

语言文化数字化传承的理论与方法

孔江平

(北京大学 中国语言学研究中心, 北京 100871)

摘要: 语言文化传承是人类文明最重要的组成部分,传承的方法也随着社会和科技的进步在不断发展。当今全球一体化的进展,使得世界有声语言和口传文化的多样性濒临消亡的境地。论文从传统文字记录语言文化的性质和缺陷以及现今音频、视频记录语言文化的不精确和不完整性,讨论了有声语言和口传文化数字化传承的基本方法,并提出了一种基于语音多模态的数字化传承方法。最后,论文还从有声语言和口传文化认知的角度讨论了全面精确传承语言文化的理论基础。

关键词: 有声语言; 口传文化; 数字化方法; 文化传承

中图分类号: H0 摇 **摇 摇 文献标识码:** A 摇 **摇 摇 文章编号:** 1000-5919 (2013) 03-0089-09

一、引 摇 言

人类的文明和进步基于人类优秀文化的积累和传承,而文化的传承又基于语言文化的记录方式和载体,目前世界文化传承的主要形式是语言文字和口传文化。人类语言产生于人类漫长的进化过程,自从有了语言,人们就开始创造文化,特别是口传文化。时至今日,世界上大部分的语言都没有自己的原创文字,其语言文化主要是以口传的方式传承,例如基于各种民族有声语言的史诗、口传经文、原生态民歌等。因此只有部分民族的有声语言和口传文化被记录下来,然而,随着部分有声语言的消亡,基于这些有声语言和口传文化的知识体系也随之消亡。

中国是一个多民族的国家,有丰富的语言及口传文化资源(谭克让,麦克康奈尔主编,1995;《中国戏曲音乐集成》编辑委员会,2002)。但随着经济全球化趋势的加剧,大众传媒的高度发达,使用不同语言的人群接触频繁,致使一些使用人数较少的语言交际功能日渐衰退,消亡速度日益加快。事实上,语言不仅是交际的工具,同时也是文化的载体。每一种语言都承载着一种独特的文化,凝聚

着一套知识体系,是人类文明最重要的组成部分。语言,尤其是占绝大多数的无文字语言一旦消亡,大量人类的知识体系也将随之消失,这无疑将给文化的多样性和人类文明的传承带来不可挽回的损失。任何一种民族的口传文化都是在百万年人类进化的长河中逐步形成的,因此,这种“文化物种”的产生和发展的过程并不亚于自然物种,都是人类智慧的宝贵财富。但目前各国列出的语言文化遗产,大都处境艰难,濒临绝灭,这说明全球都面临着如何保护优秀传统文化的问题。应对自然物种的消亡,目前主要采用建立种子库和基因库的方式加以保护。而对有声语言及口传文化这种“文化物种”的保护,世界各国还没有形成共识,主要方法有书面记录、录音和录像,然而,这些方法达不到全面保护和传承的目的。如何有效地对有声语言及口传文化进行抢救和保护是人类面临的一个迫切需要解决的问题。

从理论上讲,文化遗产具有多样性。其中,语言或称有声语言和口传文化就是世界文化遗产中最为重要、特殊和具有价值的一种,之所以这样说是因为:1)语言是人类思维的工具,存在于人的大脑中,人类所有非物质文化都基于人类的语言;

收稿日期:2013-03-18

作者简介:孔江平,男,山东齐河人,北京大学中文系教授,中国语言学研究中心研究员。

基金项目:国家社会科学基金重大项目“中国有声语言及口传文化保护与传承的数字化方法研究和基础理论研究”资助,项目批准号:10&ZD125。

2)有声语言和口传文化是以生理信号(脑电、声门阻抗信号等)和物理信号(音频、视频等)的方式来承载的,瞬时即失;3)活的语言一旦消亡,口传文化就随之消失。这些特殊性质对有声语言和口传文化保护和传承的方法提出了不同的要求。目前从世界语言和口传文化记录与传承上看,还缺少基本的理论框架和有体系的科学方法。

随着科学技术的进步和社会发展,特别是现代通讯技术的兴起和应用,语言研究已经涉及自然科学的许多领域,如生理和声学的语音数字化研究、有声语言的口传文化研究、语音信息量演化计算的语言认知和编码研究、语言发音生理的人类语言进化研究、语言生理和语言情感的歌曲、戏剧和原生态声乐研究等。显然,语言和口传文化研究已经逐步科学化。本文在有声语言生理、声学 and 数字化研究以及具有独特性质的口传文化生理和声学研究的基础上,对有声语言和口传文化数字化传承的基本理论和方法展开讨论。

