

中国学习者语调范畴感知学习机制探究

纪晓丽, 张 辉, 李爱军

摘 要: 本文借助实验语音学手段合成实验语料, 通过语调范畴感知学习实验, 考察不同记忆系统对中国学习者英语语调范畴感知学习的影响, 探索二语者语调范畴感知的心理机制。结果发现: 1) 经过语调范畴感知训练, 学习者可以形成对应的语调范畴, 1ST 变化下的测试新语料和训练中出现的语料感知结果无差异, 但该结果受到音高扰动变化、音高重音个数和位置变化的影响; 2) 训练后的语调范畴感知结果和程序性记忆、工作记忆正相关, 只有下降调的范畴感知结果和陈述性记忆负相关。本文发现两点与 COVIS 模型和 DP 模型不一致的地方: 第一, 学习者的语调范畴感知学习倾向于基于信息整合的范畴学习, 确实依赖于程序性记忆。但和 COVIS 模型不同, 学习者语调范畴感知学习也会受到工作记忆的影响, 工作记忆中的注意可以帮助学习者提取与范畴感知学习有关的信息, 抑制无关信息; 第二, 和 DP 模式句法训练结果不同, 学习者在课堂中得到的显性指导, 使学习者对边界调的范畴感知学习效果要好过音高重音, 二语语调的范畴感知学习, 可能需要的是一种显性指导和隐性训练结合的方式, 但这一点需要进一步的实验验证。

关键词: 语调范畴感知学习; 陈述性记忆; 程序性记忆; 工作记忆

中图分类号: H319 **文献标识码:** A **文章编号:** 1000-5544(2023)01-0030-09

Abstract: Using one English sentence resynthesized with different intonation contours, the paper investigates rapid perceptual learning of intonation contour categories by Chinese EFL learners and examines the relationships between procedural memory, declarative memory, working memory and L2 learners' intonation-category learning to explore L2 learners' intonation category learning mechanism. It is found that: 1) After a short intonation category training, the learners could form abstract representation of intonation contours; previous encountered and novel exemplars are categorized equally when the pitch disturbance is 1 semitone away from the prototype, yet the result can be affected by the degree of pitch disturbance, the number and location of pitch accent; 2) The standard, linear multiple regression results prove procedural memory and working memory to be statistically significant predictors of intonation category learning while declarative memory has been found to be only negative correlated with the final-fall prototype. This paper postulates two different proposition from COVIS and DP model: firstly, L2 learners' intonation category learning inclines to use information-integration strategy which relies on procedural memory as COVIS proposed. However, this process is also mediated by working memory which can help learners to extract related information and inhibit unrelated information for category learning. Secondly, unlike the syntax training under DP model, the explicit instruction of boundary tones in classroom grants an advantage to category learning of boundary compared to category learning of pitch accent, which indicates intonation training may need a combination of explicit instruction and implicit training.

Key words: perceptual learning of intonation contour category; declarative memory; procedural memory; working memory

1. 引言

范畴学习 (category learning) 是人类一项重要的基本认知技能, 婴儿在不到一岁的时候就开始学会识别不同的音段范畴, 进而才能开始习得单词的意义 (Kuhl *et al.* 1992)。国内外学者对二语音段的范畴习得进行了深入的研究 (Diaz *et al.* 2012; Quam *et al.* 2018), 并提出了相关理论 (Flege 1995; Best 1995; 陈莹 2013)。近年来, 有少量研究开始考察超音段层面的范畴习得, 但多以声调为主 (Francis *et al.* 2008; 张林军 2010; 姚尧、陈晓湘 2021), 少有实证关注二语者的语调范畴感

知学习。

然而, 语调一直是中国学习者英语习得中的难点, 中国学习者的英语语调在调群切分、句重音位置、调型等方面都存在问题 (陈桦 2006; 杨军 2006; 孟小佳、王红梅 2009)。有些研究在考察中国学习者语调习得的影响因素, 如二语水平 (纪晓丽等 2018)、习得年龄 (薛锦等 2019) 和性别 (林秋茗 2011) 等; 有些研究则开始探索改进学习者语调的方法, 如输入、输出频次 (陈桦、孙欣平 2010)、反复聆听模仿 (杨晋 2010)、讲解+跟读模仿 (高薇等 2015) 和模仿朗读 (田方 2018) 等, 但少有研究探讨学习者语调习得背后的心理机制。鉴于此, 本

文通过语调范畴感知学习实验,考察不同类型记忆对二语语调习得的影响,探究二语语调习得的深层机制。

2. 文献综述

有关范畴感知学习, Ashby 等人基于神经生物学证据提出 COVIS (The Competition Between Verbal and Implicit Systems Model) 模型 (Ashby *et al.* 1998; Ashby & Maddox 2005; Ashby & O'Brien 2005), COVIS 模型认为,不同类型的范畴学习依赖两种不同的系统:基于规则的范畴学习 (optimal rule-based category learning) 和基于信息整合的范畴学习 (optimal information-integration category learning)。人们通过前者学习的规则往往可以用言语总结出来,学习的过程是有意识的,以显性的、反应性的系统 (an explicit, reflective hypothesis-testing system) 为基础,该系统依赖于工作记忆和执行注意 (executive attention), 涉及到的脑回路包括背外侧前额皮层 (dorsolateral prefrontal cortex)、前尾状核 (anterior caudate nucleus)、前扣带回 (anterior cingulate) 和内颞叶结构 (medial temporal lobe); 而通过后者学习的内容则涉及两个或两个以上维度的信息整合,难以用言语总结规则或规律,学习的过程是无意识的,以隐性的、反射性的、程序性学习系统 (an implicit, reflexive, procedural-based learning system) 为基础,涉及的脑结构主要为后尾状核 (posterior caudate nucleus) 和壳核 (putamen)。COVIS 模型提出,两种系统有自己独立、又彼此交互的神经基础,在范畴学习的过程中,两者之间互相竞争,最后其中一种获得控制权。比如说,基于信息整合的范畴学习,在学习的初期,人们会倾向于依赖外显系统,经过训练和练习后,逐渐转向内隐系统。

