

ATR地域別英語音声データベース

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「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英語音声データベースです。

英語圏6地域（米国北東部/中西部/西部/南部、イギリス、オーストラリア）の男女比と年齢分布をバランスさせた母語話者による複数アクセントを含む英語発話を収録しています。それぞれに、旅行会話文や音素バランス文などの「読み上げ発話」音声と、話者の旅行体験談による「自然発話」音声があります。

特徴

- ◆ 米国4地域（北東部/中西部/西部/南部）、イギリス、オーストラリアの母語話者による音声を収録。
- ◆ 音声ファイルと、対応した正解（テキスト）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。
- ◆ 年代グループ別に話者数をバランスさせています。（話者の地域別年齢情報も提供可能です）

話者数/発話セット/発話量

- ◆ 話者数：400名（男女比1：1）

地域名		話者数(人)
米国	北東部	50
	中西部	50
	西部	50
	南部	50
イギリス		100
オーストラリア		100
計		400

- ◆ 年代別話者数内訳

年代グループ	話者数（男女別）
15 - 20歳	95 (45/50)
21 - 30歳	113 (55/58)
31 - 40歳	98 (49/49)
41 - 50歳	94 (51/43)
計	400 (200/200)

- ◆ 発話セット

<読み上げ>

- ・旅行会話文（約150文）
- ・音素バランス文（約150文）
- ・日本の固有名詞（約150単語）
- ・模擬対話書き起こし文（10文程度）

<自然発話>

- ・旅行体験談の自由発話

- ◆ 発話時間長

- ・読み上げ発話：約200時間
（平均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：182,611（平均：456/話者）
- ・自然発話数（総数）：約800（2/話者）

収録条件

- ◆ 収録環境：残響のない静音環境
- ◆ 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
- ◆ 正解テキスト
- ◆ 発話プロンプト
- ◆ Transcriptionファイル（書き起こしテキスト）
- ◆ Latticeファイル

納品物

- ◆ 構成：発話音声/テキストデータ、ドキュメント（マニュアル、話者情報リスト）
- ◆ 納入メディア：HDD（予定）

価格（消費税別）

商用利用：¥5,000,000（消費税別） 研究利用：¥2,000,000（消費税別）

※ アカデミック価格についてはお問合せ下さい。

※ 話者グループ、発話内容別にサブセットをご用意しております。詳細はお問合せ下さい。

購入方法

株式会社ATR-Promotionsより直接販売致します。

ホームページまたは下記よりご連絡ください。