二、记录语言和口传文化的传统方法

摇摇人类传统记录语言和口传文化的方法主要是文字和乐谱,由于语言的不同和口传文化的差异,文字系统和记录乐谱的形式也十分不同,这里以中国语言和口传文化为例,简要阐述语言文字系统和口传文化记录的形式、性质和缺陷。

1. 传统语言文字记录

人类自从有了文化,就试图记录下来。从中国各地不同时期的岩画中,除了图画外,可以很容易地发现很多抽象的符号,如广西的华山岩画。显然人们早期使用图画来记录生产和生活事件,然后简化图画逐步符号化。从中华文明的历史轨迹看,除了早期的岩画和符号外,比较系统的图画符号体系是东巴文,东巴文是介于语言文字系统和图画符号系统之间的一种形式和阶段。从语言学的角度看,中国最早的文字体系是殷商的甲骨文。后来经过大篆、小篆、金文、楷书等发展阶段,

最终形成了汉语的文字系统。除了汉字系统以外,中国主要的传统文字还有藏文、蒙文、维吾尔文、朝鲜文、彝文等。所有这些传统文字最终构成了中华民族的语言文字记录系统。

从语言文字记录文化的发展历史看,文字系统有许多重要的性质。首先,文字记录是以语言为基础,因此文字系统是语言系统的一种特殊形式。由于语言结构上的差异,文字系统在结构上也有很大的差异,例如汉藏语系语言的单音节性和声调,使得汉字采用表示单音节的方块字形式。其次,文字系统具有鲜明的民族性,因此,文字系统中包含了一个民族特有的文化背景,如汉语的诗歌主要是句尾韵母押韵,而蒙语的诗歌主要是句首重音押韵,这和语言的结构有密切关系。第三,文字系统随着语言的演化而改变。众所周知,语言随着时间会发生变化,文字系统也会随语言产生相应的变化。

2. 传统口传文化记录

在口传文化的记录和传承上,不同的民族和语言采用不同的方式,但理论上一一般采用语言文字加乐谱的形式,因此更为复杂。中国传统的记谱法主要有十三种,分别是:工尺谱、古琴谱、燕乐半字谱、弦索谱、管色谱、俗字谱、律吕字谱、方格谱、雅乐谱、曲线谱、央移谱、查巴谱和锣鼓经(王耀华,2006年)。这些记谱方法除了可以记录音乐外,部分可以用来记录口传文化。如律吕字谱和工尺谱,前者通常使用宫(do)、商(re)、角(mi)、清角(和)(fa)、徵(sol)、羽(la)、变宫(变)(si)来表示不同的音高,中国古代律学的理论体系已经十分健全。^①而工尺谱用合、四、一、上、尺、工、凡、六、五、乙等字表示音高,可相当于sol、la、si、do、re、mi、fa、sol、la、si。高八度将谱字末笔向上挑,或加偏旁亅,低八度将字的末笔向下撇。另外,用“、”或“x、一”等符号作为节拍符号。昆曲是中国一种最重要的口传文化,包含了中国古代音乐和文学的精华,有六百年的历史,昆曲直至今还在用工尺谱来记录。^②除了表示音高

① 中国的古代的律学的理论体系已经很完善,王骥德的《曲律》,是中国律学的经典著作。

② 中国的昆曲一直沿用工尺谱,大量的曲本,2009年由广陵书社出版的《昆剧手抄曲本一百册》收集和整理了昆曲的主要曲目,整体上反映了工尺谱记录昆曲的面貌。

外,还能记录许多唱腔和唱腔的组合。^① 昆曲的特点是字行腔,所以在旋律上和汉语的声调有更密切的关系。

口传文化主要依靠口头流传,如古代传说、史诗、歌曲、戏剧、说唱等,它们各有自己的特点。首先口传文化是有声语言的一种特殊形式,通过特殊的声音来表达有声语言无法表达的特殊情感。在形式上主要是靠韵律、旋律、重音和节奏的变换表达特殊情感。其次,记录口传文化除了要用到语言外,还要用特殊的记谱方法来记录具有音乐性的旋律和节奏,即用文字记录内容和利用乐谱记录旋律,如记录昆曲除了用文字记录词以外,还要用工尺谱记录旋律和节奏。

