验中,12组刺激随机播放,每次每个刺激连续播放两遍。声音播放的同时屏幕上会同时出现“搭”“达”“打”“大”中与该组刺激相对应的两个字,播放完毕后,被试必须在5秒的时间内通过按键“←”和“→”判断这个语音是哪个汉字。区分实验采用AX型,每个刺激间隔500ms成对出现(Pisoni 1973),被试需要判断这对刺激“相同”还是“不同”。不同对中的两个刺激间隔两个步长,正序和倒序都要播放,12组刺激组成了 $18 \times 12 + 9 \times 12 = 324$ 个样本对,每个样本对在实验中播放2次,被试共需做出 $324 \times 2 = 648$ 次反应。所有刺激由索尼MDR-7506耳机播放出来,双声道,72dB SPL。实验和数据采集使用E-PRIME软件进行。

## 2.4 数据分析

本文统计了每个被试的数据,根据范畴感知的标准来考察“识别曲线边界分明与否、区分曲线是否存在波峰且与识别边界对应”(Xu et al. 2006; 于谦 2017)。每个刺激的识别结果是被试将其识别为每个汉字的百分比,我们由此构造了二分变量逻辑回归模型(Binary logistic regression):  $\log_e(P_i / (1 - P_i)) = b_0 + b_1 x$ ,其中 $P_i$ 为每个刺激的识别结果, $x$ 为声音连续体的步长个数, $b_0$ 和 $b_1$ 是回归模型函数中自变量 $x$ 的回归系数和常数,识别率为0.5时所对应的 $x$ 值即为识别边界: $x_{cb} = -b_0 / b_1$ ,边界宽度则为识别率为25%和75%之间的线性距离。区分率采用公式 $P = P(\text{“S”}|\text{S}) \cdot P(\text{S}) + P(\text{“D”}|\text{D}) \cdot P(\text{D})$ 进行计算,其中 $P(\text{“S”}|\text{S})$ 为相同样本

对被判断为相同的概率,  $P("D"|D)$  为不同样本对被判断为不同的概率,  $P(S)$  和  $P(D)$  是相同样本对和不同样本对在整个区分实验中的比例, 本次实验分别为 1/3 和 2/3。对每一个样本对得到的区分率数据, 通过 One-way ANOVA 单因素方差分析和 Tukey HSD post-hoc 事后检验进行两两比较, 确定哪个或哪几个样本对的区分率显著高于其他样本对, 以判断区分波峰的位置<sup>①</sup>。此外, 我们还对载瓦语母语被试与普通话母语被试感知普通话声调的范畴化程度进行了对比。数据统计分析采用 Excel 2010, IBM SPSS Statistics 20.0 及 The R Project for Statistical Computing (一种统计软件) 进行。

### 三 实验结果

图4是T1-T2声调对立组的感知结果。统计得出, 母本为T1时, 普通话母语被试的识别边界和宽度为3.66、0.89, 载瓦语母语被试分别为4.91、3.21。普通话母语被试的区分波峰范围位于2-4、3-5样本对, 峰值出现在2-4样本对, 达90.67%, 对应识别边界。载瓦语母语被试的区分波峰范围在1-3、2-4、3-5样本对, 峰值出现于2-4样本对, 达78.51%, 未能完全对应识别边界。母本为T2时, 普通话母语被试的识别边界和宽度分别为4.24、1.39, 载瓦语母语被试分别为5.11、2.63。普通话母语被试的区分峰值出现于3-5样本对, 达到90%, 落在识别边界范围内。载瓦语母语被试区分峰值出现于3-5样本对, 达73.8%, 未能完全对应识别边界。

