

合成webapi接口

- [1. 简介](#)
- [2. 接口说明](#)
 - [2.1 接口地址](#)
 - [2.2 请求说明](#)
 - [2.3 IP白名单](#)
 - [2.4 参数](#)
 - [2.5 返回值](#)
- [3. 请求示例](#)
- [4. 错误码](#)
- [5. 常见问题](#)
- [6. 附录](#)
 - [6.1 发音人列表](#)

1. 简介

讯飞开放平台通过 *REST API* 的方式给开发者提供语音合成的 HTTP接口，基于该接口，开发者可以实现将文字信息转化为声音信息，方便开发者使用自己熟悉的编程语言快速集成。

2. 接口说明

2.1 接口地址

POST [http\[s\]://api.xfyun.cn/v1/service/v1/tts](http[s]://api.xfyun.cn/v1/service/v1/tts)

2.2 请求说明

```
POST http[s]://api.xfyun.cn/v1/service/v1/tts HTTP/1.1
Host: api.xfyun.cn
Content-Type : application/x-www-form-urlencoded; charset=utf-8
X-Appid: abcd1234
X-CurTime: 1502607694
X-Param: eyJzY2VuZSI6Im1haW4ifQ==
X-Checksum: e0e6bb78a426cf38dfe4b0f3de931b6b
Content-Length: 224
```

1. 授权认证，调用接口需要将 AppId，CurTime, Param 和 CheckSum 信息放在 HTTP 请求头中；
2. 接口统一为 UTF-8 编码；
3. 接口支持 http 和 https；
4. url说明：[http\[s\]://api.xfyun.cn/{version1}/service/{version2}/tts](http[s]://api.xfyun.cn/{version1}/service/{version2}/tts)
 - HTTP/1.1：Http协议1.1版本，目前最主流使用最广泛的http协议版本
 - api.xfyun.cn：域名
 - {version1}：一级版本号，v1，v2等
 - service：统一服务接口名称
 - {version2}：二级版本号，v1，v2等（一级版本号相同时，二级版本号向前兼容）
 - tts：语音合成
5. 请求方式：POST

2.3 IP白名单

在调用所有业务接口时，授权认证通过后，检查调用方 ip 是否在讯飞开放平台配置的 ip白名单 中。存在通过，否则拒绝提供服务。

注：

- 1. IP白名单 由第三方在页面中配置；
- 2. 拒绝提供服务返回值：

```
{
  "code": "10105",
  "desc": "illegal access|illegal client_ip",
  "data": null,
  "sid": "xxxxxx"
}
```

2.4 参数：

在Http Request Header配置以下参数：

参数	格式	说明	必须
X-Appid	string	讯飞开放平台注册申请应用的应用ID(appid)	是
X-CurTime	string	当前UTC时间戳，从1970年1月1日0点0 分0 秒开始到现在的秒数	是
X-Param	string	参数，各参数组成的JSON串经Base64编码后的字符串，见详细说明	是
X-Checksum	string	令牌，计算方法：MD5(apiKey + curTime + param)，三个值拼接的字符串，进行MD5哈希计算（32位小写），其中apiKey由讯飞提供，调用方管理。	是

在Http Request Body配置以下参数：

参数	类型	必须	说明	示例
text	string	是	待合成文本（使用 utf-8 编码），建议长度小于400字节（中文小于200 汉字），合成后音频时长小于40s（见常见问题说明）	科大讯飞是中国最大的智能语音技术提供商

注：

- apiKey：接口密钥，创建应用时自动生成，调用方注意保管，如果泄露，可在页面中进行重置；
- checksum 有效期：出于安全性考虑，每个 checksum 的有效期为 5 分钟(用 curTime 计算)，同时 curTime 要与标准时间同步，否则，时间相差太大，服务端会直接认为 curTime 无效；
- BASE64 编码采用 MIME 格式，字符包括大小写字母各26个，加上10个数字，和加号 + ，斜杠 / ，一共64个字符。

checksum 生成示例：

```
String apiKey="abcd1234";
String curTime="1502607694";
String param="eyJhYXVmIjogImF1ZGlvL0wxNjtyYXR...";
String checksum=MD5(apiKey+curTime+param);
```

- 请求头中X-Param参数说明：

X-Param：各配置参数组成的JSON串经BASE64编码之后的字符串，其原始JSON串各字段说明如下：

参数	类型	必须	说明	示例
auf	string	否	音频采样率，可选值：audio/L16;rate=8000（8k16bit），audio/L16;rate=16000（16k16bit），默认为audio/L16;rate=16000	audio/L16;rate=16000
aue	string	否	音频编码，可选值：raw（线性pcm音频），默认为raw	raw

