

# TVS Device SDK Android 版本接入说明

拟制： kangrong

日期： 2017 年 8 月

审核：

日期：

深圳腾讯计算机系统有限公司

版权所有 不得复制

## 版本修订记录

[illegible]

目录

TVS Device SDK Android 版本接入说明 .....1

    版本修订记录.....2

1. 简介.....4

    1.1. 名词解释 .....4

2. 入门.....4

    2.1. SDK 使用步骤.....4

        2.1.1. 复制库.....4

        2.1.2. SDK 初始化 .....5

    2.2. 语音唤醒.....6

        2.2.1. 功能.....7

        2.2.2. 回调.....7

        2.2.3. 交互流程.....7

    2.3. 语音语义识别 .....8

        2.3.1. 功能.....8

        2.3.2. 回调.....9

        2.3.3. 参数设置.....9

        2.3.4. 交互流程.....9

    2.4. 语音合成 .....12

        2.4.1. 功能.....12

        2.4.2. 回调.....13

        2.4.3. 参数设置.....13

        2.4.4. 交互流程.....13

# 1. 简介

## 1.1. 名词解释

**TVS DEVICE SDK:** 提供腾讯语音唤醒、语音识别、语音合成、语义理解、语义服务技术解决方案的软件开发工具包。

语音唤醒: 智能硬件/应用在休眠状态下通过个性化语音唤醒词被唤醒。

语音识别: 将语音转变为对应的语句文本。

语义理解: 将语句文本转变成结构化的实体、领域、意图、服务数据。

语音合成: 将语句文本转变为流利的语音。

# 2. 入门

## 2.1. SDK 使用步骤

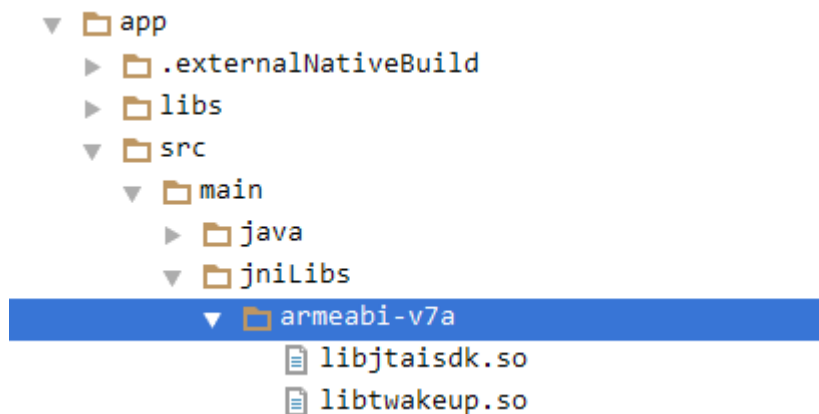
### 2.1.1. 复制库

#### 1. so 库

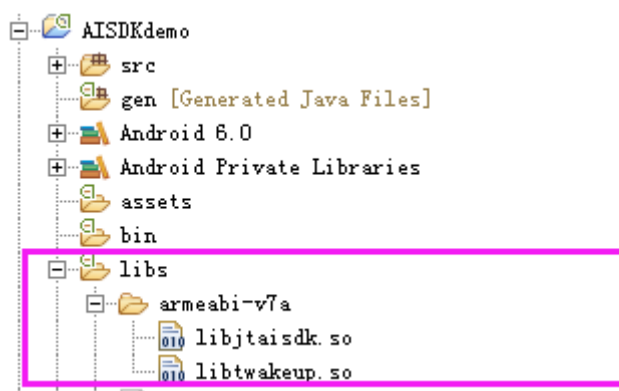
首先, 需要把 SDK zip 包内的 so 库放置在项目 `libs` 目录下:

如果是 Android Studio 开发的项目, 放置在 `jniLibs/armeabi-v7a` 下面 (放置的位置与你项目设置有关, 总而言之, 把 so 放到应用里)

