

# Table of Contents

- [1 技术概览](#)
  - [1.1 文献阅读路径](#)
    - [1.1.1 拼接法](#)
    - [1.1.2 统计参数语音合成](#)
    - [1.1.3 端到端](#)
  - [1.2 拼接法](#)
  - [1.3 统计参数](#)
    - [1.3.1 前端](#)
  - [1.4 端到端](#)
- [2 语料库TODO](#)
  - [2.1 汉语语音合成系统中的拼音](#)
    - [2.1.1 拼音概述](#)
    - [2.1.2 拼音标准风格](#)
      - [2.1.2.1 实际使用的拼音标注风格](#)
      - [2.1.2.2 参考资料](#)
      - [2.1.2.3 声韵母表](#)
    - [2.1.3 解决多音字问题](#)
  - [2.2 语音合成难点与研究方向](#)
  - [2.3 语音合成相关会议&比赛](#)
- [3 概述](#)
  - [3.1 术语表](#)
- [4 行业概况](#)
  - [4.1 国内公司](#)
    - [4.1.1 出门问问](#)
    - [4.1.2 云知声](#)
    - [4.1.3 科大讯飞](#)
    - [4.1.4 思必驰](#)
    - [4.1.5 猎户星空](#)
    - [4.1.6 简单总结](#)
  - [4.2 国外公司](#)
  - [4.3 知名实验室](#)
  - [4.4 知名学者](#)
- [5 产业链](#)
- [6 语音合成工程师可能需要的技能](#)

## 技术概览

## 文献阅读路径

1. 系统入门: [爱丁堡大学TTS课程 \(http://www.speech.zone/courses/speech-synthesis/\)](http://www.speech.zone/courses/speech-synthesis/)
2. 语音合成所需声学基础知识: [语音产生原理与特征参数提取 \(http://blog.csdn.net/u010451580/article/details/51178190\)](http://blog.csdn.net/u010451580/article/details/51178190)
3. 语音合成概述: [牛津大学+google语音合成教学pdf \(https://github.com/oxford-cs-deepnlp-2017/lectures/blob/master/Lecture%2010%20-%20Text%20to%20Speech.pdf\)](https://github.com/oxford-cs-deepnlp-2017/lectures/blob/master/Lecture%2010%20-%20Text%20to%20Speech.pdf)

4. 中文语音合成概述: [Mandarin speech synthesis 2006](http://www.speakit.cn/Group/file/TTS_CSLPBook06.pdf)  
([http://www.speakit.cn/Group/file/TTS\\_CSLPBook06.pdf](http://www.speakit.cn/Group/file/TTS_CSLPBook06.pdf))
5. 如何评测一个语音合成系统? —— [语音合成测评系统介绍](https://testerhome.com/topics/11123) (<https://testerhome.com/topics/11123>)

## 拼接法

todo

## 统计参数语音合成

- A beginners' guide to statistical parametric speech synthesis
  - [英文](http://www.cstr.ed.ac.uk/downloads/publications/2010/king_hmm_tutorial.pdf) ([http://www.cstr.ed.ac.uk/downloads/publications/2010/king\\_hmm\\_tutorial.pdf](http://www.cstr.ed.ac.uk/downloads/publications/2010/king_hmm_tutorial.pdf))
  - [中文](https://shartoo.github.io/texttospeech/) (<https://shartoo.github.io/texttospeech/>)
- 如果使用Merlin工具包, 可看[哥伦比亚大学的tts教程](http://www.cs.columbia.edu/~ecooper/tts/) (<http://www.cs.columbia.edu/~ecooper/tts/>)

## 端到端

推荐的开源代码: [github.com/mozilla/TTS](https://github.com/mozilla/TTS)

推荐笔记: <http://www.erogol.com/text-speech-deep-learning-architectures/> (<http://www.erogol.com/text-speech-deep-learning-architectures/>) (mozilla/TTS开源库作者总结的关于端到端系统的笔记, 非常实用)

应该阅读的论文: 上述笔记中已经列出

## 拼接法

## 统计参数

## 前端

前端概述可见 <http://sentiment-mining.blogspot.com/2017/10/ssmltts.html> (<http://sentiment-mining.blogspot.com/2017/10/ssmltts.html>)

## 端到端

# 语料库TODO

## 汉语语音合成系统中的拼音

出于实际需求, 在语音合成系统中的汉语拼音通常与[国家教育部发布的汉语拼音方案](http://www.moe.edu.cn/s78/A19/yxs_left/moe_810/s230/195802/t19580201_186000.html) ([http://www.moe.edu.cn/s78/A19/yxs\\_left/moe\\_810/s230/195802/t19580201\\_186000.html](http://www.moe.edu.cn/s78/A19/yxs_left/moe_810/s230/195802/t19580201_186000.html)) (下称标准拼音方案) 有所区别。

注意下面使用的韵母v等价于韵母ü.