3. 语言文字和乐谱记录的缺陷

虽然文字可以很精确地记录语言要表达的内容,但文字在记录时有一个致命的弱点,即文字记录文化和人们认识事物的程度相关,当认识不足时记录的内容只是人们当时认识事物的水平。众所周知,人们认识世界和创造文化都有一个逐步深入的过程,这就是说,古人记载的所有事物和文化现象都是古人当时的认识水平,并不一定全面和正确,有些可能是假象。这就造成了后人很难理解古人的许多文献纪录。例如,我们现在可以看到古代留下来的韵书,比较清楚地了解声韵调的类别,但我们还是很难知道古代汉语确切的读音。我们有《尔雅》和《说文解字》^②,但我们对古代很多汉字的确切含义仍有质疑。因此中国小学中的音韵学、训诂学等都要研究和考证语言文字记录的不精确性带来的后果,这说明目前文字并不能很精确全面地记录人类的文化。同样对口传文化进行记录也是如此,虽然古人采用了特殊的方式记录口传文化,实际上还是不能精确记录它们。例如,我们今天能看到李白的诗歌,但我们还是无法知道李白吟诵诗歌时的字音和旋律。我们能看到姜夔《白石道人歌曲》中的十七首词曲^③,

但仍然很难知道准确旋律和节奏。

从以上的讨论可以看出,目前人类大多数文化遗产和口传文化主要靠语言文字和音乐符号记录,但基于人们当时的认识水平,并不是完全精确和客观的记录。人们认识事物的错误和不精确也都被记录了下来。因此随着人们对事物和自然界规律认识的深入,我们不得不去考证古人留给我们的文献,去鉴别真伪。

三、语言记录的理论基础

语言是思维的工具,极为复杂,同时语言又具有民族性,并形成了各民族不同的语言文化,在此基础上最终构成了一个民族的知识体系。下面从语言记录的形式、语言的认知范畴来讨论语言记录的理论基础。

1. 语言记录的形式

从形式上看,语言记录可分为记录信号和记录介质两大部分,第一部分可以分为:1)图画类(图画,绘画);2)图形类(东巴文);3)符号类(甲骨文,大篆,小篆,金鼎文,楷书,文字的数字编码)。符号类主要是文字,又可分为:a)象形;b)音符;c)象形音符组合。学界也有将文字按结构类型分为“非字母文字”和“字母文字”两大类的。从理论上讲,文字只有达到了记录语音的程度才能算作真正的文字系统。第二部分可以分为:1)硬介质(骨料,石料,竹简,木简);2)纸张介质(锦帛,纸张);3)电子介质(磁盘,光盘,硬盘)。在电子介质中只有达到完全数字化后,才可以无损耗复制。因此,目前人类在这方面也可以认为已经达到了传承方式的极限。

2. 语言的认知基础

现在人类认识自然和人文活动的能力和深度还都十分有限,这就决定了我们只利用语言文字和一些记录口传文化的符号系统不能很完善地记

① 昆曲的声腔和节奏在俞振飞的《粟庐曲谱》中有详细的记录和论述,这些论述对现今昆曲的打谱和研究都有重要的参考价值。

② 郭璞的《尔雅》和许慎《说文解字》是中国最早的有系统地解释汉字字义的字典,这两部典籍对古汉的解读和研究至关重要。

③ 在中国宋词文献中,只有姜夔的《白石道人歌曲》中有17首标注有工尺谱,是我们现在研究宋词吟唱旋律和节奏的珍贵资料。

录我们的语言和口传文化,这是不是说明我们就不能全面精确记录人类的语言和口传文化了呢?答案是否定的。虽然人类认识自然和人文活动的深度还需要很长的时间,但并不是说我们不能用科学的方法来记录我们的语言和口传文化用于将来的研究。那么怎样证明我们记录的语言和口传文化是全面精确和没有丢失信息呢?这就涉及人类语言的认知能力和极限。如果我们知道人类语言的认知能力、范围和极限,就可以采用更多的科技手段和方法来记录语言和文化,达到全面和精确的程度,以便后人研究。

人类语言认知能力和极限一直都是语言学家、心理学家和脑神经学家研究和探索的目标,虽然近十年来在这个交叉领域有了巨大的进展,但距离了解人类大脑的语言能力还相去甚远。但不用这些先进的仪器和设备,从语言学和统计的角度人们也可以对人类语言的认知能力有很好的了解。语言学家的研究告诉我们,在人类的上千种语言中,虽然有一千多个音素可以成为音位(phoneme),但对于每一个语言来说,其音位一般只有几十个左右(Ladefoged et al, 1996)。在词汇方面,人们在口语中常用的词汇有三千左右,而且能覆盖语言交际和普通文献的百分之九十五以上。在语言词汇的习得方面,郑锦全先生经过多年对不同语言词汇掌握情况的研究提出了“词涯八千”的理论^①,即一个人在学习一种语言时,能够掌握和应用的词汇不会超过八千个。例如汉语目前已经有三十万左右的词汇,但每个人掌握的词汇从古至今不会超过八千,是有限的。在语法方面,每种语言的句法结构大都在二百种左右,并可以用语言学的方法分析得到(萨丕尔,1985)。以上的讨论告诉我们,虽然从大脑神经的角度人类对语言的认知能力的研究还有很长的路要走,但人类语言的认识能力和极限已经勾画了出来。

3. 语言记录的相关理论

从语言文字的形式和语言认知的角度看,无论是什么文字形式,文字记录语言主要是将符号系统和语言的音位系统联系起来。从语音学和音

位学层面看,虽然人类能分辨出上万种声音,也能将上千种声音用于语言的交际,但对于一种语言来说,只有大约几十个能用于语言的音位,这说明人类的大脑对语位的感知范畴是有限的。以汉语普通话为例,声韵调的组合可以构成1200个左右的音节,将声母(22个)、韵母(38个)和声调(4个)组成一个三维空间,标出1200个坐标点就构成了汉语普通话最基本语音发音和感知空间。这个空间的理论坐标点为3344个,但实际只用了1200个左右,巨大的冗余度保证了汉语普通话交际的稳定性。如果将声韵调和基本词汇联系起来,其对立关系又体现出每个音位的结构对立负担量。如果再将其放在大规模文本中,又能计算出汉语普通话的功能负担量(functional load)(王士元,1967)和信息熵(Information Entropy)(Shannon et al, 1949; Shannon, 1951)。语言信息和结构的研究表明,无论是什么语言,其信息结构具有共性,如语言词汇的分布会遵循齐普夫定律(Zipf's Law)(Zipf, 1935, 1949)。以上的分析可以看出,语音学、音位学、文字学、语音感知空间、语言冗余度、语言对立的结构负担、语言功能负担量、信息论、文献统计分布定律等构成了精确记录和传承一种语言系统的理论基础,这就为全面记录和传承语言奠定了基础。

四、口传文化记录的理论基础

口传文化是有声语言的一种特殊形式,从功能上看,口传文化如歌曲、戏曲等,都是在表达比有声语言更为强烈、独特和难以言表的情感,由于这些特殊的情感超出了语言表达的范围,因此需要用语言之外的表达方法。下面从口传文化的形式和认知范畴来讨论口传文化记录和传承的理论基础。

1. 口传文化记录的形式

在口传文化记录和传承的形式方面,除了需要语言文字记录语言内容外,其形式主要是曲谱,可以分为两类:1)文字谱,用语言文字记录口传文化的音乐旋律,例如碣石调最初是民歌,叫《碣

^① “词涯八千”的理论是台湾郑锦全教授经过多年对不同语言不同作者著作的统计,发现了人类语言认知在词汇上的极限,该理论的提出对认识人类的语言习得和认知能力有重要意义。

西行》,后来用它唱曹操的诗《碣石篇》,就改叫《碣石调》,显然早期音乐的曲谱是用文字来记录的;2)符号谱,用符号来记录口传文化的音乐旋律,例如,五线谱、简谱、工尺谱等。在中国传统口传文化中影响最大和使用最普遍的符号谱是工尺谱,其符号来自汉字,而中国传统曲谱的记录介质主要是纸张。总之,从形式上看,口传文化的记录除了包含语言的记录形式外,还增加了口传文化音乐的记录形式。

2. 口传文化的认知基础

人类的语言有其认知的基础,并具有一定的民族性,同样口传文化也具有认知的基础和一定的民族性。在音乐层面,口传文化的性质更具有普遍意义。因此,一种口传文化除了民族性以外,人类对音乐的认知更为趋同。比如,音乐的旋律通常主要由基频决定,根据心理物理学研究,人们对基频音高的感知是基本相同的。但现代声学的研究告诉我们,人们对音高的感知并不局限于基频这一物理量,比如,泛音(overtone)同样能被感知为音高,形成音乐的旋律。这时旋律不是依据这个声音的基频,而是由声音某个较强的谐波决定,一个最典型的例子就是口弦琴。口弦琴是一种乐器和语言发音器官结合的口传文化,演奏者通过调节声道形状将能量集中在某个谐波上,从而产生一个很高的旋律。另外,共振峰同样能被感知为音高,从而成为声乐的旋律,最典型的例子是蒙古族的呼麦。人们感知到的呼麦旋律非常复杂,通常呼麦会被感知为两个旋律,一个是由声带产生的非常低的旋律,称为喉音唱法(throat singing)(Lindestad, 2001)通常人们的声带不可能产生这样低的振动,但呼麦是通过真声带和假声带同时振动,并在声带的开合上有一个相位差,从而产生一个低于原声带振动一倍的次谐波降低音高的感知,这种发声的类型被称为“喉室发声”(ventricular voice)或“双声带发声”。另一个旋律是由声道调节产生一个很强的共振峰而形成了一个远高于声带振动的旋律,也可以说是泛音的一种形式。从以上的分析可以看出,人们在感知旋律时非常复杂,并不是单一的物理量在起作用,同时我们也可以看到人们对声乐的认知基础和极限,在了解了声乐的认知基础和极限后,我们就能