如



如果是 eclipse 开发的项目, 放置在 `libs/armeabi-v7a` 下



## 2. jar 包

AISDK.jar 包放置在项目的 libs 下面即可。

### 2.1.2. 增加权限

在 AndroidManifest.xml 中增加以下权限：

```
<uses-permission android:name="android.permission.RECORD_AUDIO" >
</uses-permission>
<!-- 用于读取手机当前的状态 -->
<uses-permission
android:name="android.permission.READ_PHONE_STATE" >
</uses-permission>
<!-- 写入扩展存储，向扩展卡写入数据，用于写入离线定位数据 -->
<uses-permission
android:name="android.permission.WRITE_EXTERNAL_STORAGE" >
</uses-permission>
<!-- 访问网络，网络定位需要上网 -->
<uses-permission android:name="android.permission.INTERNET" />
<!-- 获取运营商信息，用于支持提供运营商信息相关的接口 -->
<uses-permission
android:name="android.permission.ACCESS_NETWORK_STATE" />
```

### 2.1.3. SDK 初始化

使用 SDK 任何接口前，必须完成 SDK 初始化。

SDK 初始化需要调用 com.tencent.ai.sdk.control.SpeechManager 的一系列方法。

初始化代码如下：

```
//初始化函数
```

SpeechManager.getInstance().startUp(Context context, String info);

info 是 JSON 格式数据，其格式如下：

```
{
  "info": {
    "appkey": "填入应用的 appkey",
    "token": "填入应用的 access token",
    "deviceName": "固定，填入 CAR 或者 TV 或者 SPEAKER 或者 PHONE",
    "productName": "自定义, 英文产品名, 不要有空格",
    "vendor": "厂商英文名"
  },
}
```

对于 deviceName 的指定，按照如下原则：

产品类型	deviceName
车机类	CAR
电视类	TV
音箱类	SPEAKER
手机 app 类	PHONE

其他注意的方法：

//设置当前位置坐标，不仅当初初始化的设置，当用户位置变化的时候，也需要设置。

SpeechManager.getInstance().setCurrentLocation(Location location)

//设置语音识别手动模式

SpeechManager.getInstance().setManualMode(true);

## 2.2.输入音频要求

SDK 的唤醒识别和在线语音识别对输入音频的要求是一致的，要求如下表所示：

项目	要求
音频格式	PCM 格式
采样精度	16 位
采样率	16000Hz
声道	1 声道（单声道）
字节序	小端

## 2.3.语音唤醒

### 2.3.1. 功能

唤醒的接口，可以识别音频流中的唤醒词。上层可以根据识别结果进行进一步的操作。

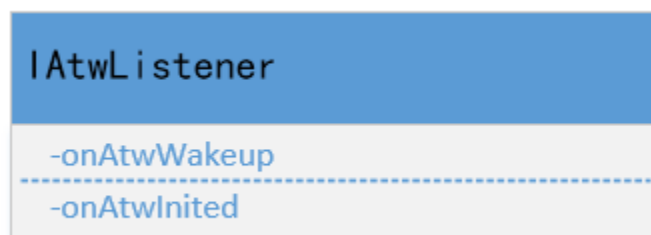
SDK 不能进行录音，所以需要上层把录音的数据实时传过来。对输入语音格式要求：16K 采样率、S16-LE、单声道的 PCM 录音。

