

# 口语评测（基础版）

## 产品简介



腾讯云

## 【 版权声明 】

©2013–2025 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

## 【 商标声明 】



及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

## 【 服务声明 】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

## 【 联系我们 】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或 95716。

# 文档目录

## 产品简介

产品概述

产品优势

应用场景

评测模式

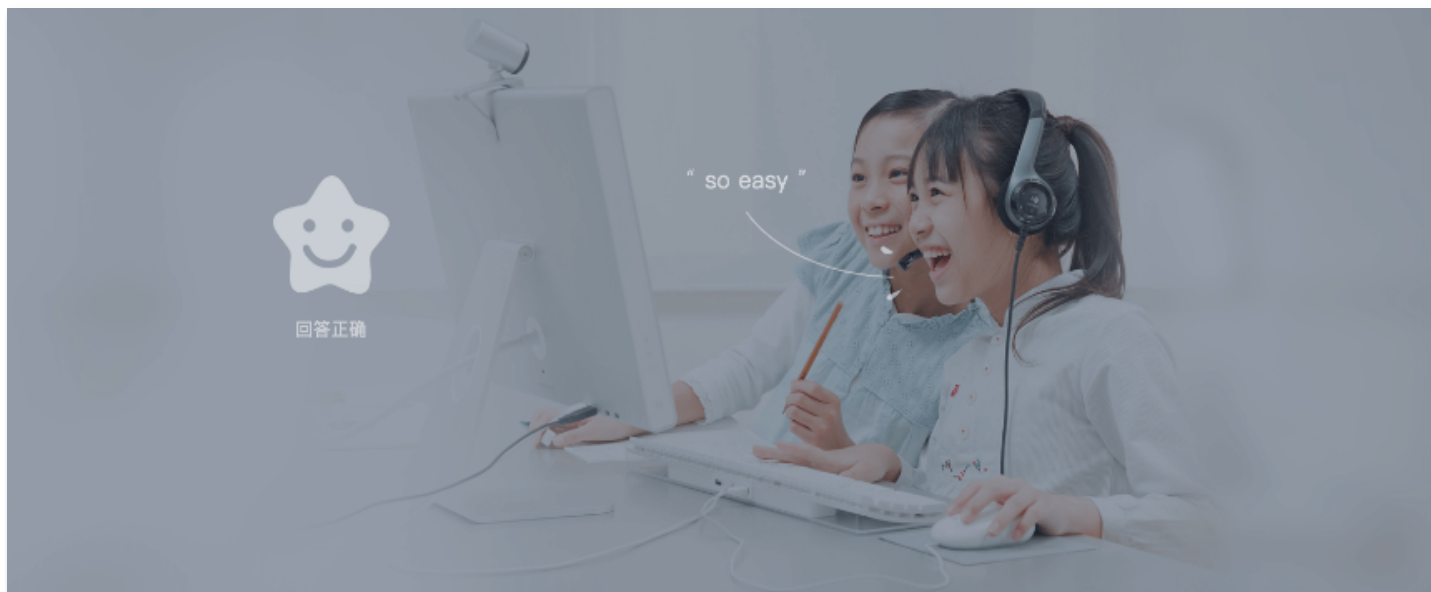
使用限制

# 产品简介

## 产品概述

最近更新时间：2024-06-26 16:52:11

腾讯云智聆口语评测（基础版）（Smart Oral Evaluation-Basic, SOE-B）是腾讯云推出的中英文语音评测产品。支持全年龄段覆盖评测，提供字、词、句、段落、自由说、多分支、关键词等多种评测模式，可返回准确度、流畅度、完整度、重音、声调、音素等细颗粒度评测结果。被广泛应用于中英文口语对话、教学、练习等场景中。



## 功能简介

### 基础功能

| 功能名称       | 描述                         |
|------------|----------------------------|
| 中英文自定义文本评测 | 通过预设文本，获取与预设文本匹配情况。        |
| 中英文无文本评测   | 不通过预设文本，获取用户发音情况。          |
| 拼音评测       | 直接对拼音发音进行评测。               |
| 重音检测       | 检测单词是否为重音，用户是否发了重音。        |
| 标准音素       | 同时展示用户发音和标准发音的匹配情况，指出具体错误。 |
| 原始单词       | 数字类型会转换为对应中英文，文本会返回原始数字文本。 |

|        |   |
|--------|---|
| 实时评测   | 实时获取评测结果，渲染 UI 界面。                        |
| 异步评测   | 传递数据后释放，不需要等待最终结果。                        |
| 查询结果   | 查询10分钟内评测结果。                              |
| 音频存储路径 | 通过设置音频存储路径，可将音频存储至当前账号 COS 桶内，当前不支持跨主账号存入 |

