

ATR地域別中国語音声データベース

概要

本データベースは、(株)国際電気通信基礎技術研究所(ATR)音声言語コミュニケーション研究所における「大規模コーパスベース音声対話翻訳技術」の研究開発のために制作された音声認識用中国語音声データベースです。中国の4地域（北京、広東、上海、台湾）の男女比と年齢分布をバランスさせた母語話者による中国語（普通話）発話（注）を収録しています。

（注）各地域の方言発話ではありませんので、ご注意ください。

特徴

- ◆ 中国4地域（北京/広東/上海/台湾）の母語話者による音声を収録。
- ◆ 読み上げテキストと、話者毎の音声ファイルにより音声認識開発（モデル学習/評価）に利用できます。
- ◆ 各地域共、男女比1:1で、年代グループ別に話者数をバランスさせています。

話者/発話文リスト/発話量

- ◆ 話者数：400名（男女比1：1）

地域	人数 (男性/女性)	1話者の発話数
北京	100 (50/50)	約405
広東		約445
上海		約420
台湾		約420
合計	400人	177,596

- ◆ 年代別話者数内訳

年代グループ	話者数	(男/女)
15-20 歳	102	(49/53)
21-30 歳	101	(51/50)
31-40 歳	98	(50/48)
41-50 歳	99	(50/49)
計	400	(200/200)

- ◆ 発話文リスト

<読み上げ>

- ・ 旅行会話文（約150文）
- ・ 音素バランス文（110～150文）
- ・ 日本の固有名詞（約150単語）
- ・ 模擬対話書き起こし文（約9文）

<自然発話>

- ・ 旅行体験談の自由発話（1～2文）

- ◆ 発話時間長

- ・ 読み上げ発話：約200時間
(発話前後の無音区間含む)
- ・ 自然発話：約40時間

- ◆ 発話数

- ・ 読み上げ発話数：177,036（平均：442/話者）
- ・ 自然発話数（総数）：約560（平均：1.4/話者）

収録条件

- ◆ 収録環境：残響の少ない静音環境
- ◆ 収録機材：
 - ・マイクロフォン：接話型ヘッドセット（Sennheiser HMD-460）
 - ・A/D converter：Roland UA-5
 - ・Recording Media: HardDisk
- ◆ サンプリング周波数：16 kHz（48kHzも提供可）
- ◆ 量子化ビット数：16 bit
- ◆ ファイル形式：Wav形式
- ◆ 無音区間：発話の前後に約500msの無音区間あり

データ構成

- ◆ 発話音声（16kHz：モノラル、48kHz：ステレオ）
- ◆ 読み上げテキスト
 - ・ 文字コード GB2312（簡体字中国語）
 - ・ 改行コード 0x0A(LF)
 - ・ 発音(ピンイン)表記含む

納品物

- ◆ 構成：データ構成物、マニュアル
- ◆ 納品メディア：(16kHz) DVD 約27GB
(48kHz) BD/HDD 約160GB

価格（消費税別）

商用利用：¥5,000,000 研究利用：¥2,000,000

※ アカデミック価格についてはお問合せ下さい。

※ 特定地域のみご利用の場合はご相談下さい。

ご利用方法

株式会社ATR-Promotionsより用途に応じた条件によるライセンス（利用許諾）をいたします。
ホームページまたは下記よりご連絡ください。

- 本データベースは、(株)国際電気通信基礎技術研究所(ATR)音声言語コミュニケーション研究所の研究成果を、(株)ATR-Promotionsが改変し製品化したものです。
- 本データベースの知的財産権は(株)国際電気通信基礎技術研究所および(株)ATR-Promotionsに帰属します。
- 本データベースの仕様/価格は予告なしに変更されることがあります。

株式会社 **ATR-Promotions**
<https://www.atr-p.com>

〒619-0288 京都府相楽郡精華町光台2-2-2
TEL:0774-95-1300 FAX:0774-95-1191
E-mail:voice@atr-p.com