功能由 `IAtwSession` 类提供。

### 2.3.2. 回调

上层应用可以注册 `IAtwListener` 对象接收唤醒的异步回调，初始化的返回和唤醒结果都是通过该对象回调回来。其类成员为：

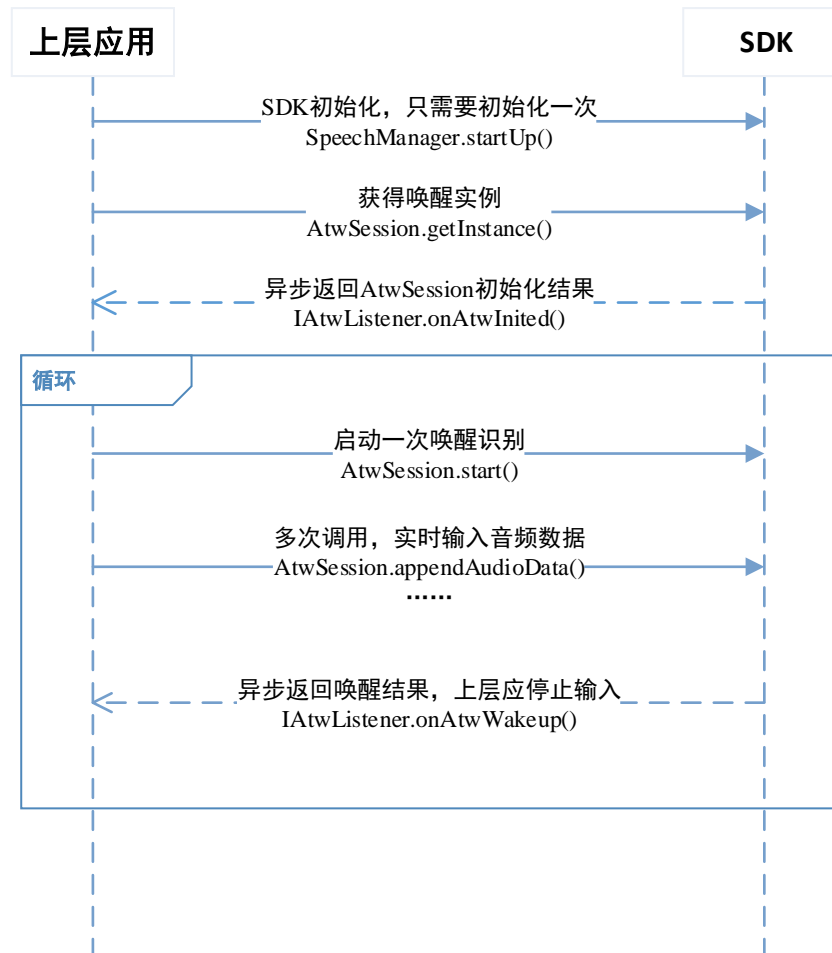
注意：初始化成功后才能进行其他操作。



### 2.3.3. 交互流程

注意：初始化的时候需要传入 `keywords_model`（唤醒词模型）的路径。

## Android SDK唤醒流程



## 2.4. 语音语义识别

### 2.4.1. 功能

语音识别的接口，可以识别音频流中的文字，并返回到上层，方式为流式识别（例如用户还在说话中，就返回一部分文本）。语义的接口，可以将文本识别为的 `domain` 和 `intent`、语义实体，并返回对应的服务数据。例如把“我想听周杰伦的歌曲”识别为 `domain` 为 `music`、`intent` 为 `play`，带有的语义实体是歌手名字为“周杰伦”，服务数据为周杰伦的歌曲列表。语音/语义的功能都由 `TrSession` 提供。

**SDK 不能进行录音，所以需要上层把录音的数据实时传过来。对输入语音格式要求：16K 采样率、S16-LE、单声道的 PCM 录音。**

`TrSession` 返回的语义结果为 JSON 格式，具体请参考《SDK 语义格式》



语音识别支持自动模式和手动模式。自动模式下，SDK 自动检测语音结束，并返回语音识别结果。手动模式下，上层控制语音的结束。

### 2.4.2. 回调

上层应用可以注册 `com.tencent.ai.sdk.tr.ITrListener` 对象接收语音语义识别的异步回调，初始化的返回和结果都是通过该对象回调回来。