以此建立口传文化记录和传承的理论框架。

3. 口传文化记录的相关理论

在了解了人们对各种口传文化旋律认知的基础上,就可以讨论口传文化记录和传承的理论基础。我们知道旋律的感知基础是音高和节奏,而音高的物理量是基频、泛音以及共振峰的综合体。传统的记谱方法对于大多数口传文化来说是有效的,这使得我们能够传承古代遗留下来的许多文化遗产。但从现代科学的角度看,只用这些记谱方法记录丰富多彩的口传文化,并不是很精确也有很多局限性。人的音高和频率的感知关系可以用数字来精确描述,钢琴键盘体现了基频和音高之间的数学关系。

人对音高的感知也有范围和极限,最低在20赫兹左右,最高在20000赫兹左右;颤音在一定的范围内人是感知不到,比如,声音以平均每秒6~7次的周律在半音限度内平稳地上下摆动形成颤音不会被察觉音高的变化,但音质会发生改变。传统的记谱方式很难对微小基频的变化所导致的音色改变去进行精确记录,例如昆曲通常也被称为“水磨腔”,但是古代和现代文献很少涉及水磨腔发音生理和声学的科学解释。从声学分析看,一定频率范围内的颤音是昆曲唱腔幽雅婉转、清丽悠远的主要声学特征,这些声学特征在生理上体现为发声类型的反复变化。显然普通的曲谱很难记录下这些信息。虽然我们传承下来了很多曲谱,但仍然很难想象出古人在吟唱时那优美的旋律和使用的嗓音发声类型。另外,中国的许多原生态民歌中存在大量的装饰音,但目前的许多记录方法主要是以骨干音为主,看看一首歌曲的基频曲线我们就知道丢掉了多少信息,虽然有些装饰音不能被明确感知出来,但它们对音色已经产生了微妙的改变。

目前人们对旋律、音色和声音物理量之间关系的了解还不是很深入,但在知道了声学的基本特性和声乐的感知基础后,就可以从理论上建立声乐记录的基本理论框架。从以上的分析可以看出,心理物理学、音乐声学、心理学等构成了声乐精确记录和传承的理论基础。在这个基础上,可以先利用现代科学技术,记录和传承那些我们目前还无法解释的现象和信号。

五、语言和口传文化记录 的基本方法

传统的有声语言和口传文化记录方式主要是文字和乐谱,上个世纪中晚期人类又发明录音机、摄影机和录像机,从而用文字无法记录的信息被记录了下来。信号的存储方式也从纸张发展到模拟电子介质和数字电子介质。数字信号在传承过程中已经达到了一个极限。这样问题就集中在记录信号上,即用什么样的方法采集什么样的信号才能精确和全面的记录语言和口传文化,而不用考虑承载形式和介质的问题。

1. 现有信号记录的缺陷

从方法上看,国际国内对有声语言和口传文化的保护基本是利用文字、语音和录像三种方法。现代科学的研究表明,简单地使用文字、录音和录像的方法会丢失很多信息。首先,文字记录语言和口传文化的性质,其中最主要的是在人们对事物认识不足的情况下不能精确和全面的记录语言和文化信息。其次,现在有了录音机和录像机是否就能精确记录语言和口传文化了呢?答案同样是否定。录音是将声学信号转变成数字信号记录和保存起来,的确能够记录语言和口传文化的大部分信息,但听录音和听现场实际的声音有时会有很大的不同,特别是在听早期的录音时,会明显发现语言的辅音会混淆,辨别率较低。虽然现在录音机从早期的模拟录音机发展到高采样频率的数字录音机,已经到达了人类听觉的极限,解决了录音技术问题,但是录音还有其他的问题,例如,录音有频响的问题,因此听到的现场声音和回放的录音是有差别的,即使是微小的差别也会导致人对声音感知的错误。第三,关于视频信号,虽然现在的录像质量已经能达到高清,但和人现场的视觉感知还是存在差异,例如,一幅实际的场景由于物体的形状和光线不同会产生阴影,人眼能从阴影处提取很好的和有用的视频信息,但是现在高质量的录像机也无法像人眼的视网膜一样录下阴影处有用信息。另外,视频信号的采集还涉及三维立体的问题,其中三维空间的信息会丢失的更多。总之,拿录音机和录像机的性能与人的听觉和视觉认识能力相比,前者的性能还差得很远,

