中国移动通信集团公司 CMCC

语音评测参数设置、试题格式、 结果说明文档





目 录

1.	参数设置	. 1
	试题格式	
2.1.	汉语评测试题	1
2.2.	英语评测试题	2
2	评测结果	_
3.1.	精简结果	.2
3.2.	详细结果	3
	1. 汉语评测	
	2. 英语评测	





1. 参数设置

在开始评测之前,需要设置评测语种、题型等参数,这些参数如下:

参数 说明 是否必需 评测语种,可选值: en_us(英语)、zh_cn(汉语)。 是 language 评测题型,可选值: read_syllable (单字朗读,汉语专有)、read_word 是 category (词语朗读)、read_sentence (句子朗读)。 上传的试题编码格式,可选值: gb2312、utf-8。当进行汉语评测时, text_encoding 是 必须设置成 utf-8, 建议所有试题都使用 utf-8 编码。 前端点超时,默认 5000ms。 vad_bos 否 后端点超时,默认 1800ms。 否 vad eos 录音超时, 当录音达到时限将自动触发 vad 停止录音, 默认-1(无超 speech_timeout 否 时)。 评测结果等级,可选值: plain(仅英语评测支持)、complete, 默认为 否 result_level complete o

表 1 评测可设置参数

2. 试题格式

评测试题需要符合一定的格式,汉语试题和英语试题有所不同,同语种的不同题型也有差异。

2.1. 汉语评测试题

评测题型 格式 示例 支持两种试卷格式: read syllable <customizer: interphonic> 1) 拼音标注试卷。以 <customizer: 丰 interphonic>开头,字单独一行,紧接着 feng1 的一行为字的拼音(用拼音后加数字代 表声调, 1-4 分别代表一到四声, 5 表 呈. 示轻声); cheng2 2) 纯文本试卷,每个字间用逗号隔开; 3) 最大长度 180 字节,超长则会报 10109 政 错误。 zheng4 或者 丰,呈,政

表 2 汉语评测试题格式





read_word	支持两种试卷格式:	<pre><customizer: interphonic=""></customizer:></pre>
	1) 拼音标注试卷。与字的拼音标注试卷一	宁可
	样,只不过将字换成词语;	ning4 ke3
	2) 纯文本试卷,每个词间用逗号隔开;	
	3) 最大长度 180 字节,超长则会报 10109	非难
	错误。	fei1 nan4
		灾难
		zai1 nan4
		或者
		宁可,非难,灾难
read_sentence	普通中文文本,最大长度为180字节,超长	这是中文语句评测示例。
	则会报 10109 错误。	

2.2. 英语评测试题

表 3 英语评测试题格式

评测类型	格式	示例
read_word	一个单词占一行,首行用[word]标记,最大	[word]
	长度 180 字节,超长则会报 10109 错误。	apple
		banana
		orange
read_sentence	普通英文文本,最大长度为180字节,超长	This is an example of sentence test.
	则会报 10109 错误。	

3. 评测结果

SDK 返回 XML 格式的评测结果,可以分成两种:一种是十分精简的结果,另一种是详细结果。

3.1. 精简结果

在英语评测中,将 result_level 设置成 plain 将返回精简结果(注:汉语评测返回的都是详细结果),格式如下:

<?xml version="1.0" ?>

<FinalResult>





<ret value="0"/>

<total score value="4.496416"/>

</FinalResult>

结果中包含一个返回值 ret(0 表示评测正常)和总分 total_score(分值 0-5 分,满分为 5 分)。

3.2. 详细结果

当 result_level 设置为 complete 或不设置时,返回详细的 XML 评测结果,结果中包含的主要标签如下:

** ** *** *** **** **** **************			
标签	说明		
rec_paper	<rec_paper>和</rec_paper> 中间的内容即为评测结果,是开发者需要关心的部分。		
read_syllable	表明这是单字朗读评测(汉语评测专有)结果。		
read_word	表明这是单词朗读评测结果。		
read_sentence	表明这是句子朗读评测结果。		
sentence	句子。		
word	单词,句子的组成部分。		
syllable	音节,单词发音的组成部分。对于汉语,一个音节对应一个字的发音。		
phone	音素,基本发音元素,音节的组成部分。		