## 拼音概述

汉语中实际使用的有超过1300个有调音节，400多个无调音节，23个声母，38个韵母，5个声调。总共有超过70,000个汉字，经常使用的有3,500个汉字（覆盖了99.5%的实际使用情况）。

标准拼音方案中规定了23个声母和35个韵母（韵母表中列出的），其中：

- 23个声母包括了21个常规声母和y, w两个零声母，因为y, w实际上并不发音。（另一说为22个声母，即将y, w当做一个零声母看待）
- 35个韵母。
- 五个声调：一声（阴平）、二声（阳平）、三声（上声）、四声（去声）和轻声

实际汉语语音合成系统中可能采用38个韵母，增加了如下3个新的韵母-i（前）、-i（后）、er：

- er 是在[ə]的基础上加上卷舌动作而成。发音例词：而且érqiě 儿歌érgē 耳朵ěrduō 二胡èrhú 二十èrshí 儿童értóng
- -i(前) 指zi/ci/si中的i 发音例词：私自sīzì 此次cǐcì 次子cìzǐ 字词zìcí 自私zìsī 孜孜zīzī
- -i(后) 指zhi/chi/shi/ri中的i 发音例词：实施shíshī 支持zhīchí 知识zhīshi 制止zhìzhǐ 值日zhírì 试制shìzhì

还有一些系统使用了39个韵母，这是因为增加了韵母ê：

- ê[ɛ] 在普通话中，ê只在语气词“欸”中单用【因此一些项目忽略了这个单韵母，即38个韵母】。ê不与任何辅音声母相拼，只构成复韵母ie、üe，并在书写时省去上面的附加符号“ˊ”。

## 拼音标准风格

日常使用的为**标准拼音方案**

([http://www.moe.edu.cn/s78/A19/yxs\\_left/moe\\_810/s230/195802/t19580201\\_186000.html](http://www.moe.edu.cn/s78/A19/yxs_left/moe_810/s230/195802/t19580201_186000.html))，继续往下阅读前请熟悉此标准拼音方案，其特别规定了：

1. “知、蚩、诗、日、资、雌、思”等七个音节的韵母用i，即“知、蚩、诗、日、资、雌、思”等字拼作zhi, chi, shi, ri, zi, ci, si。
2. 韵母er用作韵尾时写成r。例如：“儿童”拼作 ertong，“花儿”拼作 huar。
3. 韵母ei单用写成
4.
  - i行的韵母，前面没有声母的时候，写成yi, ya, ye, yao, you, yan, yin, yang, ying, yong。
  - u行的韵母，前面没有声母的时候，写成wu, wa, wo, wai, wei, wan, wen, wang, weng
  - v行的韵母，前面没有声母的时候，写成yu, yue, yuan, yun, v上两点省略。
  - v行的韵母跟声母j, q, x拼的时候，写成ju, qu, xu, v上两点省略；但是跟声母n, l拼的时候，仍然写成nv, lv。
5. iou, uei, uen前面加声母的时候，写成iu, ui, un。例如 niu, gui, lun。

## 实际使用的拼音标注风格

实际使用的拼音标注风格可能有如下改变

1. 还原被标注为u的韵母v：即将['jv', 'qv', 'xv', 'yv', 'yvan', 'yve', 'yvn', 'jve', 'jvan', 'jvn', 'qve', 'qvn', 'qvan', 'xvan', 'xvn', 'xve']还原回['ju', 'qu', 'xu', 'yu', 'yuan', 'yue', 'yun', 'jue', 'juan', 'jun', 'que', 'qun', 'quan', 'xuan', 'xun', 'xue']
2. 还原niu, gui, lun：即还原niu, gui, lun 为 niou, guei, luen。
3. 还原缩写的r为er：即还原诸如 huar 为 hua er。
4. 将轻声标注为5：轻声原本是不带标注的，如 ba（吧）标注为 ba5，其他声调依次标注为1234
5. 还原yw的实际发音（少见）：即将yi, ya, ye, yao, you, yan, yin, yang, ying, yong, wu, wa, wo, wai, wei, wan, wen, wang, weng, yu, yue, yuan, yun 还原为 i, ia, ie, iao, iou, ian, in, ang, ing, ong, u, ua, uo, uai, uei, uan, uen, uang, ueng, v, ve, van, vn
6. 标注真实发音：这个稍有些复杂，下面详细说明

