云知声模式说明文档

1. 模式总体说明

模式	模式说明	适用场景
A	性能最高模式，输出单词打分，不输出音素及音素打分情况	1.适用于单词、句子评测，且不需要返回音标信息的 2.适用于半开放题型，如三问五答，话题简述、看图说话等 3.该模式强制按照文本顺序匹配，如果存在读多遍的情况，按照第一遍进行评分，不适用于篇章
B	可以输出单词打分及音素，但不输出音素打分	1.该模式可以适用于只需要输出单词或句子打分，需要音素音标但不需要音素打分的场景 2.多读扣分
C	json 输出跟 A 模式相同，非强制按文本顺序匹配	适用于篇章评测
E	输出单词打分、音素及音素打分，同时输出空格和标点信息	1.适用于输出文本需要和输入文本一致的场景 2.多读不扣分
G	输出单词打分、音素及音素打分	适用于单词、句子评测且需要返回音标即音标打分详细信息的场景
H	选择题模式	有多个选项，选项之间用\n 或是换行隔开，通过朗读任意文本可以自动匹配上答案

2. A 模式输出

```
{
    "EvalType":"general",          general：普通朗读题型；askandanwser：问答题型；composition：作文题型
    "lines":[
        {
            "begin":0.142,            音频开始时间
            "end":3.052,              音频结束时间
            "fluency":80.612,         流利度
```

"integrity":100,   完整度

"pronunciation":85.156,   标准度

"sample":"do one thing at a time , and do well.",   评测文本

"score":84.929,   总分，百分制

"usertext":"do one thing at a time and do well",   识别的用户文本

"words":[单词详细信息

    {

        "begin":0.142,   开始时间

        "end":0.682,   结束时间

        "score":3.742,   分数，10 分制

        "text":"sil",   文本

        "type":4,   0 多词；1 漏词；2 正常词；3 错误词；4 静音；5 重复词；7 空格 or 标点；8 生词

        "volume":1.476   音量大小，范围 0-10

    },

    {

        "StressOfWord":1,   单词重音标志，0：该单词重音朗读错误；1：该单词重音朗读正确

        "begin":0.682,   音频中单词开始时间

        "end":0.902,   音频中单词结束时间

        "phonetic":"'du:",   单词音标

        "score":9.825,   单词分数，10 分制

        "text":"do",   文本

        "type":2,   0 多词；1 漏词；2 正常词；3 错误词；4 静音；5 重复词；7 空格 or 标点；8 生词

        "volume":9   音量大小，范围 0-10

    },

    {

        "StressOfWord":0,

        "begin":0.902,

        "end":1.162,

        "phonetic":"wʌn",

        "score":8.304,

        "text":"one",

        "type":2,

        "volume":7.001

    },

    {

        "StressOfWord":0,

        "begin":1.162,

        "end":1.482,

        "phonetic":"θɪŋ",

        "score":8.7,

        "text":"thing",

        "type":2,
        "volume":7.895
    },
    {
        "StressOfWord":1,
        "begin":1.482,
        "end":1.662,
        "phonetic":"ˌæt",
        "score":8.841,
        "text":"at",
        "type":2,
        "volume":8.515
    },
    {
        "StressOfWord":0,
        "begin":1.662,
        "end":1.722,
        "phonetic":"ə",
        "score":8.446,
        "text":"a",
        "type":2,
        "volume":2.307
    },
    {
        "StressOfWord":1,
        "begin":1.722,
        "end":2.142,
        "phonetic":"'taɪm",
        "score":6.782,
        "text":"time",
        "type":2,
        "volume":7.85
    },
    {
        "StressOfWord":1,
        "begin":2.142,
        "end":2.382,
        "phonetic":"'ænd",
        "score":8.436,
        "text":"and",
        "type":2,
        "volume":5.796
    },
    {

```
            "StressOfWord":1,
            "begin":2.382,
            "end":2.522,
            "phonetic":"'du:",
            "score":6.258,
            "text":"do",
            "type":2,
            "volume":9
        },
        {
            "StressOfWord":1,
            "begin":2.522,
            "end":3.042,
            "phonetic":"'wel",
            "score":9.303,
            "text":"well",
            "type":2,
            "volume":6.175
        },
        {
            "begin":3.042,
            "end":3.052,
            "score":2.253,
            "text":"sil",
            "type":4,
            "volume":1.149
        }
        ]
    }
    ],
    "score":84.929,                          最外层总分，百分制
    "version":"full 1.0"
}
```