表 4 评测 xml 结果标签

read_syllable、read_word 和 read_sentence 标签中的属性记录了本次评测的各项得分、整体时长等信息。sentence、word、syllable 和 phone 标签记录的是云端引擎对评测语音的分析结果,如哪些是静音,静音时长多少,哪些是有效发音,每个字、词的音节发音是否正确等等。

3.2.1. 汉语评测

以下为汉语评测结果的属性说明(注:一帧长度为10ms,汉语评测得分为5分制):

评测题型	层级	属性	含义
		total_score	总分
	read_syllable	beg_pos	始/末位置(单位: 帧)
mand avillable		/end_pos	
read_syllable		content	试卷内容
		time_len	朗读时长(单位: 帧)
	sentence	time_len	朗读时长(单位: 帧)

表 5 汉语评测结果属性说明





content			beg_pos/end_pos	始末位置(单位: 帧)
word			content	试卷内容
word			beg_pos/end_pos	始末位置(单位: 帧)
vord content itime_len 时长(单校:帧) beg_pos synla dp_message 如素位置(单位:帧) 如素位置(单位:帧) 如素位置(单位:帧) 如果设置的,64(回读),128(替换) symbol 拼音:数字代表声调,5和5以上表示经声 content 试卷内容(sil. silv 表示静音,由表示噪音) 时长(单位:帧) 可需ikcles. 0(正常),16(漏读),32(增 读),64(回读),128(替换) content 试卷内容(sil. silv 表示静音,由表示噪音) 时长(单位:帧) 可需ikcles. 0(正常),16(漏读),32(增 读),64(回读),128(替换) content 试卷内容(sil. silv 表示静音,自表示噪音) 时长(单位:帧) 时长(单位:帧) content 试卷内容(sil. silv 表示静音,自表示噪音) 时长(单位:帧) content 试卷内容 如果在图 如				拼音:数字代表声调,5和5以上表示轻声
beg_pos		word	-	试卷内容
beg_pos			time len	时长(单位: 帧)
Part				
read_word 由力message 增端液信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) symbol 拼音、数字代表声调, 5和 5 以上表示轻声 content 试卷内容(sil、silv 表示静音, fil 表示噪音) time_len 时长 (单位: 帧) beg_pos/end_pos 始末位置 (单位: 帧) dp_message 增漏液信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) content 试卷内容(sil、silv 表示静音, fil 表示噪音) time_len 时长 (单位: 帧) content 试卷内容(sil、silv 表示静音, fil 表示噪音) time_len 时长 (单位: 帧) content 试卷内容 time_len 时长 (单位: 帧) time_len 时长 (单位: 帧) beg_pos/end_pos 始末位置 (单位: 帧) content 试卷内容 time_len 时长 (单位: 帧) word 数字Os*end_pos 始末位置 (单位: 帧) content 试卷内容 time_len 时长 (单位: 帧) beg_pos/end_pos 始末位置 (单位: 帧) dp_message 增漏液信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) symbol 拼音: 数字代表声调, 5和 5以上表示轻声 content 试卷内容(sil、silv 表示静音, 自表示噪音) time_len 时长 (单位: 帧) beg_pos/end_pos 始末位置 (单位: 帧) 如清流信息: 0 (正常), 128 (替换)				始/末位置(单位: 帧)
gyll dp_message 读), 64 (同读), 128 (替换) symbol 拼音:数字代表声调, 5和 5 以上表示轻声 content 试卷内容 (sil, silv 表示静音, fil 表示噪音) 时长 (单位: 帧) phone beg_pos/ end_pos 始末位置 (单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增强), 32 (增强), 64 (同读), 128 (替换) content 试卷内容 (sil, silv 表示静音, fil 表示噪音) time_len 时长 (单位: 帧) read_word total_score 总分 beg_pos/end_pos 始末位置 (单位: 帧) content 试卷内容 time_len 时长 (单位: 帧) sentence beg_pos/end_pos 始末位置 (单位: 帧) content 试卷内容 time_len 时长 (单位: 帧) symbol 拼音:数字代表声调,5和5以上表示经声 content 试卷内容 time_len 时长 (单位: 帧) beg_pos/ end_pos 始/未位置 (单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (普换) symbol 拼音:数字代表声调,5和5以上表示经声 content time_len 时长 (单位: 帧) beg_pos/ end_pos 始末位置 (单位: 帧) 增加速度度度度度度度度度度度度度度度度度度度度度度度度度度度度度度度度度度度度				增漏读信息: 0(正常), 16(漏读), 32(增
Symbol 拼音: 数字代表声调, 5 和 5 以上表示轻声 content 试卷内容 (sil, silv 表示静音, fil 表示噪音) time_len 时长 (单位: 帧) 如 如 如 如 如 如 如 如 如		svll	dp_message	
Content 试卷内容 (sil, silv 表示静音, fil 表示噪音)		3,11	symbol	