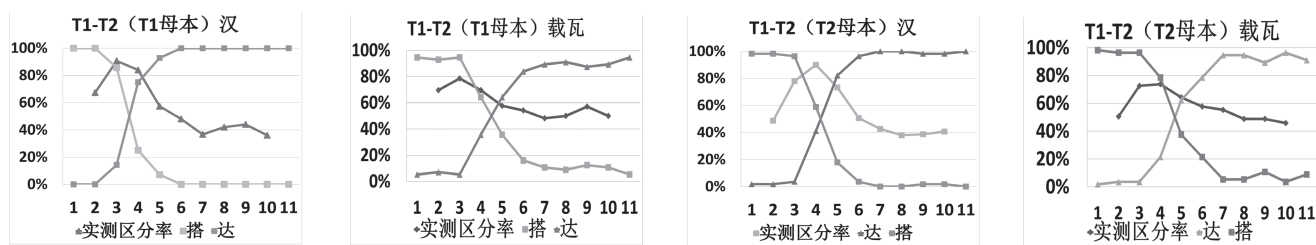


图4 T1-T2连续统感知曲线

图5是T1-T3声调对立组的感知结果。母本为T1时, 普通话母语被试的识别边界和宽度分别为5.12、1.83, 载瓦语母语被试分别为5.32、2.68。普通话母语被试各样本对区分率在统计上不存在区分波峰, 峰值位于3-5样本对, 达88.67%, 落在识别边界范围内。载瓦语母语被试区分波峰范围位于1-3、2-4、3-5、4-6样本对, 但波峰非常平缓, 峰值位于3-5样本对, 达83.3%, 未完全对应识别边界。母本为T3时, 普通话母语被试的识别边界和宽度分别为5.04、1.48, 载瓦语母语被试分别为4.89、2.96。普通话母语被试区分波峰范围位于1-3、2-4、3-5、4-6、5-7样本对, 非常平缓, 峰值位于3-5样本对, 达84%, 未落在识别边界内。载瓦语母语被试区分率均值最高组内有6个样本对: 1-3、2-4、3-5、4-6、5-7、6-8, 即区分波峰范围, 也是一个非常平缓的波峰。峰值位于2-4样本对, 为75.6%, 未能对应识别边界。

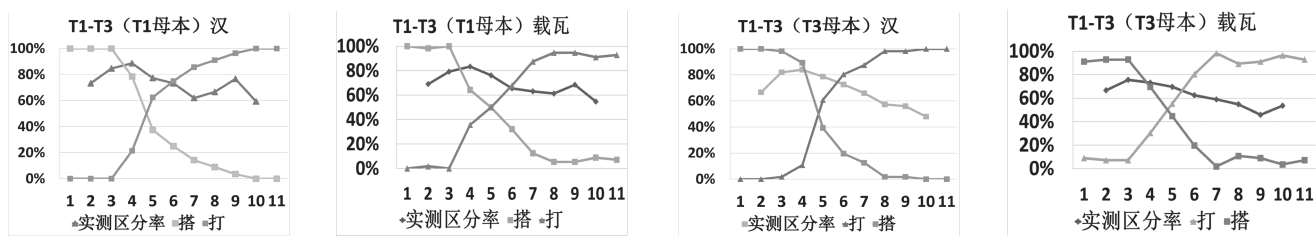


图5 T1-T3连续统感知曲线

<sup>①</sup> 两两比较常见的统计方法有LSD、Scheffe、Dunnett、Tukey等, 其中Tukey的HSD (Honestly Significant Difference) 检验是在各组样本含量相同同时进行两两比较的方法, 适用于本文的研究。

图 6 是 T1-T4 声调对立组的感知结果。母本为 T1 时, 普通话母语被试识别边界和识别宽度分别为 3.6、0.85, 载瓦语母语被试分别为 4.26、1.81。普通话母语被试区分波峰范围位于 1-3、2-4 样本对, 峰值出现于 2-4 样本对, 达 90%, 落在识别边界内。载瓦语母语被试区分波峰范围位于 1-3、2-4、3-5 样本对, 峰值出现于 2-4 样本对, 达 82.14%, 未落在识别边界内。母本为 T4 时, 普通话母语被试识别边界和宽度分别为 3.54、0.91。载瓦语母语被试分别为 4.45、2.04。普通话母语被试区分波峰范围位于 1-3、2-4 样本对, 峰值出现于 1-3 样本对, 为 78%, 对应识别边界。载瓦语母语被试区分波峰范围位于 1-3、2-4 样本对, 峰值出现于 1-3 样本对, 为 70.24%, 未能完全对应识别边界。