3. B 模式输出

```
{
    "EvalType":"general",                    general：普通朗读题型；askandanwser：问
                                             答题型；composition：作文题型
    "lines":[
        {
            "begin":0.142,                   音频开始时间
```

```json
"end":3.052,                                    音频结束时间
"fluency":80.612,                               流利度
"integrity":100,                                完整度
"pronunciation":85.156,                         标准度
"sample":"do one thing at a time , and do well.",   评测文本
"score":84.929,                                 总分，百分制
"usertext":"do one thing at a time and do well",    识别的用户文本
"words":[                                        单词详细信息
    {
        "begin":0.142,                          开始时间
        "end":0.682,                            结束时间
        "score":3.742,                          分数，10 分制
        "subwords":[                            音素信息
            {
                "begin":0.142,
                "end":0.682,
                "subtext":"sil",                sil 代表静音
                "volume":1.476
            }
        ],
        "text":"sil",
        "type":4,
        "volume":1.476
    },
    {
        "StressOfWord":1,                       单词重音标志，0：该单词重音朗读错误；
                                                1：该单词重音朗读正确
        "begin":0.682,                          音频中单词开始时间
        "end":0.902,                            音频中单词结束时间
        "phonetic":"'du:",                      单词音标
        "score":9.825,                          单词分数，10 分制
        "subwords":[
            {
                "begin":0.682,
                "end":0.792,
                "subtext":"'",                  重音符号
                "volume":9                      音量大小，范围 0-10
            },
            {
                "begin":0.682,                  音素开始时间
                "end":0.792,                    音素结束时间
                "subtext":"d",                  音素文本
                "volume":9                      音量大小，范围 0-10
            },
```

```json
                {
                    "begin":0.792,
                    "end":0.902,
                    "subtext":"u:",
                    "volume":9
                }
            ],
            "text":"do",

            "type":2,

            "volume":9
        },
        {

            "StressOfWord":0,

            "begin":0.902,

            "end":1.162,

            "phonetic":"wʌn",

            "score":8.304,

            "subwords":[
                {

                    "begin":0.902,

                    "end":0.982,

                    "subtext":"w",

                    "volume":9
                },
                {

                    "begin":0.982,

                    "end":1.092,

                    "subtext":"ʌ",

                    "volume":9
                },
                {

                    "begin":1.092,

                    "end":1.162,

                    "subtext":"n",

                    "volume":3.002
                }
            ],
            "text":"one",

            "type":2,

            "volume":7.001
        },
        {

            "StressOfWord":0,
```

单词文本

0 多词；1 漏词；2 正常词；3 错误词；4 静音；5 重复词；7 空格 or 标点；8 生词

```
            "begin":1.162,
            "end":1.482,
            "phonetic":"θɪŋ",
            "score":8.7,
            "subwords":[
                {
                    "begin":1.162,
                    "end":1.242,
                    "subtext":"θ",
                    "volume":6.971
                },
                {
                    "begin":1.242,
                    "end":1.312,
                    "subtext":"ɪ",
                    "volume":9
                },
                {
                    "begin":1.312,
                    "end":1.482,
                    "subtext":"ŋ",
                    "volume":7.713
                }
            ],
            "text":"thing",
            "type":2,
            "volume":7.895
        },
        {
            "StressOfWord":1,
            "begin":1.482,
            "end":1.662,
            "phonetic":"ˌæt",
            "score":8.841,
            "subwords":[
                {
                    "begin":1.482,
                    "end":1.562,
                    "subtext":"ˌ",
                    "volume":8.167
                },
                {
                    "begin":1.482,
                    "end":1.562,
```