## 高级功能

| 功能名称      | 描述                   |
|-----------|----------------------|
| 音素到字母映射   | 音素同步字母，同时渲染音素和字母。    |
| 音素到国际音标转换 | 可选择返回国际音标或音素。默认返回音素。 |
| 声调评测      | 评测拼音声调是否准确。          |
| 标点符号      | 返回识别的标点符号。           |
| 多组文本分支    | 提供多个正确选项文本。          |
| 候选单词数组    | 定义对多组相似文本的简写形式。      |
| 主题词和关键词   | 设置关键词，查看是否满足关键词。     |
| 负向关键词     | 设置负向关键词，命中直接0分。      |
| 关键词组      | 设置关键词组，定义多种情景。       |
| 指定发音      | 自定义单词发音。             |
| 指定国际音标    | 自定义单词的国际音标。          |

## 客户端 SDK 功能

| 核心功能 | 描述  |
|------|---|
| 内部录制 | 提供语音录制，静音检测，流式分片，分片缺失重试，评测超时重试，本地存储，签名鉴权。 |
| 外部录制 | 提供语音数据传入，支持多种音频格式，签名鉴权。                   |

## 平台支持

智聆口语评测 SDK 支持 Andorid、iOS、Web、微信等主流平台，也支持 Python，Java 等多种服务端语言。具体兼容性见下表。

| 平台/语言   | SDK | Demo体验 | 兼容性   |
|---------|-----|--------|---|
| Android | 支持  | 支持     | Android4.0 以上   |
| iOS     | 支持  | 支持     | iOS8.0 以上   |
| Web     | 支持  | 支持     | 详细参考 <a href="#">平台和兼容性</a>                             |
| 微信H5    | 支持  | 支持     | 微信浏览器   |
| 小程序     | 支持  | 支持     | 微信小程序   |
| Python  | 支持  | 支持     | 安装 Python2.7,3.6~3.9 版本                                 |
| Java    | 支持  | 支持     | 安装 JDK7 版本及以上   |
| PHP     | 支持  | 支持     | PHP5.6.0 版本及以上  |
| Go      | 支持  | 支持     | Go1.9 版本及以上（如使用 go mod 需要 Go 1.14）                      |
| Node.js | 支持  | 支持     | NODEJS 10.0.0 版本及以上                                     |
| .NET    | 支持  | 支持     | .NET standard 2.0（.NET Framework 4.5+ 或者 .NET Core 2.1） |

# 产品优势

最近更新时间：2024-06-26 16:52:11

## 算法模型精细调优

海量成人及儿童数据积累，打分准度高、速度快、误差小。

## 全年龄段覆盖

支持从儿童到成人，全年龄段覆盖的中英文语音评测打分,针对儿童提供个性化、差异化的评测能力支持。

## 细颗粒度评测

提供精准度、流利度、完整度、重音、发音时间、切题程度，关键词等多维度评测结果。提供最小可达音素级别的纠音能力。

## 全平台覆盖，集成简单

全面覆盖主流平台，客户端，前端，后端均可简单快速完成接入。前后端分工协作，应对不同业务场景需求。

## 支持多种环境

提供完备的在线口语评测服务。同时也可提供国际站口语评测服务。

# 应用场景

最近更新时间：2024-06-26 16:52:11

## 中英口语练习

为课内外口语练习提供细颗粒度语音评测，多维度反馈练习情况。

## 口语作业评测

多模式支持多题型评测，满足教师出题需求，多方位考察学生口语学习情况。

## 绘本名著跟读

中外名著、绘本跟读，评分实时反馈学习效果，持续激发阅读兴趣。

## 幼教发音启蒙

针对早教场景下的语言学习，对音素、汉字等的发音情况进行语音评测，打好基础。

## 口语能力测评

多维度测试评估学生口语水平，为课程安排提供依据。

## 人机对话考试

多题型评测可满足考试题型要求，不同的考试场景可以设定不同的情景方案。

## 企业招聘内训

企业招聘中对应聘者口语能力考察；企业培训中服务话术练习。

## 教育硬件应用

为教育平板、故事机、学习机、点读笔等硬件提供口语评测应用。



# 评测模式

最近更新时间：2024-11-28 10:34:41

腾讯云智聆口语评测（基础版）（Smart Oral Evaluation-Basic, SOE-B）目前支持8种英文和8种中文评测模式，每种模式的使用场景和参数限制如下表所示：

