微软学生俱乐部实践空间站

基于认知服务语音API的创新性科研拓展课堂语音记录助手

武汉大学

2019

目录

[成员介绍 2](#_Toc26048997)

[项目背景 3](#_Toc26048998)

[解决的问题 4](#_Toc26048999)

[创新解决方案及竞品分析 5-6](#_Toc26049000)

[手机语音输入 5](#_Toc26049001)

[语音翻译 5](#_Toc26049002)

[成果说明和截图 7](#_Toc26049003)

[在使用语音API过程中遇到的问题 9](#_Toc26049004)

# 成员介绍

|  |  |  |
| --- | --- | --- |
| 姓名 | 学校 | 专业 |
| 郑子杰 | 武汉大学 | 计算机科学与技术 |
| 叶博辰 | 武汉大学 | 计算机科学与技术 |
| 江景文 | 武汉大学 | 计算机科学与技术 |
| 熊立谦 | 武汉大学 | 计算机科学与技术 |

# 项目背景

痛点：解决语音信息记录的难题。

大学课堂较之中学的课堂，老师授课速度加快，课堂内容大大增加，内容难度提升……课堂上，许多授课重点会在学生仍在思考课堂上新的概念而来不及记录思考及学习。

本次微软实践空间站的项目，我们希望通过调用微软的语音API，把老师课堂授课的语音转化为文本，并对文本进行分词、提取关键词。用户对照生成的文本文档复习和理解课堂内容，从而让课堂的内容利用率最大化。

# 解决的问题

在如今这个节奏飞快地社会，所有人都追求高效率。工作，学习，生活都是如此，每个人都想用最短的时间完成任务。作为学生，我们每天有大量的课程。以前，记录信息我们习惯用笔和纸，赶笔记，记录采访、演讲内容。但现在，有了语音录音转文字成为现实，还能随时存储为文本，彻底解放双手，再也不用手动速记。同时，提取文本中的关键词，抓住老师讲课要点，提升学习效率，关键词作为文本储存，方便课后复习。