最外层总分

```json
                    "subtext":"æ",
                    "volume":8.167
                },
                {
                    "begin":1.562,
                    "end":1.662,
                    "subtext":"t",
                    "volume":8.863
                }
            ],
            "text":"at",
            "type":2,
            "volume":8.515
        },
        {
            "StressOfWord":0,
            "begin":1.662,
            "end":1.722,
            "phonetic":"ə",
            "score":8.446,
            "subwords":[
                {
                    "begin":1.662,
                    "end":1.722,
                    "subtext":"ə",
                    "volume":2.307
                }
            ],
            "text":"a",
            "type":2,
            "volume":2.307
        },
        {
            "StressOfWord":1,
            "begin":1.722,
            "end":2.142,
            "phonetic":"'taɪm",
            "score":6.782,
            "subwords":[
                {
                    "begin":1.722,
                    "end":1.862,
                    "subtext":"'",
                    "volume":9
```

```json
            },
            {
                "begin":1.722,
                "end":1.862,
                "subtext":"t",
                "volume":9
            },
            {
                "begin":1.862,
                "end":2.022,
                "subtext":"aɪ",
                "volume":9
            },
            {
                "begin":2.022,
                "end":2.142,
                "subtext":"m",
                "volume":5.551
            }
        ],
        "text":"time",
        "type":2,
        "volume":7.85
    },
    {
        "StressOfWord":1,
        "begin":2.142,
        "end":2.382,
        "phonetic":"'ænd",
        "score":8.436,
        "subwords":[
            {
                "begin":2.142,
                "end":2.242,
                "subtext":"'",
                "volume":7.929
            },
            {
                "begin":2.142,
                "end":2.242,
                "subtext":"æ",
                "volume":7.929
            },
            {
```

```json
                    "begin":2.242,
                    "end":2.342,
                    "subtext":"n",
                    "volume":7.803
                },
                {
                    "begin":2.342,
                    "end":2.382,
                    "subtext":"d",
                    "volume":1.656
                }
            ],
            "text":"and",
            "type":2,
            "volume":5.796
        },
        {
            "StressOfWord":1,
            "begin":2.382,
            "end":2.522,
            "phonetic":"'du:",
            "score":6.258,
            "subwords":[
                {
                    "begin":2.382,
                    "end":2.442,
                    "subtext":"'",
                    "volume":9
                },
                {
                    "begin":2.382,
                    "end":2.442,
                    "subtext":"d",
                    "volume":9
                },
                {
                    "begin":2.442,
                    "end":2.522,
                    "subtext":"u:",
                    "volume":9
                }
            ],
            "text":"do",
            "type":2,
```

```json
            "volume":9
        },
        {
            "StressOfWord":1,
            "begin":2.522,
            "end":3.042,
            "phonetic":"'wel",
            "score":9.303,
            "subwords":[
                {
                    "begin":2.522,
                    "end":2.662,
                    "subtext":"'",
                    "volume":7.654
                },
                {
                    "begin":2.522,
                    "end":2.662,
                    "subtext":"w",
                    "volume":9
                },
                {
                    "begin":2.662,
                    "end":2.862,
                    "subtext":"e",
                    "volume":7.654
                },
                {
                    "begin":2.862,
                    "end":3.042,
                    "subtext":"l",
                    "volume":1.872
                }
            ],
            "text":"well",
            "type":2,
            "volume":6.175
        },
        {
            "begin":3.042,
            "end":3.052,
            "score":2.253,
            "subwords":[
                {
```

```json
                    "begin":3.042,
                    "end":3.052,
                    "subtext":"sil",
                    "volume":1.149
                }
            ],
            "text":"sil",
            "type":4,
            "volume":1.149
        }
    ]
}
],
"score":84.929,                                    最外层总分，百分制
"version":"full 1.0"
}
```