| 评估语言 | 评测模式 | 功能支持   | 文本和音频长度限制              | 评测维度   | 使用场景                             |
|------|------|--|------------------------|--|----------------------------------|
| 英文   | 单词   | 重音检测<br>原始单词<br>音素到字母映射<br>音素到国际音标转换<br>指定发音<br>指定国际音标 | 文本=1个单词<br>且 音频≤60秒    | <ol style="list-style-type: none"> <li>1. 单词发音的精准度、流利度、时间。</li> <li>2. 音素的精准度、重音、时间。</li> </ol>                            | 对单词和音标进行评测，适用于线上教学，翻译跟读。         |
|      | 句子   | 原始单词<br>实时评测<br>音素到字母映射<br>音素到国际音标转换<br>指定发音<br>指定国际音标 | 文本≤30个单词<br>且 音频≤60秒   | <ol style="list-style-type: none"> <li>1. 句子发音的精准度、流利度、完整度。</li> <li>2. 单词发音的精准度、流利度、时间。</li> <li>3. 音素的精准度、时间。</li> </ol> | 对词组和句子进行评测，适用于线上教学，翻译跟读。         |
|      | 段落   | 原始单词<br>实时评测<br>多组文本<br>主题词和关键词                        | 文本≤120个单词<br>且 音频≤300秒 | <ol style="list-style-type: none"> <li>1. 段落发音的精准度、流利度、完整度。</li> <li>2. 单词发音的精准度、流利度、时间。</li> </ol>                        | 对文章进行评测，适用于线上教学，翻译跟读。            |
|      | 自由说  | 实时评测<br>标点符号   | 文本字数不限<br>且 音频≤300秒    | <ol style="list-style-type: none"> <li>1. 句子发音的精准度、流利度。</li> <li>2. 单词发音的精准度、流利度。</li> </ol>                               | 适用于不在意说话内容，在意发音准确，吐词清晰的场景。例如：面试。 |
|      | 单词纠错 | 重音检测<br>标准音素   | 文本=1个单词                | <ol style="list-style-type: none"> <li>1. 单词标准音标的音素序列。</li> </ol>  | 对单词进行评测，展示音素差异，适用于线上             |

|       |   |  |  |   |
|-------|---|--|--|---|
|       | <p>音素到字母映射</p> <p>音素到国际音标转换</p> <p>指定发音</p> <p>指定国际音标</p>                         | <p>且 音频 ≤ 60 秒</p>                     | <p>2. 用户真实发音的音素序列。</p> <p>3. 音素精准度、音素序列完整度、时间。</p>   | <p>教学，翻译跟读。</p>                                 |
| 情景    | <p>原始单词</p> <p>多组分支</p> <p>单词数组</p> <p>主题词和关键词</p> <p>负向关键词</p> <p>关键词组</p>       | <p>每个分支 ≤ 30个单词</p> <p>且 音频 ≤ 60 秒</p> | <p>1. 有正向主题词：完整度代表切题度，输出与答案最大匹配句子发音的精准度，流利度，完整度</p> <p>2. 无正向主题词：输出与答案最大匹配句子发音的精准度，流利度，完整度</p> <p>3. 有负向主题词：命中后返回精准度分，总分为0，无单词发音结果</p> <p>4. 单词发音的精准度，流利度，时间</p> | <p>适用于需要多选一，答案正确，错误，无关的场景。例如：问答题，故事复述，看图说话。</p> |
| 句子多分支 | <p>多组文本</p> <p>音素到字母映射</p> <p>音素到国际音标转换</p>                                       | <p>每个分支 ≤ 30个单词</p> <p>且 音频 ≤ 60 秒</p> | <p>1. 输出与参考文本最大匹配句子发音的精准度、流利度、完整度。</p> <p>2. 单词发音的精准度、流利度、时间。</p> <p>3. 音素的精准度、时间。</p>   | <p>适用于需要多选一的场景。例如：选择题。</p>                      |
| 单词实时  | <p>实时评测</p> <p>多组文本</p> <p>音素到字母映射</p> <p>音素到国际音标转换</p> <p>指定发音</p> <p>指定国际音标</p> | <p>每个分支1个单词</p> <p>且 音频 ≤ 60 秒</p>     | <p>1. 单词发音的精准度、流利度、时间。</p> <p>2. 音素的精准度、时间。</p> <p>3. 用户发音相对预设参考文本的完整度。</p>  | <p>适用于无序评测多个单词场景。例如：英语消除游戏。</p>                 |