read_word time_len 时长(单位: 帧) phone beg_pos/end_pos 始末位置(单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) content 试卷内容 (sit, sitv 表示静音, fil 表示噪音) time_len 时长(单位: 帧) beg_pos/end_pos 始末位置(单位: 帧) content 试卷内容 time_len 时长(单位: 帧) sentence beg_pos/end_pos 始末位置(单位: 帧) content 试卷内容 beg_pos/end_pos 始末位置(单位: 帧) content 试卷内容 beg_pos/end_pos 始末位置(单位: 帧) content 试卷内容 time_len 时长(单位: 帧) symbol 拼音: 数字代表声调,5和5以上表示轻声(声) content 试卷内容 jamix信息: 0(正常), 16(漏读), 32(增读), 64 (回读), 128 (替换) content 试卷内容(sit), sitv 表示静音, fil 表示噪音) time_len 时长(单位: 帧) 增加强读信息: 0(正常), 16 (漏读), 32 (增速), 64 (回读), 128 (替换) content 试卷内容(sit, sitv 表示静音, fil 表示噪音) time_len 时长(单位: 帧) phone 增加或信息: 0(正常), 16 (漏读), 32 (增速) time_len 时长(单位: 帧) <td></td> <td></td> <td>-</td> <td></td>			-	
read_word beg_pos/ end_pos 始末位置(单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) content 试卷内容(sil, silv 表示静音, fil 表示噪音) time_len 时长 (单位: 帧) beg_pos/ end_pos 始;末位置 (单位: 帧) content 试卷内容 time_len 时长 (单位: 帧) sentence beg_pos/ end_pos 始末位置 (单位: 帧) beg_pos/ end_pos 如末位置 (单位: 帧) content 试卷内容 beg_pos/ end_pos 如末位置 (单位: 帧) symbol 拼音: 数字代表声调, 5 和 5 以上表示轻声 (beg_pos/ end_pos) 如末位置 (单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) symbol 拼音: 数字代表声调, 5 和 5 以上表示轻声 (content) 试卷内容 time_len 时长 (单位: 帧) beg_pos/ end_pos 如清被信息: 0 (正常), 16 (漏读), 32 (增滤), 64 (回读), 44 (回读), 42 (增强) phone 如果实验内容(可能) 如素不位置 (单位: 帧) phone 如果实验内容(可能) 如素不位置 (单位: 帧) time_len 时长 (单位: 帧) time_len 时长 (单位: 帧) time_len 时长 (单位: 帧) time_len 如素不位置 (单位: 帧) time				
read_word dp_message 增屬读信息: 0 (正常), 16 (屬读), 32 (增读), 64 (回读), 128 (替换) read_word time_len 时长 (单位: 帧) read_word total_score 总分 beg_pos/ end_pos 始/末位置 (单位: 帧) content 试卷内容 time_len 时长 (单位: 帧) sentence beg_pos/ end_pos 始末位置 (单位: 帧) content 试卷内容 tontent 试卷内容 tontent 试卷内容 time_len 时长 (单位: 帧) symbol 拼音: 数字代表声调, 5 和 5 以上表示轻声 content 试卷内容 time_len 时长 (单位: 帧) beg_pos/ end_pos 始/末位置 (单位: 帧) phone 均message 增屬读信息: 0 (正常), 16 (屬读), 32 (增读), 64 (回读), 128 (替换) symbol 拼音: 数字代表声调, 5 和 5 以上表示轻声 content 试卷内容 (sil、silv 表示静音, fil 表示噪音) time_len 时长 (单位: 帧) phone 均message 始末位置 (单位: 帧) 增屬读信息: 0 (正常), 16 (屬读), 32 (增读), 64 (回读), 128 (替换) content 试卷内容 (sil、silv 表示静音, fil 表示噪音) time_len 时长 (单位: 帧) phone 增屬读信息: 0 (正常), 16 (屬读), 32 (增读), 64 (回读), 128 (替换) content 试卷内			_	
phone dp_message 读), 64 (回读), 128 (替换) content 试卷内容 (sil, silv 表示静音, fil 表示噪音) time_len 时长 (单位: 帧) read_word total_score 总分 beg_pos/ end_pos 始/未位置 (单位: 帧) content 试卷内容 time_len 时长 (单位: 帧) beg_pos/ end_pos 始未位置 (单位: 帧) content 试卷内容 word beg_pos/ end_pos 始未位置 (单位: 帧) content 试卷内容 time_len 时长 (单位: 帧) content 试卷内容 time_len 时长 (单位: 帧) dp_message 始/未位置 (单位: 帧) ighaig信息: 0 (正常), 16 (漏读), 32 (增速), 64 (回读), 128 (替换) content 试卷内容 (sil, silv 表示静音, fil 表示噪音) time_len 时长 (单位: 帧) phone beg_pos/ end_pos 始未位置 (单位: 帧) dp_message 始末位置 (单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增速) imp_len 时长 (单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增速) imp_len 时长 (单位: 帧) time_len 时长 (单位: 帧) time_len 时长 (单位: 帧)				
read_word content 试卷内容(sil、silv 表示静音, fil 表示噪音) read_word time_len 时长(单位: 帧) beg_pos/end_pos 始/末位置(单位: 帧) content 试卷内容 time_len 时长(单位: 帧) sentence beg_pos/end_pos 始末位置(单位: 帧) content 试卷内容 beg_pos/end_pos 始末位置(单位: 帧) symbol 拼音: 数字代表声调,5和5以上表示轻声 content 试卷内容 time_len 时长(单位: 帧) beg_pos/end_pos 始/末位置(单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) symbol 拼音: 数字代表声调,5和5以上表示轻声 content 试卷内容(sil、silv 表示静音, fil 表示噪音) time_len 时长(单位: 帧) dp_message 始末位置(单位: 帧) dp_message 始末位置(单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) content 试卷内容(sil、silv 表示静音, fil 表示噪音) time_len 时长(单位: 帧) time_len 时长(单位: 帧)		phone	dp_message	