4. C 模式输出

跟 A 模式相同，参考 A 模式

5. E 模式输出

```json
{
    "version":"full 1.0",
    "score":84.929,                                最外层总分，百分制
                                                   general：普通朗读题型；askandanwser：问答题型；
    "EvalType":"general",                          composition：作文题型
    "lines":[
        {
            "sample":"do one thing at a time , and do well.",    评测文本
            "usertext":"do one thing at a time and do well",     识别的用户文本
            "begin":0.142,                         开始时间
            "end":3.052,                           结束时间
            "score":84.929,                        总分，百分制
            "fluency":80.612,                      流利度
            "integrity":100,                       完整度
            "pronunciation":85.156,                标准度
            "words":[                              单词详细信息
                {
```

```json
        "StressOfWord":1,                              单词重音标志，0：该单词重音朗读错误；1：该单词重音朗
                                                       读正确
        "phonetic":"'du:",                             单词音标
        "text":"do",                                   单词文本
                                                       0 多词；1 漏词；2 正常词；3 错误词；4 静音；5 重复
        "type":2,                                      词；7 空格 or 标点；8 生词
        "begin":0.682,                                 单词开始时间
        "end":0.902,                                   单词结束时间
        "volume":9,                                    音量大小，范围 0-10
        "score":9.825,                                 单词打分，10 分制
        "subwords":[                                   音素详细信息
            {
                "subtext":"'",                         重音符号
                "volume":9,                            音量大小，范围 0-10
                "begin":0.682,                         开始时间
                "end":0.792,                           结束时间
                "score":10                             0：重音错误；10：重音正确
            },
            {
                "subtext":"d",                         音素文本
                "volume":9,
                "begin":0.682,
                "end":0.792,
                "score":9.532                          音素分数，10 分制
            },
            {
                "subtext":"u:",
                "volume":9,
                "begin":0.792,
                "end":0.902,
                "score":10
            }
        ]
    },
    {
        "text":" ",                                    空格
        "type":7,
        "begin":0,
        "end":0,
        "volume":0,
        "score":0,
        "subwords":[
```

6. G 模式输出

```
{
    "EvalType":"general",
    "lines":[
        {
            "begin":0.142,
            "end":3.052,
            "fluency":80.612,
            "integrity":100,
            "pronunciation":85.156,
            "sample":"do one thing at a time , and do well.",
            "score":84.929,
            "usertext":"do one thing at a time and do well",
            "words":[
                {
                    "begin":0.142,
                    "end":0.682,
                    "score":3.742,
                    "subwords":[
                        {
                            "begin":0.142,
                            "end":0.682,
                            "subtext":"sil",
                            "volume":1.476
                        }
                    ],
                    "text":"sil",

                    "type":4,
                    "volume":1.476
                },
                {

                    "StressOfWord":1,
                    "begin":0.682,
                    "end":0.902,
                    "phonetic":"'du:",
                    "score":9.825,
                    "subwords":[
                        {
                            "begin":0.682,
```

JSON	说明
"EvalType":"general",	general：普通朗读题型；askandanwser：问答题型；composition：作文题型
"lines":[	
"begin":0.142,	开始时间
"end":3.052,	结束时间
"fluency":80.612,	流利度
"integrity":100,	完整度
"pronunciation":85.156,	标准度
"sample":"do one thing at a time , and do well.",	评测文本
"score":84.929,	总分，百分制
"usertext":"do one thing at a time and do well",	识别的用户文本
"words":[	单词详细信息
"subwords":[	单词音标
{	单词文本
"subtext":"sil",	静音
"volume":1.476	音量大小，范围 0-10
],	音素详细信息
"type":4,	0 多词；1 漏词；2 正常词；3 错误词；4 静音；5 重复词；7 空格 or 标点；8 生词
"volume":1.476	音量大小，范围 0-10
"StressOfWord":1,	单词重音标志，0：该单词重音朗读错误；1：该单词重音朗读正确
"begin":0.682,	单词开始时间
"end":0.902,	单词结束时间
"phonetic":"'du:",	单词音标
"score":9.825,	单词打分，10 分制
"subwords":[	音素详细信息

```
            "end":0.792,
            "score":10,
            "subtext":"'",
            "volume":9
        },
        {
            "begin":0.682,
            "end":0.792,
            "score":9.532,
            "subtext":"d",
            "volume":9
        },
        {
            "begin":0.792,
            "end":0.902,
            "score":10,
            "subtext":"uː",
            "volume":9
        }
    ],
    "text":"do",
    "type":2,
    "volume":9
},
{
    "StressOfWord":0,
    "begin":0.902,
    "end":1.162,
    "phonetic":"wʌn",
    "score":8.304,
    "subwords":[
        {
            "begin":0.902,
            "end":0.982,
            "score":9.615,
            "subtext":"w",
            "volume":9
        },
        {
            "begin":0.982,
            "end":1.092,
            "score":8.877,
            "subtext":"ʌ",
            "volume":9
        }
```

0：重音错误；10：重音正确
重音符号

0：重音错误；10：重音正确
重音符号

            },
            {
                "begin":1.092,
                "end":1.162,
                "score":6.921,
                "subtext":"n",
                "volume":3.002
            }
        ],
        "text":"one",
        "type":2,
        "volume":7.001
    },
    {