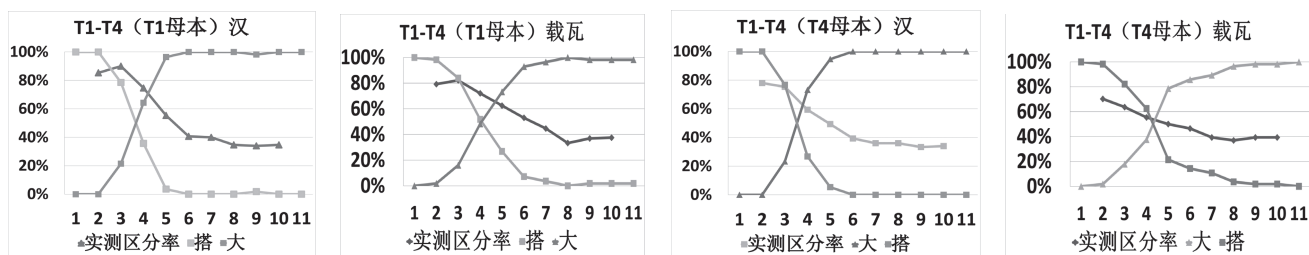


图6 T1-T4连续统感知曲线

图 7 是 T2-T3 声调对立组的感知结果。母本为 T2 时, 普通话母语被试的识别边界和宽度分别为 7.38、1.61, 载瓦语母语被试分别为 6.96、2.74。普通话母语被试在统计上不存在区分波峰, 峰值位于 7-9 样本对, 仅为 50%。载瓦语母语被试各样本对区分率组间差异不显著 ( $F(8,243)=0.264, p=0.977$ ), 不存在区分波峰, 峰值位于 7-9 样本对, 仅为 42.9%。母本为 T3 时, 普通话母语被试的识别边界和宽度分别为 7.21、2.15, 载瓦语母语被试分别为 6.28、14.83。普通话母语被试区分率均值最高组内有 5-7、6-8、7-9、8-10 样本对, 即区分波峰范围, 波峰较为平缓, 峰值位于 7-9 样本对, 为 54%, 与识别边界对应。载瓦语母语被试各样本对区分率组间差异不显著 ( $F(8,243)=0.945, p=0.480$ ), 不存在区分波峰, 峰值位于 7-9 样本对, 仅 45.24%。

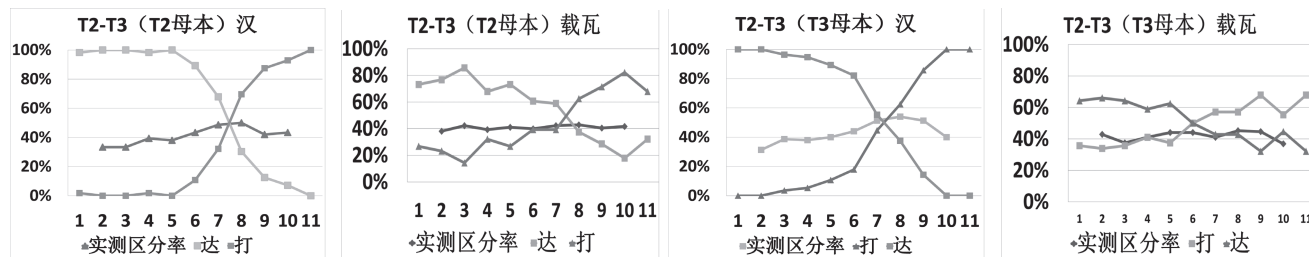


图7 T2-T3连续统感知曲线

图 8 是 T2-T4 声调对立组的感知结果。母本为 T2 时, 普通话母语被试的识别边界和宽度分别为 5.74、1.12, 载瓦语母语被试分别为 6.36、1.67。普通话母语被试区分波峰范围位于 3-5、4-6、5-7、6-8 区分对, 峰值出现于 4-6 样本对, 高达 92%, 与识别边界相对应。载瓦语母语被试区分波峰范围位于 3-5、4-6、5-7、6-8、7-9 样本对, 峰值出现于 6-8 样本对, 为 81.54%, 落在识别边界范围内。母本为 T4 时, 普通话母语被试的识别边界和宽度分别为 5.6、0.78, 载瓦语母语被试分别为 6.84、1.67。普通话母语被试区分波峰范围位于 3-5、4-6、5-7、6-8 样本对, 峰值出现于 4-6 样本对, 高达 92%, 与识别边界对应。载瓦语母语被试区分波峰范围位于 3-5、4-6、5-7、6-8、7-9、8-10 样本对, 峰值出现于 4-6 样本对, 为 64.88%, 未能对应识别边界。四组感知结果的波峰均为一个显著的平台状波峰。