read_word time_len 时长(单位: 帧) beg_pos/ end_pos 始/末位置(单位: 帧) content 试卷内容 time_len 时长(单位: 帧) beg_pos/ end_pos 始末位置(单位: 帧) beg_pos/ end_pos 始末位置(单位: 帧) content 试卷内容 time_len 好长(单位: 帧) symbol 拼音: 数字代表声调,5和5以上表示轻声 time_len 时长(单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) symbol 拼音: 数字代表声调,5和5以上表示轻声 content 试卷内容(sil、silv 表示静音, fil 表示噪音) time_len 时长(单位: 帧) beg_pos/end_pos 始末位置(单位: 帧) #漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) content 试卷内容(sil、silv 表示静音, fil 表示噪音) time_len 时长(单位: 帧) #漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) content 试卷内容(sil、silv 表示静音, fil 表示噪音) time_len 时长(单位: 帧)			content	试卷内容(sil、silv表示静音, fil表示噪音)
read_word total_score 总分 beg_pos/end_pos 始末位置(单位: 帧) content 试卷内容 time_len 时长(单位: 帧) beg_pos/end_pos 始末位置(单位: 帧) content 试卷内容 word beg_pos/end_pos 始末位置(单位: 帧) symbol 拼音: 数字代表声调, 5 和 5 以上表示轻声 content 试卷内容 time_len 时长(单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) symbol 拼音: 数字代表声调, 5 和 5 以上表示轻声 content 试卷内容(sil、silv表示静音, fil 表示噪音) time_len 时长(单位: 帧) beg_pos/end_pos 始末位置(单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) content 试卷内容(sil、silv表示静音, fil 表示噪音) time_len 时长(单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) content 试卷内容(sil、silv表示静音, fil 表示噪音) time_len 时长(单位: 帧)			time_len	时长(单位: 帧)
read_word			total_score	
content 试卷内容			beg_pos/ end_pos	始/末位置(单位: 帧)
sentence beg_pos/end_pos beg_pos/end_po		read_word	content	试卷内容
sentence beg_pos/ end_pos 始末位置(单位: 帧) read_word word beg_pos/ end_pos 始末位置(单位: 帧) symbol 拼音: 数字代表声调,5 和 5 以上表示轻声 content 试卷内容 time_len beg_pos/ end_pos 始/末位置(单位: 帧) dp_message 描演读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) symbol 拼音: 数字代表声调,5 和 5 以上表示轻声 content 试卷内容(sil、silv 表示静音, fil 表示噪音) time_len 时长 (单位: 帧) 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) content 试卷内容(sil、silv 表示静音, fil 表示噪音) time_len 时长 (单位: 帧) read_sentence read_sentence			time_len	时长(单位: 帧)
read_word word beg_pos/ end_pos			time_len	时长(单位: 帧)
read_word beg_pos/ end_pos 始末位置(单位: 帧) symbol 拼音: 数字代表声调, 5和5以上表示轻声 content 试卷内容 time_len 时长(单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) symbol 拼音: 数字代表声调, 5和5以上表示轻声 content 试卷内容 (sil、silv表示静音, fil 表示噪音) time_len 时长(单位: 帧) beg_pos/ end_pos 始末位置(单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) content 试卷内容 (sil、silv表示静音, fil 表示噪音) time_len 时长(单位: 帧) read_sentence read_sentence		sentence	beg_pos/ end_pos	始末位置(单位: 帧)
read_word 新mbol 拼音:数字代表声调,5和5以上表示轻声 content tread_word bed_pos/ end_pos 始/末位置(单位:帧) syll beg_pos/ end_pos 始/末位置(单位:帧) dp_message 対策では、5和5以上表示轻声でいます。数字代表声调、5和5以上表示轻声でいます。数字代表声调、5和5以上表示轻声でいます。数字代表声调、5和5以上表示轻声的情况表示操音)时长(单位:帧) phone beg_pos/ end_pos 始末位置(单位:帧) dp_message 均漏读信息:0(正常)、16(漏读)、32(增读)、64(回读)、128(替换) content 试卷内容(sil、silv表示静音, fil 表示噪音) time_len 时长(单位:帧) read_sentence read_sentence			content	试卷内容
read_word read_sentence		word	beg_pos/ end_pos	始末位置(单位: 帧)