因此我们说在现有技术下,文字、录音和录像都会丢失有用信息。另外,从人的认知角度看,我们感受到的世界也未必都是完全真实的。现代科学证明,不同的人感受同样的事物,结果可能不同。我们通常说“眼见为实,耳听为虚”,但实际上眼见不一定就为实。认知科学的进步,为语言文化数字化传承带来了更广阔的视野。

2. 语言信息领域的新方法

在语音学、语言学、信息科学和神经科学领域,语言和口传文化的研究在近十年已经有了很大的发展,人们从语音的发音到声音的物理特性,从对语音信息的感知到大脑神经的活动,使用了许多新的研究方法和得到了很多语言信息传递的知识,下面对这些新的方法进行简要地介绍。

传统语音学和语言学的方法主要通过语音和语言学家听音,用国际音标记音记录有声语言和口传文化的语音、研究音位系统和创制文字系统来进行有声语言和口传文化的记录,例如,无文字语言的语音系统、语音和词汇的关系、口传文化的特殊发音方式等。这种方法是传统语音学、音位学、语言学和声乐学的研究方法,记录语言和声乐的田野调查工作是语音学和语言学家、语言人类学家和音乐学家的基本功。

实验语音学的方法主要是利用声学和生理学的一些方法来对有声语言的语音和口传文化的特殊发音方式进行录音,然后对语言的元音、辅音、声调以及韵律进行声学分析,对口传文化的特殊发音方式进行生理学的研究等,最终达到记录、研究、保护和传承的目的。

言语科学的方法主要是利用各种科学仪器对有声语言和口传文化进行各种信号的采集,然后进行信号处理,对提取出的参数建立数据库和进行各种计算。目前,国际和国内根据目的的不同,建立了大量有声语言的声学语料库和数据库,主要集中在语音的合成和识别方面。

模型化研究的保护方法主要根据对语音多模态的研究,提出模型化研究的保护方法。这种方法主要是对某一种有声语言或口传文化进行建模,并在建立相关声学 and 生理模型的基础上对有声语言和口传文化进行有效的保护和传承。

很显然,随着时代和科技的进步,记录有声语

言和口传文化的介质、记录形式和记录方法都在不断地改进,为中国和世界有声语言和口传文化的数字化传承奠定了基础。

3. 言语链和信号的不可逆性

语言的交际过程是一个链条,称之为言语链(邓斯,1963),它从一个人的大脑思维活动、神经驱动言语生理肌肉运动、产生声波到听觉声学神经转换、听话人的大脑感知及思维活动,构成了语言传递和交际的一个链条,在这个链条中都包含有语言的信息。但是语言链条之间的转换并不是完全可逆的,或者说目前我们的科学技术还没有找到完全可逆的方法,这样对我们来说就是未知的,如果我们只记录言语链的一个环节,就有可能丢失许多信息,不能精确和完全记录语言,也就无法有效地传承。如,声道和语音共振峰之间的关系不是一对一的,而是一对多的关系,因此,从共振峰很难精确推出声道的形状,目前从语音中还不能很好地计算出声带的振动方式,如果不记录声带振动的方式,只从声音很难了解一种语言的发声类型。呼吸是语言的动力来源,它同时还能反映出语言交际时的情感和规划,这些信息很难从语音声波中得到。另外,言语者当时的真实情感,单从声音很难得到真实的情感信息。

4. 基于语音多模态的数字化方法

根据本文的讨论和作者在这个领域的长期研究和实践,从信号的角度提出一种适合目前技术水平的有声语言和口传文化的数字化方法。第一是语音信号,采用指向性强的高保真话筒,话筒的频响性能要好。采样频率每秒不小于 22050 赫兹,量化至少用 16 比特,录音环境的本底噪音要保持在 20 分贝左右。第二是嗓音信号,这种信号要用喉头仪来采集,它对采集声带振动信息非常有用,特别是声带闭合相的振动信息,对于保留语言发声类型的信息十分有用。第三是录像信号,对于有声语言来说,唇形和面部的动作是十分有用的信息,众所周知,言语的交际信息不只是包含在语音信号中,唇形也包含有语言交际的信息,McGurk 效应已充分证明了这一点 (McGurk, H. et