汉语中，拼音并不一定对应着固定的读音，这是因为汉语中还有一些变调规则(或者说语流音变),虽然在书面拼音上不做标识，但是读的时候还是要求变调的。**简单而言**，有下面几种情况：

- 两个三声连在一起，第一个读二声，如老鼠 (lao3 shu3) 标记成真实发音 (lao2 shu3)
- “一”和“不”的变调，
  - “一”的原调是第一声。当“一”在表示数目时，以及在词尾出现时，念为原调。例如：一二三四、第一等等。当“一”与量词搭配时，如量词的声调是第一声、第二声或第三声，“一”念成第四声，例如“一根”、“一钱”和“一种”。如量词的声调是第四声，“一”念成第二声，例如“一个”。
  - “不”只有在去声音节前才变调为阳平声。在词语之间，“不”读轻声。
- “这”的发音，通常读作 zhe，但是有时候读作 zhei，例如这个发作“zhei ge”
- 有时一些拼音在实际情况可能读轻声 (TODO)

### 如果仔细探讨，变调规则更为复杂（非专业读者可略过）

首先理解术语“五度标记法”，五度标记法是将人的音调（区别于汉语的声调）变化分成5个相对高度，即12345。通常汉语中四种声调的变化是，一声始终保持高音调，即555，二声为345（逐渐升调），三声为214（先降调再升调），四声为531（逐渐降调），由于轻声音调变化曲线很不稳定，因此通常不对轻声的音调曲线作判断。

变调主要有以下几种情况：（下面的数字是汉语普通话调值五度标记法）

1. 当两个上声字连读时，第一个字的声调值变为35，即等于阳平的调值。变调调值是214-35。例如“老鼠”，两字音调都是上声，但发音为阳平、半上声，和“牢鼠”的读法相近但不相同。
2. 当三个上声字连续时，则比较复杂，要分析具体情况。
  - A. 当词语首二字是双音节，而第三字是单音节时，首二字都读阳平声。如“保管好”，发音是阳平、阳平、上声。
  - B. 当词语首字是单音节，而尾二字是双音节时，首字变成214，第二字读阳平声。如“总保管”发音是半上声、阳平、上声。
3. 如果上声字后面接着非上声字，亦即阴平、阳平、去声和轻声前，且该上声字不处于句末、不处于被强调状态时，常读作半上声。（阴平的声调值是55 阳平是35 上声是214 去声是51）
4. 还有就是“一”和“不”的变调，
  - 在去声音节之前，“一”读阳平声，如“一定”。
  - 在非去声音节之前，“一”读去声，如“一天”、“一年”、“一起”。
  - 在词语之间，“一”读轻声。但表示序数时或其他情况下，“一”都读原本的阴平声。
  - “不”只有在去声音节前才变调为阳平声。在词语之间，“不”读轻声，一说，读二声。
5. 如果2个上声在一起不变调,念起来大概像这样:衣冠楚楚 和 衣冠除楚, 正确发音如蒙古, 念成萌古

### 参考资料

- [汉字的变调和轻声规则](https://zh.wikipedia.org/wiki/%E6%B1%89%E8%AF%AD%E6%8B%BC%E9%9F%B3%E8%AF%AD%I) (<https://zh.wikipedia.org/wiki/%E6%B1%89%E8%AF%AD%E6%8B%BC%E9%9F%B3%E8%AF%AD%I>)
- [拼读规则](https://baike.baidu.com/item/%E6%8B%BC%E9%9F%B3/322252#3) (<https://baike.baidu.com/item/%E6%8B%BC%E9%9F%B3/322252#3>)
- 轻声相关可查《普通话水平测试必读轻声词语表》和《普通话水平测试可轻读词语表》

### 声韵母表

声韵母表如下

声母 (23个) b p m f d t n l g k h j q x zh ch sh r z c s y w

韵母 (39个)

- \* 单韵母 a、o、e、ê、i、u、ü、-i (前)、-i (后)、er
- \* 复韵母 ai、ei、ao、ou、ia、ie、ua、uo、üe、iao、iou、uai、uei
- \* 鼻韵母 an、ian、uan、üan、en、in、uen、ün、ang、iang、uang、eng、ing、uen、g、ong、iong

