

# 口音对话语音数据库

AISHELL-ASR1050



希尔贝壳  
产品说明书  
A I S H E L L

Copyright

## 目录

1 产品概述.....	2
2 数据录制环境.....	2
2.1 录制环境.....	2
2.2 录制设备.....	2
2.3 录制方法.....	2
3 语音数据转写.....	2
4 数据文件目录.....	3
5 版权声明.....	3



Copyright

# 1 产品概述

AISHELL-ASR1050 口音对话语音数据库共 339 小时，录音语言为中文，语料内容为物流、社区服务相关。录制过程为安静室内环境，录音人通过手机，以自然对话形式进行录制。此数据库可用于声纹识别、语音识别等研究。

## 2 数据录制环境

### 2.1 录制环境

安静室内，无回音，无其他人说话。

### 2.2 录制设备

市场上常见的 Android 系统手机，如华为、oppo、vivo、小米等。数据存储格式为 8kHz 或 16bit wav 文件。

### 2.3 录制方法

录音人在安静室内通过手机自然通话录制。

## 3 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) 音频中有英文发音的应写成相应的汉字或英文。具体分为以下几种情况：

网址中包含的所有的字母或单词，均为大写。例如：发音内容为“www.abc.com”，应转写为“三 W 点 A B C 点 com”。

发音中包含的英文单词，转写时全部为小写。

发音中包含的英文字母，转写时全部为大写。

对于一些大写专有名词，或者一些英文缩写全部大写加空格，例如：CEO、CCTV 等。

4) 标注内容的完整性要与实际发音一致，不得删减。

5) 句中包含方言字词，不改变句意，用同音字代替，字前加<dialect>标记。

6) 如果两个人说话有重叠，只做主说话人内容转写，并同时加入<overlap>标记。无论是连续性重叠还是多字间断性重叠，都在每个重叠字处标记<overlap>。如下所示：

例 1: 甲说：“今天的天气好热呀！”话还没完，乙说：“嗯。”，“嗯”字正好跟“热”字重叠了，且两个人声音大小差不多，则转写为“今天的天气好热<overlap>呀！”。

例 2: 甲说：“我觉得是这样的，嗯。”，乙说：“我也觉得是。”，乙在甲没有说完“嗯”的时候说的话，覆盖了甲的“嗯”，则转写为“我觉得是这样的。”和“<overlap>我也觉得是。”。

## 4 数据文件目录

数据目录树	
<b>数据目录结构</b>	
AISHELL-ASR1050.pdf	(数据库简介)
└─TextGrid	(文本说明文件夹)
─00001.TextGrid	(标注文本)
└─SPEECHDATA	(数据文件夹)
─00001.wav	(音频文件)

## 5 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