al,1976)。第四种是电子腭位信号,这种信号通过一个电子腭采集舌和上腭的接触信息,对于研究和记录言语辅音的发音动作信息十分重要。第五是呼吸信号,呼吸信号包括腹呼吸和胸呼吸两种信号,因为目前的研究表明腹呼吸和胸呼吸在言语中常常不同,腹呼吸和言语动力更密切,而胸呼吸和发音更密切,这对研究和记录语言的韵律及风格十分有用。第六种是心率信号,这是一种心理信号,它能反映说话人的心理状况。第七种是指电压信号,这也是一种心理信号,指电压比较敏感,可以反映心情的变化状况。实际上心理信号也可以录制脑电信号 (ERP),因为脑电信号的时域精度较高,但相对其他信号的录制比较麻烦,需要有很多通道,因此,心理信号还需要进一步研究。前五种信号对于一种有声语言的信息录制、研究和传承都是必不可少的,因为这些信号可以基本建立一种语言的声学 and 生理模型,从而达到比较精确和全面记录一种有声语言和口传文化的目的。另外,录制有声语言还有其他一些仪器可以利用,如超声仪、电磁发音仪、磁共振等都可以得到言语的有用信息,而且这些信息都是从语音中无法得到的,但录制设备和技术上比较麻烦,目前不容易大量录制。

六、语言及口传文化记录和传承的理论思考

随着信息科学技术的发展,语言学和信息科学的关系越来越紧密,研究方法也越来越自然科学化。语言学和信息科学新的理论方法促使人们对语言和口传文化的数字化传承产生了许多新的思考,本节就语言和口传文化记录和传承的理论问题提出几点思考。

1. 人类知识体系是自然生物进化的高级形式

从 19 世纪达尔文发表《物种起源》开始^①,人类对自然物种的起源和进化的研究进入了科学的轨道,经过一百多年的发展,人们建立起了科学的物种起源和进化理论。在研究方法上,已经从物

^① 达尔文的原文为 The Origin of Species,最早于 1859 年发表于英国伦敦。该书的出版对人类和其他物种起源的科学研究起了巨大的推动作用。

种的考古、体质分类发展到了基因分析和基因库的层面。然而人类对语言 and 知识体系的起源和进化却知之甚少。生物的进化可以通过考古和基因图谱来追寻史前的轨迹,但人类史前知识体系的研究却很难从考古中找到线索和获取信息。虽然古人类学家已经发现了许多古人类的化石,但大多残缺不全,很难找到足够的信息来追寻发音器官和大脑语言起源和进化的轨迹。就更不用说建立史前人类知识体系起源和演化的理论框架了。然而这并不是最重要的,因为,目前世界上还有上千种活的语言,如同活的生物,这种存在于人脑中的活的知识体系是伴随着人类长期进化发展起来的,人类从走出非洲就开始了民族、语言和文化的多样性分化,但人类知识体系的多样性并没有得到应有的重视,由于人的原因,自然界的物种在急速消亡,人的语言和知识体系也在急速消亡。目前我们还不太清楚语言和人类知识体系从多样性分化转向快速消亡的时间分界点是从何时开始,但语言文化和人类知识体系的快速消亡已经是不争的事实。因此提高对人类语言和知识体系的科学研究和认识,将其看做是自然生物演化的高级形式是记录、保护和传承人类文明的重要前提。

2. 人类知识体系的理论极限

人类的语言和文化构成了人类的知识体系,从历史上看人类记录和传承知识体系存在许多缺陷,随着信息科学技术的进步,这些缺陷在不断得到弥补,那么从理论上记录和传承人类知识体系的极限在哪里?也就是说从理论上怎样才算完全精确记录和传承了人类的知识体系呢?由于人类的知识存在于语言和口传文化中,所以语言科学家更关心这个问题。从理论上讲,语言的交际和知识传递是通过语言链完成的,在早期的语言链研究中,人们只能认识到语言的发音器官动作和区别语言的最小单位“音位”。随着科学方法的进展,人们对发音器官的生理机制、物理声学性质、音位的听觉感知范畴以及脑电活动有了更新的认识。然而在大脑的语言记忆、感念形成、思维活动等方面,我们还知之甚少。虽然我们已经对人类的语言和认知能力有了新的认识,但我们还不完全了解构成人类知识体系的生物基础。在言

语链中,言语者的大脑思维活动为最底层,然后通过一系列神经活动和生理物理活动达到语言的表层“言语声波”,然后,从言语声波又经过一系列的听觉神经活动进入大脑底层的思维活动。从历史上看,人对语言和口传文化的记录和传承都是在语言的表层,即语音的层面。上个世纪中期以后,人们开始记录、研究和传承语言和口传文化的更底层的信息,如,发音器官的活动、声带的振动信号、语音听觉的神经信号等,并开始建立语言的生理、物理和听觉模型。这些模型是精确和全面记数字化传承语言和文化的基础。由于语言某些表层结构和深层结构转化之间存在不可逆性,通过模型记录和传承语言更底层的信号和信息也许是我们到达精确记录语言文化和人类知识体系的未来形式。研究表明,语言表层有其理论极限,这个极限可以通过语音学、语言学、语法学等来勾画出语言大致的表层认知范围。语言的生理、物理和听觉平面可以通过言语生理、声学和听觉研究来找出人类语言较深层的认知范围和极限,这两个层面的极限我们已经有了很多知识。但在最底层,大脑思维活动的极限可能还需要很长的时间去探索。因此,回到语言文化和人类知识体系的数字化传承上,目前我们能做的有:1)文字符号记录语言表层信息(有声语言);2)物理信号记录语言表层信息(音频和视频信号等);3)生理信号记录较深层的语言发音信息(X光、电子腭位、超声、磁共振等);4)生理信号记录较表层的语言情感信息(心率、指电压等);5)生理信号记录较深层的听觉信息(ERP等);6)生理信号记录较深层的语言理解信息(眼动、ERP、脑磁、磁共振等)。这些语言不同层面信号的研究,一定能使我们逐步认识到语言和人类知识体系的最终极限,从而找到全面精确数字化传承人类知识的方法。