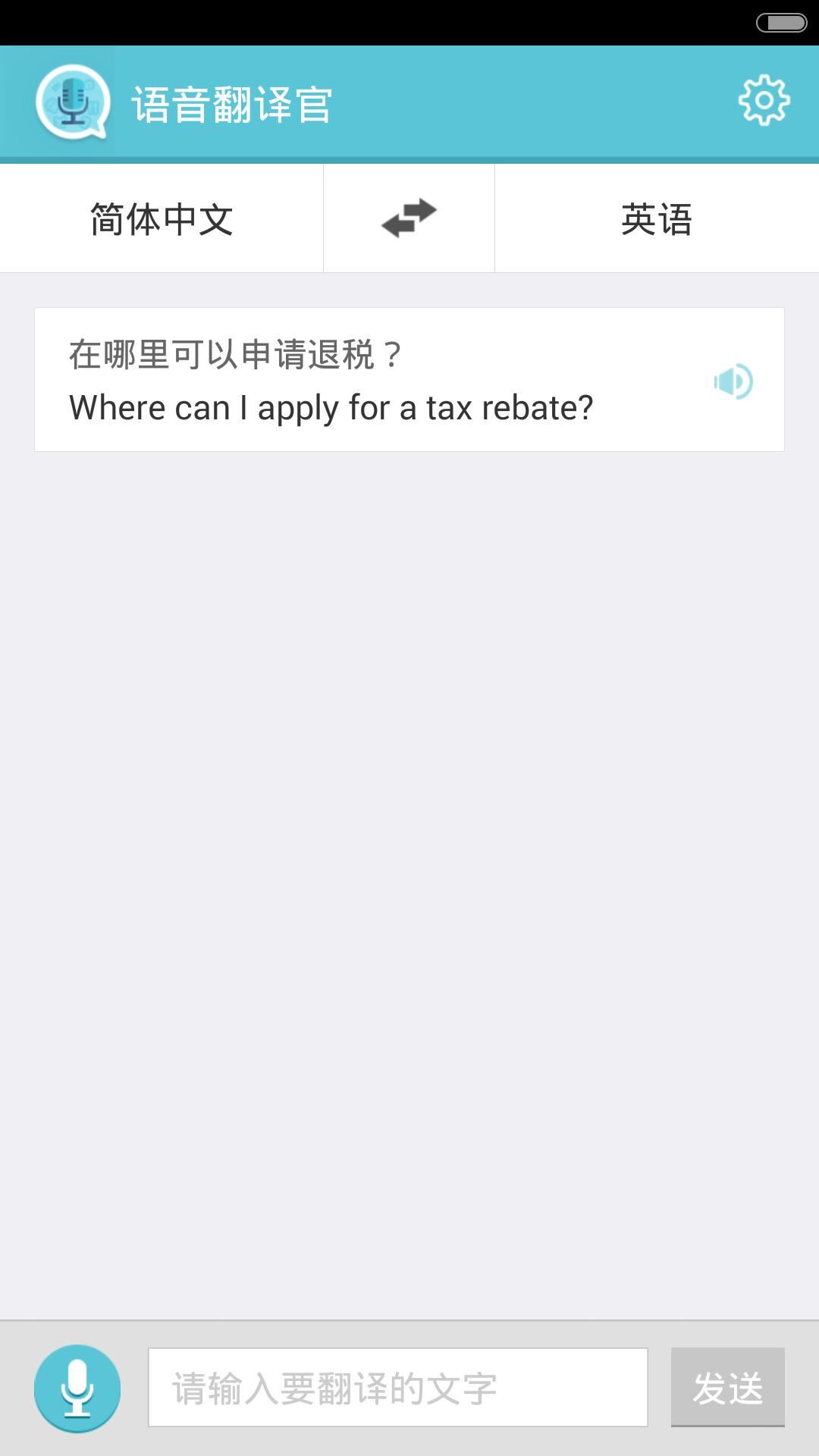
# 创新解决方案及竞品分析

目前大多数语音API的应用以呈现为主，如常见的：

手机语音输入



语音翻译



以上应用更加针对于即时、短长度文本，但当语音识别出的文本过长时（如音频时间过长，整体输入信息量过大，语速偏快）便出现以下不足：

1. 语音输入时间短，说话内容无法连贯。
2. 输出结果直接呈现，无法迅速得出梗概内容，给读者带来不便。
3. 输出结果需手动保存，降低效率。

我们更希望通过该段文章的关键词来了解文章的基本内容，这就是我们课堂语音助手的主要用途。

课堂语音助手的优势：

1. 语音输入时间长。
2. 以文本形式保存语音应用输出，便于反复利用查看。
3. 智能生成关键词，帮助读者了解演说者的大致内容。
4. 输出结果可反复查看，便于多次利用。

