

语 言 文 字 规 范

GF 0030—2026

机器合成普通话水平测评 等级标准及测评大纲

Standards and guidelines of synthesized Putonghua proficiency assessment

2026-03-02 发布

2026-07-01 实施

中华人民共和国教育部
国家语言文字工作委员会

发布

目 次

前言.....	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 总则	2
5 测评名称、性质和方式	2
6 机器合成普通话水平测评等级标准	2
7 机器合成普通话水平测评大纲	3

仅供查阅

前 言

本文件按照 GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件由教育部语言文字应用研究所提出。

本文件由教育部语言文字信息管理司归口。

本文件由国家语言文字工作委员会语言文字规范（标准）审定委员会审定。

本文件由教育部、国家语言文字工作委员会发布。

本文件起草单位：教育部语言文字应用研究所（国家语委普通话与文字应用培训测试中心）、科大讯飞股份有限公司、北京海天瑞声科技股份有限公司、安徽省普通话培训测试中心、北京语言大学、中国科学技术大学、中国传媒大学、高等教育出版社有限公司、清华大学、中教仪（北京）认证服务有限公司、中国社会科学院语言研究所、广东外语外贸大学、国家广播电视总局广播电视规划院、中国信息通信研究院、南京航空航天大学、湖北第二师范学院、江苏师范大学、北京市语言文字测试中心、北京教育督导评估院、中国教育科学研究院、上海市教师教育学院、北京市西城区教育研修学院、天津市教育科学研究院、新疆政法学院、深圳职业技术大学、上海师范大学、北京财贸职业学院、北京希尔贝壳科技有限公司、思必驰科技股份有限公司、中央民族大学、北京师范大学、语文出版社有限公司、北京外国语大学、人民教育出版社有限公司。

本文件主要起草人：王奇、李慧、刘朋建、刘培俊、于谦、陈茜、朱丽红、郝玉峰、单明、汪张龙、彭恒利、竺博、程玮、张凯波、刘慧、祖漪清、凌震华、周敏燕、陈欣、肖炜、刘晓翠、杨丽萍、李莉、喻梅、杨京峰、曹文、张劲松、杨尔弘、王东、霍扬、贾媛、王治敏、李雷雷、李荪、韩玉华、乔丽华、高雅跃、安爽、齐军华、周梅、胡明晓、杜首燕、潘泽亮、陈天同、方昕、吴奎、陈霖、于海琪、卜辉、徐昕、张辉、丁石庆、王立军、王铁琨、王翠叶、文秋芳、李宇明、陈先云、程荣。

机器合成普通话水平测评等级标准及测评大纲

1 范围

本文件规定了机器合成普通话水平的等级标准，提供了用于等级评定的测评大纲。
本文件适用于机器合成语音的普通话水平测评，可用于技术研发、产品优化及应用评价。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

《普通话水平测试等级标准》（国家语言文字工作委员会颁布）
《普通话水平测试大纲》（教育部、国家语言文字工作委员会发布）
《通用规范汉字表》（国务院公布）

3 术语和定义

下列术语和定义适用于本文件。

3.1

机器合成普通话 synthesized Putonghua
语音合成系统将输入文本转换合成为普通话语音的结果。

3.2

测评文本 test text
按照本文件的具体要求编制的以文字形式呈现的测评内容。

3.3

语音错误 pronunciation error
指普通话语音（音位）系统中，把一个音（音位）误读作另一个音（音位）。

3.4

语音缺陷 pronunciation flaw
指发音没有完全达到标准程度，介于“正确”与“错误”之间，一般不会对交际造成困难或误解。

4 总则

4.1 科学规范性原则

本文件依据国家通用语言文字规范标准制定，确保测评方案科学规范，测评结果客观可信。

4.2 技术适配性原则

本文件与语音合成技术的特性及发展趋势相适配，与机器合成普通话的应用场景和评价需求相适应。

4.3 体系兼容性原则

本文件与普通话水平测试等级标准及测试大纲有效衔接，构建人机统一协调的普通话水平评价体系。

5 测评名称、性质和方式

5.1 测评名称

测评命名为“机器合成普通话水平测评”（汉语拼音为JIQI HECHENG PUTONGHUA SHUIPING CEPING，缩写为JHPSC）。

5.2 测评性质

测评机器合成普通话的语音标准程度、自然程度及整体感受，认定其普通话水平等级，属于标准参照性测评。

5.3 测评方式

测评工作通过机器合成普通话水平测评平台实施。参测单位将应用程序编程接口（API）接入平台，根据测评文本实时合成普通话。由国家级普通话水平测试员进行评分。测评文本共分五部分，总分为100分，测评结果划分为六个等级。

6 机器合成普通话水平测评等级标准

6.1 一级

语音标准，分词处理正确，多音字、数字、符号、所用外文识读及发音准确；能正确识读《通用规范汉字表》中所有的字。语调自然，停连得当，重音准确，符合语境表达习惯。整体表达流畅，听感舒适，与人的发音几无差别。测评成绩97分及其以上。

6.2 二级

语音标准，偶有声母、韵母、声调和变调、轻声、儿化等发音偏误；分词处理及多音字、数字、符号、所用外文识读及发音偶有偏误；能正确识读《通用规范汉字表》中所有的字。语调自然，停连、重音等偶有偏误。整体表达顺畅，听感较舒适，与人的发音略有差别。测评成绩92分及其以上但不足97分。

