

人工智能语音识别和合成已经成熟投入商用，如腾讯微信和科大讯飞的语音输入等。但是目前商用产品仅覆盖普通话和比较广泛使用的方言。温州方言使用的人数相对较少，并且变种很多，发音和语法都有自己特点。

智能识别温州话的关键是建立数据库。数据就是语料库，一句温州话对应一段文字，成熟的语料库里有上千小时的语音。

温州话语料库的建立是一个循环渐进的过程，每一个阶段都可以和应用结合，产生社会效益。

“温州方言是温州人血脉的载体，是联结温州人情感纽带，温州乡土文化传承值得各界人

阶段 0. 原型。利用温州话电视节目《百晓讲新闻》《闲事婆和事老》提取语音文字素材，用最新开源语音识别和生成人工智能模型训练。

阶段 1.书面版的温州话。对着中文文字念出来的温州话,“说文”这种温州话可以用来交流,但是失去了温州话的特色,保留了语

音，但没有温州方言语法和常用语。
可以在特定场合投入使用。

阶段 2. 全汁全味的温州话，主要针对温州城区话。全面保留了温州方言语音，语法和常用语。可以作为各种人工智能软件的温州话语音出入口，如翻译，聊天，人工服务台。可以应用于教育，老年陪伴。

阶段 3. 覆盖温州地区的其他主要方言，如温州城区话、瑞安话、永嘉话、平阳话。收集这些地区的语音语料，对语音，语法和常用语作调整，相对于学会了城

区方言，再学各种地区方言。使人工智能适用于温州各种方言，相对全面的留住温州话。

温州话语料库的建立任重而道远。实现全汁全味的温州话人工智能有技术和实现上的种种挑战。现有的中文标注常常去掉温州话的独特特点，替之以中文的通用说法。但温州话的传承意义重大，值得各位同仁为之努力。

- 30%说不标准,像外地人
- 34%不会讲,听懂小部分
- 33%不会讲也听不懂
- 22%与爷爷奶奶交流困难
- 几乎不会温州童谣、儿歌

独特的语言体貌，“病昏爻罢”
副词后置于动词，“走先、吃添”
动词重叠：“这张桌摆摆动”

有一种甜叫阿嘴万端
有一种害羞叫打情人
有一种爱叫火油心
有一种灿烂叫爆白 有一种疯狂叫黄
有一种惊悸叫醒睡
有一种神经叫挂相 有一种起牌天叫能保体
有一种老不死叫老世勒黄
有一种YES I DO叫 一脚踏进少，一脚命死了，我命偷
有一种生龙活叫冤隔 有一种婆婆叫地瓜娘
有一种菜头星脑叫蛋蛋 有一种大风叫蛋蛋
有一种对不起叫罪过事
有一种无蛋叫有解 有一种死鬼叫吓的时
有一种神气叫消灭 有一种糊涂叫昏瞎
有一种HOLD住叫台型扎牢
有一种刺男刺女叫独自人
有一种找二奶叫撞豪车
有一种呼啦啦叫雨丽叨叨
有一种漂亮叫干俏

◎ 温明

温州话传承关键在于建立温州方言语料库