通过课堂语音助手，学生可以：

1、快速了解课程主题，便于梳理学习进程。

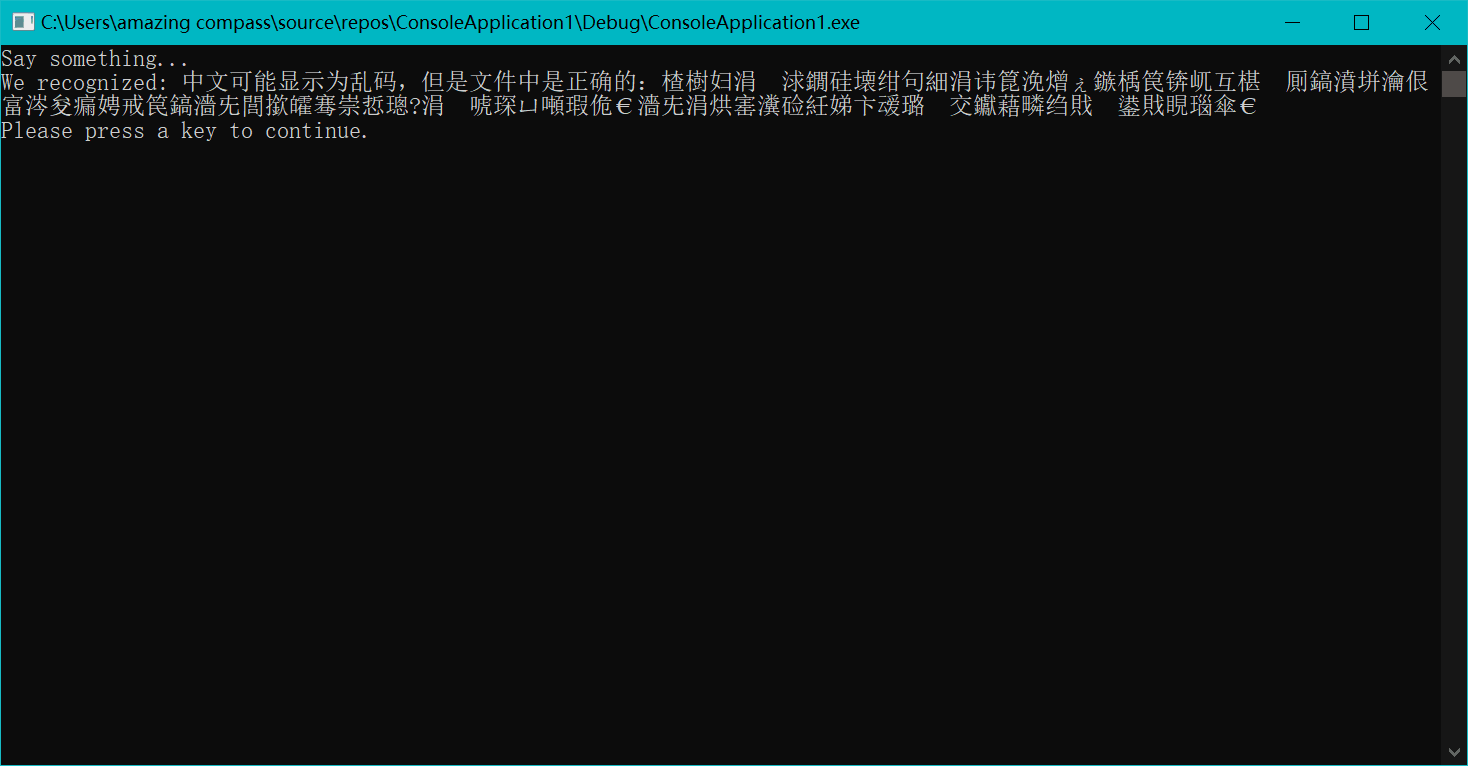
2、教学内容自动保存，便于复习查看。

3、无需手动反复操作，确保一堂课的内容连贯录入。

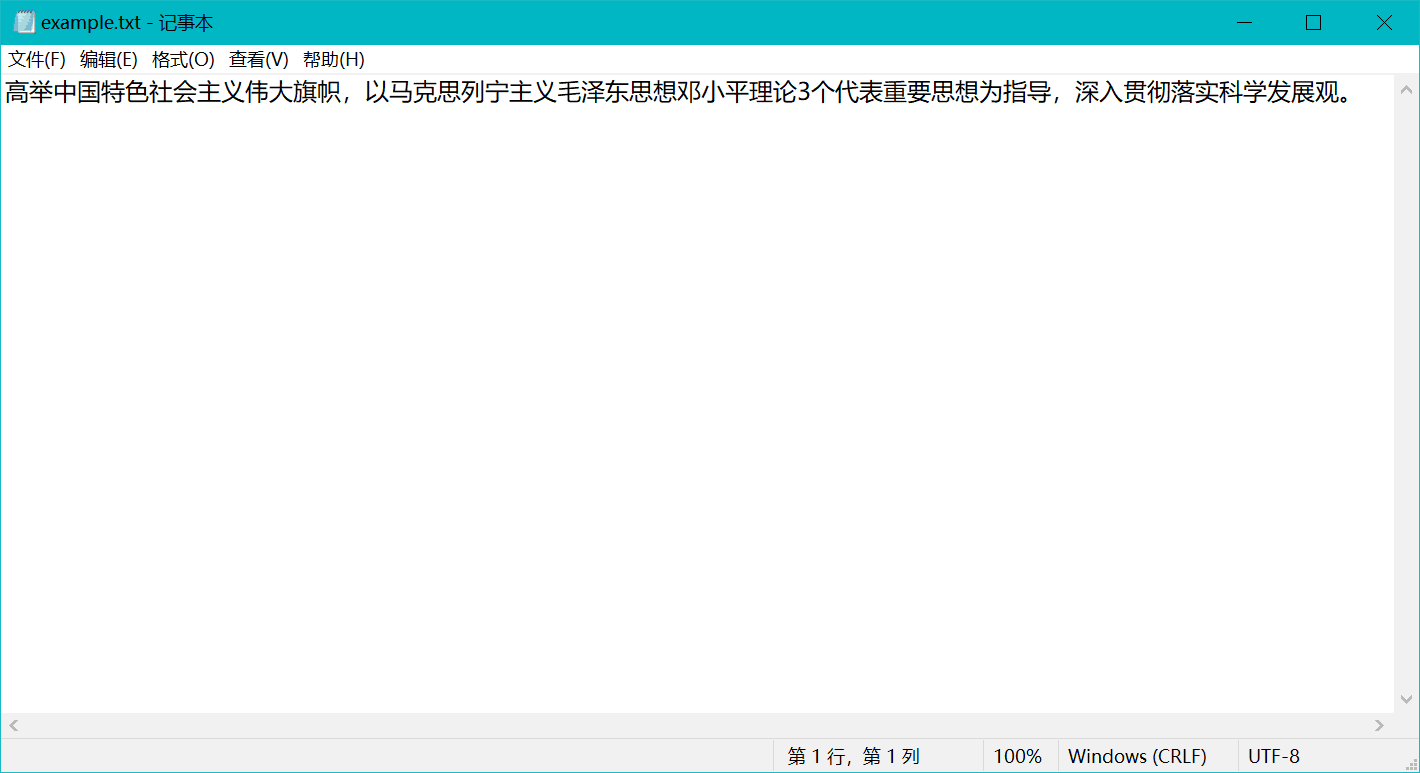
