

# ATR多言語音声データベース

## カナダフランス語/米国スペイン語/ドイツ語/フランス語

### 概要

本データベースは、各言語地域在住の母語話者による音声認識開発/評価用の読み上げ音声データベースです。(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音声翻訳研究のために作成されました。カナダフランス語、米国スペイン語、ドイツ語、フランス語があります。特に、カナダフランス語(おもにカナダケベック州在住話者)、米国スペイン語(スペイン語を母語とする米国在住話者)は、多文化共生社会の多様な言語状況に対応する音声データベースといえます。

- ・フランス語は英語とともにカナダの公用語
- ・米国のスペイン語母語話者人口は4100万人! 米国人口の13%に達する(2015年)<sup>[1]</sup>

[1] Wikipedia「Spanish language in the United States」より

### 特徴

- ◆ カナダフランス語、米国スペイン語などの他に例の少ない**多様な言語バリエーション**
- ◆ 出身地域、男女比と年齢分布をバランスさせた約**200名の話者**による**現地(話者居住地)**での収録音声
- ◆ 書き起こしテキストから作成したSAMPA<sup>[2]</sup>標準による**発音辞書付き**
- ◆ HTK<sup>[3]</sup>による**音声認識ベースライン評価データ**を付属(カナダフランス語、米国スペイン語のみ)

[2] SAMPA : Speech Assessment Methods Phonetic Alphabet

[3] HTK : Hidden Markov Model Tool Kit

### 話者数/発話文リスト/発話量

- ◆ 話者数/出身地域数

| 言語名      | 話者数 (男性/女性)   | 話者出身地域数 |
|----------|---------------|---------|
| カナダフランス語 | 200 (99/101)  | 4       |
| 米国スペイン語  | 200 (98/102)  | 4       |
| ドイツ語     | 209 (102/107) | 4       |
| フランス語    | 206 (102/104) | 6       |

- ◆ 年代別話者数内訳

| 言語名      | 18~20歳 | 21~30歳 | 31~40歳 | 41~50歳 | 51~60歳 | 計   |
|----------|--------|--------|--------|--------|--------|-----|
| カナダフランス語 | 4      | 58     | 55     | 44     | 39     | 200 |
| 米国スペイン語  | 16     | 60     | 58     | 38     | 28     | 200 |
| ドイツ語     | 25     | 109    | 27     | 27     | 21     | 209 |
| フランス語    | 26     | 48     | 42     | 55     | 35     | 206 |

- ◆ 発話文リスト

- ・音素バランス文 : 100/話者
- ・音素バランス単語 : 20/話者
- ・固有名詞(州/市/会社/ストリート) : 4/話者
- ・各種数字(1桁、2桁、電話番号など) : 99/話者
- ・ハンズフリー基本コマンド : 26/話者  
計249発話/話者

- ◆ 発話時間

- ・カナダフランス語 : 51.1時間
- ・米国スペイン語 : 56.7時間
- ・ドイツ語 : 58.5時間
- ・フランス語 : 62.5時間  
(発話前約400ms/発話後約800msの無音区間含む)

- ◆ 総発話数

- ・カナダフランス語 : 49,800
- ・米国スペイン語 : 49,800
- ・ドイツ語 : 52,040
- ・フランス語 : 51,290

## 収録条件 (全言語共通、以下同様)

- ◆ 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-25-1)
  - ・ A/D converter：M-Audio MobilePre USB pre-amplifier
  - ・ Recording Media：HardDisk
- ◆ サンプリング周波数：48kHz
- ◆ 量子化ビット数：16bit
- ◆ ファイル形式：ヘッダ付PCM RAWファイル(WINDOWS WAV形式)

## データ構成

- ◆ オーディオデータ(WAV形式、48kHz、16 bit、モノラル)：収録音声、基準音、背景音
- ◆ テキストデータ(TSV(タブ区切り)形式、文字コード UTF-8(BOM無し)、改行コード CR+LF)：  
書き起こしデータ(タグ付き)、読み上げ原稿、収録セッション情報一覧、収録条件一覧、  
話者情報一覧、音声データ情報一覧、発音辞書 (SAMPA音素表記による)
- ◆ HTKベースライン評価データ(ただしカナダフランス語、米国スペイン語のみ)

## 納品物

- ◆ 構成：データ構成物、マニュアル
- ◆ 納品メディア：DVD、BD又はHDD  
(カナダフランス語：17.6GB、米国スペイン語：19.5GB、ドイツ語：20.3GB、フランス語：21.3GB)

## 価格 (1言語あたり、消費税抜)

商用利用：¥3,000,000

研究利用：¥1,200,000

※ 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より用途に応じたライセンス条件による(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