3. 中国的语言和口传文化保护

中国是一个多民族和多语言的国家,语言有五个语系一百多种语言,经语言学家确认的也有八十几种。众多的语言和丰富多彩的口传文化构成了中华民族独有的知识体系,是人类不可多得的知识宝库。然而,随着全球经济一体化的进展,语言文化和认知的多样性受到了极大威胁,面临濒危的境地。一些民族的语言一旦消亡,经过千

万年和人类生理一同进化形成的知识体系也会随之消亡。因此中国在记录和传承语言和口传文化方面,面临更巨大和艰苦的工作。

中国是一个注重文化保护和传承的国家,文化典籍的整理和编纂是我们的文化传统。在历史上,有明代的“永乐大典”和清代的“四库全书”,这两部大典在保护和传承中华民族文化和知识体系方面起到了巨大的作用。在现代这个民族和语言大融合的时代,我们有责任利用基于现代语言信息技术的数字化传承的理论与方法记录有声语言和口传文化,将其编纂为“中华有声语言和口传文化数字化多媒体大典”,并传承下去。

参 考 文 献:

王耀华,2006,《中国传统音乐乐谱学》,福州:福建教育出版社。
 邓斯,平森,1983,《言语链:说和听的科学》,曹剑芬、任宏谟译,北京:中国社会科学出版社。
 谭克让,格兰特 D. 麦克康奈尔主编,1995,《世界的书面语·中国卷》,魁北克:拉瓦尔大学出版社。
 《中国戏曲音乐集成》编辑委员会,2002,《中国戏曲音乐集成》,中国 ISBN 中心。

爱德华·萨丕尔,1985,《语言论》,陆卓元译,陆志韦校,北京:商务印书馆。
 The Sounds of the World's Languages, Peter Ladefoged and Ian Maddieson, 1996, Wiley-Blackwell.
 William S-Y. Wang, 1967, *The Measurement of Functional Load*, Phonetic, Vol. 16. 36 – 54.
 Shannon, C. E. and Weaver, W. , 1949, *The mathematical theory of communication*, University of Illinois Press, Urbana.
 Shannon, *Prediction and entropy of printed English*, 1951, C. E. : Bell System Technical Journal, 30: 50 – 64.
 George K. Zipf, 1935, *The Psychobiology of Language*, Houghton-Mifflin.
 George K. Zipf, 1949, *Human Behavior and the Principle of Least Effort*, Addison-Wesley.
 McGurk, H. and J. MacDonald, 1976, "Hearing lips and seeing voices. " *Nature* 264: 746 – 748.
 Per-Åke Lindestad, Maria Södersten, Björn Merker, and Svante Granqvist, 2001, *Voice Source Characteristics in Mongolian "Throat Singing" Studied with High-Speed Imaging Technique*, Acoustic Spectra, and Inverse Filtering, *Journal of Voice*, Vol. 15, No. 1, pp. 78 – 85, The Voice Foundation.

On the Theory and Methodology of the Digital Inheritance of Speech and Culture

Kong Jiangping

(Research Center for Chinese Linguistics, Peking University播100871, China)

Abstract: The inheritance of speech and oral culture is the most important component part of human civilization. The methods of the inheritance have been developing along with social and technological progress. The diversities of speeches and oral cultures in the world are on the verge of disappearance due to the advance of globalization. This paper discusses the characteristics and shortcomings of the traditional way of recording speech and oral culture by using Chinese characters, and also discusses the basic method of the digital inheritance of speech and oral culture with the analysis about the inaccuracy and imperfect records by means of audio and video, and then proposes a method of digital inheritance based on multi-speech models. Finally, the theoretical basis of the inheritance of speech and oral culture is profoundly discussed from the perspective of cognition of speech and oral culture.

Key words: speech, oral culture, digital method, cultural inheritance