read_word content 试卷内容 time_len beg_pos/ end_pos dp_message symbol symbol time_len time_len beg_pos/ end_pos dp_message symbol symbol time_len time_len beg_pos/ end_pos dp_message time_len beg_pos/ end_pos dp_message dp_messag			symbol	拼音:数字代表声调,5和5以上表示轻声
read_ word beg_pos/ end_pos 始/末位置(单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) symbol 拼音: 数字代表声调, 5 和 5 以上表示轻声 content 试卷内容 (sil、silv表示静音, fil 表示噪音) time_len 时长 (单位: 帧) beg_pos/ end_pos 始末位置(单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) content 试卷内容 (sil、silv表示静音, fil 表示噪音) time_len 时长 (单位: 帧) read_sentence cotal_score			content	试卷内容
syll beg_pos/ end_pos 始/末位置(单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) symbol 拼音: 数字代表声调, 5 和 5 以上表示轻声 content time_len 时长 (单位: 帧) beg_pos/ end_pos 始末位置 (单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) content 试卷内容 (sil、silv 表示静音, fil 表示噪音) time_len 时长 (单位: 帧) read_sentence bd			time_len	时长(单位: 帧)
syll dp_message 读), 64 (回读), 128 (替换) symbol 拼音: 数字代表声调, 5和 5 以上表示轻声 (记卷内容 (sil、silv 表示静音, fil 表示噪音) time_len 时长 (单位: 帧) beg_pos/end_pos 始末位置 (单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) content 试卷内容 (sil、silv 表示静音, fil 表示噪音) time_len 时长 (单位: 帧) read_sentence read_sentence	read_ word		beg_pos/ end_pos	始/末位置(单位: 帧)
syll 读), 64 (回读), 128 (替换) symbol 拼音:数字代表声调,5和5以上表示轻声 content 试卷内容(sil、silv 表示静音, fil 表示噪音) time_len 时长(单位:帧) beg_pos/end_pos 始末位置(单位:帧) dp_message 增漏读信息:0(正常),16(漏读),32(增读),64(回读),128(替换) content 试卷内容(sil、silv 表示静音, fil 表示噪音) time_len 时长(单位:帧) read_sentence 总分				增漏读信息: 0(正常), 16(漏读), 32(增
symbol 拼音:数字代表声调,5和5以上表示轻声 content 试卷内容(sil、silv表示静音, fil 表示噪音) time_len 时长(单位:帧) beg_pos/ end_pos 始末位置(单位:帧) dp_message 增漏读信息:0(正常),16(漏读),32(增读),64(回读),128(替换) content 试卷内容(sil、silv表示静音,fil 表示噪音) time_len 时长(单位:帧) read_sentence 总分			dp_message	读),64(回读),128(替换)
time_len 时长(单位: 帧) beg_pos/ end_pos 始末位置(单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) content 试卷内容(sil、silv表示静音, fil 表示噪音) time_len 时长(单位: 帧) read_sentence 总分		syll	symbol	拼音:数字代表声调,5和5以上表示轻声
beg_pos/ end_pos 始末位置(单位: 帧) dp_message 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) content 试卷内容(sil、silv表示静音, fil 表示噪音) time_len 时长(单位: 帧) read_sentence 总分			content	试卷内容(sil、silv表示静音, fil表示噪音)
phone 增漏读信息: 0 (正常), 16 (漏读), 32 (增读), 64 (回读), 128 (替换) content 试卷内容(sil、silv表示静音, fil 表示噪音) time_len 时长 (单位: 帧) read_sentence content			time_len	时长(单位: 帧)
phone dp_message 读), 64 (回读), 128 (替换) content 试卷内容 (sil、silv 表示静音, fil 表示噪音) time_len 时长 (单位: 帧) read_sentence read_sentence			beg_pos/ end_pos	始末位置(单位: 帧)
phone 读), 64 (回读), 128 (替换) content 试卷内容 (sil、silv表示静音, fil 表示噪音) time_len 时长 (单位: 帧) read_sentence read_sentence			1	增漏读信息: 0(正常), 16(漏读), 32(增
time_len 时长(单位:帧) read_sentence read_sentence read_sentence		phone	dp_message	读),64(回读),128(替换)
read_sentence read_sentence total_score 总分			content	试卷内容(sil、silv 表示静音, fil 表示噪音)
read_sentence read_sentence			time_len	时长(单位: 帧)
read_sentence read_sentence beg_pos/ end_pos 始/末位置(单位: 帧)		read_sentence	total_score	总分
	read_sentence		beg_pos/ end_pos	始/末位置(单位: 帧)





