# Ekho

## 简介

**Ekho（余音）**是一个免费、开源的中文语音合成软件。它目前支持粤语、普通话（国语）、广东台山话、诏安客语、藏语、雅言（中国古代通用语）和韩语（试验中），英语则通过eSpeak或Festival间接实现。Ekho支持Linux、Windows和Android平台。

Ekho采用波形拼接合成法，把汉字转换成拼音，每个拼对应一个音频文件，最后把音频文件拼接起来。

## 测试流程

环境：

Linux，gcc

准备：

下载源码：<http://www.eguidedog.net/cn/ekho_cn.php>，解压ekho-7.6.tar.xz。

编译：

打开语音合成参考网站🡪Ekho(余音) - 中文语音合成软件🡪How to Install Ekho，按照文档说明进行。

测试：

测试文本：ekho\_en.txt：Hello, I’m speaking English.

ekho\_zh.txt：你好，我在说中文。

①读取文本文件，并保存到语音文件：

ekho -f ekho\_en.txt -o ekho\_en.wav

ekho -f ekho\_zh.txt -o ekho\_zh.wav

②直接输入文本，产生语音后自动播放：

ekho “Hello, I’m speaking English.”

ekho “你好，我在说中文。”

结果：

①在当前目录下输出ekho\_en.wav和ekho\_zh.wav。

②执行测试命令后自动播放合成语音。

## 性能

实现了中英文语音合成。英文语音合成质量良好，而中文语音合成质量一般。

# Festival

## 简介

Festival提供了一个通用的框架，用于构建语音合成系统，该系统包含了各种模块示例。它提供了完整的文本转语音的API，可以通过shell、C++静态库等多种方式调用，可以移植到多种平台，支持的语言包括英语和西班牙语；系统是使用c++编写。

Festival作为一套完整的TTS系统，拥有完整的语音合成模块，从语音库的建立（festival为此专门分离出festvox工具，用户可以自行构建语音库）、文本分析、韵律处理到合成模块。Festival最初版本采用双音子拼接合成方法：用线性预测（LPC）误差合成语音，然后用PSOLA算法优化韵律。后来双音子合成器（diphone synthesizer）改进为UniSyn synthesizer，优化了拼接单元不是双音子的情况（could be used even when the units being concatenated are not diphones）。

最新festival版本采用了基于HMM的合成方法，合成质量有了明显的提升。

## 测试流程

环境：

Linux，gcc

安装：

下载源码：[http://festvox.org/packed/festival/2.5](http://festvox.org/packed/festival/2.5/)。

1.解压festival-2.4-release.tar.gz。

2.按照install文件说明下载其他包：festlex\_CMU.tar.gz，festlex\_POSLEX.tar.gz，festvox\_kallpc16k.tar.gz和speech\_tools-2.4-release.tar.gz。

3.创建festival文件夹，包含解压后的festival文件以及以上4个工具包。

编译：

参考网页<https://blog.csdn.net/drink_tea/article/details/78346225>进行编译。

测试：

测试文本：festival\_en.txt：Hello, I’m speaking English.

①读取文本文件，产生合成语音后自动播放：

festival --tts festival\_en.txt

②直接输入文本，产生合成语音后自动播放：

执行命令festival，(SayText “Hello, I’m speaking English”)

执行命令festival，(tts “festival\_en.txt”)

③读取文本文件，并保存到语音文件：

cat festival\_en.txt | text2wave –o festival\_en.wav

text2wave festival\_en.txt -o fesitval\_en2.wav

结果：

①执行测试命令后自动播放合成语音。

②执行测试命令后自动播放合成语音。

③在当前目录下输出festival \_en.wav和fesitval\_en2.wav。

## 性能

实现了英文语音合成。语音合成质量良好。

# Flite

## 简介

Flite是一个小型、快速的TTS系统，是festival-diphone合成方法的精简版本，采用C语音编写，可用于嵌入式系统，支持WinCE、Palm OS。

Flite设计中兼容Festival的voiees和models，Flite使用HRGS（heterogeneous relation graphs）作为它的基本的发音代表结构。而且合成器的大部分内容是数据，比如字典，单元数据库等(lexicons，unitdatabaseetc)，实际的代码量相当的小。对于festival，大部分数据存在在外部文件中，依需要而下载，而对于Flite来说，因为它是非常小的一个引擎，所以允许核心数据在ROM中。

从架构上看，Flite包括两个部分。

1. Flite库包含核心合成代码。

2. Flite声音，这包含三个子部分

(1)语言模型：字串处理，韵律模型text processing, prosody models etc.

(2)字典和字符到声音规则Lexicon and letter to sound rules.