注意：初始化成功后才能进行其他操作。

### 2.4.3. 参数设置

`TrSession.setParam` 可以设置语义语音识别的参数。

设置 `TrSession` 工作模式

param: `TrSession.ISS_TR_PARAM_VOICE_TYPE`

值范围：

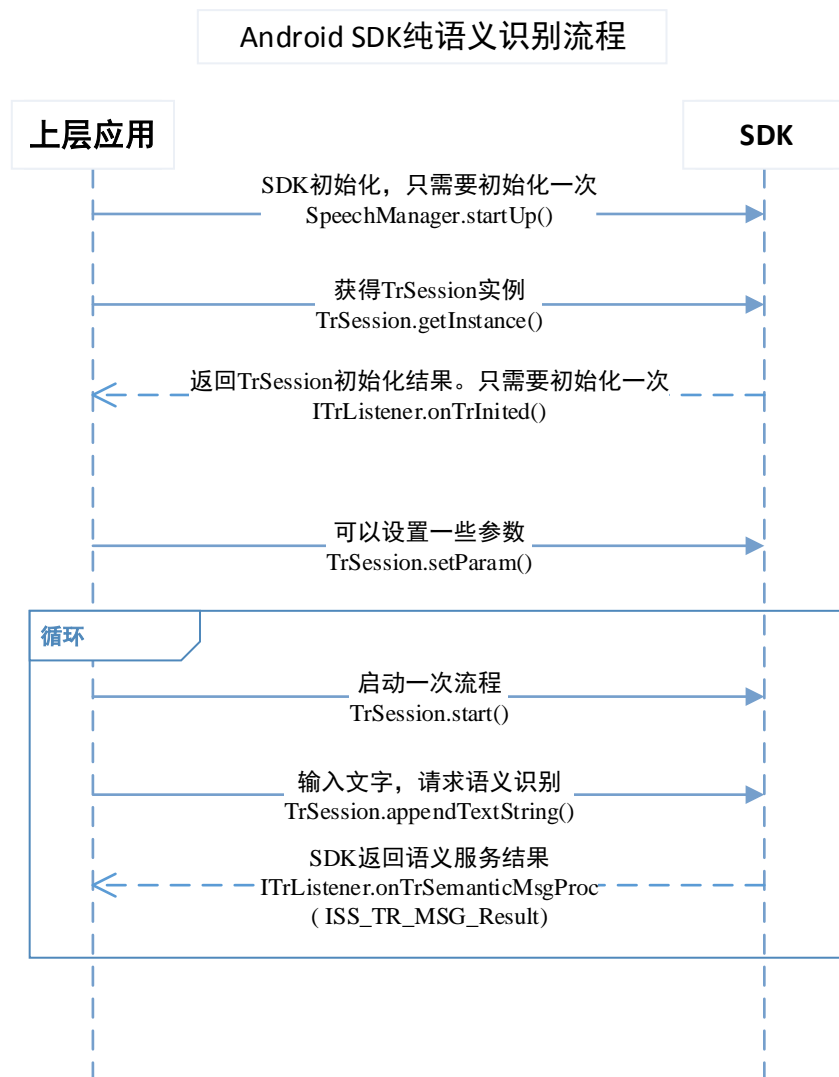
`TrSession.ISS_TR_PARAM_VOICE_TYPE_RSP_VOICE`：仅返回语音识别结果。