		content	试卷内容
		time_len	时长(单位: 帧)
		total_score	总分
	sentence	beg_pos/ end_pos	始末位置(单位: 帧)
	sentence	content	试卷内容
		time_len	时长(单位: 帧)
		beg_pos/ end_pos	始末位置(单位: 帧)
	word	symbol	拼音:数字代表声调,5和5以上表示轻声
	word	content	试卷内容
		time_len	时长(单位: 帧)
	syll	beg_pos/ end_pos	始/末位置(単位: 帧)
		dp_message	增漏读信息: 0(正常), 16(漏读), 32(增
			读),64(回读),128(替换)
		symbol	拼音:数字代表声调,5和5以上表示轻声
		content	试卷内容(sil、silv表示静音, fil表示噪音)
		time_len	时长(单位: 帧)
	phone	beg_pos/ end_pos	始末位置(单位: 帧)
		dp_message	增漏读信息: 0(正常), 16(漏读), 32(增
			读),64(回读),128(替换)
		content	试卷内容(sil、silv表示静音, fil表示噪音)
		time_len	时长(单位: 帧)

dp_message 属性说明如下:

表 6 dp_message 属性说明

取值	含义	说明
0	正常	引擎认为该单元读了,但不一定朗读正确
16	漏读	该单元没有读
32	增读	该单元是多读的文本内的内容
64	回读	该单元是重复读的相邻文本的内容
128	替换	该单元读成文本内其他的内容