        "StressOfWord":0,
        "begin":1.162,
        "end":1.482,
        "phonetic":"θɪŋ",
        "score":8.7,
        "subwords":[
            {
                "begin":1.162,
                "end":1.242,
                "score":8.713,
                "subtext":"θ",
                "volume":6.971
            },
            {
                "begin":1.242,
                "end":1.312,
                "score":9.824,
                "subtext":"ɪ",
                "volume":9
            },
            {
                "begin":1.312,
                "end":1.482,
                "score":8.33,
                "subtext":"ŋ",
                "volume":7.713
            }
        ],
        "text":"thing",
        "type":2,

```json
                    "volume":7.895
                },
                {
                    "StressOfWord":1,
                    "begin":1.482,
                    "end":1.662,
                    "phonetic":"ˌæt",
                    "score":8.841,
                    "subwords":[
                        {
                            "begin":1.482,
                            "end":1.562,
                            "score":10,
                            "subtext":"ˌ",
                            "volume":8.167
                        },
                        {
                            "begin":1.482,
                            "end":1.562,
                            "score":8.711,
                            "subtext":"æ",
                            "volume":8.167
                        },
                        {
                            "begin":1.562,
                            "end":1.662,
                            "score":9.016,
                            "subtext":"t",
                            "volume":8.863
                        }
                    ],
                    "text":"at",
                    "type":2,
                    "volume":8.515
                },
                {
                    "StressOfWord":0,
                    "begin":1.662,
                    "end":1.722,
                    "phonetic":"ə",
                    "score":8.446,
                    "subwords":[
                        {
                            "begin":1.662,
```

```json
                    "end":1.722,
                    "score":8.446,
                    "subtext":"ə",
                    "volume":2.307
                }
            ],
            "text":"a",
            "type":2,
            "volume":2.307
        },
        {
            "StressOfWord":1,
            "begin":1.722,
            "end":2.142,
            "phonetic":"'taɪm",
            "score":6.782,
            "subwords":[
                {
                    "begin":1.722,
                    "end":1.862,
                    "score":10,
                    "subtext":"'",
                    "volume":9
                },
                {
                    "begin":1.722,
                    "end":1.862,
                    "score":5.379,
                    "subtext":"t",
                    "volume":9
                },
                {
                    "begin":1.862,
                    "end":2.022,
                    "score":7.517,
                    "subtext":"aɪ",
                    "volume":9
                },
                {
                    "begin":2.022,
                    "end":2.142,
                    "score":8.661,
                    "subtext":"m",
                    "volume":5.551
```

```json
                }
            ],
            "text":"time",
            "type":2,
            "volume":7.85
        },
        {
            "StressOfWord":1,
            "begin":2.142,
            "end":2.382,
            "phonetic":"'ænd",
            "score":8.436,
            "subwords":[
                {
                    "begin":2.142,
                    "end":2.242,
                    "score":10,
                    "subtext":"'",
                    "volume":7.929
                },
                {
                    "begin":2.142,
                    "end":2.242,
                    "score":8.443,
                    "subtext":"æ",
                    "volume":7.929
                },
                {
                    "begin":2.242,
                    "end":2.342,
                    "score":8.803,
                    "subtext":"n",
                    "volume":7.803
                },
                {
                    "begin":2.342,
                    "end":2.382,
                    "score":8.183,
                    "subtext":"d",
                    "volume":1.656
                }
            ],
            "text":"and",
            "type":2,
```

```json
            "volume":5.796
        },
        {
            "StressOfWord":1,
            "begin":2.382,
            "end":2.522,
            "phonetic":"'du:",
            "score":6.258,
            "subwords":[
                {
                    "begin":2.382,
                    "end":2.442,
                    "score":10,
                    "subtext":"'",
                    "volume":9
                },
                {
                    "begin":2.382,
                    "end":2.442,
                    "score":4.929,
                    "subtext":"d",
                    "volume":9
                },
                {
                    "begin":2.442,
                    "end":2.522,
                    "score":8.107,
                    "subtext":"u:",
                    "volume":9
                }
            ],
            "text":"do",
            "type":2,
            "volume":9
        },
        {
            "StressOfWord":1,
            "begin":2.522,
            "end":3.042,
            "phonetic":"'wel",
            "score":9.303,
            "subwords":[
                {
                    "begin":2.522,
```

```json
                "end":2.662,
                "score":10,
                "subtext":"'",
                "volume":7.654
            },
            {
                "begin":2.522,
                "end":2.662,
                "score":9.428,
                "subtext":"w",
                "volume":9
            },
            {
                "begin":2.662,
                "end":2.862,
                "score":10,
                "subtext":"e",
                "volume":7.654
            },
            {
                "begin":2.862,
                "end":3.042,
                "score":8.977,
                "subtext":"l",
                "volume":1.872
            }
        ],
        "text":"well",
        "type":2,
        "volume":6.175
    },
    {
        "begin":3.042,
        "end":3.052,
        "score":2.253,
        "subwords":[
            {
                "begin":3.042,
                "end":3.052,
                "subtext":"sil",
                "volume":1.149
            }
        ],
        "text":"sil",
```

```
                    "type":4,
                    "volume":1.149
            }
        ]
    }
],
"score":84.929,                                          最外层分数，百分制
"version":"full 1.0"
}
```