(3)单元库和声音定义Unit database and voice definition

 Flite是基于diphone原理合成的，Diphone合成中所使用的是一种简约语音库，它包括一种语言中出现的所有diphone(汉语中叫做双音素)，diphone的数量是基于这种语音的因素的数目，比如:西班牙语有大约800个diphone，德语有2500左右。在diphone合成中，每个dihone只有一个例子是包含在语音库里。 在运行时，句子的韵律处理是基于这些最小单位的叠加，通过数字信号处理技术手段，诸如LPC、PSOLA或MBROA。由此而致的合成效果比单元必须系统的差一些，但是比共振峰好一点（espeak采用的方法）。 由此文本到语音的基本过程是，文字的预处理，从输入的文本到获得一个个的单词，界定出的标点符号等需停顿处，再由单词和diphone库词典，转换成相对应的diphone代表，再由录制好的diphone数据库对应的波形、韵律等属性，配合音高、音长等调节，最后采用PLOSA算法（基音同步叠加技术）拼接，确保平滑过渡。

## 测试流程

环境：

Linux，gcc

准备：

下载源码：<http://www.festvox.org/flite/download.html>。

解压flite-2.0.0-release.tar.bz2。

编译：

按照README. MD文件的编译步骤进行。

测试：

测试文本：flite\_en.txt：Hello, I’m speaking English.

①读取文本文件，并保存到语音文件：

flite -f “Hello, I’m speaking English.” –o flite\_en.wav.

②直接输入文本，产生语音后自动播放：

flite -t “Hello, I’m speaking English.”

结果：

①在当前目录下输出flite\_en.wav。（在ubuntu系统中打开文件报错：常规流错误）。

②执行测试命令后自动播放合成语音。

## 性能

实现了英文语音合成。语音合成质量良好。

# Flite+engine\_HTS

## 简介

flite+hts\_engine是一个英语TTS系统，flite+hts\_engine是以flite为前端、hts作为参数模型，由官方推出，采用CMU ARCTIC语料库。

HTS是最为著名的基于HMM统计参数合成工具，由以日本Keiichi Tokuda为中心组成团体开发，可以和多种前端处理工具配合使用festival，MaryTTS等。

## 测试环境

环境：

Linux，gcc

安装；

下载源码：<http://hts-engine.sourceforge.net/>。

1.解压flite+hts\_engine-1.07.tar.gz

2.下载hts\_voice\_cmu\_us\_arctic\_slt-1.06（由CMU训练好的HTS语音），并放在flite+hts\_engine目录下。

编译：

使用一般的安装命令进行安装即可。

测试：

需要在hts\_voice\_cmu\_us\_arctic\_slt-1.06下进行测试。

测试文本：flite+hts\_engine\_en.txt：Hello, I’m speaking English.

在hts\_voice\_cmu\_us\_arctic\_slt-1.06中写入测试文本，然后在该目录下执行测试命令：

flite+hts\_engine -m cmu\_us\_arctic\_ic\_slt.htsvoice -o flite+hts\_engine\_en.wav flite+hts\_engine\_en.txt

结果：

在当前目录下输出了flite+hts\_engine\_en.wav。

## 性能

实现了英文语音合成。语音合成质量良好。

# espeak

## 简介

eSpeak是eGuideDog计划的一个部分，是C语言写的一个小型的、开放源码的语音合成系统，支持多种语言。eSpeak使用共振峰合成方法，一种典型的参数合成法，这可以使提供的语言文件非常小，但是缺点是不如采用基于人录音的语料库的声音平滑自然。该系统支持 Windows、Linux同时已被移植到其它平台，如Solaris and Mac OSX，支持命令行、动态链接库，支持Windows平台上的SAPI5。另外eSpeak可以将文本转换成音素代码，因此它也可以用于另一个语音合成引擎的前端，提供给其他TTS语音合成模块使用。

## 测试流程

环境：

Linux，gcc

准备：

下载源码：<http://espeak.sourceforge.net/download.html>。

1.解压espeak-1.48.04-source.zip。

2.在解压后的espeak文件夹中放入espeak-data包。

编译：

根据参考网页<https://blog.csdn.net/wind_of_dawn/article/details/44594757>，编译portaudio语音引擎，espeak。

测试：

测试文本：espeak \_en.txt：Hello, I’m speaking English.

espeak \_zh.txt：你好，我在说中文。

①根据参考网页<https://www.douban.com/note/585335315/>，在src目录下编写测试文件hello.c，在hello.c中设置文本，编译后运行。

②读取文本文件，产生合成语音后自动播放：

espeak “Hello, I’m speaking English.”

espeak -vzh “Hello, I’m speaking English.”

③直接输入文本，并保存到语音文件：

espeak “Hello,I’m speaking English.” -w espeak\_en.wav

espeak -vzh “你好，我在说中文。” -w espeak\_zh.wav

结果：

①执行测试命令后自动播放合成语音。

②在当前目录下输出espeak\_en.wav和espeak\_zh.wav。

③执行测试命令后自动播放合成语音。

## 性能

实现了中英文语音合成。语音合成质量一般。

# tritium

## 简介

tritium是一款免费、优质的语音合成引擎，具有便携、小巧和快速的特点。

## 测试流程

环境：

Linux，gcc

安装：

下载源码：<https://github.com/syb0rg/tritium>。

解压tritium-master.zip。

编译：

按照README.MD文件以及<http://www.bubuko.com/infodetail-2146117.html> 参考网页的说明方法进行编译。

测试：

测试文本：tritium\_en.txt：Hello, I’m speaking English.