`TrSession.ISS_TR_PARAM_VOICE_TYPE_RSP_ALL`：语音+语义识别

如果只需要语义识别，请调用 `appendTextString` 方法

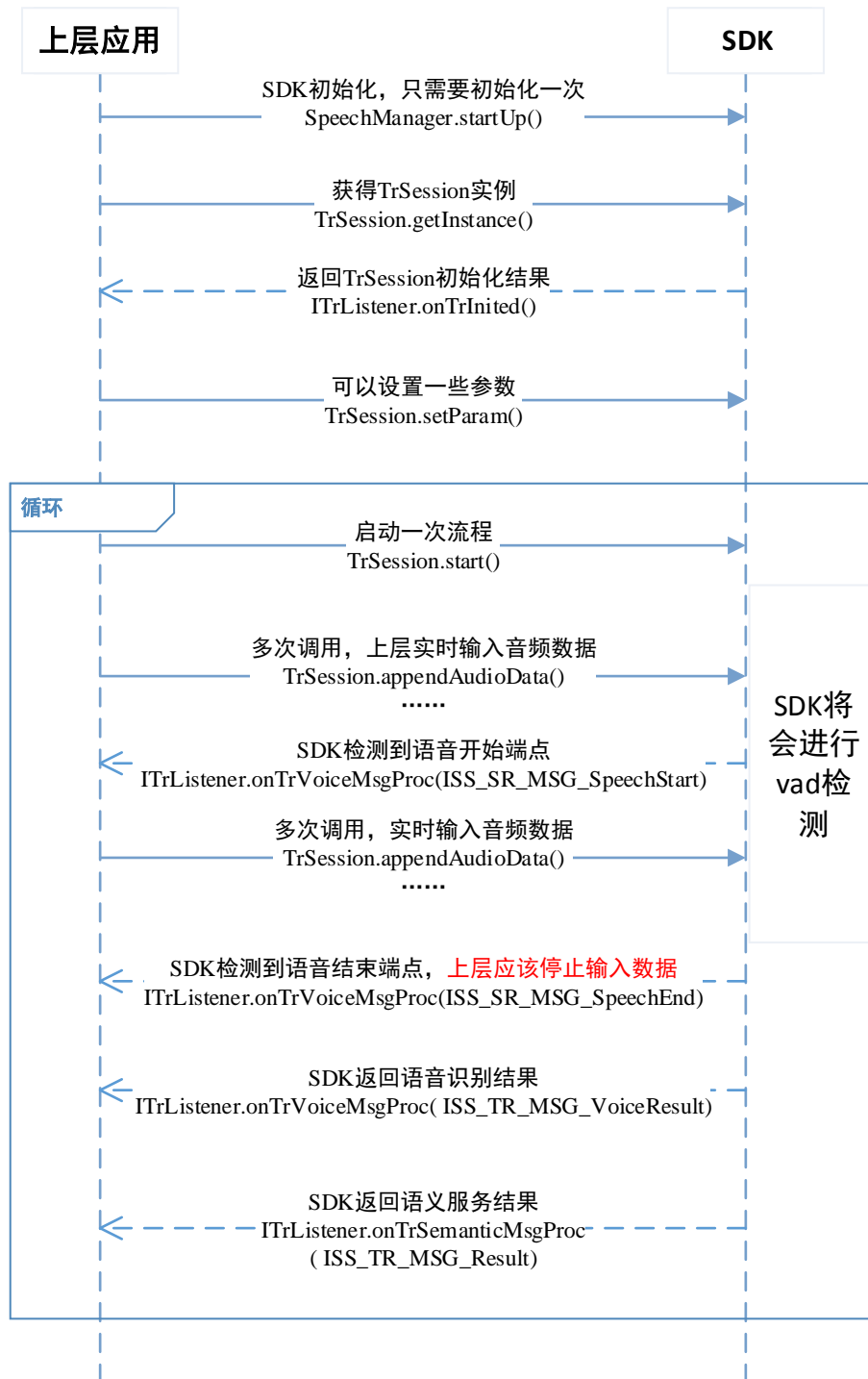
### 2.4.4. 交互流程

#### 2.4.4.1. 纯语义识别的流程



#### 2.4.4.2. 语音+语义识别的流程（自动模式，默认）

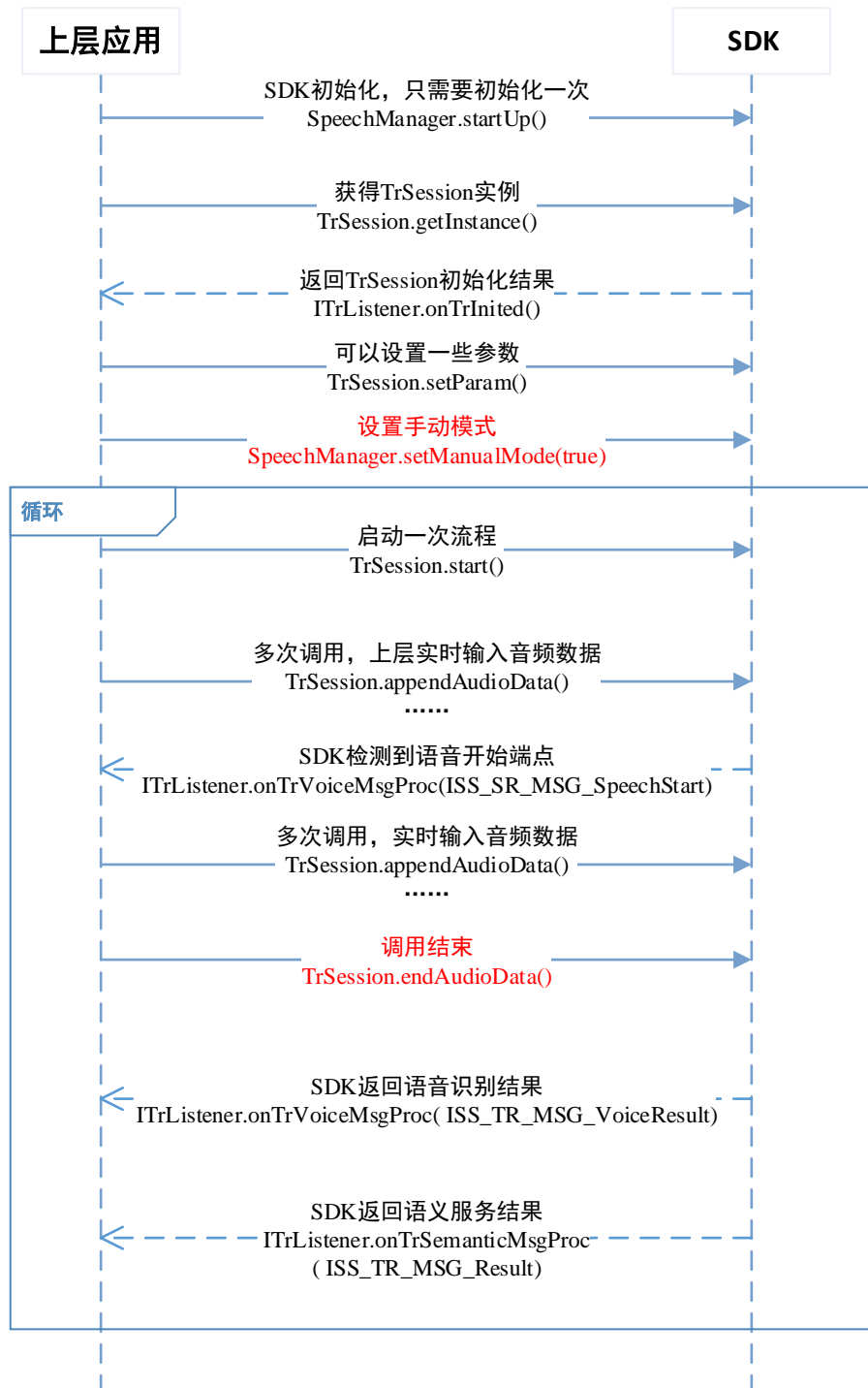
### Android SDK语音+语义识别流程



#### 2.4.4.3. 语音+语义识别的流程（手动模式）

SDK 默认情况下采用自动模式，如果开发者需要自己控制语音的结束，调用 `SpeechManager.getInstance().setManualMode(true)`；将语音识别设置为手动模式。并且结束语音时，调用 `endAudioData()`。

## Android SDK语音+语义识别流程



## 2.5. 语音合成

### 2.5.1. 功能

语音合成的接口，可以将传入的文本，转换对应为音频数据。上层可以使用返回的音频数据进行播放，也可以用使用 SDK 的自动播报（默认自动播报）。

返回的音频数据为单声道,采样率为 16000,位深度为 16bit 的 pcm 数据，对于长文本，可能分多次返回音频数据。

### 2.5.2. 回调

上层应用可以注册 `com.tencent.ai.sdk.tts.ITtsInitListener` 对象接收初始化的异步回调。注册 `com.tencent.ai.sdk.tts.ITtsListener` 对象接收播报状态的异步回调。

注意：初始化成功后才能进行其他操作。

### 2.5.3. 参数设置

`setParam` 可以设置语音合成的参数。

设置是否自动播报：

param: `TtsSession. TYPE_TTS_PLAYING`

值范围：

`TtsSession. TTS_NO_PLAYING`：不自动播报

`TtsSession. TTS_PLAYING`：自动播报，默认

### 2.5.4. 交互流程

## Android SDK语音合成流程