## 韵母 (39个) (转换标注后)

- \* 单韵母 a、o、e、ea、i、u、v、ic、ih、er
- \* 复韵母 ai、ei、ao、ou、ia、ie、ua、uo、ve、iao、iou、uai、uei
- \* 鼻韵母 an、ian、uan、van、en、in、uen、vn、ang、iang、uang、eng、ing、uen  
g、ong、iong

## 解决多音字问题

参考文献:

- 解决多音字字-音转换的一种统计学习方法 (<http://jcip.cipsc.org.cn/CN/article/downloadArticleFile.do?attachType=PDF&id=1181>)

一种实现汉字转拼音编码的方法

汉语语音合成中的多音字处理

<http://cpfd.cnki.com.cn/Article/CPFDTOTAL-ZGZR200408001052.htm>

(<http://cpfd.cnki.com.cn/Article/CPFDTOTAL-ZGZR200408001052.htm>) 一种新的基于规则的多音字自动注音方法

## 语音合成难点与研究方向

1. 合成的速度, 工业界要求快速, 但目前有些技术并不能做到实时合成
2. 不够自然
3. 工业界和学术界的分别 (工业界依然还是拼接合成, 但是学术界已经不主要研究这个了)
4. 目前大量的语言仍然无法合成, 而且针对不同的语言需要独立分析, 例如需要寻找一个合适的phoneset
5. 工业界中的语音合成仍然需要大量的人工介入, 例如专业的发音员, 需要在录音棚录音, 人工转换成文本, 人工做align的检查, 添加特定的词汇到发声词典
6. proper intonation 合适的语调 (韵律)
7. 个性化语音合成 (快速地模仿一个人的声音)
8. 歌唱以及情感语音合成
9. 儿童的语音合成

## 语音合成相关会议&比赛

- 会议: ICASSP and Interspeech
- 比赛: Blizzard Challenge

## 概述

## 术语表

语音合成文献中以及本文可能出现的术语及其解释

### 语音合成相关工具包

- HMM: Hidden Markov Model 输出序列用于描述语音的特征向量, 状态序列表示相应的文字
- HTS: HMM-based Speech Synthesis System 语音合成工具包

- HTK: Hidden Markov Model Toolkit 语音识别的工具包
- SPTK: speech signal precessing toolkit
- POS part of speech
- praat: 标记文本在音频中的对应发音位置的语音分析软件
- Merlin: 开源的统计参数语音合成系统, <https://github.com/CSTR-Edinburgh/merlin>  
(<https://github.com/CSTR-Edinburgh/merlin>).
- tacotron: 使用端到端技术的语音合成系统

## 声学相关参数

- mcep Mel-Generalized Cepstral Reprfresentation
- mcc mel cepstral coefficients
- mfcc Mel Frequency Cepstral Coefficients
- LSP: Line Spectral Pair线谱对参数
- F0 Fundamental Frequency, 基音频率
- mgc 声学参数
- pitch 音高: 表示声音(基本)频率的高低
- Timbre 音色
- Zero Crossing Rate 过零率
- Volume 音量
- bap band aperiodicity \footnote{(<http://blog.csdn.net/xmdxcjsj/article/details/72420051>)}  
(<http://blog.csdn.net/xmdxcjsj/article/details/72420051>).

## 语言学

- initial final 声母和韵母
- sil silence
- syllable 音节
- intonation 声调, 语调, 抑扬顿挫
- 多个音素的命名规则\footnote{(<http://wiki.c2.com/?NumericalPrefixes>)} (<http://wiki.c2.com/?NumericalPrefixes>)
  - monophone 单音素
  - biphone diphone 两音素
  - triphone 三音素
  - quadphone 四音素
- TOBI韵律符号系统ToBI(Tone and Break Index)

## Merlin系统中的文件含义

- .wpa word to phonetic alphabet
- .cmp Composed acoustic features
- .scp system control program
- .mlf master label file
- .pam phonetic alphabets to model
- .mgc mel generalized cepstral feature
- .lf0 log f0 a representation of pitch (音高) 音高用基频表示
- .mgc 声学参数
- .utt .utt files are the linguistic representation of the text that Festival outputs (full context training labels)

## 技术术语

- TTS text-to-speech 语音合成系统
- SPSS : 统计参数语音合成statistical parametric speech synthesis