|    |     |  |                            |   |  |
|----|-----|--|----------------------------|---|--|
| 中文 | 单词  | 指定发音                                     | 文本=1个汉字<br>且 音频≤60秒        | <ol style="list-style-type: none"> <li>1. 单词发音的精准度、流利度、时间。</li> <li>2. 音素的精准度、重音、时间。</li> </ol>   | 对汉字进行评测，适用于线上教学，翻译跟读。                    |
|    | 句子  | 实时评测<br>指定发音<br>声调检测                     | 文本≤30个汉字<br>且 音频≤60秒       | <ol style="list-style-type: none"> <li>1. 句子发音的精准度、流利度、完整度。</li> <li>2. 单词发音的精准度、流利度、时间。</li> <li>3. 音素的精准度、时间。</li> </ol>  | 对词语和句子进行评测，适用于线上教学，翻译跟读。                 |
|    | 段落  | 实时评测<br>多组文本<br>主题词和关键词<br>指定发音          | 文本≤120个汉字<br>且 音频≤300秒     | <ol style="list-style-type: none"> <li>1. 段落发音的精准度、流利度、完整度。</li> <li>2. 单词发音的精准度、流利度、时间。</li> </ol>   | 对文章进行评测，适用于线上教学，翻译跟读。                    |
|    | 自由说 | 实时评测<br>标点符号                             | 文本字数不限<br>且 音频≤300秒        | <ol style="list-style-type: none"> <li>1. 句子发音的精准度、流利度。</li> <li>2. 单词发音的精准度、流利度。</li> </ol>  | 适用于不在意说话内容，在意发音准确，吐词清晰的场景。例如：面试，普通话考试。   |
|    | 情景  | 多组分支<br>单词数组<br>主题词和关键词<br>负向关键词<br>关键词组 | 每个分支<br>≤30个汉字<br>且 音频≤60秒 | <ol style="list-style-type: none"> <li>1. 有正向主题词：完整度代表切题度，输出与答案最大匹配句子发音的精准度，流利度，完整度。</li> <li>2. 无正向主题词：输出与答案最大匹配句子发音的精准度，流利度，完整度。</li> <li>3. 有负向主题词：命中后返回精准度分，总分为0，无单词发音结果。</li> <li>4. 单词发音的精准度，流利度，时</li> </ol> | 适用于需要多选一，答案正确，错误，无关的场景。例如：问答题，故事复述，看图说话。 |

|           |                      |  |   |                          |
|-----------|----------------------|--|---|--------------------------|
|           |                      |  | 间。  |                          |
| 句子<br>多分支 | 多组文本<br>指定发音         | 每个分支<br>≤30个汉字<br>且 音频≤60<br>秒             | <ol style="list-style-type: none"> <li>1. 输出与参考文本最大匹配句子发音的精准度、流利度、完整度。</li> <li>2. 单词发音的精准度、流利度、时间。</li> <li>3. 音素的精准度、时间。</li> </ol> | 适用于需要多选一的场景。例如：选择题。      |
| 单词<br>实时  | 实时评测<br>多组文本<br>指定发音 | 每个分支1个<br>汉字<br>且 音频≤60<br>秒               | <ol style="list-style-type: none"> <li>1. 单词发音的精准度、流利度、时间。</li> <li>2. 音素的精准度、时间。</li> <li>3. 用户发音相对预设参考文本的完整度。</li> </ol>            | 适用于无序评测多个单词场景。例如：汉字消除游戏。 |
| 拼音        | 声调评测<br>指定发音         | 文本≤30个<br>拼音<br>声调检测≤4<br>个拼音 且 音<br>频≤60秒 | <ol style="list-style-type: none"> <li>1. 拼音序列发音的精准度、流利度、完整度。</li> <li>2. 拼音发音的精准度、流利度、时间。</li> <li>3. 音素的精准度、时间。</li> </ol>          | 对拼音进行评测，适用于线上教学，翻译跟读。    |