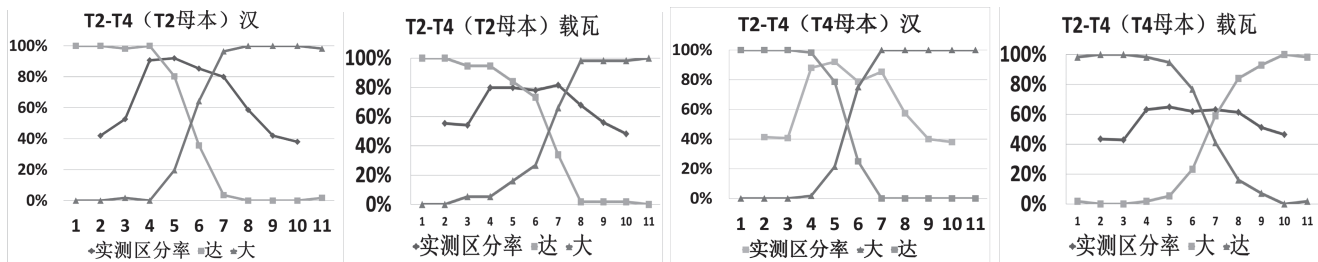


图8 T2-T4连续统感知曲线

图9为T3-T4声调对立组的感知结果。母本为T3时,普通话母语被试的识别边界和宽度分别为4.64、1.08,载瓦语母语被试分别为5.57、1.93。普通话母语被试区分波峰范围位于3-5、4-6样本对,峰值出现于3-5样本对,达76.67%,与识别边界对应。载瓦语母语被试区分波峰非常平缓,范围位于2-4、3-5、4-6、5-7、6-8这5个样本对,峰值位于3-5样本对,为63.1%,未能对应识别边界。母本为T4时,普通话母语被试的识别边界和宽度分别为5.61、2.57,载瓦语母语被试分别为6.72、1.91。普通话母语被试区分曲线组均值最高的为2-4、3-5、4-6、5-7、6-8这5个样本对,是一个较为平缓的波峰,峰值位于3-5样本对,为66%,偏离识别边界。载瓦语母语被试各样本对区分率组间差异不显著( $F(8,243)=0.722, p=0.672$ ),不存在区分波峰,峰值位于2-4样本对,仅为53.57%,完全不能对应识别边界。

