

# 语音合成(TTS) API说明

## 1. 接口描述

接口请求域名: <https://openapi-ms.mobvoi.com/api/tts/v1>

接口请求频率限制: 5次/秒

出门问问语音合成技术 (TTS) 可以将任意文本转化为语音, 实现让机器和应用张口说话。出门问问 TTS 技术可以应用到很多场景, 例如, 在视频APP中作为配音解说; 小说 App 完成有声阅读, 移动 App 语音播报新闻; 智能设备语音提醒; 车载导航语音合成的个性化语音播报等。

出门问问TTS服务提供了普通话、台湾腔、粤语、四川话、东北话等多种方言, 数百个发音人, 上千种风格, 满足客户在不同场景的选择需求。实时合成支持 SSML, 语法详见 SSML 标记语言。

## 2. 请求参数

**HTTP Method:** 支持POST请求

**调用参数及说明:**

|   | A          | B         | C         | D  |
|---|------------|-----------|-----------|--|
| 1 | <b>字段名</b> | <b>必填</b> | <b>类型</b> | <b>描述</b>  |
| 2 | text       | 是         | String    | 要合成的文本内容, 限制为1000字符。支持ssml标记说明见附录3。  |
| 3 | appkey     | 是         | String    | 开发者在AI开放平台上申请的appkey。  |
| 4 | signature  | 是         | String    | 签名, 通过“appkey+secret+timestamp”进行md5的32位MD5值。其中加号也参与MD5的计算。每次请求实时计算签名, 签名有效期为10分钟。 |
| 5 | timestamp  | 是         | Long      | 当前时间戳, 单位为秒。   |
| 6 | speaker    | 否         | String    | 合成音频指定发音人<br>默认值: cissy_meet<br>其他发音人传值及计费价格请参考声音商店。                               |
| 7 | audio_type | 否         | String    | 合成音频的格式<br>默认值: mp3<br>可选值: pcm/mp3/speex-wb-10/wav<br>只支持这四种格式中的一种                |
| 8 | speed      | 否         | Float     | 发音人合成的语速, 支持小数点后一位<br>默认值: 1.0<br>可选值: 0.5-2.0                                     |
| 9 | convert    | 否         | String    | 默认值: 无<br>可选值: robot<br>是否转化为机器声   |

|    |              |   |         |  |
|----|--------------|---|---------|--|
| 10 | rate         | 否 | Long    | 音频采样率<br>默认值：无，由speaker指定默认值<br>可选值：8000/16000/24000   |
| 11 | volume       | 否 | Float   | 合成音量<br>默认值：1.0<br>可选值：0.1-1.0   |
| 12 | pitch        | 否 | Float   | 语调参数，参数小于0则语调变低，反之则高<br>默认值：0<br>可选值：-10<pitch<10<br>(streaming接口不支持)  |
| 13 | symbol_sil   | 否 | String  | 符号停顿时长映射方法（逗号分割）见下方停顿符号映射<br>需要指定appkey开通权限<br>(streaming接口不支持)  |
| 14 | ignore_limit | 否 | Boolean | 默认值：false<br>可选值：false/true<br>是否限制字符数，如果设置true，传输的文本可以超过限制<br>需要指定appkey开通权限  |
| 15 | gen_srt      | 否 | Boolean | 可以控制是否生成对应的srt字幕文件。当ignore_limit和audio_type为wav可以返回字幕，其他类型不行。默认不生成字幕文件。生成字幕文件需要额外付费，价格详情请参考 <a href="#">价格表</a> 。字幕文件地址通过response header返回，参考下面response header样例。<br>默认值：false<br>可选值：false/true |
| 16 | merge_symbol | 否 | Boolean | 粗粒度合成参数，默认为false。可以指定为true，打点会更接近真人效果。<br>默认值：false<br>可选值：false/true  |

## Request Header设置

```
1 {"Content-Type": "application/json"}
```

## Request Body格式

body采用json格式传输

```
1 {"signature": "appkey+secret+timestamp的MD5值", "timestamp": "1665717322", "appkey": "appkey"}
```

### 3. 返回值说明

如果成功合成，则直接返回语音流。

#### Response Header格式

```
1 {"Content-Type": "audio/mpeg", "srt_address": "https://mobvoi-speech-public.cn-bj
```

body中是合成后的音频数据，可以保存成音频文件，或者交给播放器进行语音播放。

如果合成音频文件失败，返回json格式信息，包含具体的错误原因。

#### Error Response Body

```
1 {"status": "error", "errorMessage": {"code": 31002, "desc": "failed to convert tex
```

### 符号停顿映射表

|   | A   | B   |
|---|-----|---|
| 1 | 符号  | <b>设置方式</b><br><b>(名称+时长/ms)</b><br><b>时长为10的倍数，且大于等于0。</b> |
| 2 | ； ； | semi_200  |
| 3 | ！ ！ | exclamation_200   |
| 4 | ？ ？ | question_200  |
| 5 | ， ， | comma_200   |
| 6 | 。 。 | stop_200  |
| 7 | 、 、 | pause_200   |

### 4. 接口Demo

目前语音识别提供了Java、Python2、Python3、Go语言四个版本的demo，示例代码如下：

- Java 示例代码 [点击下载](#)
- Python2 示例代码 [点击下载](#)
- Python3 示例代码 [点击下载](#)
- Go 示例代码 [点击下载](#)