# 使用限制

最近更新时间：2024-06-26 16:52:11

## 音频文件规范

腾讯云智聆口语评测（基础版）（Smart Oral Evaluation-Basic, SOE-B）在流式或非流式评测下都需要开发者按以下格式上传音频数据。

### ⚠ 注意：

- 需要满足音频属性，如有不一致，可能导致评估不准确或失败。
- 比特率的控制模式推荐使用 CBR，固定码率。

| 音频文件格式 | 音频压缩格式 | 采样率 (sample rate) | 声道 (channels) | 位深 (bit depth) | 比特率 (bit rate) |
|--------|--------|-------------------|---------------|----------------|----------------|
| pcm    | pcm    | 16kHz             | 单声道           | 16bit          | 256kbps以上      |
| wav    | pcm    |                   |               | -              | 32kbps以上       |
| mp3    | MP3    |                   |               | -              | 24kbps以上       |
| speex  | speex  |                   |               | -              |                |

## 流式数据包规范

流式分片序号 (SeqId) 取值范围为 1 - 3000 的整数。

## 评测时长规范

最长支持300s的评测时长。

## 语言规范

英文：美式发音，参考剑桥词典。

中文：普通话。

## 评估文本规范

### RefText 转换标准

| 单词类型 | 组成元素 | 内部转换 | 示例 |
|------|------|------|----|
|------|------|------|----|

|       |              |  |   |
|-------|--------------|--|---|
| 普通单词  | a-z,0-9的组合   | 不转换  | apple   |
| 单词缩写  | 包含一个"'"的单词   | 不转换  | it's  |
| 常见组合词 | 由"-"连接的两个单词  | 对常见组合词不转换; 对 OOV 组合词, 单词模式不转换, 其他模式拆分为两个单词 | bye-bye   |
| 整数    | 数字序列         | 0 - 99 转换为对应单词, 其他转换为数字序列                  | 23→twenty three<br>123→one two three                        |
| 浮点数   | 包含一个"."的数字序列 | 分别将整数和小数部分转换为数字序列, 小数点转换为"point"           | 1.23→one point two three                                    |
| 数字组合词 | 包含"-"的数字序列   | 去除"-"转换为纯数字序列                              | 86-130→eight six one three zero                             |
| 序数词   | 1st - 99th   | 转换为对应的单词                                   | 1st→first   |
| 时间    | 小时+":"+分钟    | 按照英语顺序读法转换                                 | 09:00→nine o'clock<br>09:05→nine o five<br>09:40→nine forty |

## RefText 文本过滤规则

1. 我们支持 GBK 编码集内的所有文本, 并对下述符号进行了过滤 (即 RefText 中可以出现下述符号, 且不会影响评测结果):

~ ' = + " ° ~ ] & \ - ° " ` ! [ ' [= ' " \* [ ' # ] % || [ ¥ : " < -- [ ] # \* ) ( " > + [ ' " . ^ . & ... \$ % : ] ? < > ) ( ( ) ~ \_ \_ \_ \_ \_ \$ ; ° - < / \ . ° ? ! , ; : ? ! , ; . } {

### ⚠ 注意:

个别评测模式下, 上述符号可能会影响评测结果。当出现冲突时, 以上方表格“评测文本要求”一列中的描述为准。

2. 不支持当前评测模式以外的语言（如日、韩语等），不支持填写开发语言（如前端语言）。
3. {} 仅在可支持语法格式中使用，不能单独使用。
4. | 仅在可支持模式中生效。

## 评测文本组成

评测文本组成主要是由 **命令块** + **评测文本** + **发音块** 组成。**命令块**可在部分模式实现附加模块开关。**发音块**可在部分模式实现指定发音。**分割符**可在部分模式实现分割多组文本评测。