通过语音课堂助手，学生的学习将更加智能，更加轻松！

# 成果说明和截图

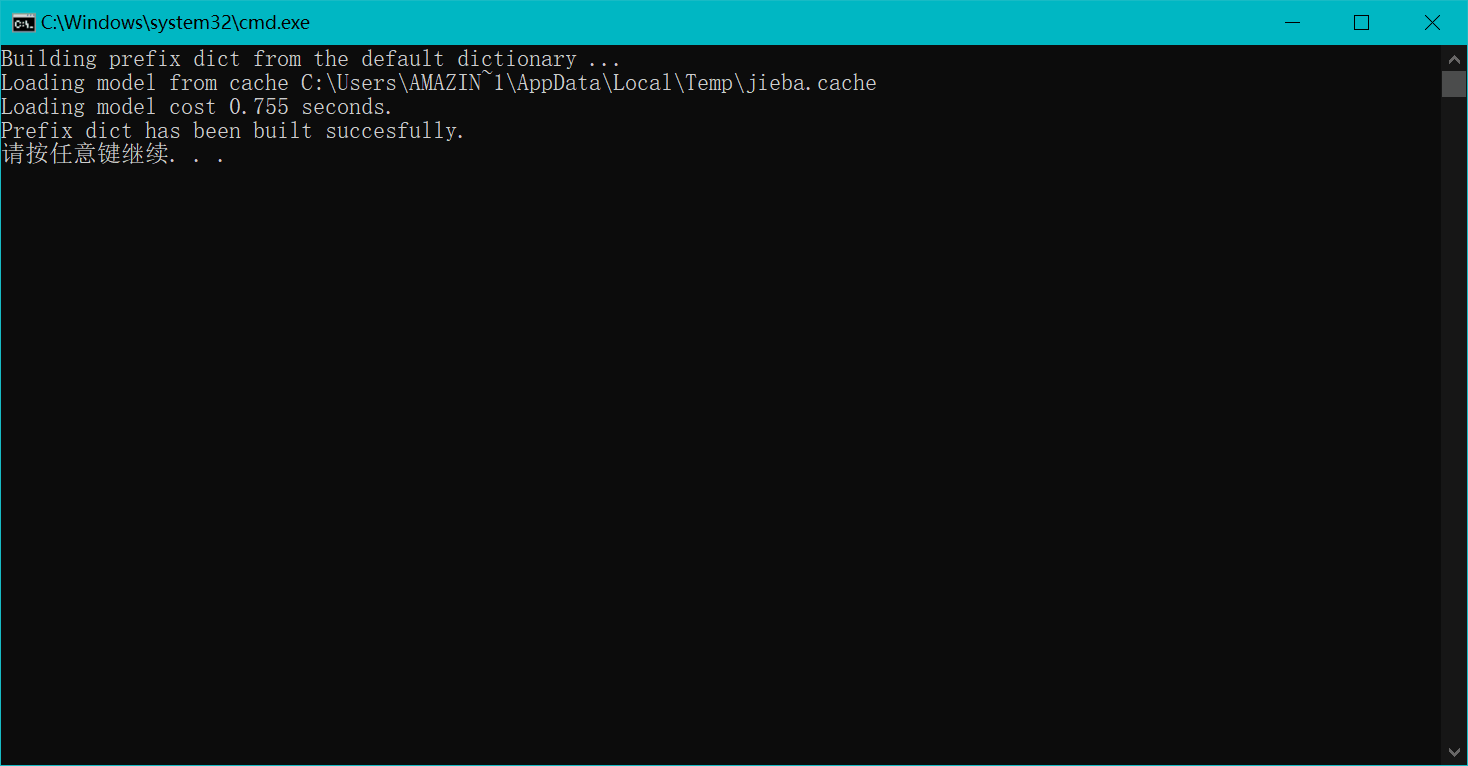
双击执行ConsoleApplication1.exe文件，出现如下页面表示开始录音。（测试版本的录音时长为15秒）