- Front end 前端 The part of a TTS system that transforms plain text into a linguistic representation is called a frontend
- vocoder 声音合成机（声码器）将声学参数转换为时域音频
- AM Acoustic Model, 声学模型
- LM Language Model, 语言模型
- ASR: Automatic Speech Recognition自动语音识别
- CART Classification and Regression Tree, 分类回归树
- CFHMM Continuous F0, 连续基频模型
- MOS Comparison Mean Opinion Score, 比较平均意见分
- CORC Correlation Coefficient, 相关系数
- DBN Dynamic Bayesian Network, 动态贝叶斯网络
- DCT Discrete Cosine Transform, 离散余弦变换
- EM Expectation Maximization, 期望最大化
- GMM Gaussian Mixture Model, 高斯混合模型
- HMM Hidden Markov Model, 隐马尔科夫模型
- HNR Harmony Noise Ratio, 谐波噪声比
- HSMM Hidden Semi-Markov Model, 隐半马尔科夫模型
- LPC Linear Prediction Coefficient, 线性预测系数
- MAP Maximum A Posterior, 最大后验概率
- MCD Mel-Cepstral Distortion, 倒谱系数失真
- MDL Minimum Description Length, 最小描述长度
- MGCC Mel-Generalized Cepstral Coefficient, 梅尔广义倒谱系数
- MLI Maximum Likelihood Increase, 最大似然增量
- MLSA Mel Log Spectral Approximation, 梅尔对数谱近似
- MLLR Maximum Likelihood Linear Regression, 最大似然线性回归
- MLPG Maximum Likelihood Parameter Generation, 最大似然参数生成
- MOS Mean Opinion Score, 平均意见分
- MSD Multi-Space Distribution, 多空间分布
- PM Prosodic Model, 韵律模型
- RMSE Root-Mean-Square-Error, 根均方误差
- SA Speaker Adaptation, 说话人自适应
- SI Speaker Independent, 说话人无关
- SSML Speech Synthesis Markup Language, 语音合成标记语言
- VC Voice Conversion, 声音转换
- VPR Voice Print Recognition, 声纹识别

## 行业概况

### 国内公司

1. 科大讯飞、百度、捷通华声、腾讯、搜狗、阿里巴巴
2. 知名行业公司：思必驰，云知声，出门问问、猎户星空、Rokid、
3. 初创公司 深声科技 灵伴科技
4. 语料库与数据标注公司：标贝科技、海天瑞声

其中提供了语音开放平台的有：讯飞语音，百度智能语音开放平台，阿里云，腾讯云，云知声，思必驰，捷通华声

### 出门问问

核心技术 语音识别 语义分析 垂直搜索 语音合成 智能推送

李志飞对雷锋网回复称，目前出门问问在国内主要定位为软件+硬件+AI 三位一体的公司，在海外则将主要战略定位为 AI消费类电子公司，同时尽量充分利用谷歌的生态系统的优势，在 2015 年 10 月，Google 出资约 5000 万美元参与了出门问问的 C 轮融资。大众汽车集团（中国）决定出资 1.8 亿美元参与出门问问的 D 轮融资，双方还将共同出资成立一家研发智能汽车产品的合资公司，各持股 50%。

## 云知声

黄伟先后就职于摩托罗拉和盛大创新院语音分院，Unisound AI Labs（云知声人工智能研究院）研究员全部拥有硕士及以上学历，其中博士占比超过45%，20170817，3亿融资，

## 科大讯飞

科大讯飞“主要”做的是2B（面向企业用户）服务，即将自己的人工智能服务授权给别的企业使用。现在和京东合作做了一些2C（面向消费者）产品，比如智能音箱。

## 思必驰

CEO：高始兴，剑桥大学商学院硕士 联合创始人/首席科学家：俞凯-剑桥大学语音博士，上海交大教授  
aispeech-上海交大实验室

## 猎户星空

猎户语音OS现已与行业多家知名企业展开合作，如小米、喜马拉雅、美的、当当、易视腾等。截止目前，小米AI音箱已经全量应用了猎户的TTS（语音合成）技术以及ASR（语音识别）技术。此外，小米电视也同步接入猎户星空的相关技术。

## 简单总结

思必驰更专注于垂直领域智能硬件的语音交互，基于软硬件一体化解决方案拓展to B方向合作，聚焦于车载、家居、机器人领域；出门问问则以智能穿戴为切入点，将AI应用场景逐步向车载、家居领域扩展，走软硬结合的路线。

