

ATR音声データベース多言語評価セット

日本語/英語/中国語/インドネシア語/タイ語/台湾華語/ヒンディ語

概要

本データベースは、ATR音声言語コミュニケーション研究所における研究開発のために制作された音声認識評価用の音声データベースです。旅行分野の読み上げ原稿を、各言語で男女約40名が読み上げています。

言語は、日本語・英語・中国語・インドネシア語・タイ語・台湾華語・ヒンディ語の計7言語です。

特徴

- ◆ 母語話者による読み上げ音声を収録
- ◆ 主に音声認識システムの評価セットとして利用可能
- ◆ 読み上げ原稿は各言語で対訳として利用可能(但し、ヒンディ語は除く)

話者数／発話文リスト／発話量

- ◆ 言語毎の話者数・原稿行数・発話数・発話時間

言語名	話者数 (男性/女性)	原稿 行数	話者あたりの発話数	総発話 数	発話 時間 [1]
日本語	40 (20/20)	1,524 [2]	304～305行/話者 A～Eの5セットのうちいずれかを読み上げ	12,191 [7]	11.0
英語	40 (20/20)	1,524 [2]	304～305行/話者 A～Eの5セットのうちいずれかを読み上げ	12,192	9.6
中国語	40 (20/20)	1,716 [3] [4]	418～431行/話者 A～Eの5セットのうちいずれかを読み上げ[3]	16,990 [7]	15.2
インドネシア語	42 (20/22)	510 [5]	510行/話者 全員同じ原稿を読み上げ	21,420	23.4
タイ語	42 (21/21)	510 [5]	510行/話者 全員同じ原稿を読み上げ (男女で原稿が異なる)	21,420	23.6
台湾華語	48 (24/24)	510 [5]	510行/話者 全員同じ原稿を読み上げ	24,480	21.9
ヒンディ語	51 (28/23)	13,994 [6]	250行～316行/話者 各話者異なる原稿を読み上げ	13,994	14.0

[1] 発話前後の無音区間を含みます。

[2] 日本語・英語の読み上げ原稿は、A～Eセットがあり、合計1,524行あります。

[3] 中国語のA～Eセットは、日本語・英語のA～Eセットの対訳に、中国語独自の読み上げ原稿102行が追加されたものです。

[4] 中国語は、日本語・英語で1行の原稿が複数行に分かれる場合があるため、日本語・英語と比べ、原稿行数が90行増えます。

[5] インドネシア語・タイ語・台湾華語の読み上げ原稿510行は、日本語読み上げ原稿1,524行に含まれる510行についての対訳です。

[6] ヒンディ語の読み上げ原稿は、男女による表現の違いが1,222行、地域による表現の違いが22行含まれています。

[7] 読み上げるべき原稿に対し、日本語は1名(1行)、中国語は2名(各1行)の発話が不足しています。

収録条件

- ◆ 収録環境：残響の少ない静音環境
- ◆ 収録機材(日本語・英語の収録機材)
 - ・ マイクロフォン：接話型ヘッドセット(Sennheiser HMD410)
 - ・ A/D converter：TASCAM TM1000
 - ・ Recording Media：HardDisk
- ◆ サンプル周波数：日本語・英語・中国語は16kHz(その他は48kHz)
- ◆ 量子化ビット数：16bit
- ◆ 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：WAV形式、16kHz、16bit、モノラル
インドネシア語・タイ語・台湾華語・ヒンディー語は48kHz(ステレオ)も提供可
- ◆ 読み上げ原稿：文字コード UTF-8(BOM無し)、改行コード LF
日本語・英語・台湾華語：TSV(タブ区切り)形式
中国語・インドネシア語・タイ語・ヒンディー語
※ ヒンディー語は、デーヴァナーガリー表記と発音表記の2種類有り
ジェンダータグ(男性/女性表現)と換言タグ(英語/ヒンディー語表現)付きで記載
- ◆ 話者情報：PDF形式
日本語 — 性別、年代、出身地、職業、5年以上滞在した都道府県 など
英語 — 性別、年代、出身地、職業、5年以上滞在した州 など
中国語 — 性別、年代、出身地 など
ヒンディー語 — 性別、年代 など
※ インドネシア語・タイ語・台湾華語は、話者情報がありません。

納品物 (1言語あたり)

構成：データ構成物、マニュアル

納品メディア：日本語・英語・中国語 DVD (約1.2~1.6GB)

インドネシア・タイ・台湾華語・ヒンディー語 16kHzの場合は、DVD (約1.6~2.7GB)

48kHzの場合は、BD (約9.3~15.7GB)

価格 (1言語あたり、消費税抜)

商用利用：¥500,000

研究利用：¥300,000

※ アカデミック価格についてはお問合せ下さい。

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。