参数	类型	必须	说明	示例
voice_name	string	是	发音人，可选值：详见附录 发音人列表	xiaoyan
speed	string	否	语速，可选值：[0-100]，默认为50	50
volume	string	否	音量，可选值：[0-100]，默认为50	50
pitch	string	否	音高，可选值：[0-100]，默认为50	50
engine_type	string	否	引擎类型，可选值：aisound（普通效果），intp65（中文），intp65_en（英文），mtts（小语种，需配合小语种发音人使用），x（优化效果），默认为inpt65	intp65
text_type	string	否	文本类型，可选值：text（普通格式文本），默认为text	text

X-Param生成示例：

```
原始JSON串：
{
  "auf": "audio/L16;rate=16000",
  "aue": "raw",
  "voice_name": "xiaoyan",
  "speed": "10",
  "volume": "10",
  "engine_type": "intp65",
  "text_type": "text",
  "pitch": "10"
}
BASE64编码（即X-Param）：
ewoJICAgICJhdWYiOiAiYXVkaW8vTDE2O3JhdGU9MTYwMDAiLAoJICAgICJhdWUiOiAiAicmF3IiwKCSAgICAidm9pY2VfbmFtZSI6ICJ4aWFvewFuIiwKC
```

2.5 返回值：

需要根据 Content-type的头部来确定是否服务端合成成功。

若**合成成功**，响应body为二进制音频流，响应header为 Content-type: audio/mpeg；sid: hts0000bb3f@ch3d5c059d83b3477200;

该音频保存类型由入参的aue决定：

- aue=raw, 音频文件类型为 wav

若**合成出现错误**，响应body为json字符串，具体header信息为：Content-type: text/plain。json字段如下：

参数	类型	说明
code	string	结果码(具体见错误码)
data	string	空
desc	string	描述
sid	string	会话ID

其中sid字段主要用于DEBUG追查问题，如果出现问题，可以提供sid给讯飞技术人员帮助确认问题。

3. curl请求示例

请求：

```
curl -X POST http[s]://api.xfyun.cn/v1/service/v1/tts -H "X-Appid: 594b62c3 " -H
"X-CurTime: 1502184180" -H "X-Checksum: 001388491350e266fab5e15da9aea749" -H
"X-Param: ewoJICAgICJhdWYiOiAiYXVkaW8vTDE2O3JhdGU9MTYwMDAiLAoJICAgICJhdWUiOiAiAicmF3IiwKCSAgICAidm9pY2VfbmFtZSI6ICJ4aWF
"text=科大讯飞是中国最大的智能语音技术提供商"
```

响应体：

失败：

```
header:
  Content-Type:text/plain
body:
  {
    "code": "10106",
    "desc": "invalid parameter|invalid speed",
    "data": "",
    "sid": "hts0000bb3f@ch3d5c059d83b3477200"
  }
```

成功：

```
header:
  Content-Type:audio/mpeg
  Connection:keep-alive
  Cache-Control:no-cache
  sid:hts0000bb3f@ch3d5c059d83b3477200
body:
  音频流数据（二进制）
```

4. 错误码：

错误码	描述	说明	处理方式
0	success	成功	
10105	illegal access	没有权限	检查apiKey，ip，checkSum等授权参数是否正确
10106	invalid parameter	无效参数	上传必要的参数，检查参数格式以及编码
10107	illegal parameter	非法参数值	检查参数值是否超过范围或不符合要求
10109	illegal text length	文本长度非法	检查上传文本长度是否超过限制
10110	no license	无授权许可	提供请求的 appid、auth_id 向服务商反馈
10114	time out	超时	检测网络连接或联系服务商
10700	engine error	引擎错误	提供接口返回值，向服务商反馈
11201	appid authorize number not enough	服务单调用次数超过限制	确认服务单调用次数是否超过限制
11200	no vcn authorize	无该发音人授权	检查是否获取该发音人效果授权

5. 常见问题：

1. 常见错误码查询及解决方法参见 讯飞语音云错误码 <https://shimo.im/sheet/w3yUy39uNKs0J7DT>
2. 接口超时限制 30s，超过返回504。合成1.5s音频引擎处理约耗时1s。建议待合成文本不要超过400字节（长文本可切割成若干短文本），合成后音频时长不超过40s，每秒内单个IP请求数在20以内。长文本和高并发会影响服务器性能、增加超时概率。

6. 附录

6.1 发音人列表

访问讯飞开放平台，获取最新最全的发音人列表：http://www.xfyun.cn/services/online_tts?tab_index=1

小语种发音人、优化效果权限需单独申请。