相比之下，云知声则是从前期提供开放语音云平台服务到后续聚焦物联网，试图将技术能力从云端迁移到终端，通过AI芯、AIUI端、AI Service三大解决方案构建“云端芯”产品战略，重点布局家居、车载、医疗领域，从而形成完整的生态闭环。

## 国外公司

todo

1. 微软、google、亚马逊、apple、IBM、Nuance

## 知名实验室

todo

1. 中科院声学所
2. [自动化所人机语音交互课题组-陶建华 \(http://www.speakit.cn/show/index\\_introduce?typeid=1\)](http://www.speakit.cn/show/index_introduce?typeid=1)
3. 中科大
4. [西北工业大学-音频、语音与语言处理研究组 \(http://www.npu-aslp.org/\)](http://www.npu-aslp.org/)



# 知名学者

陶建华

## 产业链

与语音合成相关的技术有：

1. 语音识别
2. 语音合成
3. 语义理解/分析
4. 垂直搜索
5. 智能推荐
6. 语音助手

## 语音合成工程师可能需要的技能

### 基础能力

- 数学——良好的数学功底
- 机器学习
- 软件工程能力
- 数据结构
- Linux——awk sed perl
- 精通类Linux平台下的C/C++语言开发，精通数据结构算法，能够迅速实现算法

### 专业知识

- 深入理解开源代码如merlin HTS tensorflow HTK unit selection；
- 有语音合成、识别、评测、自然语言处理等专业知识者优先；
- 自然语言处理如：分词、词性标注、NER、句法分析等相关经验；
- 语音相关知识如：语音编码、滤波器、平滑算法等信号处理经验；
- 深度学习模型如RNN/LSTM/CNN等深度学习模型，及其应用于NLP或speech的相关经验。

### 算法类能力

- 对语音识别/合成的算法细节有深刻的了解
- 熟悉语音识别/模式识别/神经网络/信号处理/自然语言处理/语音合成等相关算法，对算法细节有深刻的了解，了解语音识别的各模块关键技术

### 语音信号

- 具有音频信号处理和分析经验，精通音频分析的常用方法和工具，如语音频谱分析。
- 熟练掌握音频信号处理算法，包括AEC、AGC、去混响、噪声抑制、BeamForming等；

### 经验能力

- 主导过商用的项目，有模式识别和机器学习研究（完整项目）
- 有语音领域的工作或研究经验，搭建过语音合成系统或熟悉HTS者优先

### 虚能力

- 具有很强的分析问题和解决问题的能力，对解决具有挑战性问题充满激情

- 有强烈的上进心和求知欲，自我管理能力强，有良好的时间意识
- 善于交流和分享
- 钻研能力，团队合作能力

## 论文/科研能力

- 在Interspeech, ICASSP等语音学术会议中有论文发表者优先

## 可能的具体工作

- 前端
- 负责前端文本分析中分词和韵律模型训练及优化算法的研究；
- 声码器
- 从事声码器改进工作
- 从事语音合成声学建模研发工作
- 合成系统
- 负责语音合成系统搭建与改进
- 负责离在线语音合成系统的搭建与改进
- 承担语音合成在互联网平台的应用项目。
- 从事垂直领域定制合成的效果提升工作

## 算法研究

- 探索基于拼接的合成方法，提高参数合成方法的合成效果
- 负责大数据下基于深度学习的声学模型训练方法的研究；
- 从事基于深度神经网络的端到端语音合成研发工作
- 从事小数据自适应研发工作
- 负责跟踪国内外最新的端到端、参数化语音合成技术或者拼接语音合成技术；
- 负责高质量后处理算法和合成器的研究或者单元挑选算法的研究；

## 对于语音识别工程师

- 声学模型、语言模型、解码器三个方向之一的相关工作

## 举例

思必驰语音合成-本科- <https://www.lagou.com/jobs/3590699.html?source=pl&i=pl-5>  
(<https://www.lagou.com/jobs/3590699.html?source=pl&i=pl-5>)

专业/方向：计算机、电子、计算机语言

学历：本科以上

工作内容：

1. 语音合成相关算法研究和应用；
2. 语音合成相关数据搜集、分析及质量控制。 任职资格：
3. 本科及以上学历，专业与语音、模式识别、自然语言处理相关，对语音语言处理领域有浓厚的兴趣，有较好的理论基础和快速的学习能力。
4. 熟练掌握C/C++/perl，熟悉常用的数据结构，能够迅速实现算法；
5. 严谨有责任心，具有钻研精神和团队合作精神。

In [ ]: