项目语言/方言点录音原则与方法说明

1、简要情况

项目计划在中国境内采录200个点各300个扩展版核心词的语料。200个点考虑语种平衡和地域平衡。每个地点选择2名发音人，分别是老年男性和老年女性，年龄均为50-65岁之间。发音人应说地道本地语言/方言，不会说其他语言/方言。

2、所有录音一定要在本地进行，不能在外地寻找某地发音人录音（如：不能在大学中寻找来自某地的学生录音。原因是后期可能分析不同地区的jitter、shimmer差异等，因此一定需要在本地录音）。

3、项目组成员应充分利用自己的亲戚朋友关系，在自己的家乡录制本地的发音。

4、录制某地发音时需要先将当地最地道的说法整理在Excel表格中（参见附件样本Excel文件：扩展版300核心词表-西南官话重庆巫溪文峰（参考样本）），对于汉语方言来说这一步非常重要，否则就会造成使用的是普通话的说法，只是用当地的声、韵、调读出来而已（例如以西南官话为例，“知道”不是用当地语音读为“tsɿ55 tau24”，这个词在当地大多说的是“晓得”，因此应该采录“晓得”的发音；其他还如以后-二天；结冰-起凌……等等，不一一说明）。因此这一步一定要耐心细致，确保准确无误。

切忌只是用当地的声、韵、调把普通话的词说一遍。

如果某个概念在该方言中没有相应的说法，则在汉字栏、IPA栏均用英文模式的感叹号“!”表示。如果某个概念有多种说法，最多记录3个，按常用程度降序排列。多个说法的国际音标之间用英文模式的分号（;）隔开。多个说法的顺序男、女发音人应该保持一致。

音节与音节之间使用空格分开。如果有变音、变调等，只记变音、变调之后的实际发音。

国际标音字体请使用IPAPANNEW。

塞擦音的音标尽量使用合体符号，如使用ʦ而不是ts（t+s）。其他合体符号还有ʨ、ʣ、ʧ、ʥ、ʤ。实在没有合体符号的使用两个符号相加，如tʂ（t+ʂ）。

轻声请用数字“0”表示，例如“知道”tsɿ55 tau0。

5、选择的发音人应无口、咽、喉、哮喘等疾病。发音时声音洪亮，吐字清晰，语速适中（不可太快）。

6、条件允许的，建议使用电脑、麦克风在实验室或其他安静的环境录音（采用此种录音方式建议使用byly软件，具体可以咨询冉启斌：13920215480（微信同号））。

无法使用电脑录音的，建议使用本项目合作成员熊子瑜老师编写的“九州音集”微信小程序进行录音（好处是项目所需录制的300个核心词词项已经导入其中。具体请与冉启斌联系。如使用其他APP或小程序等录音，也请提前与冉启斌联系）。经测试，不同的手机对声音文件音质可能影响，录音时请采用华为、vivo、苹果等主流品牌手机（2000元以上。如发音人手机不符合要求，由联络人使用自己或亲友的手机注册后录音。多人可用同一台手机录音，但务必提前通知项目组，将原来已录的某个发音人的语音下载下来，否则后面发音人的录音将覆盖前面发音人的录音）。

7、使用“九州音集”手机录音的，录音之前请在小程序中按照提示注册发音人个人信息，然后进行录音。采样率一律请设为44100。时长每个词一般设为4秒。如某个词的音节特别多，可以设为更长。每个词确定地道的母语说法后录音1遍即可。

8、录音时最重要的是选择安静的环境，建议晚上在安静的卧室进行。避免房间外面的汽车声、刮风下雨声等。也要尤其避免房间里的开关门声、走路声等。

录音过程中要严格控制噪声。鞋在地上的摩擦声、衣服的摩擦声等极细微的声音都会产生噪声，因此一定要由联络人提醒发音人，进行严格控制。

9、发音人如文化程度较低，请联络人逐一提示发音人300个词的说法。每个词请用地道的当地语言/方言的表达方式说出。

10、发音人时手机麦克风放在嘴角约5-8厘米左右处，不要正对着麦克风，以免喷麦。

11、发音人以正常语速朗读，切忌发音人急着想读完而读得太快。当然也不宜过于缓慢。发音时请控制发音人不要着急，词语前后留出1秒左右时间。英语的in等，读为“在里面”。

12、男、女2名发音人所有词的发音请完全一致，即同一个词有几种说法、每种说法各有几个音节、每个音节的声韵调等是完全相同的。否则后期自动标注会很麻烦。

13、各地联络人以参与重大项目的方式进行此项工作。如需要参与重大项目的证明请联系项目组。每个点建议用一个星期左右的时间耐心、仔细地进行核心词IPA转写标音、音系录入以及录音工作。

每个点最后需要提交的文件：

（1）扩展版核心词（300词）当地的IPA注音（参见“参考样本”Excel表）；

（2）当地音系（含声母、韵母、声调（如果当地语言有声调的话），参见“参考样本”Excel表）；

（3）录制完成的300扩展版核心词录音文件；

（4）发音人信息登记表（参见“参考样本”Excel表）。

注意：声音文件的有无、有多少种说法，和词表里的标注和注音一定要完全对应起来。

14、所有参考样本的格式请勿改动。

15、每个点采录的过程中如有任何问题请随时联系下面的项目组人员：

冉启斌：13920215480（微信同），ranqibin@126.com

“中国境内语言核心词汇声学数据库及计算研究”（19ZDA300）

重大项目课题组

2020年1月拟，5月修订