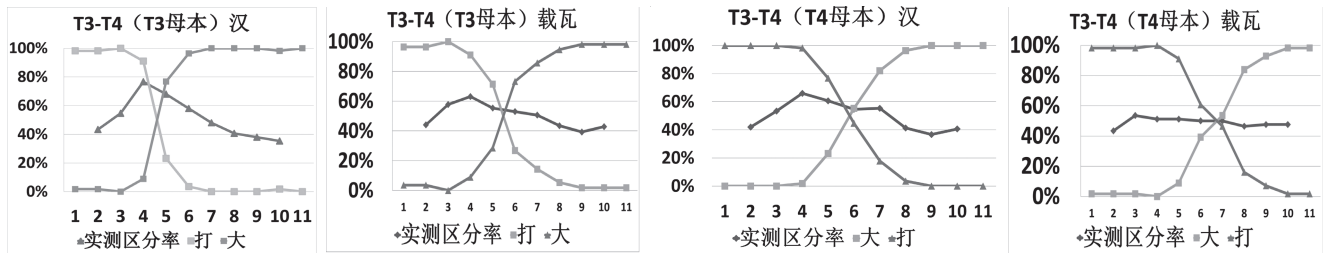


图9 T3-T4连续统感知曲线

## 四 讨论

### 4.1 载瓦语母语被试与普通话母语被试对普通话声调的感知模式对比

前文提到,学界普遍将语音感知模式二分为“范畴感知”与“连续感知”两种,但是,范畴感知与连续感知仅仅是行为学意义上的两种不同感知范式,不能由此确定被试是否能区分音位范畴。已有学者(刘掌才等2016)明确指出,无论是作为音段音位的元、辅音,还是作为超音段音位的声调,一个音位本身就应该是一个范畴,特定的音位在发音人的语言系统中应该都能够区别意义,在某种意义上它们就应该是感知范畴。根据实验数据,本文认为,从语言学的意义上来说,只要感知结果的识别曲线呈现较为清晰的“X”型,有陡峭的识别边界,即可反映出其能够区分两个对立的音位范畴。但从行为学的意义上来说,范畴感知模式除了需要有陡峭的识别边界,区分曲线也应当在识别边界处出现区分峰值。不过,识别宽度究竟窄到什么程度才算“陡峭”,区分曲线的波峰究竟该如何判定,目前学界并无定论,需要根据实验数据具体情况具体分析。

从本文对识别实验数据的统计来看,参照之前的研究(陆尧、孔江平2020),将识别宽度在0~2范围内定义为非常陡峭,2~4范围内属于较为陡峭,这两类都可以说明被试能够将一组声调对立组中的两个声调在音位学意义上区分为不同的范畴,而如果识别宽度超过4,就定义为识别边界不陡峭,被试无法区分两

个音位范畴。因此,本文数据显示,在T2-T3声调对立组中,当合成母本为T3时,载瓦语母语被试的识别宽度高达14.83,可见无法区分T2和T3这两个音位范畴,图中也可以观察到两条识别曲线靠得极近。

从本文对区分实验数据的统计来看,虽然被试的部分区分结果通过统计学的方法不能确定波峰范围,但可以观测到区分曲线的起伏及峰值所在处,结合统计及观察可见,普通话母语被试对普通话声调感知的区分曲线均有波峰的出现,只不过波峰的类型和显著性有所差别。而载瓦语母语被试对T2-T3(T2母本)、T2-T3(T3母本)、T3-T4(T4母本)这三组声调对立的区分曲线几乎为一条直线,统计数据和观察都显示没有波峰存在。

综上,普通话母语被试对普通话声调的感知模式均是典型的范畴感知,而载瓦语母语被试在行为学意义上对普通话T2-T3(T2母本)、T3-T4(T4母本)声调对立组的感知范式是连续感知,对其他各声调对立组的感知范式是范畴感知,对T2-T3(T3母本)声调对立组完全无法区分两个声调音位范畴。

#### 4.2 载瓦语声调系统对载瓦语母语者感知普通话声调的影响

载瓦语母语被试为何对T2-T3声调对立组的感知模式与其他各组显著不同,也与普通话母语被试有显著区别呢?我们认为,这显然受到了载瓦语自身声调系统的影响。普通话的四个调类与调值分别为:阴平55、阳平35、上声214、去声51。普通话阳平和上声这两个声调在调型上都有“升”的相似特征。也就是说,普通话声调中存在1个降调、1个平调与2个升调。将载瓦语声调系统与普通话声调系统进行对比,可以发现,载瓦语声调系统中存在与普通话类似的高降调和高平调,有1个升调35,但缺失类似上声214的调型。Best(1995)的知觉同化模型(Perceptual Assimilation Model,简称PAM)认为,二语习得者如果将两个非母语语音范畴同化为一个母语范畴(Single Category,简称SC),且这两个非母语语音与母语中一个语音相似度一致,那么对这两个非母语语音的区分能力就会较低,因此被试对于母语中本身缺失、目标语中又比较相似的两个声调T2和T3,自然难以感知。