7. H 模式输出

这个模式的特点是输入文本有差异

在设置 H 模式时，输入 4 个不同的文本，文本之间是 \n 或是换行分割，如：
Good\nperfect\nnice\nsmart

当朗读 perfect 时，可以自动返回 perfect 的打分结果，如果业务层对应 4 个选项，则自动匹配上 B 答案，打分结果的层次结构跟 G 模式相同

```
{
                                                        general：普通朗读题型；askandanwser：
    "EvalType":"general",                               问答题型；composition：作文题型

    "lines":[
        {
            "begin":0,                                  开始时间
            "end":1.821,                                结束时间
            "fluency":62.083,                           流利度
            "integrity":100,                            完整度
            "pronunciation":77.888,                     标准度
            "sample":"perfect",                         评测文本
            "score":77.097,                             总分，百分制
            "usertext":"perfect",                       识别的用户文本
            "words":[
                {
                    "begin":0.081,
                    "end":0.801,
                    "score":3.612,
                    "subwords":[
                        {
                            "begin":0.081,
                            "end":0.801,
                            "subtext":"sil",
                            "volume":1.464                音量大小，范围 0-10
                        }
                    ],
                    "text":"sil",
```

```
        "type":4,

        "volume":1.464
},
{

        "StressOfWord":1,

        "begin":0.801,

        "end":1.661,

        "phonetic":"'pɜːfɪkt",

        "score":7.707,

        "subwords":[
            {

                "begin":0.801,

                "end":0.941,

                "score":10,

                "subtext":"'",

                "volume":6.922
            },
            {

                "begin":0.801,

                "end":0.941,

                "score":8.21,

                "subtext":"p",

                "volume":9
            },
            {

                "begin":0.941,

                "end":1.101,

                "score":3.971,

                "subtext":"ɜː",

                "volume":6.922
            },
            {

                "begin":1.101,

                "end":1.201,

                "score":9.103,

                "subtext":"f",

                "volume":6.809
            },
            {

                "begin":1.201,

                "end":1.321,

                "score":9.308,
```

0 多词；1 漏词；2 正常词；3 错误词；4 静音；5 重复词；7 空格 or 标点；8 生词

单词重音标志，0：该单词重音朗读错误；1：该单词重音朗读正确

单词音标
得分，10 分制

0 多词；1 漏词；2 正常词；3 错误词；4 静音；5 重复词；7 空格 or 标点；8 生词

                    "subtext":"ɪ",
                    "volume":6.333
                },
                {
                    "begin":1.321,
                    "end":1.441,
                    "score":9.636,
                    "subtext":"k",
                    "volume":6.923
                },
                {
                    "begin":1.441,
                    "end":1.661,
                    "score":8.691,
                    "subtext":"t",
                    "volume":7.05
                }
            ],
            "text":"perfect",
            "type":2,
            "volume":7.173
        },
        {
            "begin":1.661,
            "end":1.821,
            "score":4.498,
            "subwords":[
                {
                    "begin":1.661,
                    "end":1.821,
                    "subtext":"sil",
                    "volume":1.014
                }
            ],
            "text":"sil",
            "type":4,
            "volume":1.014
        }
        ]
    }
    ],
    "score":77.097,
    "version":"full 1.0"