6.3 三级

语音基本标准，声母、韵母、声调和变调、轻声、儿化等发音偏误较少；有少量分词处理及多音字、数字、符号、所用外文识读及发音偏误；能正确识读《通用规范汉字表》中所有的字。语调自然，停连、重音等有少量偏误。整体表达较顺畅，听感尚好，合成语音特征较明显。测评成绩 87 分及其以上但不足 92 分。

6.4 四级

语音基本标准，声母、韵母、声调和变调、轻声、儿化等发音偏误较多；有较多分词处理及多音字、数字、符号、所用外文识读及发音偏误；未正确识读少量《通用规范汉字表》中的字。语调不够自然，停连、重音等有较多偏误。整体表达不够顺畅，听感一般，合成语音特征明显，但基本不影响内容理解。测评成绩 80 分及其以上但不足 87 分。

6.5 五级

声母、韵母、声调和变调、轻声、儿化等发音偏误多；多有分词处理及多音字、数字、符号、所用外文识读及发音偏误；未正确识读部分《通用规范汉字表》中的字。语调不自然，停连、重音等偏误多。整体表达不顺畅，听感较差，合成语音特征很明显，影响内容理解。测评成绩 70 分及其以上但不足 80 分。

6.6 六级

有大量声母、韵母、声调和变调、轻声、儿化等发音偏误；有大量分词处理及多音字、数字、符号、所用外文识读及发音偏误；未正确识读部分《通用规范汉字表》中的字。语调机械、不自然，屡有停连不当、重音失准。整体表达和听感差，合成语音特征突出，导致内容理解有困难。测评成绩 60 分及其以上但不足 70 分。

7 机器合成普通话水平测评大纲

7.1 测评文本编制

测评文本由教育部语言文字应用研究所（国家语委普通话与文字应用培训测试中心）根据本文件编制。

7.2 测评文本构成及评分办法

第一部分：单音节字词，共 15 分。

1. 目的

测评机器合成普通话声母、韵母、声调的语音标准程度及汉字识读能力。

2. 要求

(1) 测评单音节字词发音，共 100 个音节，覆盖全部的声母、韵母、声调。其中，每个声母出现次数一般不少于 3 次，不超过 6 次；每个韵母出现次数一般不少于 2 次，不超过 4 次。四个声调出现次数大致均衡。

(2) 测评《通用规范汉字表》识读，共计 8105 个字。

3. 评分办法

(1) 单音节字词发音，满分 10 分。语音错误，每个音节扣 0.1 分；语音缺陷，每个音节扣 0.05 分。

(2) 识读《通用规范汉字表》，满分 5 分。根据未识读音节数量扣分。

一档：未识读 1—200 个音节，扣 1 分。

二档：未识读 201—400 个音节，扣 2 分。

三档：未识读 401 个及其以上音节，扣 5 分。

第二部分：双音节词语，共 10 分。

1. 目的

测评机器合成普通话声母、韵母、声调和变调（包含上声变调和“一”“不”变调）、轻声、儿化的语音标准程度。

2. 要求

测评双音节词语发音，共 50 个词语、100 个音节，声母、韵母、声调出现的次数与单音节字词要求相同。上声与上声相连的词语不少于 3 个，上声与非上声相连的词语不少于 4 个，轻声不少于 3 个，儿化不少于 4 个（应为不同的儿化韵母）。

3. 评分办法

语音错误，每个音节扣 0.1 分；语音缺陷，每个音节扣 0.05 分。

第三部分：句子，共 40 分。

1. 目的

重点测评机器合成普通话分词处理及多音字、变调、轻声、儿化、数字、符号、所用外文识读及发音情况，同时测评语音标准程度。

2. 要求

测评 200 个句子发音。句子长短大致均衡，一般不超过 30 个音节。

3. 评分办法

分词处理及多音字、变调、轻声、儿化、数字、符号、所用外文识读及发音错误，每处扣 0.2 分；如有上述以外的语音错误，每处扣 0.1 分。每个句子扣完 0.2 分为止。

第四部分：句段，共 30 分。

1. 目的

重点测评机器合成普通话停连、重音、语调的自然程度，同时测评语音标准程度。

2. 要求

测评 6 个句段发音。每个句段为不同体裁，200 个音节左右。

3. 评分办法

每个句段 5 分，其中停连 2 分、重音 2 分、语调 1 分；如有语音错误，每处扣 0.1 分。每个句段扣完 5 分为止。

(1) 停连（每个句段 2 分）：

一档：1—2 处停连不当，扣 0.5 分。

二档：3—5 处停连不当，扣 1 分。

三档：超过 5 处停连不当，扣 2 分。

(2) 重音（每个句段 2 分）：

一档：1—2 处重音不当，扣 0.5 分。

二档：3—5 处重音不当，扣 1 分。

三档：超过 5 处重音不当，扣 2 分。

(3) 语调（每个句段 1 分），视偏误程度扣 0.5 分、1 分。

第五部分：篇章，共 5 分。

1. 目的

综合测评机器合成普通话的表达、听感等整体感受。

2. 要求

测评 1 个篇章发音。节选自现代汉语经典篇目，400 个音节左右。

3. 评分办法

一档：表达流畅，听感舒适，与人的发音几无差别，不扣分。

二档：表达顺畅，听感较舒适，与人的发音略有差别，偶能发现合成语音特征，扣 1 分。

三档：表达较顺畅，听感尚好，合成语音特征较明显，扣 2 分。

四档：表达不够顺畅，听感一般，合成语音特征明显，但基本不影响内容理解，扣 3 分。

五档：表达不顺畅，听感较差，合成语音特征很明显，影响内容理解，扣 4 分。

六档：表达和听感差，合成语音特征突出，导致内容理解有困难，扣 5 分。

仅供查阅