更值得注意的是,当T2-T3声调对立组合成母本不同时,载瓦语母语被试的感知结果存在显著差异。由图7可见,母本为T2时,载瓦语母语被试是能够将T2和T3识别为两个不同的音位范畴的。但当合成母本是T3时,被试却几乎不能区分这两个音位范畴。这种感知结果差异的原因是什么?我们可以再从发声的角度来进行讨论。众所周知,普通话的发声虽然在语音学上不具有区别意义的功能,但在声调高低不同时,发声也存在变化,孔江平(2001)、Keating and Esposito(2006)等认为,普通话声调最重要的发声特征是上声在中间的低调部分会出现“挤喉噪音”(creaky voice)。前文也提到,载瓦语的元音有发声类型的对立,且与声调有一定的配合关系,紧噪音的发声类型与上声的挤喉噪音类似,声学上体现为高速度商。而且,载瓦语在35调中不会出现紧噪音,即在载瓦语母语被试的认知空间中,升调只有一个35调,且发声类型一定是正常噪音。在陆尧、孔江平(2019)进行的载瓦语声调感知研究中发现,载瓦语的基频和发声类型对载瓦语自身声调的感知都有贡献,都具有语言学意义,当发声类型缺失时,载瓦语母语被试无法区分两个声调音位范畴。所以,在本文的研究中,当T2-T3声调对立组的合成母本是正常噪音的T2调时,载瓦语母语被试可以清晰地识别出“基频+发声”均与其母语35调类似的T2调,很容易就将其与T3调区分。而当合成母本为带有挤喉噪音的T3时,虽然基频逐步变为35调,但发声类型却与载瓦语的35调不同,被试就无法区分T2和T3这两个声调音位范畴了。

再对照T1-T3组的感知结果来看,无论母本为何,识别边界都非常陡峭,当母本为T1时,基频从T1变化到T3,被试对T1的识别率从接近100%降至接近0%,对T3的识别率从接近0%升至接近100%,当母本为T3时亦是如此,被试能够清楚区分T1和T3。这是因为载瓦语的高平调中可以出现正常噪音,也可以出现紧噪音。所以当T1为母本时,被试很容易识别与区分正常噪音的高平调与正常噪音的升调,当T3

为母本时,挤喉嗓音的高平调也同样可以轻易被识别出来,将其与母语中不存在的挤喉嗓音 T3 调区分。这些感知结果可以证明,载瓦语母语经验中的发声类型和声调基频对感知结果都有贡献,且发声类型缺失时,就会影响对普通话的声调感知。

对于 T3-T4 (T4 母本) 声调连续统,被试的区分任务完成较差,没有找到区分波峰,本文认为也与 T3 的发声类型相关,具体原因还需要进一步分析研究。

### 4.3 载瓦语母语被试与普通话母语被试感知普通话声调范畴化程度对比

为了对比两组被试范畴感知的范畴化程度,本文对每组声调连续统中,普通话母语被试和载瓦语母语被试的识别宽度、区分峰值进行了 One-way ANOVA 单因素方差检验。由于不是范畴感知,T2-T3 (T2 母本)、T2-T3 (T3 母本)、T3-T4 (T4 母本) 声调对立组的数据未做进一步统计。被试的识别宽度越大,识别曲线越陡峭,区分峰值越高,范畴化程度越高。(Xu et al. 2006; 于谦 2017; Chen et al. 2017)

表2 载瓦语母语被试与普通话母语被试感知普通话声调范畴化程度对比

| 声调    | 母本 | 识别宽度                  | 区分峰值                  |
|-------|----|-----------------------|-----------------------|
| T1-T2 | T1 | 载瓦语 > 普通话, $p < 0.05$ | 载瓦语 < 普通话, $p > 0.05$ |
|       | T2 | 载瓦语 > 普通话, $p < 0.05$ | 载瓦语 < 普通话, $p < 0.05$ |
| T1-T3 | T1 | 载瓦语 > 普通话, $p < 0.05$ | 载瓦语 < 普通话, $p > 0.05$ |
|       | T3 | 载瓦语 > 普通话, $p < 0.05$ | 载瓦语 < 普通话, $p > 0.05$ |
| T1-T4 | T1 | 载瓦语 > 普通话, $p < 0.05$ | 载瓦语 < 普通话, $p > 0.05$ |
|       | T4 | 载瓦语 > 普通话, $p < 0.05$ | 载瓦语 < 普通话, $p > 0.05$ |
| T2-T4 | T2 | 载瓦语 > 普通话, $p > 0.05$ | 载瓦语 < 普通话, $p < 0.05$ |
|       | T4 | 载瓦语 > 普通话, $p < 0.05$ | 载瓦语 < 普通话, $p < 0.05$ |
| T3-T4 | T3 | 载瓦语 > 普通话, $p > 0.05$ | 载瓦语 < 普通话, $p < 0.05$ |

如表2所示,虽然部分数据不具有显著差异,但载瓦语母语被试的各组识别宽度均大于普通话母语被试,区分峰值均小于普通话母语被试,证明其感知范畴化程度弱于普通话母语被试,这说明了载瓦语的声调背景并没有提高被试感知普通话声调的范畴化程度。也就是说,即使是具备声调语言母语背景的被试,在习得另外一种声调语言时,也往往不能达到母语者的感知水平。

## 五 结论

本文通过声调感知实验,探讨了载瓦语母语者对普通话声调的感知情况。结果发现载瓦语自身的声调系统对载瓦语母语者感知普通话声调产生了显著影响,具体表现为:1. 载瓦语母语者对普通话大部分声调对立组的感知能达到近似普通话母语者的范畴感知模式,但范畴化程度小于普通话母语者;2. 载瓦语母语者在行为学意义上对 T2-T3 (T2 母本)、T3-T4 (T4 母本) 的感知模式是连续感知,对其他各组声调对立组的感知模式是范畴感知,且在合成母本为 T3 时完全无法区分 T2 和 T3 这两个声调;3. 当我们使用与载瓦语 35 调发声类型不同的母本 T3 合成到普通话 T2 (基频与载瓦语 35 调相同) 时,被试无法区分 T2 和 T3 这两个声调音位范畴,可见载瓦语中的发声类型和基频都对载瓦语母语者感知普通话声调有着重要贡献,如果发声类型缺失,就会影响对普通话的声调感知。只有同时具备“基频+发声”的相似特征,载瓦语母语者才能较好地识别与母语声调对应的普通话声调。以上研究表明,除了基频模式,声调中的发声类型对二语声调感知与习得也具有重要的语言学意义。

## 参考文献

- 戴庆厦 (1989) 载瓦语声调研究, 《中央民族学院学报》第 1 期。
- 高云峰 (2004) 《声调感知研究》, 上海师范大学博士学位论文。
- 何勒腊 (2016) 《载瓦语语音研究》, 上海师范大学博士学位论文。
- 康 宇、吴辰琛、黄丽芳、李 利 (2017) 声调语言经验对第二语言声调习得的影响, 《现代语言学》第 5 期。
- 孔江平 (1992) 紫云苗语五平调系统的声学及感知研究, 马学良主编《民族语文研究新探》, 四川民族出版社, 成都。
- 孔江平 (1995) 藏语 (拉萨话) 声调感知研究, 《民族语文》第 3 期。
- 孔江平 (2001) 《论语言发声》, 中央民族大学出版社, 北京。
- 李洪彦、蓝庆元、孔江平 (2006) 壮语龙州话声调的声学分析, 《民族语文》第 6 期。
- 刘掌才、秦 鹏、石 锋 (2016) 汉语普通话基础元音的听感格局初探, 《南开语言学刊》第 1 期, 商务印书馆, 北京。
- 陆 尧、孔江平 (2019) 载瓦语声调的声学及感知研究, 《民族语文》第 1 期。
- 陆 尧、孔江平 (2020) 藏语拉萨话母语者普通话声调感知研究, 《语言学论丛》第六十二辑, 商务印书馆, 北京。
- 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》第 9 期。
- 王韞佳、李美京 (2011) 韩语母语者对普通话阳平和上声的知觉, 《语言教学与研究》第 1 期。
- 徐悉艰、徐桂珍 (1984) 《景颇族语言简志 (载瓦语)》, 民族出版社, 北京。
- 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。
- Abramson, A. S. (1979) The noncategorical perception of tone categories in Thai. In Lindblom, B. and S. Ohman (eds.) *Frontiers of Speech Communication Research*, 91-100. London: Academic Press.
- Best, Catherine T. (1995) A direct realist view of cross-language speech perception. In Winifred Strange (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, 171-204. Timonium, MD: York Press.
- Chen, Ao (2012) *Universal Biases in the Perception of Mandarin Tones: From Infancy to Adulthood*. Amsterdam: LOT.
- Chen, Fei, Peng, Gang, Yan, Nan and Wang, Lan (2017) The development of categorical perception of Mandarin tones in four-to seven-year-old children. *Journal of Child Language* 44(6): 1413-1434.
- Francis, Alexander L., Ciocca, Valter and Ng, Brenda Kei Chit (2003) On the (non) categorical perception of lexical tones. *Perception and Psychophysics* 65(7): 1029-1044.
- Francis, Alexander L., Ciocca, Valter, Ma, Lian and Fenn, Kimberly (2008) Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics* 36: 268-294.
- Fry, D. B., Abramson, A. S., Eimas, P. D. and Liberman, A. M. (1962) The identification and discrimination of synthetic vowels. *Language and speech* 5(4): 171-189.
- Hallé, Pierre A., Chang Yueh-Chin and Best, Catherine T. (2004) Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics* 32(3): 395-421.
- Keating, Patricia and Esposito, Christina (2006) Linguistic voice quality. *Paper presented at the 11th Australasian International Conference on Speech Science and Technology*, from <http://www.linguistics.ucla.edu/people/keating/keating.htm>.
- Liberman, A. M., Harris, K. S., Hoffman, H. S. and Griffith, B. C. (1957) The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* 54(5): 358-368.

- Lieberman, A. M., Harris, K. S., Kinney, J. A. and Lane, H. (1961) The discrimination of relative onset-time of the components of certain speech and nonspeech patterns. *Journal of Experimental Psychology* 61(5): 379.
- Peng, Gang, Zheng, Hong-Ying, Gong, Tao, Yang, Ruo-xiao, Kong, Jiang-ping and Wang, William S.-Y. (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.
- Qin, Zhen and Mok, Peggy (2011) Discrimination of Cantonese tones by Mandarin, English and French speakers. *International Congress of Phonetic Sciences*: 17-21.
- Schouten, M. and van Hessen, A. (1992) Modeling phoneme perception. I: Categorical perception. *The Journal of the Acoustical Society of America* 92(4):1841-1855.
- Studdert-Kennedy, M., Liberman, A. M., Harris, K. S. and Cooper, F. S. (1970) Motor theory of speech perception: A reply to Lane's critical review. *Psychological Review* 77(3): 234-249.
- Wang, William S.-Y. (1976) Language change. *Annals of the New York Academy of Sciences* 280(1): 61-72.
- Wang, Xinchun (2006) Perception of L2 tones: L1 lexical tone experience may not help. From [https://www.iscaspeech.org/archive/speechprosody\\_2006/wang06b\\_speechprosody.html](https://www.iscaspeech.org/archive/speechprosody_2006/wang06b_speechprosody.html).
- Wang, Xinchun (2013) Perception of Mandarin tones: The effect of L1 background and training. *The Modern Language Journal* 97(1): 144-160.
- Wayland, Ratree P. and Guion Susan G. (2004) Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Language Learning* 54(4): 681-712.
- Xu, Yisheng, Gandour Jackson T. and Francis Alexander L. (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 Categorical Perceptual Study on Mandarin Tones by Zaiwa Language Speakers

LU Yao

**Abstract:** Mandarin is the national language of our country. However, it is difficult for many speakers with ethnic language backgrounds to learn the tones of Mandarin. This paper intends to investigate how the speakers of the Zaiwa language perceive Mandarin tones. Identification and discrimination experiments were conducted of the four Mandarin tone continuums with Zaiwa language speakers. The results suggest that Zaiwa speakers cannot establish a categorical perception for the continua of T2-T3(T2) and T3-T4(T4); nor are they able to clearly divide the T2-T3 continuum into two phoneme categories when the original synthesis sample was T3. On the other hand, they perceive other continua categorically, although the degree of categorization is much weaker than that observed with native Mandarin speakers. The above findings suggest that phonation and F0 should be considered crucial cues for the categorical perception of tone. Lack of phonation greatly affects the perception of Mandarin tones. This finding is enlightening. The above research indicates that, in addition to the fundamental frequency, the phonation type also holds significant linguistic significance for the perception and acquisition of second language tones.

**Keywords:** Mandarin tones, Zaiwa language tones, categorical perception, tone acquisition