COVIS 模型一开始多用来解释视觉范畴学习, Maddox *et al.* (2013) 将其拓展到听觉范畴学习,考察了不同年龄的被试对汉语声调的范畴感知学习,结果发现在训练结束时,使用内隐的、反射性学习策略的学习者的正确率要高于使用外显的、反应性学习策略的学习者,成功的学习者在一开始的时候也使用反应性学习策略,在训练后期开始转为使用反射性学习策略。另外,该研究还对受试的工作记忆和程序性记忆进行测量,考察这两种记忆和两种范畴学习类型的关系,结果显示,依赖外显系统的反应性学习消耗资源,对工作记忆和执行注意要求高;但依赖内隐系统的、基于

信息整合的范畴学习则不受工作记忆的影响,而是和程序性记忆相关。

尽管 Maddox 团队最近才开始将 COVIS 模型拓展到二语习得领域,记忆对习得的影响一直是二语习得研究的热点问题,其中 Ullman 团队提出的陈述/程序性模式 (Declarative/Procedural Model, 以下简称 DP 模式) 也涉及了两种不同的记忆和学习系统:陈述性记忆和程序性记忆 (张辉、卞京 2016)。Ullman (2020) 将陈述性记忆定义为以内部颞叶区域 (medial temporal lobe regions) 和相关脑回路为基础的学习和记忆,负责学习任意性的符号关系,是学习、表征和使用事实 (semantic knowledge 语义知识) 和事件 (episode knowledge 情景知识) 的基础。程序性记忆 (procedural memory) 则是以基底神经节 (basal ganglia) 和相关脑回路为基础的学习和记忆,该系统包括连接基底神经节和前额叶回路的多个脑区网络,主要负责习惯 (habits)、感知运动技能 (perceptual-motor skills)、感知序列 (perceptual sequences)、范畴 (categories) 和路线 (routes) 的学习和加工。陈述性记忆学习的知识可以是显性的,也可以是隐性的;程序性记忆学习的知识只能是隐性的。DP 模式认为句法是隐性的,基于此, Ullman 的团队考察了显隐性训练对句法习得的影响,结果发现,隐性训练组的学习者对句法的加工可以呈现和母语者类似的模式,也就是说,基于程序性记忆的知识通过隐性训练的习得效果较好 (Morgan-Short *et al.* 2012)。

随着 DP 模式的发展,该模式开始对心理词库 (lexicon) 和句法以外的知识,如言语—声音表征 (speech-sound representation)、发声和言语产出 (articulation and speech production) 和言语感知 (speech perception) 等语音相关的能力进行预测,提出语调感知和产出依赖程序性记忆。但到目前为止, Ullman 团队并未进行语调方面的实证实验,仅基于 Maddox 团队的核磁共振结果 (Yi *et al.* 2014; Ullman *et al.* 2020; Arthur *et al.* 2021) 提出,学习者范畴学习的成功取决于从反应性系统到反射性系统的转换, Ullman *et al.* (2020: 389) 认为,这就是从陈述性记忆到程序性记忆的转换,也就是说,反应性系统和陈述性记忆相关。

有关听觉范畴感知学习的研究主要集中在音段和声调方面 (Francis *et al.* 2008; Diaz *et al.* 2012; Maddox *et al.* 2013; Quam *et al.* 2018), 目前,还没有实验考察不同记忆系统对二语者英语语调的范畴学习机制的影响。根据 AM 理论,英语语调

由音高重音 (pitch accent)、短语重音 (phrase accent) 和边界调 (boundary tone) 三种音高事件组成,这三种音高事件都可以用高 (H)、低 (L) 两种相对音高值来描述 (Pierrehumbert 1980; 马秋武、贾媛 2009)。Ladd & Morton (1997) 实施了一系列实验,讨论英语母语者对语调的重音感知是范畴化的还是连续的。结果证明母语者在语用的作用下,对重音的解读是范畴化的。Kapatsinski *et al.* (2017) 考察了美国成年人和儿童对英语语调的范畴感知,结果显示,成年人和儿童都能形成语调范畴,儿童的范畴比成年人宽,他们推测可能是因为儿童工作记忆低于成年人,无法整合多维的信息来识别语调。还有学者认为从语音、语调以及语言结构上看,英语本质上是一种“听音知意”的听觉语言,具有适听特质 (苏锦平 2022)。国内外有少量研究考察二语者的英语语调感知,但主要通过不同语言背景受试者的对比,来探讨语调感知是具有跨语言普遍性,还是受到受试母语的影响 (Mennen *et al.* 2012)。到目前为止,还没有看到相关研究对二语者英语语调范畴感知进行考察。二语学习者能否形成语调范畴,如果可以的话,二语语调的范畴感知学习依赖哪种记忆是本研究将要探讨的问题。

3. 实验方法

3.1 受试

43 名 (18 男, 25 女) 中国英语学习者参与本实验,为了排除语言背景和语言学习经历的影响,所有的受试都是来自吴方言区的大一非英语专业学生,年龄在 18~20 之间 ($M = 18.95$),没有受过专业的音乐或英语语音训练,也没有超过国外一个月的旅居史。所有的受试均通过了四级考试 ($M = 541.9$, $SD = 42.0$),自述没有听说问题,完成实验后受试获得一定的报酬。

3.2 实验刺激

本实验原始语料取自语料库“Online Speech Accent Archive”中一名男性美国母语者朗读的语料 (Weinberger 2013),时长 3.43 秒,用 MBROLA 将句子中的辅音和元音分别用 /m/ 和 /i/ 代替,排除句子意义和节奏特征的影响 (Dutoit *et al.* 1996)。然后,将这句话的语调曲拱进行调整,合成三种语调范畴原型 (prototype): 第一种是平调 (flat prototype),音高曲拱的基频保持在 250 赫兹左右不变;第二种是下降调 (final fall prototype),音高曲拱前半段和平调一样,基频保持在 250 赫兹,

