

# 中文多人合成数据库

[AISHELL-TTS0017]



希尔贝壳  
A I S H E L L

Copyright

## 目录

1 产品概述.....	2
2 录音语料.....	2
3 发音人信息.....	2
3.1 基本信息记录.....	2
3.2 发音人结构特征.....	3
3.2.1 性别比例 .....	3
3.2.2 年龄比例 .....	3
4 数据录制环境.....	3
4.1 录制环境.....	3
4.2 录制设备.....	3
4.3 录制方法.....	3
5 语音数据转写.....	4
6 数据文件目录.....	4
6.1 目录结构.....	4
6.2 命名规则.....	5
6.2.1 目录命名规则（图表 6-2-1） .....	5
6.2.2 文件命名规则（图表 6-2-2） .....	5
7 版权声明.....	5



# 1 产品概述

希尔贝壳中文多人语音合成数据库 AISHELL-TTS0017 共 12.76 小时，12519 条语音。录音语言，中文；录音地区，中国。邀请 41 名专业发音人在专业录音棚进行录制。录音设备为高保真，音频采样率为 48kHz，比特率 16bit。

此数据库经过专业语音校对人员转写标注，标注包含字、拼音以及韵律，并通过严格质量检验，文本正确率达到 99%。

## 2 录音语料

此数据库录音文本采用每份平均 300 句，相关处理如下：

- 脱敏处理。删除政治敏感、个人隐私、色情暴力等内容。
- 音素覆盖全面。
- 删除 <, >, [, ], ~, /, \, = 等符号。
- 删除含有中文和英文以外语言的内容。
- 删除单句含有 30 字以上的内容。
- 统一格式。

## 3 发音人信息

### 3.1 基本信息记录

发音人信息记录内容包括任务编号、性别、年龄区间。（图表 3-1）

任务编号	性别	年龄区间
W0001	男	A

图表 3-1

任务编号：每个发言人领取 1 个任务编号，每个任务编号对应 1 份录音文本。每个发言人只能参加一次录制。

年龄区间：A(18-25 岁)、B(26-36 岁)。

## 3.2 发音人结构特征

### 3.2.1 性别比例

数据库总人数为 41 人，男 14 人，女 27 人。（图表 3-2-1）

性别	男性	女性	合计
比例	34%	66%	100%

图表 3-2-1

### 3.2.2 年龄比例

A(18-25 岁)；B(26-40 岁)。（图表 3-2-2）

	年龄段	人数	比例
A	18-25 岁	17	41.5%
B	26-36 岁	24	58.5%
	合计	100	100%

图表 3-2-2

## 4 数据录制环境

### 4.1 录制环境

专业的录音棚进行录制，环境底噪小于 35dB SPL，混响时间小于 150ms。

### 4.2 录制设备

录制设备为高保真。

### 4.3 录制方法

发音人距离设备 20~30 厘米，正常音量、语速自然朗读录音文本。语音振幅不高于-3dBFS，不低于-9dBFS。无截幅和喷麦。

## 5 语音数据转写

数据转写人员根据所听到的音频写出内容，力求使文本内容与音频发音内容保持一致。

一般准则如下：

- 1) 转写的内容必须和听到的语音完全一致，不能多字、少字、错字。
- 2) 数字要转写为汉字形式，如“一二三”，而不是“123”。注意区分“一”和“幺”，“二”和“两”。
- 3) 标注内容的完整性要与实际发音一致，不得删减。
- 4) 拼音以及声调根据实际发音转写，拼音和文本需一致，并考虑变调以及儿化音，如

以下情况：

拼音应与实际发音一致，如“这”分清楚读的是“zhe4”还是“zhei4”等。

上声的变调，33 变调：了解 23；333 变调：展览馆 223、很勇敢 323。

多个上声 3 要根据紧密程度，划分出语义停顿，再根据上述规则进行变调。

“一、不”变调：单念或语句末尾声调不变；去声 4 前变为 2 等。

儿化音在原来音节加上“r”即可，如花儿 -huar。

5) 韵律标注包含韵律词(#1)、韵律短语(#2)、语调短语(#3)以及语调组边界(#4)，音频与语法不一致的情况，以实际音频发音为准；相邻两级之间难以确定的情况，采取低一级标注。

## 6 数据文件目录

### 6.1 目录结构

数据目录树	
<b>数据目录结构</b>   AISHELL-TTS0017 中文多人合成数据库.pdf └─DOC    └─all_wav_list.txt    └─spkrinfo.xlsx    └─content.txt 	<b>(数据库简介)</b> (文本说明文件) (音频列表) (录音人信息) (录音人信息)

└─SPEECHDATA	(数据文件夹)
─P0001	(录音人信息)
P0001W0001.wav	(音频文件)

图表 6-1-1 数据目录结构

## 6.2 命名规则

### 6.2.1 目录命名规则（图表 6-2-1）

<USAGE>/<Gender>/<SpeakerID>/<mood>/<wav\_ID>

e. g. SPEECHDATA/P0001/P0001W0001. wav

目录	内容	备注
<b>USAGE</b>	SPEECHDATA	文件夹名称
<b>SpeakerID</b>	P0001	录音人 ID
<b>wav_ID</b>	P0001W0001. wav	音频文件

图表 6-2-1

### 6.2.2 文件命名规则（图表 6-2-2）

<SPEAKER\_ID><WAV\_NUM>

e. g. P0001W0001. wav

文件	内容	备注
<b>SPEAKER_ID</b>	P0001	录音人 ID
<b>WAV_NUM</b>	W0001.wav	WAV 编号

图表 6-2-2

## 7 版权声明

本文内容禁止转载，AISHELL (北京希尔贝壳科技有限公司) 对本文拥有修改权、更新权及最终解释权。



**Copyright**



**Copyright**