在build目录下，执行测试命令：

①读取文本文件，产生合成语音后自动播放

./tritium\_main -f tritium\_en.txt

②读取文本文件，并保存到语音文件

./tritium\_main -f tritium\_en.txt -o tritium\_en.wav

结果：

①执行测试命令后自动播放合成语音。

②在当前目录下输出tritium\_en.wav。

## 性能

实现了英文语音合成。语音合成质量一般，有颤音。

# MaryTTS

## 简介

Mary是一个采用Java开发的、多语种的文本转语音平台，它支持：德语、英语、美式英语、泰卢固语、土耳其语和俄语。由德国DFKI语言技术实验室等单位推出，支持基于HMM的统计参数合成方法。

## 测试流程

环境：

windows，jdk1.8

安装：

下载源码：<https://github.com/marytts/marytts-txt2wav/tree/maven>。

1.解压marytts-txt2wav-maven.zip。

2.eclipse导入该maven项目，使用maven install命令将java程序打包成jar。

测试：

测试文本：mary\_en.txt：Hello, I’m speaking English.

采用java -jar命令执行程序：

java -jar target/txt2wav-1.0-SNAPSHOT.jar -i mary\_en.txt -o mary\_en.wav

结果：

在当前目录下输出mary\_en.wav。

## 性能

实现了英文语音合成。语音合成质量良好。

# Cainteoir-Engine

## 简介

Cainteoir-Engine 是一款读取和记录不同文件格式（如：ePub，HTML，MHT，RTF，email等）到多种音频输出格式的工具，音频输出格式包括pulseaudio、WAV 和 Ogg/Vorbis。它提供了以下命令行工具：cainteoir、前端的Cainteoir文本到语音库、metadata、提取元数据文件的RDF元组和tagcloud、生成标签和标签云数据。采用C++编写，只支持Linux平台。

Cainteoir支持共振峰参数，双音素两种合成方法，前端能将文本转化为这两种方法处理的音素格式，系统目前用espeak引擎实现了共振峰合成模块。系统也可以扩展参数合成模块（合成器）。

## 测试流程

环境：

Linux，gcc

安装：

下载源码：<https://github.com/rhdunn/cainteoir-engine/>。

1.解压cainteoir-engine-master.zip，ucd-tools.zip。

2.将ucd-tools放在cainteoir-engine-master目录下。

编译：

按照INSTALL文档说明，进行编译。

测试：

测试文本：cainteoir\_en.txt：Hello, I’m speaking English.

执行测试命令：cainteoir cainteoir\_en.txt --output=cainteoir\_en.wav

结果：

在当前目录下输出cainteoir\_en.wav。

## 性能

实现了英文语音合成。语音合成质量良。

# HanTTS

## 简介

HanTTS是一个基于波形拼接的汉语TTS系统，自带录制语音库，python编写，利用python丰富的类库资源简洁实现了TTS功能，采用jieba分词库和pypinyin汉语专文字库，将汉语句子变成拼音，然后在语音库中查找语音单元，拼接合成。

## 测试流程

环境：

Windows，Python3.x

安装：

下载源码：<https://github.com/junzew/HanTTS>。

1.按照requires.txt安装第三方库：pypinyin，jieba，pydub和pyaudio。

2.解压HanTTS-master.zip，根据README.zh.md文件，下载语音库syllables.zip，并解压安装到HanTTS-master目录下。

测试：

运行main.py，在控制台输入文本：你好，我在说中文。

结果：

输出文本对应的拼音，并自动播放合成语音。

## 性能

实现了中文语音合成。语音合成质量良好。

# FreeTTS

## 简介

FreeTTS是完全采用 Java 开发的语音合成系统，它是卡内基梅隆大学基于 [Flite](http://www.oschina.net/p/flite) 这个小型的语音合成引擎开发的。

基于FreeTTS的语音合成技术研究：

https://wenku.baidu.com/view/439acfdb50e2524de5187e80.html?from=search

## 测试流程

环境：

Windows，jdk1.8

安装：

下载源码：<https://freetts.sourceforge.io/#download_and_install>。

1.解压freetts-1.2.2-bin.zip，MBROLA zip、还有声音文件。

2.按照<https://freetts.sourceforge.io/demo/JSAPI/HelloWorld/README.html>官方网站进行环境安装，以便进行JSAPI demo的测试。

测试：

在freetts-1.2目录下，参考<https://freetts.sourceforge.io/#run_demo>。

执行java -jar bin/HelloWorld.jar等命令。

问题（未解决）：

System property "mbrola.base" is undefined.  Will not use MBROLA voices.

<https://stackoverflow.com/questions/2486985/exception-with-freetts-when-using-kevin-or-mbrola>，根据以上参考网页进行安装、设置，但依旧报错。

结果：

按照命令测试bin目录下的jar文件后，自动播放合成语音。

## 性能

实现了英文语音合成。语音合成质量良好。