从第 11 个音节开始下降,到句尾下降至 110 赫兹;第三种是 M 调 (multifeature “M” prototype),音高曲拱从 250 赫兹开始缓慢上升,到第四个音节上升至 350 赫兹,然后开始下降,到第 6 个音节下降至 215 赫兹,之后保持不变到第 11 音节,接着曲拱再次上升,到第 13 个音节上升到 340 赫兹,最后降到 175 赫兹 (Kapatsinski *et al.* 2017)。下降调代表了英语陈述句边界调的语调曲拱变化;M 调代表了英语陈述句有两个音高重音时的语调曲拱变化 (Pierrehumbert 1980),平调不是英语中会出现的语调,主要用于做对比参考。

训练刺激用 Praat5.314,是基于上述三种语调范畴原型生成的 (Boersma & Weenink 2001),把三种语调范畴原型中部分元音的音高值进行扰动变化,上升或下降 1 个半音 (以下简称 1ST),生成语调变体。每个变体 (exemplar) 都会有一个镜像变体,音高扰动变化和其原变体完全对称,每种范畴生成 6 对变体。

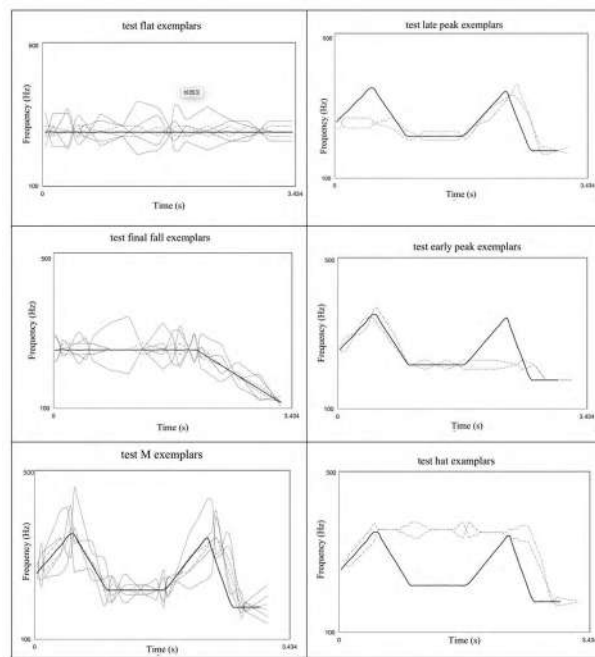


图 1. 测试刺激示例

测试刺激包括了三种范畴的原始音频、每个范畴中两对在训练中出现的 1ST 变体、两对训练中没有出现的 1ST 变体、两对 3 个半音 (以下简称 3ST) 扰动的变体和两对 5 个半音 (以下简称 5ST) 扰动的变体 (见图 1 左栏)。除此之外,测试语料还增加了由音高重音变化生成的 M 调变体,只保留 M 调中的左音高重音 (early peak, 简称左重) 和右音高重音 (late peak, 简称右重) 以及去掉两个音高重音间波谷的 Hat 调,对这些变体也进行 1ST 扰

动变化,每种生成两对 M 调变体(见图 1 右栏)。

3.3 实验流程

所有受试首先完成语言背景调查问卷,对受试的性别、年龄、英语学习经历、四级分数进行调查。然后在语音消声室完成语调训练实验、语调测试实验和认知测试。

语调训练实验时,受试被告知会听到三个外星人的声音,受试的任务是记住外星人说话的特征。受试在了解实验流程后开始训练。首先,受试在屏幕上看到“+”号,然后听到随机播放的训练语料,同时看到和刺激对应的外星人图片,两个刺激间隔 500 毫秒,每个刺激重复出现 3 遍。

训练结束后,受试马上进行语调测试实验。在测试实验中,受试听到每个测试刺激后,需要选择是哪位外星人在说话,如果受试认为听到的刺激不是在训练中听到的外星人,可以选择图片中的“其他外星人”(见图 2),每个刺激也是重复出现 3 遍。



图 2. 测试实验选择界面

认知测试包括三部分:工作记忆、陈述性记忆和程序性记忆。工作记忆使用南京师范大学心理学院开发的 MindTest 软件中的中文版阅读计算广度测试(Stine & Hindman 1994)。受试在电脑屏幕上看到一系列中文句子,其中一半句子是正确的,如“树叶有很多形状”,一半句子是错误的,如“狗有四只眼睛”,受试需要尽快判断句子是否正确,同时记住句尾双字词,在一定间隔后,受试需要复述之前看到句子的句尾词。受试正确复述句尾词的个数用来计算被试的工作记忆水平。陈述性记忆使用的是连续性视觉记忆测试,受试会在屏幕上看到一系列复杂抽象的图片,其中的 7 张图片会重复出现 7 次,和 63 张干扰图片一起随机出现在屏幕上,每张图片出现时间为 2 秒,受试需要判断屏幕上的图片是看到过的还是没有看到过的,判

断正确率用来计算受试的陈述性记忆分数(Larabee 2009)。程序性记忆测试使用交替序列反应任务。电脑屏幕上会出现四个空心圆圈,其中一个圆圈会随机(随机模式)或按照一定规律(固定模式)变为实心黑色圆圈,随机模式和固定模式交替出现,受试需要尽快按下实心黑色圆圈对应的按键,实验结束后,受试对固定模式和随机模式反应时的差异代表了受试的程序性记忆水平(Howard & Howard 1997)。

表 1. 学习者对不同音高扰动变化的语调范畴变体感知结果

	平调				下降调				M 调			
	平调	下降调	M 调	其他	平调	下降调	M 调	其他	平调	下降调	M 调	其他
TR1	0.651	0.035	0.238	0.076	0.076	0.843	0.064	0.017	0.273	0.163	0.512	0.052
1ST	0.616	0.029	0.227	0.128	0.064	0.831	0.064	0.041	0.256	0.134	0.523	0.087
3ST	0.477	0.047	0.203	0.273	0.064	0.547	0.093	0.297	0.174	0.122	0.337	0.366
5ST	0.302	0.029	0.203	0.465	0.058	0.407	0.099	0.436	0.105	0.047	0.140	0.709