除了特定语法格式和转换标准之外的其他字符，标点符号均会被过滤。评测文本均不区分大小写。

| 类型    | 语法格式   | 说明   | 支持模式  | 示例  |
|-------|--|--|---|---|
| 发音描述块 | 单词/汉字<br>{::pron{p1,<br>p2..},<br>{p3,p4..}..}                         | 指定发音，定义单词发音。发音单元为 <b>智聆音素/拼音</b> 。需要将国际音标通过音素映射表转换为智聆音素。 | 英文：单词，单词实时，单词 素诊断，句 模式。<br>中文：单词，单词实时，单词 素诊断，句 ，句子多分子，段落，情景，拼音模式。 | happy{::pron{hh,ae,<br>p,iy}} every<br>山{::pron{shan1}} 水   |
|       | 单词<br>{::ipapron<br>{p1,p2..},<br>{p3,p4..}..}                         | 指定国际音标，定义单词发音。发音单元为IPA（ <b>国际音标</b> ）。                   | 英文：单词，单词实时，单词 素诊断，句 模式。   | orange{::ipapron{'ɔ,r<br>,i,n,dʒ}}  |
|       | {"wordList":<br>[{"word":<br>单词/汉字,"pron":<br>[[p1,p2],<br>[p3,p4]]]}] | 音素结构，定义单词发音。发音单元为 <b>智聆音素/拼音</b> 。                       | 英文：单词、单词实时、单词 素诊断、句 模式。<br>中文：单词，单词实时，单词 素诊断，句 ，句子多分子，段落，情景，拼音模式。 | {"wordList":<br>[{"word": "spout", "pron":<br>[["s", "p", "ao", "t"]]}]<br>{"wordlist": [{"word":<br>"钢", "pron":<br>[["gang1"]]}, {"word":<br>"儿"}]} |
| 命令描述块 | {::cmd{F_IPA=true}}<br>+ 单词  | 开启 素到国际 标转换功能  | 英文：单词，单词实时，单词 素诊断，句 ，句子多分支模式。                                     | {::cmd{F_IPA=true}}<br>orange   |
|       | {::cmd{F_P2L=true}} +<br>单词  | 开启 素到字 映射功能  | 英文：单词，单词实时，单词 素诊断，句 ，句子多分支模式。                                     | {::cmd{F_P2L=true}}<br>hello  |

|        |  |                         |  |  |
|--------|--|-------------------------|--|--|
|        | <code>{::cmd{F_TDET=true}}</code><br>+ 汉字/拼音 | 开启声调检测功能                | 中文：句子，拼音模式。                                  | <code>{::cmd{F_TDET=true}}</code> shui1<br><code>{::cmd{F_TDET=true}}</code> 水 |
|        | <code>{::cmd{F_PUNC=true}}</code>            | 开启加标点功能                 | 英文：自由说。<br>中文：自由说。                           | <code>{::cmd{F_PUNC=true}}</code>  |
| 分隔符    | 单词/汉字   单词/汉字                                | 将文本用   划分为多组文本。多组文本关系为或 | 英文：段落，情景，句子多分支，单词实时。<br>中文：段落，情景，句子多分支，单词实时。 | by bus   by intercity train  |
| 候选单词数组 | <code>@@{::words{a},{b}...}</code>           | 定义对多组相似文本的简写形式          | 英文：情景模式。<br>中文：情景模式。                         | by @@{::words{bus},{intercity train}}  |
| 重音检测   | 有重音的单词                                       | 检测是否为重音                 | 英文：单词，单词纠错。                                  | orange   |
| 原始单词   | 阿拉伯数字  | 将通过转换标准为单词的数字，显示为原始数字   | 英文：单词，句子，段落，情景模式。                            | 7  |

**注意：**

- 发音描述块只能选择其中一种进行使用。
- 当音素到字母映射与音素到国际音标转换或发音描述块一起使用时，音素到字母映射功能可能会失效。
- 默认返回智聆音素。即使指定国际音标，也需要开启 音素到国际 标转换功能，才会返回国际音标。

## 拼音说明

1. 采用数字拼音的标注方式，拼音标注时用1、2、3、4标识一声、二声、三声、四声，目前不支持轻声。例如 sh ān 使用 shan1 表示。
2. 拼音之间用空格区分，例如：sh an1 shan1 算作3个拼音。
3. 支持声母、韵母 + 音调、声母 + 韵母 + 音调。韵母 ü 用 v 代替。
4. ü 需要使用 v 代替。当l, n 与 üe 拼读，需要使用 u，例如：nue1。

## 主题词与关键词文本格式

| 类型 | 语法格式 | 说明 | 支持模式 | 示例 |
|----|------|----|------|----|
|----|------|----|------|----|



|      |                      |           |                            |                          |
|------|----------------------|-----------|----------------------------|--------------------------|
| 普通文本 | 无                    | 主题词       | 英文：段落，情景模式。<br>中文：段落，情景模式。 | bus                      |
| 分隔符  |                      | 于划分不同主题词  | 英文：段落，情景模式。<br>中文：段落，情景模式。 | bus   bike               |
| 分隔符  | #                    | 于划分不同主题词组 | 英文：情景模式。<br>中文：情景模式。       | bus # bike               |
| 命令块  | {::cmd{F_UKEY=true}} | 于指定负向主题词  | 英文：情景模式。<br>中文：情景模式。       | {::cmd{F_UKEY=true}}bike |