3.2.2. 英语评测

以下为英语评测结果的属性说明(注:一帧长度为10ms,英语评测得分为5分制):

表 7 英语评测结果属性说明

评测类型	层级	属性	含义
	mand record	beg_pos	多个单词开始边界时间
read_ word	read_word	content	多个单词内容





		end_pos	多个单词结束边界时间
			异常信息,0表示无异常,常见的异常代码
		except_info	见表 8 常见 except_info 列表
			是否被拒绝: true, false。为 true 表明引擎
			检测到乱读。注: 只有当评测试题长度达
		is_rejected	到 5 个单词时才会进行乱读检测,否则该
			属性的取值总为 false
		total_score	多个单词总分的平均分
		beg_pos	句子开始边界时间
		content	句子内容
	sentence	end_pos	句子结束边界时间
		index	句子索引
		beg_pos	单词开始边界时间
		content	单词内容
		end_pos	单词结束边界时间
			增漏读信息: 0(正常), 16(漏读), 32(增
	word	dp_message	读),64(回读),128(替换)
		global_index	单词在全篇章索引
		index	单词在句子索引
		total_score	单词总分
		beg_pos	音节开始边界时间
		content	音节内容(sil、silv表示静音, fil表示噪音)
	syll	end_pos	音节结束边界时间
		serr_msg	音节错误信息
		syll_accent	音节重读标记
		beg_pos	音素开始边界时间
		content	音素内容(sil、silv表示静音, fil表示噪音)
	phone	end_pos	音素结束边界时间
		1	增漏读信息: 0(正常), 16(漏读), 32(增
		dp_message	读),64(回读),128(替换)
		beg_pos	篇章开始时间
		content	篇章内容
		end_pos	篇章结束时间
		total_score	篇章总分
	read_chapter	except_info	异常信息,0表示无异常
			是否被拒绝: true, false。为 true 表明引擎
read_sentence		is rejected	检测到乱读。注: 只有当评测试题长度达
		is_rejected	到 5 个单词时才会进行乱读检测, 否则该
			属性的取值总为 false
		word_count	篇章全部单词数量
	sentence	beg_pos	句子开始边界时间
		content	句子内容
		end_pos	句子结束边界时间





		index	句子索引
		total_score	全部分数,结构体(隐藏)
		word_count	句子全部单词数量
		beg_pos	单词开始边界时间
		content	单词内容
		end_pos	单词结束边界时间
	ad	de massaga	增漏读信息: 0(正常), 16(漏读), 32(增
	word	dp_message	读),64(回读),128(替换)
		global_index	单词在全篇章索引
		index	单词在句子索引
		total_score	单词总分
	syll	beg_pos	音节开始边界时间
		content	音节内容(sil、silv表示静音, fil表示噪音)
		end_pos	音节结束边界时间
		beg_pos	音素开始边界时间
		content	音素内容(sil、silv表示静音, fil表示噪音)
	phone	end_pos	音素结束边界时间
		dp_message	增漏读信息: 0(正常), 16(漏读), 32(增
			读),64(回读),128(替换)

常见的 except_info 代码有:

表 8 常见 except_info 列表

except_info	说明
0	无异常
28673	该评测语音为无语音或音量太小
28676	该评测语音为乱说类型
28680	该评测语音为信噪比太低
28690	该评测语音出现截幅

XML 结果中音素和音节的音标标记(即 syll 和 phone 标签中的 content 属性)并不是标准音标,以下是它们与标准音标的对应关系:

表 9 音标对照

音标标记	标准音标	音标标记	标准音标	音标标记	标准音标
aa	a:	uw	υ:	f	f
00	э	ch	t∫	g	g
ae	æ	dh	ð	p	p
ah	Λ	hh	h	r	r
ao	o:	jh	dз	S	S
aw	aυ	ng	ŋ	t	t
aw	Э	sh	ſ	v	v
ay	aı	th	θ	w	w





eh	e	zh	3	Z	Z
er	ә:	у	j	ar	eə
ey	еі	d	d	ir	iə
ih	I	k	k	ur	ບອ
iy	i:	1	1	tr	tr
ow	ອບ	m	m	dr	dr
oy	OI IC	n	n	ts	ts
uh	υ	b	b	dz	dz