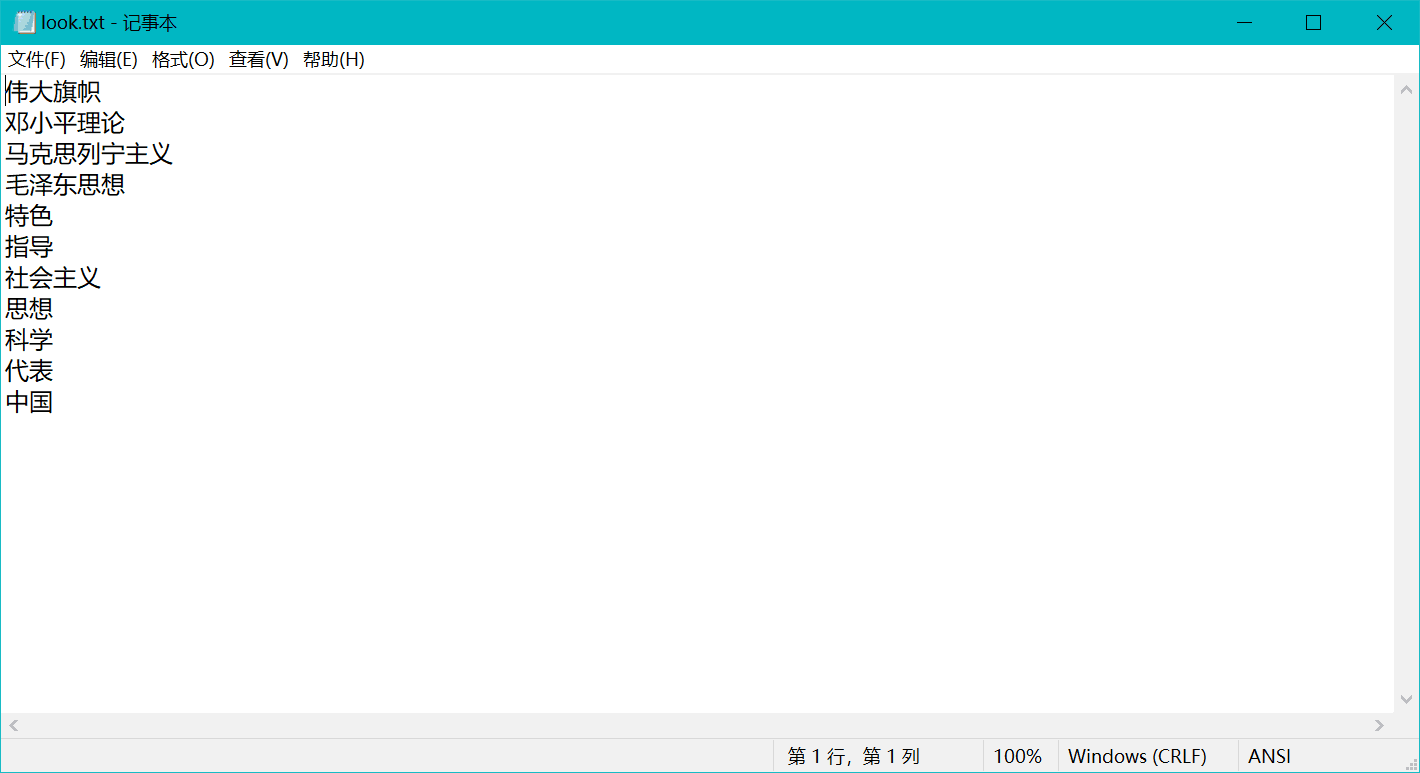
录音结束后，ConsoleApplication1.exe所在文件夹自动将识别文字转化为文本并存储在example.txt文件中。



运用python的结巴分词工具，对example.txt中的文本进行分词，依据出现频率等因素，有权重的依次输出关键词，并将其储存在look.txt文件中。



图为用户查看结巴分词工具对调用语音API识别的文字生成的文本文件进行分词的结果。



用户对照关键词，就可以有侧重的对课堂录音文本进行学习。

# 在使用语音API过程中遇到的问题

在整个项目中，调用语音API部分未出现问题；同时微软的语音API功能十分强大，对语音的识别的正确率极高且对部分词汇会智能的屏蔽。