表 2. 学习者对 1ST 变体下的所有语调范畴变体的范畴感知结果

范畴	平调	下降调	M 调	左重	右重	Hat 调
比例						
平调	0.616	0.064	0.256	0.308	0.279	0.134
下降调	0.029	0.831	0.134	0.076	0.163	0.110
M 调	0.227	0.064	0.523	0.285	0.372	0.302
其他	0.128	0.041	0.087	0.331	0.186	0.453

4. 实验结果

表 1 和表 2 分别展示了音高扰动变化和音高重音变化的感知结果,其中 TR1 指的是在训练中出现过的,音高扰动为 1ST 的测试语料,1ST 是训练中没有出现过的,扰动为 1ST 的测试语料,3ST 是扰动为 3ST 的测试语料,5ST 是扰动为 5ST 的测试语料。比例一栏是所有受试将某一测试刺激归为某一个选项的比例,以表 1 中的第一行第一列的 0.651 为例,指的是所有受试听完 TR1 情况下的平调后,将其归为平调的比例是 65.1%,比例中的“其他”选项,即受试将听到的声音归为训练中没有出现过的、新的语调范畴的比例。

表 1 中的灰色区域是语调范畴被归入原语调范畴的比例,可以看出,平调、下降调和 M 调的语调范畴感知结果受到音高扰动变化大小的影响。在音高扰动是 1 个半音时,不管是训练中出现过的旧语料(TR1),还是没有出现过的新语料(1ST),三种语调范畴被归入原范畴的比例都在 50% 以上。当音高扰动是 3 个半音时,只有下降调被归入原范畴的比例还在 50% 以上,三种语调范畴均有

30%左右的变体被归到新语调范畴。当音高扰动变化到5个半音时,三种语调范畴被归入原范畴的比例都降到50%以下,被归入新范畴的比例增加,大于归入原范畴的比例。

另外,从表1还可以看出,三种语调范畴的感知结果也有很大的不同,无论音高扰动变化大小,下降调被归入原范畴的比例都是最高的,平调次之,M调最低。表2详细对比了音高重音变化对语调范畴感知的影响,灰色区域是受试选择比例最大的选项,平调、下降调和M调被归为原范畴的比例都在50%以上,下降调被归为原范畴的比例高达83.1%。左重、右重和Hat调都是由M调音高重音变化生成的变体,它们被受试归为M调的比例都在40%以下,Hat调和左重被归入新范畴的比例超过了原范畴,左重和右重还有30%左右被归入了平调。

首先,我们以各语调范畴被归入到原范畴的比例为因变量,语调范畴类型(平调、下降调和M调)和音高扰动变化水平(TR1、1ST、3ST和5ST)为自变量做重复测量方差分析。结果显示,范畴类型的主效应显著, $F(2,41)=23.726, p<0.001$,音高扰动水平主效应显著, $F(3,40)=29.755, p<0.001$,范畴类型和音高扰动水平的交互效应不显著, $F(6,37)=0.742, p=0.619>0.05$ 。Bonferroni多重比较结果显示,范畴类型变化方面,两两范畴之间的差异显著, p 值均小于0.001,音高扰动变化方面,TR1和1ST之间差异不显著($p=1.000>0.05$),其他音高扰动变化之间均存在显著差异。

接着,我们以M调变化类型为自变量,以被归为M调的比例为因变量,进行重复测量方差分析,结果显示,语调范畴变化主效应显著, $F(3,40)=7.975, p<0.001$ 。Bonferroni多重比较结果显示,M调和三种变体的结果都存在显著性差异,三种变体两两之间差异不明显。

为了考察语调范畴感知和记忆的关系,我们对归一后的工作记忆、陈述性记忆、程序性记忆做多重线性回归。基于表1和表2的结果,我们考察音高扰动变化的语调范畴感知时,以不同音高扰动下的平调、下降调和M调被归为原范畴的比例为因变量,以工作记忆、陈述性记忆和程序性记忆为自变量,建立多重线性回归模型。

在考察音高重音变化的语调范畴感知时,以M调和M调变体被归为M调的比例为因变量,以工作记忆、陈述性记忆和程序性记忆为自变量,建立多重线性回归模型。

表3. 各语调范畴回归模型的描述性数据

语调范畴类型	Adjusted R^2	F	p 值	D-W Statistic	工作记忆 VIF	陈述性记忆 VIF	程序性记忆 VIF
平调	0.017	6.039	<0.001***	1.767	1.084	1.034	1.056
下降调	0.011	4.209	0.006**	1.538	1.084	1.034	1.056
M调	0.020	6.796	<0.001***	1.568	1.084	1.034	1.056
左重	0.041	3.451	0.018**	1.802	1.084	1.034	1.056
右重	0.036	3.143	0.027***	1.580	1.084	1.034	1.056
Hat调	0.079	5.883	<0.001***	1.846	1.084	1.034	1.056

表4. 工作记忆、陈述性记忆和程序性记忆对各语调范畴的回归系数表

语调范畴类型	工作记忆			陈述性记忆			程序性记忆		
	β	t 值	p 值	β	t 值	p 值	β	t 值	p 值
平调	0.072	4.108	<0.001***	0.002	0.094	0.925	0.003	0.182	0.855
下降调	0.032	1.962	0.050**	-0.038	-2.380	0.018**	0.037	2.314	0.021**
M调	0.074	4.309	<0.001***	0.001	0.059	0.953	0.035	2.071	0.039**
左重	0.070	1.998	0.047**	0.033	0.950	0.343	0.092	2.639	0.009***
右重	0.093	2.474	0.014	0.019	0.518	0.606	0.079	2.128	0.035**
Hat调	0.123	3.506	<0.001***	-0.013	-0.385	0.670	0.104	3.022	0.003***

从表3可以看出,尽管 R^2 的值不高,但是各语调范畴的多重线性回归模型的 p 值都小于0.05,说明各回归模型具有统计学意义。各回归模型的德宾—沃森值均在1到3之间,说明残差不存在自相关。另外,从COVIS模型和DP模式来看,工作记忆和陈述性记忆之间关系紧密,可能存在一定的关系。我们用方差膨胀因子(VIF)来检测各回归模型是否存在多重共线性的问题,结果显示,工作记忆的VIF值为1.084,陈述性记忆的VIF值为1.034,程序性记忆的VIF值是1.056,均在1—5之间,这说明工作记忆、陈述性记忆和程序性记忆是三个独立的自变量,彼此之间不存在线性相关关系(张文彤2004)。

从表4可以看出,各语调范畴回归模型主要分为三种:第一种,平调的回归模型,该模型中只有工作记忆有统计意义,标准化回归系数 $\beta=0.072$;第二种,下降调的回归模型,工作记忆、陈述性记忆和程序性记忆对下降调的感知都有影响意义,其中工作记忆系数和程序性记忆的系数为正,而陈述性记忆的系数是负数,说明下降调的感知与工作记忆、程序性记忆正相关,和陈述性记忆负相关;第三种,M调及其变体,其中M调、左重和Hat调结果都与工作记忆和程序性记忆正相关,右重只和程序性记忆正相关。

5. 讨论

5.1 中国学习者能否形成语调感知范畴

正如前文所述,语调由音高重音、短语重音和

边界调三种音高事件组成,一句话的音高曲拱除了上述三种音高事件外,其他部分的音高扰动变化对语调的意义影响不大,也就是说,我们对语调范畴进行的音高扰动变化,只要没有改变其音高重音和边界调(图 1 左栏),原则上就还属于原语调范畴。在实验中,我们对平调、下降调和 M 调三种语调范畴进行音高扰动的变化,结果发现在音高扰动变化是 1 个半音时,学习者对训练中出现过的语调(TR1)和没有出现过的新语料(1ST)的感知结果没有显著性差异,也就是说学习者可以形成对应的语调范畴,将平调、下降调和 M 调变体归入原语调范畴,并没有因为 1ST 语料和 TR1 语料音高扰动变化位置的不同,将 1ST 语料归入新的语调范畴。随着音高扰动变化增大,学习者将测试语料归为新范畴的比例逐渐增加,当音高扰动变化增加到 5 个半音时,学习者将测试语料归为新范畴的比例超过了归为原范畴的比例,但学习者并未在这三种语调范畴内部产生太多的混淆,前文我们已经提到,下降调和英语陈述句边界调的音高走势相似,M 调和英语陈述句音高重音的音高曲拱相似,也就是说,学习者基本能够对英语的边界调和音高重音形成语调范畴,但结果受到音高扰动变化的影响。

在实验中,我们还设计了由 M 调音高重音位置和个数变化生成的 M 调变体。M 调有两个音高重音,且两个音高重音的音高上升幅度相当(见图 1 右栏),按照 Gussenhoven (1983) 的焦点即重音(Focus-to-Accent)理论,核心重音标记的即是句子的焦点,M 调的焦点在第二个音高重音位置。Hat 调保留了 M 调的两个音高重音,且位置不变,但是去掉了两个重音间的波谷,改变了音高重音的实现方式,生成了英语语调中的 Hat 调型,Hat 调在英语中常常被称为呼调,用于从远处呼叫某人或者在演讲开始引起大家的注意等(Ladd 1996)。左重和右重都只保留了 M 调的一个音高重音,但是左重保留了 M 调的第一个音高重音,右重保留了 M 调的第二个音高重音,也就是说,左重的焦点在 M 调第一个音高重音,右重的焦点则和 M 调一样。从语调意义和语法句式的关系来看,左重和右重都是陈述句语调,可以和 M 调归为一个语调范畴,Hat 调是呼调,和 M 调不属于一个语调范畴。从语调的语义功能来说,左重和 M 调焦点位置不同,是新的语调范畴,右重和 M 调焦点位置一致,可以将右重和 M 调归为一个语调范畴(卜友红 2016)。

从表 2 可以看出,学习者将 Hat 调和左重归为

新范畴的比例最高,将右重归为 M 调的比例最高,学习者对 M 调变体的感知似乎是参考了语调的语义功能,但将表 1 和表 2 对比发现,Hat 调有 45.3% 被归为“其他”,有 30.2% 被归为 M 调;3ST 音高扰动变化的 M 调有 36.6% 被归为“其他”语调范畴,33.7% 被归为 M 调,音高重音实现方式改变的 M 调变体和 3ST 音高扰动变化后的 M 调变体感知结果类似,有可能学习者认为音高重音实现方式的变化和音高扰动变化类似,学习者并未考虑音高重音实现方式的改变引起的句子意义上的改变。在我们前期语调感知实验中,发现母语者对音高重音的感知基于核心重音的位置,核心重音的位置和句子的焦点相关,而低水平中国学习者对音高重音的感知依赖音高重音的个数(纪晓丽等 2018)。因此,有可能学习者对 M 调变体的感知结果仍然受到音高重音的个数的影响。Schafer (2018) 考察韵律焦点对日韩英语学习者代词指代加工的影响,结果发现,母语者能够使用句子中的体和韵律焦点信息确定先行词,但具体表现不同,这可能是因为学习者没有将韵律焦点和更高层次的语义联系起来,而是因为韵律焦点能留下更强的记忆痕迹(memory trace),增加注意(attention),从而影响学习者对先行词的选择。这也间接证明我们的推测,学习者在形成语调范畴时,音高重音的变化增加了学习者的注意,但是学习者并没有充分考虑这些音高曲线的变化与语调意义和功能的关系。

5.2 语调范畴感知和记忆的关系

根据 COVIS 模型,成功的语调范畴感知应该是基于信息整合的范畴学习,依赖程序性记忆,不受工作记忆的影响。DP 模型也推测,语调的产出和感知依赖程序性记忆。我们实验结果发现,学习者三种语调范畴在训练后的范畴感知结果确实都和程序性记忆正相关,这说明中国学习者对英语语调的范畴感知是基于信息整合的范畴感知,依赖于程序性记忆,这可能是因为语调,尤其是音高重音的规律,很难用言语总结出来,而且语调变化的维度比较多,包括了音高重音和边界调的变化,即使只考虑音高重音的变化,也包括了个数、位置等不同的变化,在相同的音高个数和位置的情况下,还有不同水平的音高扰动变化和音高重音实现方式的变化,因此学习者在训练后对英语语调的范畴感知主要是基于信息整合的范畴感知,与 COVIS 模型和 DP 模式的预测一致。

但是,本实验也有和 COVIS 模型不符的地方,

COVIS 模型提出,基于信息整合的范畴学习不受工作记忆的影响。本实验回归结果表明,除了右重调外,下降调、M 调及其变体在训练后的范畴感知结果都和工作记忆关系密切,而除了下降调外,其他语调和语调变体都和陈述性记忆不相关,回归模型中的 VIF 值也证明工作记忆、陈述性记忆和程序性记忆不存在多重共线性,这就排除了学习者英语语调范畴感知学习和工作记忆的关系是因为学习者在训练后对英语语调范畴感知是基于规则的范畴学习的可能性,也就是说,尽管学习者训练后的英语语调范畴的感知倾向于基于信息整合的范畴感知,仍然和工作记忆有着密切的关系。我们认为这可能和工作记忆的机制有关,Unsworth & Robison(2017)提出,工作记忆的个体差异是受蓝斑—去甲肾上腺素系统(locus coeruleus-norepinephrine system,以下简称 LC-NE 系统)的瞬时波动驱动的,低工作记忆者的注意(attention)会因为 LC-NE 系统中断而上下波动,LC-NE 系统接受前额叶皮层的交互输入(reciprocal input),这些交互映射可以调节对有意义的凸显事件和无意义事件的注意程度,从这个角度或许可以解释为什么本实验中二语者语调范畴感知学习和工作记忆相关,而 Maddox *et al.* (2013, 2014) 实验中二语者声调范畴感知学习和工作记忆无关。本实验需要学习者提取对语调范畴有用的信息(如音高重音和边界调处的音高变化),抑制和语调范畴无关的信息(如音高重音和边界调外的音高变化),而在 Maddox *et al.* (2013, 2014) 的实验中,声调范畴包括了两个维度,音高高低和音高走势,这两个维度都和声调的范畴有关,没有无关信息。Craig & Lewandowsky(2012)考察了 173 名正常成年人工作记忆对 4 位数字串范畴感知的作用,这些 4 位数字串的信息,只有部分和范畴感知相关,结果发现,工作记忆可能没有影响受试对视觉范畴感知策略的选择,但是影响了受试使用选择策略的成功程度,因此,我们认为,当言语范畴感知学习中需要学习者提取与范畴感知学习有关的信息,抑制无关信息时,工作记忆可能会影响或者预测受试在言语范畴感知中的策略选择,从而影响言语范畴感知学习的结果。

本实验中还发现了下降调的范畴感知学习和陈述性记忆有关,COVIS 模型认为,在基于信息整合的范畴学习过程的初期,人们会倾向于依赖外显系统,经过训练和练习后,逐渐转向内隐系统。按照 Ullman 的 DP 模型,外显系统和陈述性记忆

相关,内隐系统和程序性记忆相关,那么是不是学习者对下降调的范畴感知学习还处在学习的初期呢,从学习者训练后的语调范畴感知结果来看,并非如此,学习者训练后下降调的范畴感知结果要远远好于 M 调及其变体,排除了学习者的下降调还处在学习初期的可能性。还有一种可能是,课堂教学的影响。在中国,大部分教师对英语语调的教学只涉及边界调的教学(陈桦 2008),将英语边界调的规则用言语总结为“陈述句结尾用下降调,一般疑问句结尾用上声调”,使英语边界调的规则显性化,造成下降调范畴感知学习和陈述性记忆相关,但这种情况下,学习者对下降调的范畴感知学习效果应该和陈述性记忆正相关,而不是负相关。

Ullman *et al.* (2020: 389) 提出,陈述性记忆和程序性记忆之间存在交互关系:第一,两种记忆在学习、加工序列和范畴知识时可以互补;第二,通过陈述性记忆学习的知识表征会抑制(阻断)程序性记忆对这些知识的表征,反之亦然。很多因素可以调节人们在学习时对两种系统的依赖程度,陈述性记忆具有快速学习的能力,在序列和范畴的学习早期往往依赖陈述性记忆,程序性记忆逐渐获取类比知识(analogous knowledge),在学习后期占主导地位。显性训练或指导、延时反馈或无反馈会促进学习依赖陈述性记忆,隐性训练、即时反馈和高变化性训练刺激会导致学习依赖程序性记忆。因此,我们推测,教师在课堂上对英语边界调的教学促使学习者通过陈述性记忆快速习得了英语边界调的规则,本实验的训练环节采取了隐性的形式,促使学习者对英语边界调的范畴习得开始转向依赖程序性记忆,程序性记忆对陈述性记忆的知识表征有抑制作用,因此,本实验中的学习者在训练后对边界调的范畴感知结果和程序性记忆正相关,和陈述性记忆负相关,而且因为前期的课堂教学,代表边界调变化的下降调比代表音高重音变化的 M 调及其变体更早进入了学习的后期,所以,边界调的范畴感知学习效果要远远大于 M 调及其变体的效果。但该推测还需要进一步的实验,比如说在训练环节的前期加入 M 调的显性训练等方法,来进一步验证。

6. 结论

通过语调范畴感知学习实验,本研究发现中国学习者在短暂的语调范畴感知训练后可以建立基本的英语语调范畴,但学习者对英语语调边界

调的范畴感知学习的效果要好于音高重音,我们认为,学习者在形成语调范畴的过程中,主要依靠的线索是音高曲线的声学变化,没有充分考虑这些音高曲线的变化,尤其是音高重音的音高变化,对语调意义和功能的影响,也就是说,尽管学习者的对语调的范畴感知在训练后是基于信息整合的范畴学习,但学习者的信息整合可能主要考虑的是音高曲线维度上的变化,没有充分考虑英语语调中形式和意义的联系。

与 COVIS 模型和 DP 模式的推测一致,本研究发现,学习者在训练后对英语语调的范畴感知主要依赖程序性记忆。但是,本研究还提出两点与 COVIS 模型和 DP 模式不同的观点:第一,COVIS 模型认为,基于信息整合的范畴学习和程序性记忆相关,不受工作记忆的影响。我们研究发现,工作记忆对学习者的语调范畴感知学习有促进作用,随着工作记忆的提高,学习者对英语语调范畴感知学习效果越好。因此,我们认为,工作记忆中的注意可以帮助学习者提取与范畴感知学习有关的信息,抑制无关信息,影响或预测学习者对言语范畴感知中的策略选择。因此,当言语范畴感知任务需要学习者提取有效信息、抑制无关信息时,工作记忆会影响学习者的言语范畴感知学习;第二,我们发现,学习者对英语语调边界调的范畴感知学习的效果要好于对音高重音的学习效果,而且学习者边界调的范畴感知学习和陈述性记忆负相关。我们认为,英语课堂中教师对边界调的教学,使得学习者在边界调学习的早期依赖陈述性记忆快速习得了英语边界调的规则,在本实验中经过隐性训练后开始依赖程序性记忆,更早转入学习的后期,因此,学习者边界调的范畴感知学习效果要远远好于音高重音的学习效果。我们认为,二语者语调范畴感知学习和 DP 模式下的句法习得不同,可能需要的是一种显性指导和隐性训练结合的方式,但这一点需要进一步的实验验证。

参考文献

- [1] Arthur, D. T., Ullman, M. T. & F. S. Earle. Declarative memory predicts phonological processing abilities in adulthood [J]. *Frontiers in Psychology*, 2021(12): 1-12.
- [2] Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U. & E. M. Waldron. A neuropsychological theory of multiple systems in category learning [J]. *Psychological Review*, 1998(3): 442-481.
- [3] Ashby, F. G. & W. T. Maddox. Human category learning [J]. *Annual Review of Psychology*, 2005(3): 149-178.
- [4] Ashby, F. G. & J. B. O'Brien. Category learning and multiple memory systems [J]. *Trends in Cognitive Sciences*, 2005(2): 83-89.
- [5] Best, C. T. A direct realist view of cross-language speech perception [A]. In W. Strange (eds.). *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* [C]. Timonium, AD: New York Press, 1995: 171-204.
- [6] Boersma, P. & D. Weenink. Praat, a system for doing phonetics by computer [J]. *Glott International*, 2001(5): 341-345.
- [7] Craig, S. & S. Lewandowsky. Whichever way you choose to categorize, working memory helps you learn [J]. *The Quarterly Journal of Experimental Psychology*, 2012(3): 439-464.
- [8] Díaz, B., Mitterer, H., Boersma, M. & N. Sebastián-Gallés. Individual differences in late bilinguals' L2 phonological processes: From acoustic-phonetic analysis to lexical access [J]. *Learn. Individ. Differ*, 2012(6): 680-689.
- [9] Dutoit, T., Pagel, V., Pierret, N., Bataille, F. & O. Van Der Vrecken. The MBROLA project: Towards a set of high quality speech synthesizers free of use for non commercial purposes [A]. In H. T. Bunnell & W. Idsardi (eds.). *Proceeding of Fourth International Conference on Spoken Language Processing* [C]. New York: Citation Delaware, 1996: 1393-1396.
- [10] Flege, J. E. Second language speech learning: Theory, finding and problems [A]. In W. Strange (eds.). *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* [C]. Timonium, AD: York Press, 1995: 233-277.
- [11] Francis, A. L., Ciocca, V., Ma, L. & K. Fenn. Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers [J]. *Journal of Phonetics*, 2008(2): 268-294.
- [12] Gussenhoven, C. Focus, mode and the nucleus [J]. *Journal of Linguistics*, 1983(2): 377-417.
- [13] Howard, J. H. & D. V. Howard. Age differences in implicit learning of higher order dependencies in serial patterns [J]. *Psychology and Aging*, 1997(4): 634-636.
- [14] Kapatsinski, V., Olejarczuk, P. & M. A. Redford. Perceptual learning of intonation contour categories in adults and 9-to-11 year-old children: Adults are more narrow-minded [J]. *Cognitive Science*, 2017(2): 383-415.
- [15] Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N. & B. Lindblom. Linguistic experience alters phonetic perception in infants by 6 months of age [J]. *Science*, 1992(6): 606-608.
- [16] Ladd, D. R. *Intonational Phonology* [M]. Cambridge: Cambridge University Press, 1996.
- [17] Ladd, D. R. & R. Morton. The perception of intonational emphasis: Continuous or categorical [J]. *Journal of Phonetics*, 1997(3): 313-342.
- [18] Maddox, W. T., Chandrasekaran, B., Smayda, K. & H. G. Yi. Dual systems of speech category learning across the lifespan [J]. *Psychology and Aging*, 2013(4): 1042-1056.
- [19] Maddox, W. T., Chandrasekaran, B., Smayda, K., Yi, H. G., Koslov, S. & C. G. Beevers. Elevated depressive symptoms enhance reflexive but not reflective auditory category learning [J]. *Cortex*, 2014(4): 186-198.
- [20] Mennen, I., Schaeffler, F. & G. Docherty. Cross-language

- differences in fundamental frequency range: A comparison of English and German[J]. *Journal of the Acoustical Society of America*, 2012(3): 2249-2260.
- [21] Morgan-Short, K., Steinhauer, K., Sanz, C. & M. T. Ullman. Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns[J]. *Journal of Cognitive Neuroscience*, 2012(4): 933-947.
- [22] Pierrehumbert, J. B. *The Phonology and Phonetics of English Intonation*[D]. MIT, 1980.
- [23] Quam, C., Wang, A., Maddox, W. T., Golisch, K. & A. Lotto. Procedural-memory, working-memory, and declarative-memory skills are each associated with dimensional integration in sound-category learning[J]. *Frontiers in Psychology*, 2018(9): 1-15.
- [24] Schafer, A. J., Camp, A., Rohde, H. & T. Grueter. Contrastive prosody and the subsequent mention of alternatives during discourse processing[A]. In K. Carlson, Clifton, C. & J. D. Fodor (eds.). *Grammatical Approaches to Language Processing: Essays in Honor of Lyn Frazier*[C]. Switzerland AG: Springer, 2018: 29-44.
- [25] Stine, E. A. L. & J. Hindman. Age differences in reading time allocation for propositionally dense sentences[J]. *Aging and Cognition*, 1994(1): 2-16.
- [26] Larrabee, G. J. Malingering scales for the continuous recognition memory test and the continuous visual memory test[J]. *Clinical Neuropsychologist*, 2009(1): 167-180.
- [27] Ullman, M. The declarative/procedural model: A neurobiologically motivated theory of first and second language[A]. In B. Van Patten & J. Williams (eds.). *Theories in Second Language Acquisition: An Introduction*[C]. New York: Routledge, 2020: 135-158.
- [28] Ullman, M. T., Earle, F. S., Walenski, M. & K. Janacek. The neurocognition of developmental disorders of language[J]. *Annual Review of Psychology*, 2020(71): 389-417.
- [29] Unsworth, N. & M. K. Robison. A locus coeruleus-norepinephrine account of individual differences in working memory capacity and attention control[J]. *Psychonomic Bulletin and Review*, 2017(4): 1282-1311.
- [30] Weinberger, S. Speech accent archive[OL]. George Mason University. (2014-1-10). <http://accent.gmu.edu>.
- [31] Yi, H. G., Maddox, W. T., Mumford, J. A. & B. Chandrasekaran. The role of corticostriatal systems in speech category learning[J]. *Cerebral Cortex*, 2014(4): 1409-1420.
- [32] 卜友红. 中国英语学习者语调习得问题研究[J]. *外语教学与研究*, 2016(4): 569-582.
- [33] 陈桦. 中国学生朗读口语中的英语调型特点研究[J]. *现代外语*, 2006(4): 418-425.
- [34] 陈桦. 中国学生英语语调模式研究[M]. 上海: 上海外语教育出版社, 2008.
- [35] 陈桦, 孙欣平. 输入, 输出频次对英语韵律特征习得的作用[J]. *外语研究*, 2010(4): 1-8.
- [36] 陈莹. 第二语言语音感知研究的理论基础和教学意义[J]. *外国语*, 2013(3): 68-76.
- [37] 高薇, 许毅, 穆凤英. 中国英语学习者韵律焦点教学的实验研究[J]. *外语教学与研究*, 2015(6): 867-873, 960.
- [38] 纪晓丽, 张辉, 李爱军, 龚箭. 不同水平学习者对英语语调感知的实证研究[J]. *外语教学与研究*, 2018(3): 393-406, 480-481.
- [39] 林秋茗. 中国学生英语韵律习得与性别的关联研究[J]. *现代外语*, 2011(2): 195-201.
- [40] 马秋武, 贾媛. 语调音系学综览[J]. *南开语言学刊*, 2009(1): 85-92.
- [41] 孟小佳, 王红梅. 中国英语学习者朗读口语的边界调模式研究[J]. *外语教学与研究*, 2009(6): 447-451.
- [42] 苏锦平. 英汉语的适听适视特质及对语境文化的影响[J]. *西安外国语大学学报*, 2022(1): 51-55.
- [43] 田方. 二语模仿朗读中重音输入对重音输出的影响研究[J]. *外语与外语教学*, 2018(3): 55-64, 144.
- [44] 杨晋. 反复聆听模仿和中国英语学生重音模式习得研究[J]. *外语研究*, 2010(4): 9-16.
- [45] 姚尧, 陈晓湘. 音乐训练对4~5岁幼儿普通话声调范畴感知能力的影响[J]. *心理学报*, 2021(4): 456-468.
- [46] 杨军. 中国大学生英语朗读中的语调短语划分不当[J]. *现代外语*, 2006(4): 409-417.
- [47] 薛锦, 聂亚如, 李斑斑. 英语韵律特征和初始习得年龄的关系研究[J]. *外语学刊*, 2019(6): 79-86.
- [48] 张辉, 卞京. 二语习得和加工假说与模式: 主要观点与分歧[J]. *外语与外语教学*, 2016(4): 10-20, 147.
- [49] 张林军. 日本留学生汉语声调的范畴化知觉[J]. *语言教学与研究*, 2010(3): 9-15.
- [50] 张文彤. SPSS 统计分析高级教程[M]. 北京: 高等教育出版社, 2004.
- 基金项目:** 本文系国家社科基金重大项目“中国方言区英语学习者语音习得机制的跨学科研究”(项目编号: 15ZDB103)以及江苏省人文社科基金青年项目“中国学习者英语韵律加工的ERP研究”(项目编号: 19YYC001)的阶段性研究成果。
- 作者简介:** 纪晓丽, 江苏科技大学外国语学院副教授, 南京师范大学博士研究生, 研究方向: 实验语音学、二语习得等。
- 张辉(通讯作者), 南京师范大学外国语学院教授, 博士生导师, 研究方向: 认知语言学、二语习得。
- 李爱军, 中国社会科学院语言所教授, 博士生导师, 研究方向: 实验语音学、语音习得等。

责任编辑 薛旭辉